

1N-64
193178
76P

Accurate Upwind Methods for the Euler Equations

Hung T. Huynh
Lewis Research Center
Cleveland, Ohio

November 1993

(NASA-TM-106388) ACCURATE UPWIND
METHODS FOR THE EULER EQUATIONS
(NASA) 76 p

N94-16882

Unclass



G3/64 0193178

ACCURATE UPWIND METHODS FOR THE EULER EQUATIONS

Hung T. Huynh

NASA Lewis Research Center
Cleveland, Ohio 44135, USA

Abstract. A new class of piecewise linear methods for the numerical solution of the one-dimensional Euler equations of gas dynamics is presented. These methods are uniformly second-order accurate, and can be considered as extensions of Godunov's scheme. With an appropriate definition of monotonicity preservation for the case of linear convection, it can be shown that they preserve monotonicity. Similar to Van Leer's MUSCL scheme, they consist of two key steps: a reconstruction step followed by an upwind step. For the reconstruction step, a monotonicity constraint that preserves uniform second-order accuracy is introduced. Computational efficiency is enhanced by devising a criterion that detects the 'smooth' part of the data where the constraint is redundant. The concept and coding of the constraint are simplified by the use of the median function. A slope-steepening technique, which has no effect at smooth regions and can resolve a contact discontinuity in four cells, is described. As for the upwind step, existing and new methods are applied in a manner slightly different from those in the literature. These methods are derived by approximating the Euler equations via linearization and diagonalization. At a 'smooth' interface, Harten, Lax, and Van Leer's one-intermediate-state model is employed. A modification for this model that can resolve contact discontinuities is presented. Near a discontinuity, either this modified model or a more accurate one, namely, Roe's flux-difference splitting, is used. The current presentation of Roe's method, via the conceptually simple flux-vector splitting, not only establishes a connection between the two splittings, but also leads to an admissibility correction with no conditional statement, and an efficient approximation to Osher's approximate Riemann solver. These reconstruction and upwind steps result in schemes that are uniformly second-order accurate and economical at smooth regions, and yield high resolution at discontinuities.

1. Introduction. Among methods for the numerical solution of conservation laws in general, and the Euler equations in particular, upwind schemes are very popular due to their accurate shock-capturing capability. Review articles by the pioneers of these methods are readily available (Harten, Lax, and Van Leer 1983, Roe 1986, 1989). Several formulations of upwind schemes, which generally extend that of Godunov (1959), exist [see also Boris and Book (1973)]. Here, we essentially employ the projection-evolution approach advocated by Van Leer (1979, 1985). In this approach, an upwind scheme consists of two key steps: a reconstruction (projection) step where, within each cell, the data are approximated by polynomials; and an upwind (evolution or approximate Riemann solver) step where the average fluxes at each interface are evaluated by a procedure that takes into account the ‘wind’ (or ‘wave’) directions. The Godunov scheme employs the simplest reconstruction—the piecewise constant function—and the most accurate upwind procedure—the exact solution of the Riemann problem. Although this method is only first-order accurate and computationally costly, it is the basis of most modern upwind schemes. A great deal of effort has been spent to enhance accuracy by using higher-order polynomials in the reconstruction step, and to improve efficiency by approximating the solution of the Riemann problem in the upwind step.

First, we discuss the reconstruction step. For simplicity, we consider only linear reconstructions here. Accurate interpolations (linear and higher order) are derived by assuming that the data are smooth. If the data have a discontinuity, these interpolations always create oscillations in the solutions. To prevent such oscillations, Van Leer (1974, 1977, 1979) introduced a monotonicity constraint for his MUSCL scheme: the reconstruction must preserve the monotonicity of the data. At a smooth region with nonzero slope, the constraint is redundant provided the mesh is fine enough, and the MUSCL scheme is second-order accurate. At a discontinuity, the constraint takes effect to prevent oscillations. A major drawback of the constraint is that it also takes effect adjacent to an extremum and causes accuracy to degenerate to first order. In other words, high resolution at discontinuities is obtained at the expense of accuracy near extrema. The popular total variation diminishing (TVD) schemes (Harten 1983, Sweby 1984, Roe 1985) share the same advantages as well as disadvantages since they satisfy Van Leer’s monotonicity constraint.

Harten and Osher (1987) solved the loss of accuracy problem by introducing the nonoscillatory piecewise linear reconstruction, which involves no constraint. Their UNO scheme is uniformly second-order accurate and does not create oscillations, but it is diffusive, e.g., it smears contact discontinuities. The reason for this smearing effect is that, roughly speaking, the slope of the UNO reconstruction is closest to zero among all the slopes that are second-order accurate ($O(h^2)$); therefore, its error is one of the largest in this class (Huynh 1993). Modifications of the UNO and ENO (Harten et al. 1987) schemes to improve the resolution at contact discontinuities can be found in Harten (1989), Yang (1990), and Mao (1991). These modifications, however, are somewhat complicated. For more developments of these schemes, see Shu and Osher (1989), Rogerson and Meiburg (1990), and Shu (1990). For a technique that senses an extremum and turns off the constraint there, see Leonard and Niknafs (1991).

In Huynh (1989, 1993), the author presented a geometric framework in which Van Leer's constraint can be combined with Harten and Osher's nonoscillatory interpolation. For scalar conservation laws, the resulting SONIC schemes are uniformly second-order accurate and nonoscillatory. In addition, each TVD scheme corresponds to a SONIC counterpart, and the UNO scheme is the most diffusive member of the SONIC class.

In this paper, the author's approach is extended to the system of Euler equations. Similar to the UNO scheme, we employ a five-point stencil for the reconstruction. While the UNO slope is 'closest' to zero, we start with the most accurate one for this stencil: the slope of the quartic through these five points. Next, since monotonicity constraints are derived for the convection equation, and characteristic variables are the quantities that are being convected, the constraints are applied to these variables. Our constraint consists of two limits (bounds): the lower one, which assures uniform second-order accuracy, is defined by a slope similar to that of the UNO scheme; the upper one, which controls oscillations, is derived by making use of the upper limit for the MUSCL scheme. The constraint requires the final slope to lie between these two limits. This requirement is conveniently enforced by using the median function: the final slope is the median of the quartic slope and the above two limits. As such, the algorithm is costlier than that of the UNO scheme. To save computing time, we present a simple criterion that detects the 'smooth' part of the data: if a cell is in the

‘smooth’ region, then the monotonicity constraint has no effect, and the slope reduces to the quartic formula for the primitive variables. With this criterion, the characteristic variables and the constraints are necessary only for about six to eight cells around each discontinuity. The resulting algorithm for the reconstruction step is therefore accurate and efficient; however, it still smears a contact discontinuity, although to a lesser extent than the MUSCL or UNO schemes. To improve the resolution in this situation, we present a simple slope-steepening technique that has no effect at the smooth part of the data and generally can resolve a contact discontinuity in four cells or less.

Once the slopes (space partial derivatives) are defined, the time partial derivatives can be calculated by the non-conservation form of the Euler equations (Harten et al. 1987, Roe 1989). At each interface, we can march half a time step via Taylor series expansions, but the two adjacent cells yield two sets of results (states); we need to derive one set of fluxes from these two states. This is accomplished by one of the upwind procedures discussed below. Note that this half-step calculation is a simplification of an observation by Hancock (Van Albada, Van Leer, and Roberts 1982). It helps simplify our algorithm at the ‘smooth’ part of the data (below); in addition, it makes the treatment of source terms straightforward.

While the reconstruction step is mainly numerics, the upwind step is where physics plays its role. There are several upwind procedures in the literature (Roe 1981, Steger and Warming 1981, Osher and Solomon 1982, Van Leer 1982, Harten, Lax, and Van Leer 1983, Roe and Pyke 1984, Roe 1986, Einfeldt 1988, Toro 1991, Liou and Steffen 1993). The most popular method is perhaps that of Roe, but the perfect one is yet to be found (Roe 1989, Roberts 1990). Upwind procedures are generally perceived as difficult to formulate and implement. In this paper, existing and new methods employed are derived from a simple and unified point of view: the Euler equations in conservation form are approximated by linearization and diagonalization. For the flux-vector splitting introduced below, what is crucial is that the fixed state for the linearization is chosen so that the resulting fluxes depend continuously on the data, and the left and right flux vectors are diagonalized by the same matrices, i.e., those formed by the eigenvectors of the fixed state.

Since our reconstruction step already distinguishes between the ‘smooth’ and ‘non-smooth’ parts, a different upwind flux can be applied to each part. At a ‘smooth’

interface, the left and right states are close to each other, and a simple model is sufficient; in fact, a sophisticated one does not significantly improve the results. Here, we make use of Harten, Lax, and Van Leer's one-intermediate-state model, which obtains the flux via a scalar calculation. Combining with the quartic formula, at a smooth region, the computing effort is therefore only slightly more than that of a central-difference scheme with artificial viscosity (Jameson, Schmidt, and Turkel 1981). Using information from the reconstruction step, we also present a modification for the one-intermediate-state model that can resolve contact discontinuities. Near a discontinuity, either this modified model or a more accurate one, namely, Roe's flux-difference splitting (1981), is employed. Our presentation of Roe's method as a flux-vector splitting (Steger and Warming 1981) has the advantage of conceptual simplicity. It also establishes a connection between the two splittings. This connection, in turn, facilitates the derivation of an admissibility (or entropy) correction with no conditional statement for the flux-difference splitting. In addition, it leads to an efficient approximation to Osher's approximate Riemann solver (Osher and Solomon 1982), which is named the simple-wave upwind flux.

At smooth regions—and generally, most of the flow field is smooth—the scheme is accurate and economical due to the quartic formula and the one-intermediate-state model. Near a discontinuity, the monotonicity constraint takes effect to prevent oscillations without causing a loss of accuracy. Moreover, the upwind step for this case, via either Roe's flux-difference splitting with admissibility condition or the simple-wave upwind flux, yields high resolution at shocks as well as contact discontinuities, and non-physical solutions are also excluded. The resulting schemes are efficient and, as will be shown by numerical experiments, their accuracy compares favorably with higher-order schemes.

This paper is essentially self-contained and is organized as follows. In §2, we review the various forms of the Euler equations and the corresponding characteristic variables. The discretization and the half time step is carried out in §3. The two main sections 4 and 5 are independent of each other: §4 deals with the reconstruction step; §5, the upwind step. The algorithm is extended to the case of flows with area variations in §6. Numerical experiments are shown in §7. Finally, the discussion and conclusions are presented in §8.

2. Forms of the Euler equations of gas dynamics. The Euler equations are derived from the physical principles of conservation of mass, momentum, and energy. In this form, called the conservation form, the equations are valid whether the flow is smooth or has discontinuities. To assure correct shock speed (Lax 1954), therefore, we solve the equations in conservation form. The problem thus reduces to a flux estimate at each interface. To calculate this flux, we make use of the systems of equations in non-conservation form for both the primitive and conservative variables. The characteristic variables associated with each of these two systems also play an essential role: upwind algorithms are generally derived for a scalar convection equation; for systems, the quantities that are being convected are the characteristic variables. We review the various forms of the Euler equations in this section [see also Hirsch (1990)]. Different approaches that make use of these forms can be found in Roe and Pyke (1984), Roe (1986, 1989), and Toro (1991).

2.1. Conservation form. The one-dimensional flow of an inviscid and compressible gas obeys the conservation laws for mass, momentum, and energy:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} = 0, \quad (2.1a)$$

$$\mathbf{U} = \begin{pmatrix} \rho \\ m \\ e \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} m \\ (m^2/\rho) + p \\ (e + p)m/\rho \end{pmatrix}, \quad (2.1b, c)$$

where t is time, x distance, ρ density, m momentum, e total energy per unit volume, and p pressure. Let γ be the ratio of specific heats, then for a perfect gas, the system is completed by the equation of state

$$p = (\gamma - 1)[e - m^2/(2\rho)]. \quad (2.2a)$$

Denote by u the velocity, then $m = \rho u$, and the above yields

$$e = \frac{p}{\gamma - 1} + \frac{1}{2}\rho u^2. \quad (2.2b)$$

The flux vector is more commonly written as

$$\mathbf{F} = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ (e + p)u \end{pmatrix}; \quad (2.2c)$$

the above expression is also convenient for the purpose of coding.

2.2. Primitive variables. With superscript T representing the transpose, let \mathbf{V} be the column vector of primitive variables,

$$\mathbf{V} = (\rho, u, p)^T. \quad (2.3a)$$

At smooth regions of \mathbf{U} , one can derive from (2.1-2) that

$$\frac{\partial \mathbf{V}}{\partial t} + \mathbf{A}_p \frac{\partial \mathbf{V}}{\partial x} = 0, \quad (2.3b)$$

where

$$\mathbf{A}_p = \begin{pmatrix} u & \rho & 0 \\ 0 & u & 1/\rho \\ 0 & \gamma p & u \end{pmatrix}. \quad (2.3c)$$

Equation (2.3b) for the primitive variables is of non-conservation (or non-divergent) form. Let c be the speed of sound,

$$c = \sqrt{\frac{\gamma p}{\rho}}. \quad (2.4)$$

Then the eigenvalues of \mathbf{A}_p are

$$\lambda^{(1)} = u - c, \quad \lambda^{(2)} = u, \quad \lambda^{(3)} = u + c. \quad (2.5)$$

One can easily verify that the columns of the following matrix are the right eigenvectors corresponding to the above eigenvalues,

$$\mathbf{R}_p = \begin{pmatrix} 1 & 1 & 1 \\ -c/\rho & 0 & c/\rho \\ c^2 & 0 & c^2 \end{pmatrix}. \quad (2.6a)$$

Now set

$$\mathbf{L}_p = \mathbf{R}_p^{-1}; \quad (2.6b)$$

then

$$\mathbf{L}_p = \begin{pmatrix} 0 & -\rho/(2c) & 1/(2c^2) \\ 1 & 0 & -1/c^2 \\ 0 & \rho/(2c) & 1/(2c^2) \end{pmatrix}, \quad (2.6c)$$

and the rows of \mathbf{L}_p are the left eigenvectors of \mathbf{A}_p . It follows from the definitions of \mathbf{R}_p and \mathbf{L}_p that

$$\mathbf{L}_p \mathbf{A}_p \mathbf{R}_p = \mathbf{\Lambda}, \quad \mathbf{A}_p = \mathbf{R}_p \mathbf{\Lambda} \mathbf{L}_p, \quad (2.7a, b)$$

where

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda^{(1)} & 0 & 0 \\ 0 & \lambda^{(2)} & 0 \\ 0 & 0 & \lambda^{(3)} \end{pmatrix}. \quad (2.7c)$$

Note that we do not need the subscript p for $\mathbf{\Lambda}$ because, as shown below, this matrix is identical to that of the conservative variables. Also observe that the matrices \mathbf{L}_p and \mathbf{R}_p are functions of ρ and c only (or ρ and p); they do not depend on u .

Next we define the characteristic variables associated with \mathbf{V} . Let $\hat{\mathbf{V}} = (\hat{\rho}, \hat{u}, \hat{p})^T$ be a fixed state. Let $\hat{\mathbf{A}}_p$, $\hat{\mathbf{L}}_p$, $\hat{\mathbf{R}}_p$, and $\hat{\mathbf{\Lambda}}$ be the corresponding matrices given by (2.3c), (2.6c,a), and (2.7c). By linearizing (2.3b) at $\hat{\mathbf{V}}$,

$$\frac{\partial \mathbf{V}}{\partial t} + \hat{\mathbf{A}}_p \frac{\partial \mathbf{V}}{\partial x} = 0. \quad (2.8a)$$

Since $\hat{\mathbf{A}}$ is a constant matrix, the above equation is linear; it is an approximation to the nonlinear equation (2.3b) in a neighborhood of $\hat{\mathbf{V}}$ provided that the solution \mathbf{V} is smooth. By (2.7b),

$$\hat{\mathbf{A}}_p = \hat{\mathbf{R}}_p \hat{\mathbf{\Lambda}} \hat{\mathbf{L}}_p.$$

In addition, because $\hat{\mathbf{L}}_p$ is a constant matrix, and it is the inverse of $\hat{\mathbf{R}}_p$, equation (2.8a) is equivalent to

$$\frac{\partial \hat{\mathbf{L}}_p \mathbf{V}}{\partial t} + \hat{\mathbf{\Lambda}} \frac{\partial \hat{\mathbf{L}}_p \mathbf{V}}{\partial x} = 0. \quad (2.8b)$$

We can now define the characteristic variables. Set

$$\mathbf{W}_p = \hat{\mathbf{L}}_p \mathbf{V}; \quad (2.9a)$$

then it follows from (2.6c) that

$$\mathbf{W}_p = \begin{pmatrix} -[\hat{\rho}/(2\hat{c})] u + [1/(2\hat{c}^2)] p \\ \rho - [1/\hat{c}^2] p \\ [\hat{\rho}/(2\hat{c})] u + [1/(2\hat{c}^2)] p \end{pmatrix}, \quad (2.9b)$$

and (2.8b) takes the form

$$\frac{\partial \mathbf{W}_p}{\partial t} + \hat{\mathbf{\Lambda}} \frac{\partial \mathbf{W}_p}{\partial x} = 0. \quad (2.9c)$$

Because $\hat{\mathbf{\Lambda}}$ is a diagonal matrix, the above system is decoupled, and each component corresponds to a convection equation of speed $\hat{u} - \hat{c}$, \hat{u} , and $\hat{u} + \hat{c}$, respectively. For this reason, the components of \mathbf{W}_p are called characteristic variables. Note that all

components of \mathbf{W}_p have units of density; as a result, a comparison among them makes sense. The second component corresponding to the speed \hat{u} is called the linear characteristic; the other two are the nonlinear ones. The above characteristic variables are simpler than those defined from the conservative variables below.

2.3. Conservative variables. At smooth regions of \mathbf{U} , equation (2.1) can be written in non-conservation form via the chain rule:

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{A}_c \frac{\partial \mathbf{U}}{\partial x} = 0, \quad (2.10a)$$

where \mathbf{A}_c is the Jacobian matrix,

$$\mathbf{A}_c = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \left(\frac{\partial \mathbf{F}}{\partial \rho}, \frac{\partial \mathbf{F}}{\partial m}, \frac{\partial \mathbf{F}}{\partial e} \right). \quad (2.10b)$$

A straightforward calculation using (2.2a) yields

$$\mathbf{A}_c = \begin{pmatrix} 0 & 1 & 0 \\ (\gamma - 3)u^2/2 & (3 - \gamma)u & \gamma - 1 \\ (\gamma - 1)u^3 - \gamma ue/\rho & -3(\gamma - 1)u^2/2 + \gamma e/\rho & \gamma u \end{pmatrix}. \quad (2.10c)$$

To diagonalize \mathbf{A}_c , we make use of the transformation from the primitive variables to the conservative ones (Warming, Beam, and Hyett 1975). This transformation will also play a key role in the relations between the flux-vector and flux-difference splittings later. Denote the Jacobian matrix of this transformation by \mathbf{M} :

$$\mathbf{M} = \frac{\partial \mathbf{U}}{\partial \mathbf{V}} = \left(\frac{\partial \mathbf{U}}{\partial \rho}, \frac{\partial \mathbf{U}}{\partial u}, \frac{\partial \mathbf{U}}{\partial p} \right). \quad (2.11a)$$

Then it follows from the definitions of \mathbf{U} and \mathbf{V} and (2.2b) that

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 0 \\ u & \rho & 0 \\ u^2/2 & \rho u & 1/(\gamma - 1) \end{pmatrix}. \quad (2.11b)$$

As a result,

$$\mathbf{M}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -u/\rho & 1/\rho & 0 \\ (\gamma - 1)u^2/2 & -(\gamma - 1)u & \gamma - 1 \end{pmatrix}. \quad (2.11c)$$

Next, we relate \mathbf{A}_c and \mathbf{A}_p by using \mathbf{M} . Apply the chain rule to (2.10a),

$$\mathbf{M} \frac{\partial \mathbf{V}}{\partial t} + \mathbf{A}_c \mathbf{M} \frac{\partial \mathbf{V}}{\partial x} = 0. \quad (2.12)$$

As a consequence of the above and (2.3b),

$$\mathbf{A}_p = \mathbf{M}^{-1} \mathbf{A}_c \mathbf{M}. \quad (2.13)$$

This fact can also be verified by carrying out the right hand side matrix multiplications using (2.10c) and (2.11b,c); the result is identical to (2.3c). Substitute (2.13) into (2.7a),

$$(\mathbf{L}_p \mathbf{M}^{-1}) \mathbf{A}_c (\mathbf{M} \mathbf{R}_p) = \mathbf{A}. \quad (2.14a)$$

Now set

$$\mathbf{L}_c = \mathbf{L}_p \mathbf{M}^{-1}, \quad \mathbf{R}_c = \mathbf{M} \mathbf{R}_p. \quad (2.14b, c)$$

Then

$$\mathbf{L}_c \mathbf{A}_c \mathbf{R}_c = \mathbf{A}; \quad (2.15)$$

the matrices \mathbf{L}_c and \mathbf{R}_c are the inverse of each other; the columns of \mathbf{R}_c are the right eigenvectors of \mathbf{A}_c , and the rows of \mathbf{L}_c , the left. Expressions (2.14c), (2.11b), and (2.6a) yield

$$\mathbf{R}_c = \begin{pmatrix} 1 & 1 & 1 \\ u - c & u & u + c \\ H - uc & u^2/2 & H + uc \end{pmatrix}, \quad (2.16)$$

where H is the total enthalpy,

$$H = \frac{e + p}{\rho} = \frac{u^2}{2} + \frac{\gamma p}{(\gamma - 1)\rho} = \frac{u^2}{2} + \frac{c^2}{\gamma - 1}. \quad (2.17)$$

Similarly, from (2.14b), (2.11c), and (2.6c),

$$\mathbf{L}_c = \begin{pmatrix} M/2 + (\gamma - 1)M^2/4 & [-1 - (\gamma - 1)M]/(2c) & (\gamma - 1)/(2c^2) \\ 1 - (\gamma - 1)M^2/2 & (\gamma - 1)M/c & -(\gamma - 1)/c^2 \\ -M/2 + (\gamma - 1)M^2/4 & [1 - (\gamma - 1)M]/(2c) & (\gamma - 1)/(2c^2) \end{pmatrix}, \quad (2.18)$$

where the Mach number M is defined by

$$M = u/c. \quad (2.19)$$

Later, we will make use of the fact that \mathbf{L}_c and \mathbf{R}_c depend only on any two of the three variables u , H , and M ; they do not depend on ρ . In fact, the matrix \mathbf{A}_c can also be written in a form that only depends on u and H : by (2.17),

$$p = \rho H - e; \quad (2.20)$$

substituting the above into (2.2b),

$$\frac{\gamma e}{\rho} = H + \frac{\gamma - 1}{2} u^2; \quad (2.21)$$

as a result,

$$\mathbf{A}_c = \begin{pmatrix} 0 & 1 & 0 \\ (\gamma - 3)u^2/2 & (3 - \gamma)u & \gamma - 1 \\ (\gamma - 1)u^3/2 - uH & -(\gamma - 1)u^2 + H & \gamma u \end{pmatrix}. \quad (2.22)$$

The derivation and definition of the characteristic variables associated with \mathbf{U} are similar to those for \mathbf{V} in (2.9): with $\hat{\mathbf{L}}_c$ defined by (2.18) at a fixed state $\hat{\mathbf{U}}$, set

$$\mathbf{W}_c = \hat{\mathbf{L}}_c \mathbf{U}; \quad (2.23a)$$

then it follows from (2.10a) and (2.15) that

$$\frac{\partial \mathbf{W}_c}{\partial t} + \hat{\mathbf{A}} \frac{\partial \mathbf{W}_c}{\partial x} = 0. \quad (2.23b)$$

By (2.6c), $\hat{\mathbf{L}}_p$ is sparse; by (2.18) $\hat{\mathbf{L}}_c$ is dense. Consequently, \mathbf{W}_p is employed whenever possible, and \mathbf{W}_c is used only when necessary.

Before moving to the next section, we make the following remark. At smooth regions of the solution, the three forms (2.1), (2.3), and (2.10) are equivalent; at a discontinuity, however, only the conservation form (2.1) is valid. In each cell, therefore, since the solution is assumed to be smooth, we can employ (2.3), i.e., the primitive variables, for the purpose of reconstruction because it is the simplest of the three. Of course the conservative variables (2.10) can be used for this step at the cost of extra computing time. The results for these two reconstructions are essentially identical for the second-order methods presented below. (Note that for higher-order methods, the advantage of employing (2.3) over (2.10) is no longer clear because deriving accurate averages for the primitive variables is more involved.) At an interface, however, there is a discontinuity; consequently, we employ (2.1) to obtain the flux.

3. Discretization. Let $x_j = jh$, $j = 0, 1, 2, \dots$ be a uniform mesh, $x_{j+1/2} = (j + 1/2)h$ be the boundary between the j th and $j + 1$ cells, and $t^n = n\tau$ be the time level; here, h is the mesh size, and τ , the time step. Let \mathbf{U}_j^n be an approximation to the average value of the solution vector \mathbf{U} in the j -th cell at time t^n ,

$$\mathbf{U}_j^n \approx \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{U}(x, t^n) dx, \quad (3.1)$$

where the integration acts on each component of \mathbf{U} . Assume that we know \mathbf{U}_j^n , for all j ; we want to calculate \mathbf{U}_j^{n+1} , where the time step τ satisfies the CFL condition

$$\left[\max_j (|u_j| + c_j) \right] \frac{\tau}{h} \leq 1. \quad (3.2)$$

In this and the next two sections, we only consider the interior points; boundary conditions employed in this paper are standard and will be addressed as needed.

To derive second-order methods, we use the midpoint rule:

$$\frac{\mathbf{U}_j^{n+1} - \mathbf{U}_j^n}{\tau} + \frac{\mathbf{F}_{j+1/2}^{n+1/2} - \mathbf{F}_{j-1/2}^{n+1/2}}{h} = 0, \quad (3.3a)$$

or equivalently,

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n + \frac{\tau}{h} (\mathbf{F}_{j-1/2}^{n+1/2} - \mathbf{F}_{j+1/2}^{n+1/2}). \quad (3.3b)$$

For simplicity of notation, we omit the superscript n when there is no confusion; e.g., \mathbf{U}_j denotes \mathbf{U}_j^n ; the superscripts $n + 1/2$ and $n + 1$, however, are retained.

The problem, therefore, reduces to estimating $\{\mathbf{F}_{j+1/2}^{n+1/2}\}$ from the data $\{\mathbf{U}_j\}$ (the braces represent all j). To obtain these fluxes, we first calculate the primitive variables $\{\mathbf{V}_j\}$ in a straightforward manner by their definitions. Next, in each j th cell, and for $t^n \leq t \leq t^{n+1}$, we approximate $\mathbf{V}(x, t)$ by a linear function $\mathbf{R}_j(x, t)$:

$$\mathbf{R}_j(x, t) = \mathbf{V}_j + (x - x_j)\mathbf{S}_j + (t - t^n)\mathbf{T}_j, \quad (3.4)$$

where \mathbf{S}_j and \mathbf{T}_j respectively approximate $\mathbf{V}_x(x_j, t^n)$ and $\mathbf{V}_t(x_j, t^n)$. Suppose, for the moment, \mathbf{S}_j is known. We march to time $t^{n+1/2}$ as follows. Since \mathbf{V}_j is also known, \mathbf{T}_j can be estimated by using the equations for the primitive variables (2.3):

$$\mathbf{T}_j = -(\mathbf{A}_p)_j \mathbf{S}_j. \quad (3.5)$$

At time $t^{n+1/2}$ and at the two interfaces of the j th cell,

$$\mathbf{R}_j(x_{j-1/2}, t^{n+1/2}) = \mathbf{V}_j - \frac{h}{2}\mathbf{S}_j + \frac{\tau}{2}\mathbf{T}_j, \quad (3.6a)$$

$$\mathbf{R}_j(x_{j+1/2}, t^{n+1/2}) = \mathbf{V}_j + \frac{h}{2}\mathbf{S}_j + \frac{\tau}{2}\mathbf{T}_j. \quad (3.6b)$$

At each interface $j + 1/2$, we now have two values for the primitive variables: one from the Taylor series expansion in the j th cell, namely, $\mathbf{R}_j(x_{j+1/2}, t^{n+1/2})$; and one from

that in the $j + 1$ cell, $\mathbf{R}_{j+1}(x_{j+1/2}, t^{n+1/2})$. From these two values, we will calculate the flux $\mathbf{F}_{j+1/2}^{n+1/2}$ via one of the upwind procedures that employ the conservation form (2.1). It is only here that the ‘wind’ (or ‘wave’) directions come into play.

As a result of the above arguments, the problem reduces to a slope estimate and an upwind procedure for the flux. We remark that these two steps are independent of each other.

Note that the calculation of the time partial derivative from the spatial one via (3.5) was employed in the derivation of the Lax-Wendroff (1964) scheme. In the context of reconstruction, this procedure for the conservative variables was observed by Hancock (Van Albada, Van Leer, and Roberts 1982). It can be traced back to the Cauchy-Kowalewsky theorem (Courant and Hilbert 1962, Harten et al. 1987).

4. Slope estimates. This section deals only with interpolations and constraints; the ‘wind’ direction does not play any role. When the data are smooth, accurate formulas for the slope are readily available (parabolic or quartic). Near a discontinuity, however, such formulas are no longer accurate; in fact, they yield a poor representation of the data and create oscillations in the solutions. By imposing various constraints on the slope, a better representation of the data can be obtained; the resulting schemes are generally also free from oscillations. The constraints should be designed so that at smooth regions of the data, they have no effect; i.e., they do not alter the original slope. The monotonicity constraint presented below consists of two limits (bounds): a lower one, which is ‘closest’ to zero among all slopes that are accurate to $O(h^2)$; and an upper one, which prevents the slope from becoming too steep. The constraint requires that the final slope lies between these two limits.

4.1. Constraints for the slopes. To be precise, we need a few definitions. Let z_1, \dots, z_k be real numbers, and let

$$\alpha = \min(z_1, \dots, z_k), \quad \beta = \max(z_1, \dots, z_k). \quad (4.1)$$

We define $I[z_1, \dots, z_k]$ to be the smallest closed interval containing z_1, \dots, z_k ; i.e.,

$$I[z_1, \dots, z_k] = [\alpha, \beta].$$

Next, let the median of three numbers be the one that lies between the other two. Then it follows that the median lies in the interval defined by any two of the three

arguments; e.g., if x , y , and z are real numbers, then

$$\text{median}(x, y, z) \in I[x, y], \quad \text{median}(x, y, z) \in I[y, z]. \quad (4.2)$$

It is very convenient to enforce a constraint by using the median function. As an example, suppose y_1 and y_2 are the two limits for a quantity y ; i.e., the constraint is that y lies in the interval $I[y_1, y_2]$. Then this constraint is satisfied by y_f (f for final) defined by

$$y_f = \text{median}(y, y_1, y_2). \quad (4.3)$$

Note that if y lies between y_1 and y_2 , then $y_f = y$, and the constraint does not alter the original estimate. If y lies outside the interval $I[y_1, y_2]$, then between the two values y_1 and y_2 , y_f equals the one that is closer to y .

We now make an observation concerning accuracy: if two of the three arguments are accurate to a certain order, then the median yields a value that is accurate to the same order. More precisely, suppose x and y approximate an exact value \bar{y} to $O(h^k)$; then $\text{median}(x, y, z)$ is also accurate to $O(h^k)$ regardless of the value z :

$$\text{median}(x, y, z) = \bar{y} + O(h^k). \quad (4.4)$$

Indeed, since x and y are accurate to $O(h^k)$, so is any value between x and y . By (4.2), $\text{median}(x, y, z)$ lies between x and y ; as a result, it is accurate to the same order.

With z_1, \dots, z_k being real numbers, let $\text{minmod}(z_1, \dots, z_k)$ be the function which returns the smallest argument if all arguments are positive, the largest if they are all negative, and 0 otherwise. This definition is equivalent to

$$\text{minmod}(z_1, \dots, z_k) = \text{median}(\alpha, \beta, 0) \quad (4.5a)$$

where α and β are given by (4.1). For the purpose of coding,

$$\text{minmod}(z_1, \dots, z_k) = s \min(|z_1|, \dots, |z_k|) \quad (4.5b)$$

where

$$s = \frac{1}{2}[\text{sgn}(z_1) + \text{sgn}(z_2)] \frac{1}{2}[\text{sgn}(z_1) + \text{sgn}(z_3)] \dots \frac{1}{2}[\text{sgn}(z_1) + \text{sgn}(z_k)], \quad (4.5c)$$

and $\text{sgn}(z) = 1$ if z is positive; $\text{sgn}(z) = -1$ if z is negative. In the above formula for s , the expression within the square brackets yields the appropriate sign. Note that

there is only one set of square brackets, but there are several sets of absolute value sign. Clearly, if any two arguments are of opposite sign, then one of them must be of opposite sign with z_1 ; consequently, $s = 0$. In general, s equals either -1 , 0 , or 1 . Also observe that if any argument z_i equals 0 , then it does not matter whether $\text{sgn}(z_i)$ is defined as -1 , 0 , or 1 in (4.5c) since the resulting minmod , by (4.5b), is 0 .

With two arguments x and y , $\text{minmod}(x, y)$ is simply the median of x , y , and 0 . In this case, expressions (4.5b,c) reduce to

$$\text{minmod}(x, y) = \frac{1}{2} [\text{sgn}(x) + \text{sgn}(y)] \min(|x|, |y|). \quad (4.6)$$

The minmod function is more commonly defined as

$$\text{minmod}(x, y) = \begin{cases} \text{sgn}(x) \min(|x|, |y|) & \text{if } xy > 0, \\ 0 & \text{otherwise} \end{cases}$$

(e.g., Harten 1983, Sweby 1984, Harten and Osher 1987).

For convenience of programming, we express the median function in terms of minmod :

$$\begin{aligned} \text{median}(x, y, z) &= x + \text{minmod}(y - x, z - x) \\ &= y + \text{minmod}(x - y, z - y) \\ &= x + \frac{1}{2} [\text{sgn}(y - x) + \text{sgn}(z - x)] \min(|y - x|, |z - x|). \end{aligned} \quad (4.7)$$

Let the extended-minmod function $\text{xm}(x, y)$ be defined by

$$\text{xm}(x, y) = \text{median}(x, y, -x - y). \quad (4.8a)$$

(In terms of Sweby's limiter function (1984), $\text{xm}(x, y) = x \text{ median}(1, r, -1 - r)$ where $r = y/x$.) It follows from (4.7) that

$$\text{xm}(x, y) = -(x + y) + \frac{1}{2} [\text{sgn}(2x + y) + \text{sgn}(2y + x)] \min(|2x + y|, |2y + x|). \quad (4.8b)$$

As explained below, the above function is slightly different from the minmod function. If x and y are of the same sign, then the two functions return the same value, namely, the one with smaller modulus. If x and y are of opposite sign, however, the minmod function returns 0 , while the extended-minmod returns the value with smaller modulus provided that $|x/y| > 2$ or $|x/y| < 1/2$ [see also expression (2.27) in Huynh (1993)]. Loosely speaking, the extended minmod function returns the argument of

smaller modulus nearly all the time, and next to a discontinuity, it preserves the slope of the smooth side more often than the minmod function.

We are now ready for the constraint. Since monotonicity constraints are derived for the convection equation (Van Leer 1977), and the characteristic variables are the quantities that are being convected, we interpolate \mathbf{W}_p , the characteristic variables associated with \mathbf{V} defined in (2.9) [see also Roe and Pyke (1984), Roe (1986), Harten and Osher (1987)]. In the rest of this section, we only deal with \mathbf{W}_p ; as a result, for simplicity of notation, we drop the subscript p .

Let j be a fixed index. Let $\hat{\mathbf{V}} = \mathbf{V}_j$ in (2.8); that is, we linearize the equations for the primitive variables at \mathbf{V}_j . Since we use a five-point stencil, for $-2 \leq l \leq 2$, employing (2.9b), set

$$\mathbf{W}_l = \hat{\mathbf{L}} \mathbf{V}_{j+l} = \begin{pmatrix} -[\rho_j/(2c_j)] u_{j+l} + [1/(2c_j^2)] p_{j+l} \\ \rho_{j+l} - [1/c_j^2] p_{j+l} \\ [\rho_j/(2c_j)] u_{j+l} + [1/(2c_j^2)] p_{j+l} \end{pmatrix}; \quad (4.9)$$

again, $\hat{\mathbf{L}}$ is the matrix of left eigenvectors at the state \mathbf{V}_j . Let w_l be any component of \mathbf{W}_l , say, the first component, for $-2 \leq l \leq 2$. We obtain the final slope for this component as follows. Let s_+ and s_- be slopes of the linear interpolations:

$$s_+ = (w_1 - w_0)/h; \quad s_- = (w_0 - w_{-1})/h. \quad (4.10a, b)$$

Let p_- , p_0 , and p_+ be the slopes at $l = 0$ of the three quadratics defined respectively by $\{w_{-2}, w_{-1}, w_0\}$, $\{w_{-1}, w_0, w_1\}$, and $\{w_0, w_1, w_2\}$:

$$p_- = (w_{-2} - 4w_{-1} + 3w_0)/(2h), \quad (4.11a)$$

$$p_0 = (w_1 - w_{-1})/(2h), \quad (4.11b)$$

$$p_+ = (-w_2 + 4w_1 - 3w_0)/(2h). \quad (4.11c)$$

Note that the three slopes above are accurate to $O(h^2)$. Next, between $l = 0$ and $l = 1$, the slope s_+ does not oscillate, but it approximates the exact slope at $l = 0$ only to $O(h)$. To obtain a slope that is accurate to $O(h^2)$ and does not oscillate, we simply use the median function:

$$q_+ = \text{median}(s_+, p_+, p_0), \quad (4.12a)$$

and by symmetry,

$$q_- = \text{median}(s_-, p_-, p_0). \quad (4.12b)$$

For further details, see §3 of Huynh (1993). Between the two slopes above, we take, roughly speaking, the less steep one,

$$q_* = \text{xm}(q_+, q_-). \quad (4.12c)$$

The slope q_* is clearly accurate to $O(h^2)$ because by (4.8a) and (4.2), it lies between q_- and q_+ . It behaves very much like the UNO slope of Harten and Osher (1987): the resulting scheme is uniformly second-order accurate and does not create oscillations. Because it is (loosely speaking) closest to zero among all slopes that are accurate to $O(h^2)$, the scheme shares the property of being diffusive with the first-order upwind method (slope equals zero) but to a much lesser extent. Note that if the extended minmod function in (4.12c) is replaced by the minmod function, then the result is the UNO slope. The extended minmod function, however, has the advantage of preserving the slope of the smooth side more often near a discontinuity. More importantly, while the UNO slope is the final one, the slope (4.12c) is employed only as one of the two bounds (the lower limit) for the monotonicity constraint described below.

Next, we define the other bound (upper limit). The following monotonicity constraint was presented by Van Leer (1974, 1977): for the cell $l = 0$, if the linear reconstruction takes values in $I[w_{-1}, w_0, w_1]$, then it does not create oscillations in the solution. This constraint is very simple; however, it causes a loss of accuracy near extrema. We present a constraint that preserves uniform second-order accuracy below. The argument to the right of $l = 0$ is carried out first. Van Leer's constraint can be restated as follows. The reconstruction to the right of $l = 0$ ($x_j \leq x \leq x_{j+1/2}$) takes values in $I[w_0, w_1]$; that is, the final slope q_f lies between 0 and $2s_+$ (see Fig. 4.1). It is the limit 0 that causes accuracy to degenerate near extrema. To obtain uniform second-order accuracy, we require that the final slope lies between q_* and $2s_+$, where q_* is defined by (4.12c). The requirement to the left of $l = 0$ is that q_f lies between q_* and $2s_-$. The two requirements together result in the following constraint: the final slope lies in the intersection of the two intervals $I[q_*, 2s_-]$ and $I[q_*, 2s_+]$. Clearly, one end of this intersection interval is q_* ; the other is

$$q^* = \text{median}(q_*, 2s_-, 2s_+). \quad (4.13)$$

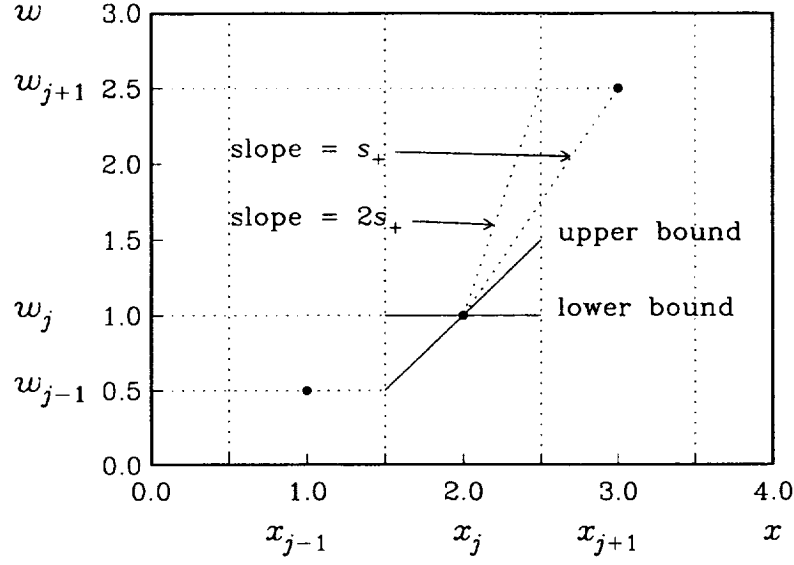


Figure 4.1. *Van Leer's constraint.*

Thus the constraint takes the form

$$q_f \in I[q_*, q^*].$$

Having obtained the two limits, we now define an accurate slope. Using the quartic (five-point) formula, set

$$q_5 = (w_{-2} - 8w_{-1} + 8w_1 - w_2)/(12h). \quad (4.14)$$

The above slope is highly accurate; however, near a discontinuity, it may have the wrong sign (see Fig. 4.2). We avoid this problem by requiring that q_5 lies between p_0 and p_m where

$$p_m = \text{median}(p_-, p_+, p_0). \quad (4.15a)$$

To bring q_5 into the interval $I[p_0, p_m]$, we once again use the median function:

$$q_6 = \text{median}(q_5, p_m, p_0). \quad (4.15b)$$

Observe that at the smooth part of the data, q_6 is generally identical to q_5 ; that is, p_m and p_0 provide plenty of 'room' for an accurate slope. Indeed, without loss of generality, we may assume $x_j = 0$, and the Taylor series expansion of the exact solution w_e takes the form

$$w_e(x) = sx + dx^2 + ex^3 + O(h^4)$$

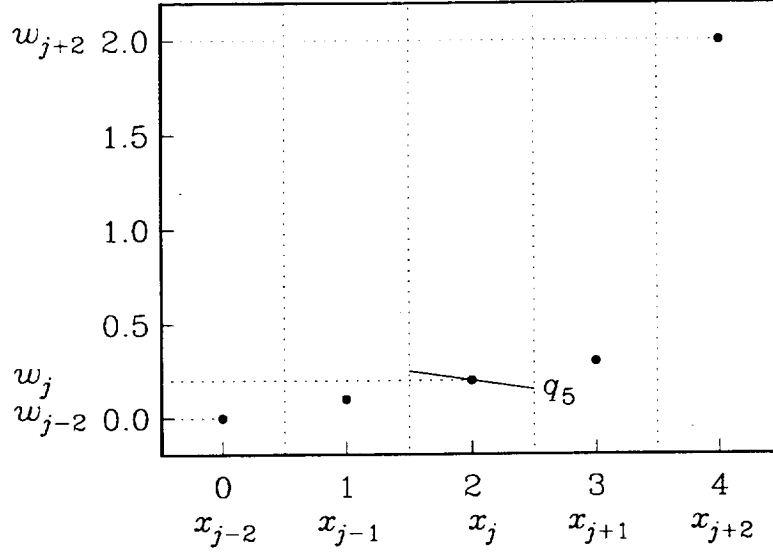


Figure 4.2. *The quartic slope may have the wrong sign.*

[see §3.5 of Huynh (1993)]. Then, by omitting higher-order terms,

$$p_- = s - 2eh^2, \quad p_0 = s + eh^2, \quad p_+ = s - 2eh^2.$$

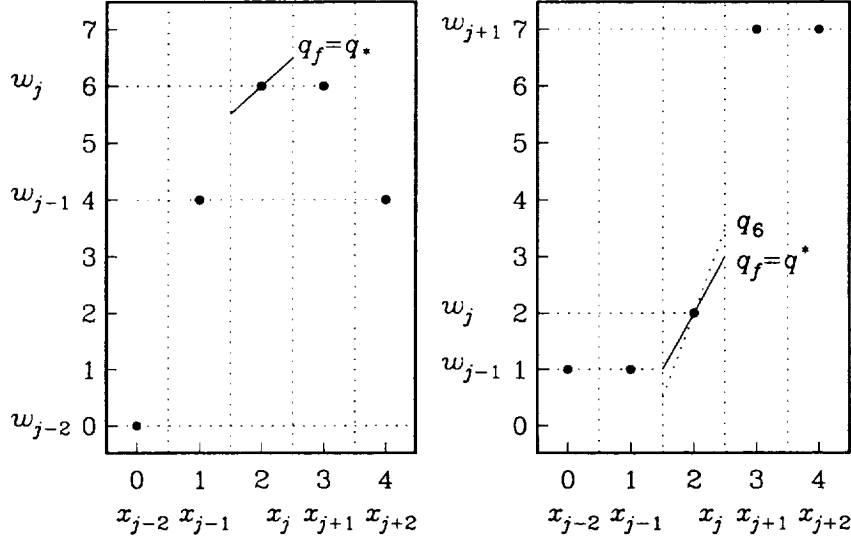
The interval $I[p_0, p_m]$ is therefore equal to $I[s - 2eh^2, s + eh^2]$, which contains the exact slope s , and the above observation follows. (Also note that instead of the requirement that q_5 lies between p_0 and p_m , we can insist that q_5 lies between q_- and q_+ , and thus save one median calculation. This latter constraint, however, does not provide as much ‘room’ as the former, i.e., it alters q_5 more often.)

Finally, using the median function, we limit q_6 ,

$$q_f = \text{median}(q_6, q_*, q^*). \quad (4.16)$$

Since q_* and q_6 are accurate to $O(h^2)$, the above q_f is also accurate to the same order by observation (4.4).

A remark concerning the effect of constraint (4.16) is in order. At the smooth part of the data where the slope is nonzero, expression (4.16) yields $q_f = q_6$ because q_* is closer to 0 than q_6 , and q^* is further from 0. Near an extremum, the interval $I[q_*, q^*]$ may reduce to the point $\{q_*\}$, and in this case, $q_f = q_*$ (see Fig. 4.3(a)). It is here that our monotonicity constraint preserves second-order accuracy while that of Van



(a) *near an extremum* (b) *near a discontinuity*

Figure 4.3. *Effects of the new constraint.*

Leer may not. Near a discontinuity, or where the data change rapidly, the slope q_6 is generally steeper than q^* , and the final slope is identical to q^* , which is either $2s_+$ or $2s_-$ (see Fig. 4.3(b)).

Let the final slopes by the algorithm (4.10–16) for the three components of \mathbf{W}_l be $q_f^{(1)}$, $q_f^{(2)}$, and $q_f^{(3)}$. Let

$$\mathbf{Q}_f = (q_f^{(1)}, q_f^{(2)}, q_f^{(3)})^T. \quad (4.17)$$

The slopes of the primitive variables are obtained by

$$\mathbf{S}_j = \hat{\mathbf{R}} \mathbf{Q}_f = \mathbf{R}_j \mathbf{Q}_f \quad (4.18)$$

where \mathbf{R}_j is the matrix formed by the right eigenvectors given in (2.6a). This expression completes the slope estimates.

Note that (4.15) and (4.16) can be replaced by

$$q_f = q_* + \min\text{mod}(q_6 - q_*, 2s_+ - q_*, 2s_- - q_*). \quad (4.19)$$

We also remark that in the case of linear convection, with an appropriate definition of monotonicity of the data in an interval (Huynh 1993), it can be shown that the

slope defined by (4.10–16) yields a scheme that preserves monotonicity and is uniformly second-order accurate (to appear).

4.2. Smooth regions. The above constraints are constructed so that they have little or no effect at the smooth part of the data. For such a region, the slope reduces to the quartic formula for the primitive variables. More precisely, suppose the constraints (4.15b) and (4.16) have no effect at the j th cell. Then, since $q_f = q_5$, and multiplication by a constant commutes with matrix multiplication, it follows from (4.13) that

$$\mathbf{Q}_f = [\hat{\mathbf{L}}\mathbf{V}_{j-2} - \hat{\mathbf{L}}(8\mathbf{V}_{j-1}) + \hat{\mathbf{L}}(8\mathbf{V}_{j+1}) - \hat{\mathbf{L}}\mathbf{V}_{j+2}]/(12h). \quad (4.20)$$

Because $\hat{\mathbf{R}}$ and $\hat{\mathbf{L}}$ are the inverse of each other, (4.18) and the above yield

$$\mathbf{S}_j = (\mathbf{V}_{j-2} - 8\mathbf{V}_{j-1} + 8\mathbf{V}_{j+1} - \mathbf{V}_{j+2})/(12h), \quad (4.21)$$

which is the quartic formula for the primitive variables.

Next, we present a simple criterion which detects the ‘smooth’ regions where the constraints have no effect, and as shown above, the slope is then given by (4.21). For a wide class of problems—including all the test problems below—a discontinuity in the flow field leads to a discontinuity in the density field. As a result, to separate discontinuities from smooth regions, it suffices to test the density field. (For highly complex problems, we may need to test more than one field.) Let $\Delta_j^2\rho$ be the second difference of ρ :

$$\Delta_j^2\rho = \rho_{j-1} - 2\rho_j + \rho_{j+1}. \quad (4.22)$$

For each index j , let k_j be a flag that is set as follows: if

$$\frac{4}{5} \leq \frac{\Delta_{j-1}^2\rho}{\Delta_j^2\rho} \leq \frac{5}{4}, \quad \text{and} \quad \frac{4}{5} \leq \frac{\Delta_{j+1}^2\rho}{\Delta_j^2\rho} \leq \frac{5}{4}, \quad (4.23a, b)$$

then $k_j = 1$, and the flow is considered to be ‘smooth’ in the j th cell; otherwise, $k_j = 0$. (We remark that in the case of linear convection, the factors $4/5$ and $5/4$ are replaced by $2/3$ and $3/2$, and it can be shown that a constraint slightly more general than (4.16) has no effect on the original slope. This general constraint was employed for cubic interpolation in §3.4 in (Huynh 1993). The factors $4/5$ and $5/4$ are chosen by numerical experiments to account for nonlinearity as well as the fact that only one out

of three fields are tested. In fact, the more relaxed factors $2/3$ and $3/2$ are sufficient for all of our test cases except the interacting-blast-wave problem of Woodward and Colella (1984).)

For the purpose of coding, observe that the following statement is equivalent to (4.23a): $\Delta_{j-1}^2 \rho$ lies between $(4/5)\Delta_j^2 \rho$ and $(5/4)\Delta_j^2 \rho$. That is,

$$\left(\frac{4}{5}\Delta_j^2 \rho - \Delta_{j-1}^2 \rho\right) \left(\frac{5}{4}\Delta_j^2 \rho - \Delta_{j-1}^2 \rho\right) \leq 0.$$

Thus, if we set

$$a_j = \left(\frac{4}{5}\Delta_j^2 \rho - \Delta_{j-1}^2 \rho\right) \left(\frac{5}{4}\Delta_j^2 \rho - \Delta_{j-1}^2 \rho\right), \quad (4.24a)$$

and

$$b_j = \left(\frac{4}{5}\Delta_j^2 \rho - \Delta_{j+1}^2 \rho\right) \left(\frac{5}{4}\Delta_j^2 \rho - \Delta_{j+1}^2 \rho\right), \quad (4.24b)$$

then conditions (4.23) are equivalent to $\max(a_j, b_j) \leq 0$. In practice, we allow for some noise in the data because they are quickly damped out by the scheme. Conditions (4.23) are coded as

$$\max(a_j, b_j) \leq \epsilon \rho_j. \quad (4.24c)$$

where ϵ is a small number, typically, $\epsilon = 10^{-5}$ or 10^{-4} . In this paper, all numerical tests employ $\epsilon = 5 \times 10^{-5}$. The corresponding expression for k_j is

$$k_j = \begin{cases} 1 & \text{if (4.24c) holds,} \\ 0 & \text{otherwise.} \end{cases} \quad (4.25)$$

Roughly speaking, if $k_j = 1$, then the data are ‘smooth’ at x_j , and the slope S_j is given by (4.21). Otherwise, the constraint may take effect, and the estimate for S_j is given by (4.9–18).

Notice that at the ‘smooth’ regions ($k_j = 1$), the quartic formula (4.21) can be replaced by the parabolic one (similar to (4.11b)), which has a three-point stencil; the results are generally slightly more diffusive.

4.3. A slope-steepening technique. The above slope estimate results in a scheme that is effective, uniformly second-order accurate, and does not create oscillations, but it still slightly smears a contact discontinuity, although to a lesser extent than the MUSCL or UNO schemes. The reason for this smearing is that near a discontinuity, the constraints help prevent oscillations by bringing the slope closer to zero

when necessary, but they do not steepen the slope to the maximum allowable value, which keeps the discontinuity from smearing. In the case of a shock, the characteristics converge; the shock remains crisp, and no modification is needed. For a contact discontinuity, however, the characteristics are parallel; there is no self-steepening mechanism, and smearing takes place. To keep contact discontinuities sharp, Colella and Woodward (1984) presented a criterion to detect them, and in that case, a steeper reconstruction is employed for their piecewise parabolic method. For other techniques, see Harten (1989), Yang (1990), and Mao (1991). Our approach is somewhat different: the slope steepening should take effect near a discontinuity but have no effect at the smooth part of the data.

The following simple algorithm can generally resolve a contact discontinuity in four cells or less; it is applied only to the linear characteristic, i.e., the second component of \mathbf{W}_l . For this component, with q_- and q_+ given by (4.12), q_6 by (4.15b), κ a fixed constant which typically equals 5, we reset q_6 to

$$q_6 \leftarrow \text{sgn}(q_6) \max(\kappa|q_+ - q_-|, |q_6|). \quad (4.26)$$

Then we apply (4.16) to obtain the final slope.

To see the effect of the above expression, observe that at the smooth part of the data, since both q_- and q_+ approximate the exact slope at $l = 0$ to $O(h^2)$, the quantity $|q_+ - q_-|$ is a small number of order $O(h^2)$. As a result, except for a neighborhood of length $O(h^2)$ around an extremum where the original q_6 is too close to 0, expression (4.26) does not alter q_6 . (We need not pay attention to this neighborhood because the constraint (4.16) will make the final slope $q_f = q_-$.) Loosely speaking, (4.26) has no effect at the smooth part of the data. Near a discontinuity, on the other hand, q_- and q_+ are far apart, and $\kappa|q_+ - q_-|$ becomes a large number; consequently, the slope q_6 by (4.26) is considerably steeper than the original q_6 .

For readers who are familiar with limiter functions, the following remark explains (4.26) from this point of view (see Fig. 4.4). Strictly speaking, the explanation is incorrect, but it conveys the idea. Consider the case of q_- and q_+ both being positive. Set $\kappa = 5$, $r = q_+/q_-$, then

$$5|q_+ - q_-| = (5|r - 1|)q_-.$$

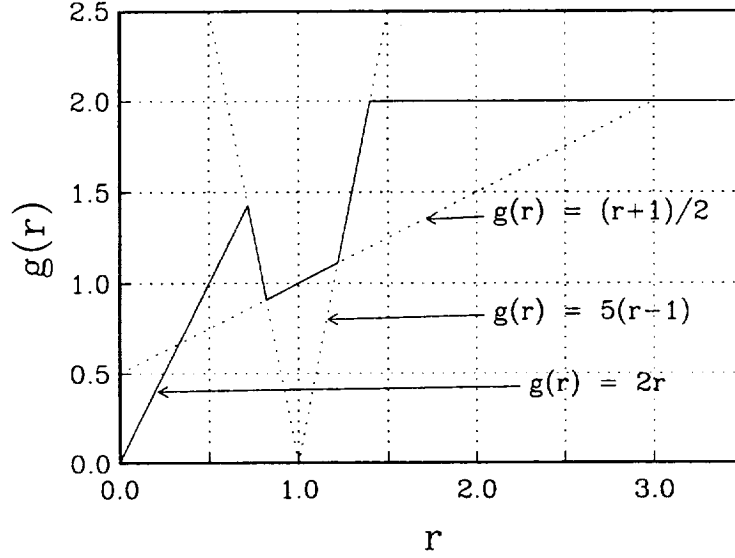


Figure 4.4. *Slope steepening explained via limiter function.*

The lines $g(r) = \pm 5(r - 1)$ intersect $g(r) = (r + 1)/2$ at $r = 9/11$ and $r = 11/9$. The lines $g(r) = 5(r - 1)$ and $g(r) = 2$ intersect at $r = 7/5$; $g(r) = -5(r - 1)$ and $g(r) = 2r$, at $r = 5/7$. Roughly, for r between $9/11$ and $11/9$, expression (4.26) does not alter the original q_6 , which is represented by $g(r) = (r + 1)/2$. But for $5/7 < r < 9/11$ or $11/9 < r < 7/5$, expression (4.26) steepens q_6 , and the result can be even more compressive than the ‘Superbee’ limiter function (Roe 1985). Also note that since both q_- and q_+ are accurate to $O(h^2)$, at most of the smooth regions, r lies between $9/11$ and $11/9$, and (4.26) has no effect. Near a discontinuity, it takes effect and steepens the slope considerably. The final constraint (4.16) then limits the slope to assure that no oscillations are created ($r \leq 5/7$ or $r \geq 7/5$).

Certain contact discontinuities, such as those resulting from a shock interacting with another shock, are extremely difficult to resolve. Even highly compressive methods like Roe’s ‘Superbee’ scheme may smear them. The above technique can be used to capture such discontinuities by setting $\kappa = 10$, or even 20. For such large κ , the slope steepening (4.26) may take effect at the smooth part of the data; consequently, it should be applied sparingly, for example, only when the strength of the linear characteristic dominates that of the other two. This condition is quantified below: for the i th component of

\mathbf{W}_l , $1 \leq i \leq 3$, set

$$\sigma^{(i)} = |w_2^{(i)} - w_0^{(i)}| + |w_0^{(i)} - w_{-2}^{(i)}|. \quad (4.27a)$$

Then the linear characteristic is said to dominate the other two if

$$\sigma^{(2)} > 2(\sigma^{(1)} + \sigma^{(3)}). \quad (4.27b)$$

(The factor 2 is found by numerical experiments; the results generally vary only slightly if a different factor is chosen.) In this case we also reset k_j :

$$k_j = -2. \quad (4.28)$$

As noted at the end of §2.3, all $\sigma^{(i)}$ have units of density; consequently, a comparison among them makes sense.

We summarize the slope estimate below.

Algorithm for the slope. *Given the conservative variables $\{\mathbf{U}_j\}$, calculate the primitive variables $\{\mathbf{V}_j\}$, and $\{\Delta_j^2 \rho\}$ by (4.22). For each index j , compute a_j and b_j by (4.24a, b). Next consider the following two cases.*

- (a) *If (4.24c) is satisfied, the cell is in the ‘smooth’ region; set $k_j = 1$; define \mathbf{S}_j by the quartic formula (4.21); then move on to the next index.*
- (b) *Otherwise, set $k_j = 0$, and for $-2 \leq l \leq 2$, calculate \mathbf{W}_l by (4.9). For each component i , $1 \leq i \leq 3$, obtain $\sigma^{(i)}$ by (4.27a). If (4.27b) holds, i.e., the linear characteristic is dominant, reset $k_j = -2$. Next, for each i , $1 \leq i \leq 3$, calculate $q_f^{(i)}$ by (4.10–16); while estimating $q_f^{(2)}$, however, after obtaining q_6 via (4.15b), if $k_j = -2$, reset q_6 by (4.26) before carrying out (4.16). Having obtained \mathbf{Q}_f , calculate \mathbf{S}_j by (4.18). Then move on to the next index.*

Note that instead of (or in addition to) (4.27b), one can employ a more sophisticated criterion to detect a contact discontinuity without much penalty in computing time (at least, for scalar computers) because generally, the flow field is mostly ‘smooth’, i.e., (4.24c) holds.

For comparison, we next present the slopes of three popular TVD (total variation diminishing) schemes. With s_- and s_+ defined by (4.10), instead of (4.11–16) the minmod scheme (Harten, 1983) is given by

$$q_f = \text{minmod}(s_-, s_+); \quad (4.29)$$

Van Leer's average limiter (1977),

$$q_f = \text{minmod}[(s_- + s_+)/2, 2s_-, 2s_+]; \quad (4.30)$$

and the 'Superbee' scheme (Roe and Baines 1984, Roe 1985),

$$q_f = \text{minmod}[\text{sgn}(s_-) \max(|s_-|, |s_+|), 2s_-, 2s_+]. \quad (4.31)$$

We remark that if s_- and s_+ are of opposite sign, the three expressions above return 0 for the slope. Expression (4.30) can also be written as

$$q_f = \frac{1}{2}[\text{sgn}(s_-) + \text{sgn}(s_+)] \min(|s_- + s_+|/2, 2|s_-|, 2|s_+|);$$

similarly, (4.31) is equivalent to

$$q_f = \frac{1}{2}[\text{sgn}(s_-) + \text{sgn}(s_+)] \min[\max(|s_-|, |s_+|), 2|s_-|, 2|s_+|].$$

Compared with the TVD schemes, which are only first-order accurate near extrema, the slope estimate for our scheme has the disadvantage of being more involved in terms of coding, but it has three distinct advantages: it is both accurate and economical at smooth regions, and it gives high resolution at contact discontinuities.

The author's SONIC-A scheme, which is a uniformly second-order extension of Van Leer's average limiter, is defined by

$$q_f = \text{median}[(q_- + q_+)/2, q_*, q^*].$$

The SONIC-B scheme, which extends 'Superbee', is given by

$$q_f = \text{median}[\text{sgn}(q_*) \max(|q_-|, |q_+|), q_*, q^*].$$

Due to the use of q_- and q_+ instead of s_- and s_+ , the above slopes yield uniformly second-order accurate schemes. Note that at the 'smooth' part of the data, these slopes do not simplify easily; in contrast, the slope presented in the algorithm summarized above does.

5. Upwind flux. With the slope $\{S_j\}$ defined in the previous section, one can calculate the time partial derivative by (3.5) and obtain the interface values at time $t^{n+1/2}$ by (3.6a,b). Let $j + 1/2$ be a fixed interface, and let

$$\mathbf{V}_L = \mathbf{R}_j(x_{j+1/2}, t^{n+1/2}), \quad \mathbf{V}_R = \mathbf{R}_{j+1}(x_{j+1/2}, t^{n+1/2}), \quad (5.1)$$

where the reconstruction values are given in (3.6). From the above left and right states, we derive a flux denoted by \mathbf{F}_U via one of the upwind procedures below.

In this section, time t is employed only for convenience of derivation; the final flux \mathbf{F}_U is independent of time. Let the conservative variables corresponding to \mathbf{V}_L and \mathbf{V}_R be denoted by \mathbf{U}_L and \mathbf{U}_R . Consider the conservation laws (2.1) with the initial condition at $t = 0$:

$$\mathbf{U} = \begin{cases} \mathbf{U}_L & \text{if } x \leq 0, \\ \mathbf{U}_R & \text{otherwise.} \end{cases} \quad (5.2)$$

The above problem is called the Riemann problem. The exact solution, which only depends on \mathbf{U}_L , \mathbf{U}_R , and x/t , is known (e.g., Godunov 1959, Van Leer 1979, Holt 1984, Gottlieb and Groth 1988, Hirsch 1990). It is a constant along each ray (half line) $x = \lambda t$ where the speed λ is fixed and $t \geq 0$. An approximation \mathbf{F}_U for the flux at $x = 0$ (or $\lambda = 0$) and $t \geq 0$, called an upwind flux (or approximate Riemann solver), is sufficient for the purpose of calculation. In fact, these approximations are sometimes called alternative Riemann solvers because they can be more diffusive than the exact solution, and may occasionally produce smoother (more accurate) solutions for (3.3) as in the case of a slowly moving shock wave (Roe 1989). Notice that the coordinates have been shifted so that $(x_{j+1/2}, t^{n+1/2})$ in (5.1) now corresponds to $(x = 0, t = 0)$.

At an interface in the smooth region, \mathbf{V}_L and \mathbf{V}_R are generally different, but close to each other. As a matter of fact, for such an interface, $\mathbf{V}_R - \mathbf{V}_L$ is an $O(h^2)$ quantity; consequently, for simplicity and economy, we may derive \mathbf{F}_U via the non-conservation form (2.3). This method is called the primitive-variable splitting, which can be considered as a simplification of Toro's approximate Riemann solver (1991); it is carried out in §5.1. On the other hand, although \mathbf{V}_L and \mathbf{V}_R are close, they are generally different; in that case, there is a discontinuity in the initial condition. Theoretically, therefore, only the conservation form (2.1) is valid. The simplest model that employs (2.1) is perhaps the one-intermediate-state model by Harten, Lax, and Van Leer (1983), which resolves the two acoustic waves, but ignores the contact discontinuity. This model turns out to be more economical than the primitive-variable splitting; in addition, at smooth regions, it performs as well as Roe's flux-difference splitting with roughly a 35 percent savings in (total) computing time.

Near a discontinuity, $\mathbf{V}_R - \mathbf{V}_L$ is $O(1)$, and the flux also has an error $O(1)$. Consequently, we have to take extra care: if the model is inadequate, this error may

propagate throughout the flow field and cause accuracy to degenerate. For such an interface, we employ a flux-vector splitting derived by linearization and diagonalization. It will be shown that this flux-vector splitting, which is conceptually simple, yields a result identical to Roe's flux-difference splitting, which is computationally simple. Our derivation also leads to an admissibility correction with no conditional statement and an economical approximation to Osher's approximate Riemann solver.

For convenience of notation, with the initial condition (5.2), denote

$$\Delta \mathbf{U} = \mathbf{U}_R - \mathbf{U}_L; \quad (5.3)$$

similarly, $\Delta \mathbf{F} = \mathbf{F}_R - \mathbf{F}_L$, and $\Delta \rho u = \rho_R u_R - \rho_L u_L$, etc. Notice that the operation Δ is linear, i.e., for any functions f , g , and constants α , β ,

$$\Delta(\alpha f + \beta g) = \alpha \Delta f + \beta \Delta g.$$

5.1. Primitive-variable splitting. Among the subsequent upwind fluxes, this is the only one that employs the nonconservation form (2.3). It is perhaps the simplest model where the 'wind' directions play their roles. As mentioned above, it should only be employed at 'smooth' interfaces. More precisely, with k_j defined by (4.25) and (4.28), set

$$k_{j+1/2} = k_j + k_{j+1}. \quad (5.4)$$

Then this flux is used only when $k_{j+1/2} = 2$. For a similar method, which employs (2.3) with a more sophisticated fixed state for the linearization and a slightly more complex calculation involving the left and right states of the contact discontinuity, see Toro (1991).

With the initial condition \mathbf{V}_L and \mathbf{V}_R , we linearize the Euler equations (2.1) at the average state

$$\hat{\mathbf{V}} = \frac{1}{2}(\mathbf{V}_L + \mathbf{V}_R). \quad (5.5)$$

Next, equations (2.8a) for the primitive variables are diagonalized as shown in (2.8b). The characteristic variables corresponding to the initial condition \mathbf{V}_L and \mathbf{V}_R are given by (2.9a),

$$\mathbf{W}_{p,L} = \hat{\mathbf{L}}_p \mathbf{V}_L, \quad \text{and} \quad \mathbf{W}_{p,R} = \hat{\mathbf{L}}_p \mathbf{V}_R. \quad (5.6a, b)$$

By (2.8b), each component of these variables is advected with speed $\hat{\lambda}^{(i)}$. Consequently, the upwind characteristic variable $\mathbf{W}_{p,U}$ is determined by the ‘wind’ direction or the sign of $\hat{\lambda}^{(i)}$:

$$w_{p,U} = \begin{cases} w_{p,L} & \text{if } \hat{\lambda} \geq 0, \\ w_{p,R} & \text{otherwise} \end{cases} \quad (5.6c)$$

where, for convenience, the superscript (i) is understood. In both cases, the above expression is equivalent to

$$\begin{aligned} w_{p,U} &= \frac{1}{2} [1 + \operatorname{sgn}(\hat{\lambda})] w_{p,L} + \frac{1}{2} [1 - \operatorname{sgn}(\hat{\lambda})] w_{p,R} \\ &= \frac{1}{2} [w_{p,L} + w_{p,R}] - \frac{1}{2} \operatorname{sgn}(\hat{\lambda}) [w_{p,R} - w_{p,L}]. \end{aligned}$$

Denote by $\operatorname{sgn}(\hat{\mathbf{A}})$ the diagonal matrix whose diagonal entries are $\operatorname{sgn}(\hat{\lambda}^{(i)})$, $i = 1, 2, 3$; then

$$\mathbf{W}_{p,U} = \frac{1}{2}(\mathbf{W}_{p,L} + \mathbf{W}_{p,R}) - \frac{1}{2} \operatorname{sgn}(\hat{\mathbf{A}})(\mathbf{W}_{p,R} - \mathbf{W}_{p,L}).$$

To obtain the corresponding primitive variable \mathbf{V}_U , multiply the above equation on the left by $\hat{\mathbf{R}}_p$, and by (5.6a,b),

$$\mathbf{V}_U = \frac{1}{2}(\mathbf{V}_L + \mathbf{V}_R) - \frac{1}{2} \hat{\mathbf{R}}_p \operatorname{sgn}(\hat{\mathbf{A}}) \hat{\mathbf{L}}_p(\Delta \mathbf{V}), \quad (5.7)$$

where the operation Δ is defined in (5.3). Finally, the upwind flux is the flux of the above state.

The following remark is in order. If $\hat{\lambda} = 0$, then should we select $w_{p,U} = w_{p,L}$ or $w_{p,U} = w_{p,R}$ in (5.6c)? Since the fluxes resulting from these two quantities are generally not identical, the final upwind flux does not depend continuously on the data. At a smooth interface, this discrepancy is not critical, and the above upwind flux performs well; near a discontinuity, however, it may create oscillations in the solutions. By using the conservation form (2.1), one can derive upwind fluxes that depend continuously on the data; in fact, all of the subsequent methods share this property.

5.2. One-intermediate-state model. This model, introduced by Harten, Lax, and Van Leer (1983), is independent from all other upwind models presented here. Instead of using diagonalization, it approximates the solution of the Riemann problem (5.2) by three constant states: \mathbf{U}_L , \mathbf{U}_R , and an intermediate state denoted by \mathbf{U}_I . Our

presentation below is slightly simpler than the original one. Similar to the primitive-variable splitting, this flux is employed at ‘smooth’ interfaces. At ‘nonsmooth’ interfaces ($k_{j+1/2} < 2$), together with a modification to resolve contact discontinuities presented below, it can also be used.

Our choice of wave speeds is most economical: with the average state $\hat{\mathbf{V}}$ defined in (5.5) and \hat{c} the corresponding speed of sound, let the minimum (subscript l) and the maximum (subscript r) wave speeds be

$$a_l = \hat{u} - \hat{c}, \quad a_r = \hat{u} + \hat{c}. \quad (5.8a, b)$$

The corresponding Mach numbers are

$$\hat{M} = \frac{\hat{u}}{\hat{c}}, \quad M_l = \frac{a_l}{\hat{c}} = \hat{M} - 1, \quad M_r = \frac{a_r}{\hat{c}} = \hat{M} + 1. \quad (5.8c, d, e)$$

For a more sophisticated choice of wave speeds, see Einfeldt (1988). Next, the top half of the (x, t) -plane is divided into three regions separated by the two half lines $x = a_l t$ and $x = a_r t$ (see Fig. 5.1). For the left region, the solution is assumed to be the constant state \mathbf{U}_L ; for the right, \mathbf{U}_R ; and for the middle region, \mathbf{U}_I , which is determined below. Let $\tau > 0$ be a fixed (provisionary only) time step. Consider the control volume whose four corners are $(a_l \tau, 0)$, $(a_l \tau, \tau)$, $(a_r \tau, \tau)$, $(a_r \tau, 0)$. By balancing the fluxes for (integrating (2.1a) over) this control volume,

$$\tau(a_r - a_l)\mathbf{U}_I = \tau(a_r \mathbf{U}_R - a_l \mathbf{U}_L) - \tau(\mathbf{F}_R - \mathbf{F}_L).$$

As a result,

$$\mathbf{U}_I = \frac{a_r \mathbf{U}_R - a_l \mathbf{U}_L}{a_r - a_l} - \frac{\mathbf{F}_R - \mathbf{F}_L}{a_r - a_l}. \quad (5.9)$$

Note that τ no longer appears in the above expression.

We are now ready to derive the upwind flux. For the moment, suppose $a_l < 0$ and $a_r > 0$. By balancing the fluxes for the control volume with corners $(a_l \tau, 0)$, $(a_l \tau, \tau)$, $(0, \tau)$, and $(0, 0)$,

$$-\tau a_l \mathbf{U}_I = -\tau a_l \mathbf{U}_L - \tau(\mathbf{F}_U - \mathbf{F}_L).$$

As a result,

$$\mathbf{F}_U = \mathbf{F}_L + a_l(\mathbf{U}_I - \mathbf{U}_L). \quad (5.10a)$$

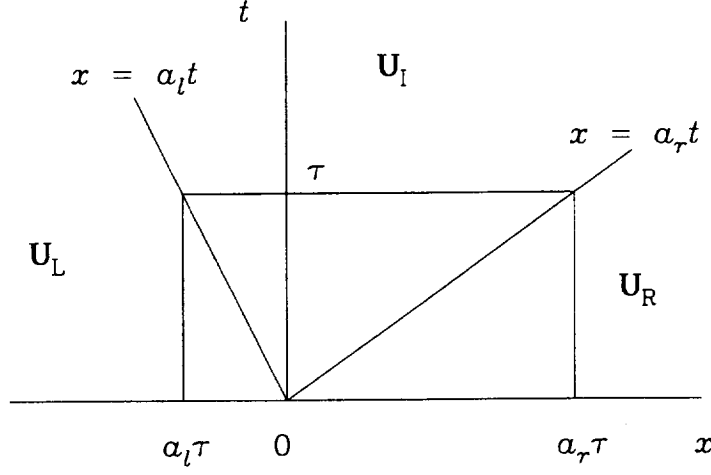


Figure 5.1. *The one-intermediate-state model.*

Similarly, for the control volume with corners $(0, 0)$, $(0, \tau)$, $(a_r\tau, \tau)$, and $(a_r\tau, 0)$,

$$\mathbf{F}_U = \mathbf{F}_R + a_r(\mathbf{U}_I - \mathbf{U}_R). \quad (5.10b)$$

Note that the above two \mathbf{F}_U are identical by (5.9). Multiply (5.10a) by a_r , (5.10b) by $-a_l$, and then add the results,

$$\mathbf{F}_U = \frac{-a_l}{a_r - a_l} \mathbf{F}_R + \frac{a_r}{a_r - a_l} \mathbf{F}_L + \frac{a_l a_r}{a_r - a_l} (\mathbf{U}_R - \mathbf{U}_L). \quad (5.10c)$$

Thus the flux \mathbf{F}_U is a weighted average of the left and right fluxes plus an ‘artificial viscosity’ term. Note that the above expression involves only scalar operations and no matrix multiplications. Next, if $a_l \geq 0$ then $\mathbf{F}_U = \mathbf{F}_L$; similarly if $a_r \leq 0$ then $\mathbf{F}_U = \mathbf{F}_R$. That is,

$$\mathbf{F}_U = \begin{cases} \mathbf{F}_L & \text{if } 0 \leq a_l, \\ \frac{-a_l}{a_r - a_l} \mathbf{F}_R + \frac{a_r}{a_r - a_l} \mathbf{F}_L + \frac{a_l a_r}{a_r - a_l} (\mathbf{U}_R - \mathbf{U}_L) & \text{if } a_l < 0 < a_r, \\ \mathbf{F}_R & \text{if } a_r \leq 0. \end{cases} \quad (5.11a)$$

Observe that if $a_l = 0$, then (5.10c) yields $\mathbf{F}_U = \mathbf{F}_L$; similarly, if $a_r = 0$, it implies $\mathbf{F}_U = \mathbf{F}_R$. As a result, the above expression for \mathbf{F}_U depends continuously on the data.

Using the Mach numbers defined in (5.8c,d,e),

$$\mathbf{F}_U = \begin{cases} \mathbf{F}_L & \text{if } 0 \leq M_l, \\ \frac{1}{2}(-M_l \mathbf{F}_R + M_r \mathbf{F}_L) + \frac{1}{2} M_l M_r \hat{c} (\mathbf{U}_R - \mathbf{U}_L) & \text{if } M_l < 0 < M_r, \\ \mathbf{F}_R & \text{if } M_r \leq 0. \end{cases} \quad (5.11b)$$

The above expression can be combined into a single formula:

$$\mathbf{F}_U = \frac{1}{2} [(M_r^- - M_l^-)\mathbf{F}_R + (M_r^+ - M_l^+)\mathbf{F}_L] + \frac{1}{2} M_l^- M_r^+ \hat{c} [\mathbf{U}_R - \mathbf{U}_L] \quad (5.11c)$$

where for any real number z , the negative and positive parts of z are

$$z^- = \min(z, 0) = \frac{1}{2}(z - |z|), \quad z^+ = \max(z, 0) = \frac{1}{2}(z + |z|). \quad (5.12)$$

We summarize the algorithm below.

Algorithm for an upwind flux via the one-intermediate-state model. At a ‘smooth’ interface where $k_{j+1/2} = 2$, with \mathbf{V}_L and \mathbf{V}_R given by (5.1), first calculate the average state $\hat{\mathbf{V}}$ by (5.5); next obtain the Mach numbers \hat{M} , M_l , and M_r by (5.8c, d, e); finally, \mathbf{F}_U is given by (5.11b) or (5.11c).

Notice that if the maximum and minimum wave speeds are defined as

$$a_l = -(|\hat{u}| + \hat{c}), \quad a_r = |\hat{u}| + \hat{c}, \quad (5.13a)$$

then (5.11a) reduces to

$$\mathbf{F}_U = \frac{1}{2}(\mathbf{F}_L + \mathbf{F}_R) - \frac{1}{2}(|\hat{u}| + \hat{c})(\mathbf{U}_R - \mathbf{U}_L), \quad (5.13b)$$

which is a ‘central difference’ plus an ‘artificial viscosity’ term. This simple formula is due to Rusanov (1961) [see also Anderson, Thomas, and Van Leer (1985), Yee (1987), and Davis (1988)]. Generally, it saves about 6 percent of total computing time compared with (5.11), and it performs well for all of our tests except for the Lax problem in §7.

The one-intermediate-state model (5.11) is economical and produces good results at smooth regions. It also handles shocks well; however, it cannot resolve a contact discontinuity accurately since the model does not contain a contact.

We remedy this problem by the following simple technique. At an interface $j + 1/2$, the linear characteristic is said to be dominant if

$$k_{j+1/2} < 0 \quad (5.14a)$$

where $k_{j+1/2}$ is defined by (5.4). In this case, we reset ρ_R to ρ_L if $\hat{u} \geq 0$, and reset ρ_L to ρ_R if $\hat{u} < 0$:

$$\text{if } \hat{M} \geq 0, \quad \rho_R \leftarrow \rho_L; \quad \text{if } \hat{M} < 0, \quad \rho_L \leftarrow \rho_R. \quad (5.14b, c)$$

In other words, near a contact discontinuity, density is determined by the flow direction. Note that velocities and pressures are not altered. The flux \mathbf{F}_U is then calculated by (5.5), (5.8), and (5.11b) or (5.11c).

With the above modification, the model works well for all the standard tests except the Mach 3 expansion shock in §7; i.e., it may admit non-physical solutions (also known as solutions that violate the entropy condition). Further modifications can be made so that non-physical solutions are excluded; e.g., at a ‘nonsmooth’ index $j + 1/2$, i.e., $k_{j+1/2} \leq 1$, instead of (5.5), a_l and a_r can be defined by

$$a_l = \min(u_L - c_L, u_R - c_R), \quad a_r = \max(u_L + c_L, u_R + c_R).$$

Such calculations involve two speeds of sound and take more computing time than (5.5). We will not elaborate on different definitions for a_l and a_r any further; instead, we present below an estimate for \mathbf{F}_U that has two intermediate states, and with an admissibility condition, it excludes non-physical solutions.

5.3. Flux-vector splitting. The flux-vector splitting developed by Sanders and Prendergast (1974) and Steger and Warming (1981) is conceptually simple; however, the resulting scheme smears a contact discontinuity and may become inaccurate at sonic points [see also Hirsch (1990)]. On the other hand, Roe’s flux-difference splitting (1981, 1986) is conceptually more involved, but it is highly accurate. The models for these two splittings are very different: for the former, the interaction between the left and right states is accomplished through mixing of pseudo-particles (the Boltzman approach); for the latter, through finite-amplitude waves (the Riemann approach) [see Harten, Lax, and Van Leer (1983)].

We present below a flux-vector splitting that turns out to be identical to Roe’s flux-difference splitting. The advantage of this presentation is its conceptual simplicity. The method, in fact, is very similar to that of §5.1, but instead of splitting the primitive variables, we split the fluxes. The two key techniques are, again, linearization and diagonalization. What is crucial, however, is that the fixed state for the linearization must be derived carefully in order that the solution depends continuously on the data. Our derivation of this state, which is Roe’s ‘square root averaging’ state, follows the remark concerning M. Brio in Roe and Pyke (1984).

As expressed in (2.10), after a linearization, the Euler equation can be written as

$$\frac{\partial \mathbf{U}}{\partial t} + \tilde{\mathbf{A}}_c \frac{\partial \mathbf{U}}{\partial x} = 0;$$

here $\tilde{\mathbf{A}}_c = (\partial \mathbf{F} / \partial \mathbf{U})(\tilde{\mathbf{U}})$, and $\tilde{\mathbf{U}}$ is yet to be determined. With the initial condition (5.2), let $\tilde{\mathbf{U}}$ be a state that satisfies

$$\Delta \mathbf{F} = \tilde{\mathbf{A}}_c \Delta \mathbf{U} \quad (5.15)$$

where Δ is defined in (5.3). (As will be shown later, the above condition implies the upwind flux depends continuously on the data.) Loosely speaking, we wish to find a solution for the mean value theorem (which says a solution exists in the scalar case). Using expression (2.22) for $\tilde{\mathbf{A}}_c$,

$$\begin{pmatrix} \Delta \rho u \\ \Delta(\rho u^2 + p) \\ \Delta(e + p)u \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ (\gamma - 3)\tilde{u}^2/2 & (3 - \gamma)\tilde{u} & \gamma - 1 \\ (\gamma - 1)\tilde{u}^3/2 - \tilde{u}\tilde{H} & -(\gamma - 1)\tilde{u}^2 + \tilde{H} & \gamma\tilde{u} \end{pmatrix} \begin{pmatrix} \Delta \rho \\ \Delta \rho u \\ \Delta e \end{pmatrix}. \quad (5.16)$$

The first component of the above vector equation yields $\Delta \rho u = \Delta \rho u$, which is useless; as a result, the solution of the above system cannot be unique. Next, using expression (2.2b) for e , the second component can be simplified to

$$\tilde{u}^2 \Delta \rho - 2\tilde{u} \Delta(\rho u) + \Delta(\rho u^2) = 0. \quad (5.17)$$

The two solutions of the above quadratic in \tilde{u} are

$$\tilde{u} = \frac{\Delta(\rho u) \pm [\{\Delta(\rho u)\}^2 - \Delta \rho \Delta(\rho u^2)]^{1/2}}{\Delta \rho}.$$

For simplicity of notation, set

$$r_L = \sqrt{\rho_L}, \quad r_R = \sqrt{\rho_R}, \quad (5.18)$$

then, by expanding Δ via (5.3),

$$\tilde{u} = \frac{r_L u_L \pm r_R u_R}{r_L \pm r_R}. \quad (5.19)$$

The solution corresponding to the negative sign is meaningless when $\rho_L = \rho_R$, and is therefore discarded. The one with the positive sign always makes sense and yields

$$\tilde{u} = \frac{r_L u_L + r_R u_R}{r_L + r_R}. \quad (5.20)$$

The above expression is Roe's well-known 'square root averaging'. For the purpose of coding and for later use, denote

$$\beta_L = \frac{r_L}{r_L + r_R}, \quad \beta_R = \frac{r_R}{r_L + r_R}, \quad (5.21a, b)$$

and

$$\rho_* = \sqrt{\rho_L \rho_R}. \quad (5.22)$$

Then

$$\beta_L = \frac{\rho_L}{\rho_L + \rho_*}, \quad \beta_R = 1 - \beta_L, \quad (5.23a, b)$$

and

$$\tilde{u} = \beta_L u_L + \beta_R u_R; \quad (5.24)$$

that is, \tilde{u} is a weighted average of u_L and u_R .

Consider now the third component of (5.16). Since \tilde{u} is already determined above, this component yields the total enthalpy \tilde{H} . Indeed, by definition (2.17) for H ,

$$e + p = \rho H, \quad (5.25a)$$

and by (2.2b),

$$e = \frac{1}{\gamma} \rho H + \frac{\gamma - 1}{2\gamma} \rho u^2. \quad (5.25b)$$

Using the above expressions, (5.17), and the linearity of Δ , the third component can be simplified to

$$\Delta(\rho u H) = -\tilde{u} \tilde{H} \Delta \rho + \tilde{H} \Delta(\rho u) + \tilde{u} \Delta(\rho H).$$

As a consequence,

$$\tilde{H} = \frac{\Delta(\rho u H) - \tilde{u} \Delta(\rho H)}{\Delta(\rho u) - \tilde{u} \Delta \rho}. \quad (5.26)$$

By expanding Δ in the above expression and factoring out $r_L r_R (u_R - u_L)$,

$$\tilde{H} = \frac{r_L H_L + r_R H_R}{r_L + r_R}; \quad (5.27)$$

that is, with β_L and β_R by (5.23a,b),

$$\tilde{H} = \beta_L H_L + \beta_R H_R \quad (5.28)$$

(note the similarity between (5.28) and (5.24)).

The argument (5.16–28) yields the following important fact: equation (5.15) is satisfied with \tilde{u} by (5.24), \tilde{H} as above, and $\tilde{\rho}$ arbitrary.

Next, we split the fluxes. As noticed after (2.20), we do not need $\tilde{\rho}$ for the purpose of diagonalization because it does not appear in the expressions of $\tilde{\mathbf{A}}_c$, $\tilde{\mathbf{\Lambda}}_c$, $\tilde{\mathbf{L}}_c$, and $\tilde{\mathbf{R}}_c$. Multiplying (2.1) on the left by $\tilde{\mathbf{L}}_c$, and since $\tilde{\mathbf{L}}_c$ is a constant matrix,

$$\frac{\partial \tilde{\mathbf{L}}_c \mathbf{U}}{\partial t} + \frac{\partial \tilde{\mathbf{L}}_c \mathbf{F}}{\partial x} = 0. \quad (5.29a)$$

Let the characteristic variable be $\mathbf{W} = \mathbf{L}_c \mathbf{U}$ (for convenience, we omit the subscript c in \mathbf{W}), and the characteristic flux, $\mathbf{G} = \mathbf{L}_c \mathbf{F}$. Then, the above is equivalent to

$$\frac{\partial \mathbf{W}}{\partial t} + \frac{\partial \mathbf{G}}{\partial x} = 0. \quad (5.29b)$$

Corresponding to the initial condition (5.2),

$$\mathbf{W}_L = \tilde{\mathbf{L}}_c \mathbf{U}_L, \quad \mathbf{W}_R = \tilde{\mathbf{L}}_c \mathbf{U}_R, \quad (5.30)$$

and more importantly,

$$\mathbf{G}_L = \tilde{\mathbf{L}}_c \mathbf{F}_L, \quad \mathbf{G}_R = \tilde{\mathbf{L}}_c \mathbf{F}_R. \quad (5.31)$$

Since $\tilde{\mathbf{L}}_c$ is a constant matrix, (5.30) and (5.31) respectively imply

$$\Delta \mathbf{W} = \tilde{\mathbf{L}}_c \Delta \mathbf{U}, \quad \Delta \mathbf{G} = \tilde{\mathbf{L}}_c \Delta \mathbf{F}. \quad (5.32a, b)$$

We now make use of the crucial expression (5.15): $\Delta \mathbf{F} = \tilde{\mathbf{A}}_c \Delta \mathbf{U}$. Substitute this into (5.32b),

$$\Delta \mathbf{G} = \tilde{\mathbf{L}}_c \tilde{\mathbf{A}}_c (\tilde{\mathbf{R}}_c \tilde{\mathbf{L}}_c) \Delta \mathbf{U}.$$

By (2.15) and (5.32a),

$$\Delta \mathbf{G} = \tilde{\mathbf{\Lambda}} \Delta \mathbf{W}. \quad (5.32c)$$

Let $g^{(i)}$ be the i th component of \mathbf{G} , then the upwind characteristic flux \mathbf{G}_U at $x = 0$ for (5.29b) is determined by the sign of $\tilde{\lambda}^{(i)}$:

$$g_U = \begin{cases} g_L & \text{if } \tilde{\lambda} \geq 0, \\ g_R & \text{otherwise} \end{cases} \quad (5.33)$$

where, once again, the superscript (i) is omitted. Expression (5.32c) shows that if $\tilde{\lambda} = 0$, then $g_L = g_R$, and it does not matter whether the left or the right flux is selected in (5.33). Finally, the upwind flux is

$$\mathbf{F}_U = \tilde{\mathbf{R}}_c \mathbf{G}_U. \quad (5.34)$$

In summary, let ρ_* be defined by (5.22); β_L, β_R by (5.23); \tilde{u} , (5.24); \tilde{H} , (5.28); $\tilde{\mathbf{L}}_c$, (2.18); and $\tilde{\mathbf{R}}_c$, (2.16). Then the upwind flux is given by (5.31), (5.33), and (5.34).

Similar to the derivation of (5.7), the above algorithm can be expressed with no conditional statement. Indeed, by considering the two cases of (5.33),

$$g_U = \frac{1}{2} \left[1 + \operatorname{sgn}(\tilde{\lambda}) \right] g_L + \frac{1}{2} \left[1 - \operatorname{sgn}(\tilde{\lambda}) \right] g_R. \quad (5.35)$$

As a result, switching to vector notation,

$$\mathbf{G}_U = \frac{1}{2}(\mathbf{G}_L + \mathbf{G}_R) - \frac{1}{2}\operatorname{sgn}(\tilde{\Lambda})(\mathbf{G}_R - \mathbf{G}_L).$$

Denote by $|\tilde{\Lambda}|$ the diagonal matrix whose diagonal entries are $|\tilde{\lambda}^{(i)}|$. Then from (5.32c), and since $[\operatorname{sgn}(\tilde{\Lambda})]\tilde{\Lambda} = |\tilde{\Lambda}|$,

$$\mathbf{G}_U = \frac{1}{2}(\mathbf{G}_L + \mathbf{G}_R) - \frac{1}{2}|\tilde{\Lambda}|\Delta\mathbf{W}. \quad (5.36)$$

By (5.32a),

$$\mathbf{G}_U = \frac{1}{2}(\mathbf{G}_L + \mathbf{G}_R) - \frac{1}{2}|\tilde{\Lambda}|\tilde{\mathbf{L}}_c\Delta\mathbf{U}. \quad (5.37)$$

Multiplying the above expression on the left by $\tilde{\mathbf{R}}_c$ and employing definition (5.31), one obtains

$$\mathbf{F}_U = \frac{1}{2}(\mathbf{F}_L + \mathbf{F}_R) - \frac{1}{2}\tilde{\mathbf{R}}_c|\tilde{\Lambda}|\tilde{\mathbf{L}}_c\Delta\mathbf{U}. \quad (5.38)$$

Thus, in the flux-vector splitting algorithm summarized after (5.34), expressions (5.31), (5.33), and (5.34) can be replaced by (5.38). Expression (5.38) is very similar to Roe's flux-difference splitting formula; in fact, in the next subsection, we will show that they are identical.

Note that in (5.31) and (5.33), we split the left and right fluxes to obtain the upwind flux; in (5.38), however, it is $\Delta\mathbf{U}$, the difference of the conservative variables, that is split. As shown above, the two splittings are identical provided (5.15) is satisfied. If a different state is chosen for the linearization, e.g., the average state $\hat{\mathbf{U}}$ defined in (5.5), the splitting (5.33) no longer yields a result identical to (5.38).

With the tilde state given by (5.15), expression (5.33) yields a flux that depends continuously on the data. If a different state is chosen for the linearization, the flux via (5.33) may not depend continuously on the data any longer. In contrast, expression (5.38) always results in a flux that depends continuously on the data no matter which

state we pick for the linearization. For this reason, (5.38) is preferred. Conceptually, however, (5.33) is simpler.

Also note the similarity between the primitive-variable splitting (5.7) and the flux splitting (5.38). What is crucial, however, is the differences between the two: to the advantage of the flux splitting, the sign function in (5.7), which is discontinuous, is replaced by the absolute value function in (5.38), which is continuous; to its disadvantage, the matrices in (5.7) associated with the primitive variables are sparse, while those in (5.38) associated with the fluxes are not.

Observe that the key difference between the flux-vector splitting of (5.33) and that of Steger and Warming is that in the latter, the left flux is split by the matrix $\mathbf{L}_{c,L}$ at the left state, and similarly, the right flux, by the matrix at the right state; whereas in the current method, both the left and right fluxes are split simultaneously by the matrix $\tilde{\mathbf{L}}_c$ at the tilde state.

5.4. Flux-difference splitting. In this subsection, we show that the above method yields a result identical to Roe's flux-difference splitting. To prove this, we first calculate $\tilde{\rho}$ by employing the Jacobian of the transformation between the primitive and conservative variables. In addition to (5.15), we require that the tilde state satisfies

$$\Delta \mathbf{U} = \tilde{\mathbf{M}} \Delta \mathbf{V}, \quad (5.39)$$

or, with $\tilde{\mathbf{M}}$ expressed by (2.11b),

$$\begin{pmatrix} \Delta \rho \\ \Delta \rho u \\ \Delta e \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \tilde{u} & \tilde{\rho} & 0 \\ \tilde{u}^2/2 & \tilde{\rho} \tilde{u} & 1/(\gamma - 1) \end{pmatrix} \begin{pmatrix} \Delta \rho \\ \Delta u \\ \Delta p \end{pmatrix}. \quad (5.40)$$

Notice that due to the use of $\tilde{\mathbf{M}}$, our derivation of the tilde state is conceptually simpler than that of Roe and Pyke (1984), where the eigenvectors are employed to decompose $\Delta \mathbf{F}$ and $\Delta \mathbf{U}$.

The first component of the above vector equation yields $\Delta \rho = \Delta \rho$, which is again useless. The second and third yield, respectively,

$$\Delta(\rho u) = \tilde{u} \Delta \rho + \tilde{\rho} \Delta u, \quad (5.41)$$

$$\Delta(\rho u^2) = \tilde{u}^2 \Delta \rho + 2\tilde{\rho} \tilde{u} \Delta u. \quad (5.42)$$

Note that these two equations are identical to (3.9–10) of Roe and Pyke (1984). Calculating $\tilde{\rho}$ from (5.41) and substituting it into (5.42), one obtains

$$\tilde{u}^2 \Delta \rho - 2\tilde{u} \Delta(\rho u) + \Delta(\rho u^2) = 0.$$

Strikingly, this quadratic is identical to (5.17), and the solution for \tilde{u} is given in (5.20). Substitute \tilde{u} into (5.41), one obtains

$$\tilde{\rho} = \sqrt{\rho_L \rho_R}. \quad (5.43)$$

The argument (5.40–43) yields the following important fact: equation (5.39) is satisfied with $\tilde{\rho}$ as above, \tilde{u} by (5.24), and \tilde{H} arbitrary.

As a result of the statement above and that after (5.28), the following crucial conclusion can be drawn. There exists a state $\tilde{\mathbf{V}}$ which satisfies the two equations

$$\Delta \mathbf{F} = \tilde{\mathbf{A}}_c \Delta \mathbf{U} \quad \text{and} \quad \Delta \mathbf{U} = \tilde{\mathbf{M}} \Delta \mathbf{V} \quad (5.44a, b)$$

simultaneously, and this state is unique: it is defined by (5.43), (5.24), and (5.28).

We make use of (5.44b) to simplify (5.38). Multiplying (5.44b) on the left by $\tilde{\mathbf{L}}_c$ and employing (5.32a) and (2.14b),

$$\Delta \mathbf{W} = \tilde{\mathbf{L}}_c \Delta \mathbf{U} = \tilde{\mathbf{L}}_p \Delta \mathbf{V}. \quad (5.45)$$

Substituting this into (5.38), we obtain

$$\mathbf{F}_U = \frac{1}{2}(\mathbf{F}_L + \mathbf{F}_R) - \frac{1}{2} \tilde{\mathbf{R}}_c |\tilde{\mathbf{A}}| \tilde{\mathbf{L}}_p \Delta \mathbf{V}. \quad (5.46)$$

The above is identical to expression (28) of Roe (1986), which is the flux-difference splitting formula. To clarify the differences in notation, note that the components of $\tilde{\mathbf{L}}_p \Delta \mathbf{V}$ are denoted by α_k by Roe; and the columns of \mathbf{R}_c , by \mathbf{e}_k . Since the flux-vector and flux-difference splittings yield identical results, when a distinction between them is not necessary, they are called the flux-splitting method.

Observe that the advantage of the flux-difference splitting (5.46) over the flux splitting (5.38) is that the matrix $\tilde{\mathbf{L}}_p$ is sparse while $\tilde{\mathbf{L}}_c$ is not.

Also note the similarity between the flux-difference splitting and the primitive-variable splitting (5.7). For comparison, suppose we replace the tilde state in (5.46)

by the average state (5.5). Then, compared to the primitive-variable splitting, the flux-difference splitting has the disadvantage of costing slightly more since $\hat{\mathbf{R}}_p$ is sparse while $\hat{\mathbf{R}}_c$ is not; on the other hand, its advantage is that the resulting flux depends continuously on the data, and consequently, it is more robust.

We remark that with the tilde state defined by (5.24), (5.28), and (5.43), one can derive \tilde{p} :

$$\tilde{p} = \beta_R p_L + \beta_L p_R + \frac{\gamma - 1}{2\gamma} \beta_L \beta_R \tilde{\rho} (\Delta u)^2.$$

From the definition of the tilde state, it is not obvious that \tilde{p} is strictly positive. But from the above expression, and since $\tilde{\rho}$ by (5.43), β_L , and β_R by (5.23) are all strictly positive, so is \tilde{p} .

As discussed by Roe (1981, 1986, 1989), the flux-splitting method may admit non-physical solutions. This problem can be corrected by an admissibility condition (Roe and Pyke 1984) or by adding numerical diffusion (Harten 1983, Harten and Hyman 1985). One can also derive an upwind flux that satisfies the entropy condition a priori (Engquist and Osher 1981, Osher and Solomon 1982, Osher 1984, Osher and Chakravarty 1984). For a recent treatment, see Roe (1992). Two types of admissibility conditions that are somewhat different from the one in Roe and Pyke (1984) are shown below. Our presentation also establishes a connection among these different approaches: they all add numerical diffusion by increasing the magnitude of the wave speeds.

5.5. Admissibility conditions for flux-vector splitting. In this subsection, we present a conceptually simple admissibility condition that excludes non-physical solutions. Due to its simplicity, it may easily extend to other systems of equations.

Again, the superscript (i) is understood. Suppose λ_L and λ_R (yet to be determined) are respectively the speeds associated with the characteristic quantities $\{w_L, g_L\}$ and $\{w_R, g_R\}$ defined in (5.30–31). Then, for the flux-vector splitting, we have effectively assumed that $\lambda_L = \lambda_R = \tilde{\lambda}$. In other words, the i th component of equation (5.29b) is a convection equation with speed $\tilde{\lambda}$. Since (5.29b) is nonlinear, if we can approximate λ_L and λ_R more accurately, in addition to the case of a convection, we can have either that of a shock where $\lambda_L > \tilde{\lambda} > \lambda_R$, or a fan where $\lambda_L < \tilde{\lambda} < \lambda_R$.

In the case of a shock (see Fig. 5.2(a)), expression (5.33) is still appropriate for

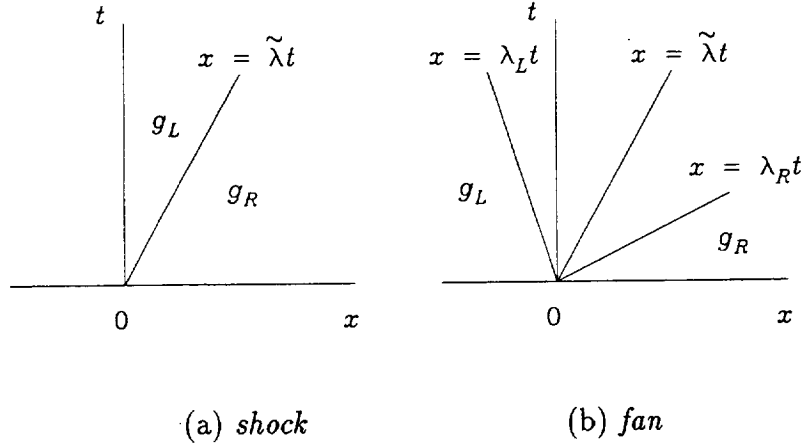


Figure 5.2. *Characteristic fluxes for a shock and a fan.*

the following reasons. The upper half of the (x, t) -plane is divided into two regions by the shock line $x = \tilde{\lambda}t$, and the characteristic flux is a constant for each region, taking the value g_L to the left of the shock, and the value g_R to the right. Expression (5.33) determines which of these two values is chosen for g_U , the characteristic flux at $x = 0$ ($x = \lambda t$, $\lambda = 0$, and $t \geq 0$).

In the case of a fan (see Fig. 5.2(b)), however, (5.33) may no longer be appropriate for the following reasons. The upper half of the (x, t) -plane is divided into three regions in this case: to the left of $x = \lambda_L t$, the flux is g_L ; to the right of $x = \lambda_R t$, it is g_R ; and for the region consisting of the rays $x = \lambda t$ where $\lambda_L \leq \lambda \leq \lambda_R$, the flux $g(\lambda)$ changes smoothly from g_L to g_R . If the fan does not spread past $x = 0$ (or $\lambda = 0$), then clearly, (5.33) is still valid. But if the fan spreads past $x = 0$, (5.33) still yields either g_L or g_R , which is a poor approximation for g_U and may result in non-physical solutions. This situation is depicted further in Fig. 5.3, where the characteristic flux g as a function of the characteristic variable w is plotted.

We estimate λ_L and λ_R by using the tilde state as a pivot, and approximate the fan by a (piecewise) linear function below. First, denote the vector of characteristic variables and fluxes at the tilde state by $\tilde{\mathbf{W}}$ and $\tilde{\mathbf{G}}$:

$$\tilde{\mathbf{W}} = \tilde{\mathbf{L}}_c \tilde{\mathbf{U}} = \tilde{\mathbf{L}}_p (\tilde{\mathbf{M}}^{-1} \tilde{\mathbf{U}}), \quad \tilde{\mathbf{G}} = \tilde{\mathbf{L}}_c \tilde{\mathbf{F}} = \tilde{\mathbf{L}}_p (\tilde{\mathbf{M}}^{-1} \tilde{\mathbf{F}}). \quad (5.47)$$

By carrying out the multiplications in the brackets using (2.11c) and then the multi-

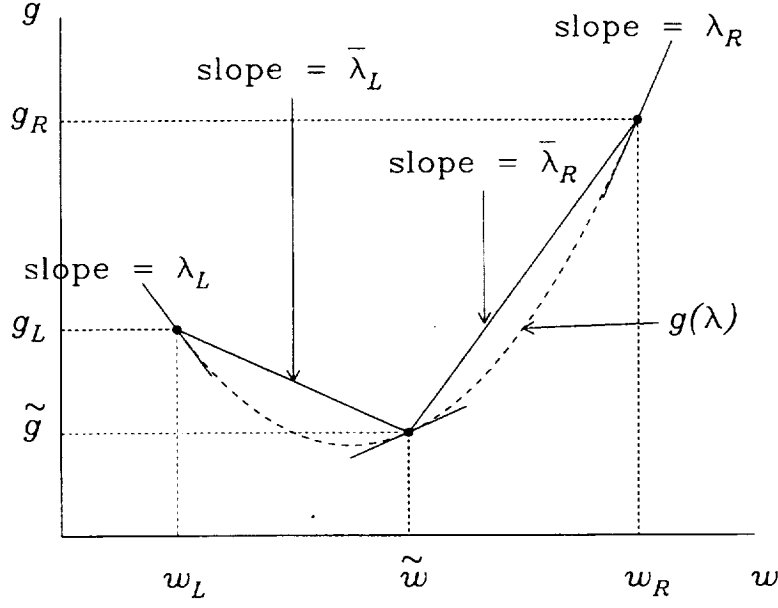


Figure 5.3. A fan spreading past $\lambda = 0$ in (w, g) -space.

plication by $\tilde{\mathbf{L}}_p$ using (2.6c), the above expressions yield

$$\tilde{\mathbf{W}} = \frac{\tilde{\rho}}{\gamma} \begin{pmatrix} 1/2 \\ \gamma - 1 \\ 1/2 \end{pmatrix}, \quad \tilde{\mathbf{G}} = \frac{\tilde{\rho}}{\gamma} \begin{pmatrix} (\tilde{u} - \tilde{c})/2 \\ \tilde{u}(\gamma - 1) \\ (\tilde{u} + \tilde{c})/2 \end{pmatrix}. \quad (5.48a, b)$$

(We remark that \mathbf{G}_L and \mathbf{G}_R defined in (5.31) can be simplified in a similar manner.) Note that $\tilde{\mathbf{G}}$ can also be derived by observing that the flux \mathbf{F} , as a function of the conservative variable \mathbf{U} , is homogenous of degree one: $\tilde{\mathbf{F}} = \tilde{\mathbf{A}}_c \tilde{\mathbf{U}}$; as a result, $\tilde{\mathbf{G}} = \tilde{\mathbf{A}} \tilde{\mathbf{W}}$. In addition, notice that each component \tilde{g} of the vector $\tilde{\mathbf{G}}$ has units of momentum, while each component \tilde{w} of $\tilde{\mathbf{W}}$ has units of density and is strictly positive. Next, let the average speeds be

$$\bar{\lambda}_L = \frac{g_L - \tilde{g}}{w_L - \tilde{w}}, \quad \bar{\lambda}_R = \frac{g_R - \tilde{g}}{w_R - \tilde{w}}. \quad (5.49a, b)$$

We can now estimate λ_L and λ_R :

$$\lambda_L = 2\bar{\lambda}_L - \tilde{\lambda}, \quad \lambda_R = 2\bar{\lambda}_R - \tilde{\lambda}. \quad (5.50a, b)$$

If $\tilde{\lambda} \geq 0$ and $\lambda_L < 0$, then we have a fan that spreads past $x = 0$. By assuming that the flux g is a linear function of the speed λ between (λ_L, g_L) and $(\tilde{\lambda}, \tilde{g})$, we can estimate

g_U , the flux corresponding to the speed 0:

$$\frac{g_U - g_L}{-\lambda_L} = \frac{\tilde{g} - g_L}{\tilde{\lambda} - \lambda_L}.$$

Thus, for this case, we reset g_U defined in (5.33) by

$$g_U \leftarrow \frac{\tilde{\lambda}g_L - \lambda_L\tilde{g}}{\tilde{\lambda} - \lambda_L}. \quad (5.51a)$$

Similarly, if $\tilde{\lambda} < 0$ and $\lambda_R > 0$, then

$$g_U \leftarrow \frac{-\tilde{\lambda}g_R + \lambda_R\tilde{g}}{-\tilde{\lambda} + \lambda_R}. \quad (5.51b)$$

By considering the limiting case of $\lambda_L \rightarrow 0$ and that of $\lambda_R \rightarrow 0$, one can easily verify that g_U defined by (5.33) and (5.51a,b) depends continuously on the data.

Notice that we need to carry out the above admissibility correction only for the nonlinear characteristics; for the linear one, (5.33) is sufficient (more on this in the next subsection). In fact, if $\tilde{u} \geq 0$, then the first component $g_U^{(1)}$ requires this correction, but not the third since the left and right speeds corresponding to a nonlinear characteristic do not spread beyond the contact speed \tilde{u} (approximately anyway; for the exact solution, this is always the case). Similarly if $\tilde{u} < 0$, then the correction is carried out only for $g_U^{(3)}$.

Also note that in general, λ_L and λ_R given by (5.50) may lie on the same side relative to $\tilde{\lambda}$. However, if the left and right states differ only by either the first or the third characteristic, then, for all practical situations, λ_L and λ_R lie on different sides relative to $\tilde{\lambda}$; i.e., we have either a shock or a fan.

5.6. Admissibility conditions for flux-difference splitting. Since the role of the admissibility condition is to break up the solution when it is not physical, a simple model may serve the purpose. In this subsection, we make use of the simple wave model employed by Roe and Pyke (1984) to approximate λ_L and λ_R , and then, derive an admissibility condition for the flux-difference splitting that involves no conditional statement. Moreover, formula (5.67) below shows that this condition produces additional ‘numerical viscosity’ when the fan spreads past $x = 0$, and thus, (5.67) provides an explanation for the connection between the approach of admissibility condition (Roe and Pyke 1984) and that of numerical diffusion (Harten 1983, Harten and Hyman 1985).

Let $\mathbf{V} = (\rho, u, p)^T$ be a fixed state, which plays the role of $\hat{\mathbf{V}}$ in (2.9). Let $\mathbf{V} + \delta\mathbf{V} = (\rho + \delta\rho, u + \delta u, p + \delta p)^T$ be a nearby state. For each i th component, we will estimate $\delta\lambda^{(i)}$, the change in the characteristic speed, in terms of $\delta w^{(i)}$, the change in the characteristic variable. By linearizing the Euler equations at \mathbf{V} , (2.9) implies

$$\delta\mathbf{W} = \mathbf{L}_p \delta\mathbf{V} = \begin{pmatrix} -[\rho/(2c)] \delta u + [1/(2c^2)] \delta p \\ \delta\rho - [1/c^2] \delta p \\ [\rho/(2c)] \delta u + [1/(2c^2)] \delta p \end{pmatrix}. \quad (5.52)$$

(See also expression (5.45).)

Next, we assume that the two states differ from each other only by the first characteristic; that is,

$$\delta w^{(2)} = \delta\rho - [1/c^2] \delta p = 0, \quad (5.53a)$$

and

$$\delta w^{(3)} = [\rho/(2c)] \delta u + [1/(2c^2)] \delta p = 0. \quad (5.53b)$$

All δ quantities are expressed in terms of $\delta\rho$ below. Equations (5.53a,b) are equivalent to

$$\delta p = c^2 \delta\rho, \quad \delta u = -(c/\rho) \delta\rho. \quad (5.54a, b)$$

As a consequence of (5.52) and (5.54),

$$\delta w^{(1)} = \delta\rho. \quad (5.55)$$

Now, we calculate δc . By definition (2.4) for the speed of sound,

$$2c\delta c = \gamma \frac{\rho\delta p - p\delta\rho}{\rho^2}.$$

Using (5.54a) and again (2.4), the above implies

$$\delta c = \frac{(\gamma - 1)c\delta\rho}{2\rho}. \quad (5.56)$$

The speed corresponding to the first characteristic of \mathbf{V} is $\lambda^{(1)} = u - c$; that of $\mathbf{V} + \delta\mathbf{V}$ is $\lambda^{(1)} + \delta\lambda^{(1)} = u + \delta u - (c + \delta c)$. Consequently, by (5.54b) and (5.56),

$$\delta\lambda^{(1)} = \delta u - \delta c = -\frac{(\gamma + 1)c\delta\rho}{2\rho}.$$

And by (5.55),

$$\delta\lambda^{(1)} = -\frac{(\gamma+1)c\delta w^{(1)}}{2\rho}. \quad (5.57)$$

A similar argument for the third characteristic with the assumption that $\delta w^{(1)} = 0$ and $\delta w^{(2)} = 0$ yields

$$\delta\lambda^{(3)} = \frac{(\gamma+1)c\delta w^{(3)}}{2\rho}. \quad (5.58)$$

For the second characteristic, $\delta w^{(1)} = 0$ and $\delta w^{(3)} = 0$ imply that $\delta u = \delta p = 0$. As a result, $\delta\lambda^{(2)} = \delta u = 0$. For this reason, the second component needs no modification.

We are now ready for the admissibility correction. With the initial condition \mathbf{V}_L and \mathbf{V}_R , let the tilde state be defined by (5.43), (5.24), and (5.28). Let $\Delta\mathbf{W} = \tilde{\mathbf{L}}_p\Delta\mathbf{V}$, and for the purpose of derivation, let \mathbf{G}_U be defined by (5.36).

We carry out the argument for the first and third characteristics below. Again, the superscript (i) is omitted unless the corresponding expressions for the two characteristics are different. Equation (5.36) implies

$$g_U = \frac{1}{2}(g_L + g_R) - \frac{1}{2}|\tilde{\lambda}|\Delta w. \quad (5.59)$$

For the purpose of estimating λ_L and λ_R , we assume that $\tilde{w} = (w_L + w_R)/2$; therefore,

$$w_R - \tilde{w} = \tilde{w} - w_L = \frac{1}{2}\Delta w.$$

Next, we employ either estimate (5.57) or (5.58): with $\tilde{\mathbf{V}}$ playing the role of the fixed state \mathbf{V} ; \mathbf{V}_L and \mathbf{V}_R , that of $\mathbf{V} + \delta\mathbf{V}$; in the case of the first characteristic, set

$$\Delta\lambda = \Delta\lambda^{(1)} = -\frac{(\gamma+1)\tilde{c}\Delta w^{(1)}}{2\tilde{\rho}}, \quad (5.60a)$$

or if it is the third one that we are dealing with, set

$$\Delta\lambda = \Delta\lambda^{(3)} = \frac{(\gamma+1)\tilde{c}\Delta w^{(3)}}{2\tilde{\rho}}. \quad (5.60b)$$

Then

$$\lambda_L = \tilde{\lambda} - \frac{1}{2}\Delta\lambda, \quad \lambda_R = \tilde{\lambda} + \frac{1}{2}\Delta\lambda. \quad (5.61a, b)$$

Note that if $\Delta\lambda < 0$, then $\lambda_L > \tilde{\lambda} > \lambda_R$, the characteristic speeds converge, and we have a shock. On the other hand, if $\Delta\lambda > 0$, then $\lambda_L < \tilde{\lambda} < \lambda_R$, the characteristic

speeds diverge, and we have a fan. Furthermore, for the first characteristic, $\Delta\lambda$ and Δw have opposite sign by (5.60a); for the third, the same sign by (5.60b).

Denote the speed 0 by λ_0 and the corresponding characteristic variable and flux by w_0 and g_0 , respectively. (Thus $g_0 = g_U$.) Suppose now we are in the situation of a fan that spreads past $x = 0$; i.e., $\lambda_L < 0 < \lambda_R$. In addition, suppose $\tilde{\lambda} \geq 0$. It follows from (5.33) that expression (5.59) yields g_L . Since the correct choice for g_U is g_0 , we should add the quantity $g_0 - g_L$ to the right hand side of (5.59).

To keep the formulas simple, we estimate $g_0 - g_L$ in terms of Δw and the various speeds below. The following estimates are independent of the assumption of a fan described above. As assumed in (5.60a–61), the speed λ is a linear function of w ; therefore,

$$\frac{w_0 - w_L}{\Delta w} = \frac{\lambda_0 - \lambda_L}{\Delta\lambda} = \frac{-\lambda_L}{\Delta\lambda}.$$

Moreover, w_L corresponds to the speed λ_L , and w_0 to 0; consequently, the average speed between w_0 and w_L can be approximated by $\lambda_L/2$:

$$\frac{g_0 - g_L}{w_0 - w_L} \approx \frac{\lambda_L}{2}.$$

As a result of the above two expressions,

$$g_0 - g_L \approx -\frac{\lambda_L^2}{2\Delta\lambda}\Delta w. \quad (5.62a)$$

A similar argument yields

$$g_0 - g_R \approx -\frac{\lambda_R^2}{2\Delta\lambda}\Delta w. \quad (5.62b)$$

Now, we turn to the situation of a fan that spreads past $x = 0$. There are two subcases: for $\tilde{\lambda} \geq 0$, set

$$\eta = \frac{\lambda_L^2}{\Delta\lambda}; \quad (5.63a)$$

for $\tilde{\lambda} < 0$, set

$$\eta = \frac{\lambda_R^2}{\Delta\lambda}. \quad (5.63b)$$

The corrected upwind flux is given by adding $g_0 - g_L$ to the right hand side of (5.59) in the former case, and by adding $g_0 - g_R$ in the latter. As a result, by (5.62–63),

$$g_U = \frac{1}{2}(g_L + g_R) - \frac{1}{2}(|\tilde{\lambda}| + \eta)\Delta w. \quad (5.64)$$

Next, we derive a formula for η that works for all cases. Instead of λ_L and λ_R , consider the quantity

$$|\tilde{\lambda}| - \frac{1}{2}\Delta\lambda.$$

In the case of a shock, i.e., $\Delta\lambda < 0$, the above expression is positive. For a fan that does not spread past $x = 0$, it is also positive. For a fan that spreads past $x = 0$, it is negative; it equals λ_L if $\tilde{\lambda} \geq 0$; but $-\lambda_R$ if $\tilde{\lambda} \leq 0$. Only when the fan spreads past $x = 0$, does the flux need a correction. Thus, with

$$\lambda_N = \{|\tilde{\lambda}| - \frac{1}{2}\Delta\lambda\}^-, \quad (5.65)$$

where z^- is the negative part of z defined in (5.12), set

$$\eta = \frac{\lambda_N^2}{|\Delta\lambda| + 10^{-15}}; \quad (5.66)$$

then the corrected flux is given by (5.64).

Let \mathbf{H} be the diagonal matrix whose diagonal entries are $\eta^{(1)}$, 0, and $\eta^{(3)}$. By multiplying the matrix expression corresponding to (5.64) on the left by $\tilde{\mathbf{R}}_c$, one obtains

$$\mathbf{F}_U = \frac{1}{2}(\mathbf{F}_L + \mathbf{F}_R) - \frac{1}{2}\tilde{\mathbf{R}}_c(|\tilde{\Lambda}| + \mathbf{H})\Delta\mathbf{W}. \quad (5.67)$$

Before summarizing the algorithm, note that the quantities $\gamma-1$, $1/(\gamma-1)$, $\gamma/(\gamma-1)$, and $(\gamma-1)/\gamma$ are often used and should be stored.

Algorithm for flux-difference splitting with admissibility correction. At a ‘nonsmooth’ interface, i.e., $k_{j+1/2} < 2$ where $k_{j+1/2}$ is defined by (5.4), with \mathbf{V}_L and \mathbf{V}_R given by (5.1), calculate: e_L and e_R via (2.2b); H_L and H_R , (2.17); $\tilde{\rho}$, (5.43); β_L and β_R , (5.23); \tilde{u} , (5.24); \tilde{H} , (5.28); \tilde{c}^2 , \tilde{c} , (2.17); $\Delta\mathbf{V}$, (5.3). With $\tilde{\mathbf{L}}_p$ by (2.6c), obtain $\Delta\mathbf{W} = \tilde{\mathbf{L}}_p\Delta\mathbf{V}$. Next, for $i = 1$ and $i = 3$, calculate $\tilde{\lambda}^{(i)}$ via (2.5); $\Delta\lambda^{(1)}$, (5.60a); $\Delta\lambda^{(3)}$, (5.60b); $\lambda_N^{(i)}$, (5.65); $\eta^{(i)}$, (5.66). Finally, with $\tilde{\mathbf{R}}_c$ by (2.16), \mathbf{F}_L and \mathbf{F}_R , (2.2c), the upwind flux is given by (5.67).

Observe that expression (5.66) can be replaced by a simpler one via an additional approximation as shown below. Rewriting (5.66) as $\eta = |\lambda_N|(|\lambda_N|/|\Delta\lambda|)$, then by (5.65), the quantity in the parenthesis is bounded between 0 and 1/2. We can replace this quantity by its upper bound 1/2:

$$\eta = \frac{1}{2}|\lambda_N|. \quad (5.68a)$$

The above expression always yields a more diffusive flux than (5.66). Instead of the factor $1/2$, we can use $1/4$, which is the average of 0 and $1/2$,

$$\eta = \frac{1}{4}|\lambda_N|. \quad (5.68b)$$

For the test problems below, the results by these two expressions are essentially identical to those by (5.66); nevertheless, only (5.66) is employed here.

5.7. Simple-wave upwind flux. We can now derive an economical approximation to Osher's approximate Riemann solver (Engquist and Osher 1981, Osher and Solomon 1982). First, in our context, instead of (5.59), this flux is expressed as

$$g_U = \frac{1}{2}(g_L + g_R) - \frac{1}{2} \int_{w_L}^{w_R} |\lambda| dw. \quad (5.69)$$

Note the key difference between Osher's method and the above expression: for Osher's flux, the integral follows a certain path in the space of state variables, and determining this path is quite involved; here, the path is already determined by the linearization and diagonalization. Observe that symbolically,

$$\frac{dg}{dw} = \lambda, \quad \int \lambda dw = g, \quad \int_{w_L}^{w_R} \lambda dw = g_R - g_L.$$

As a consequence, if $\lambda_L \geq 0$ and $\lambda_R \geq 0$, then since λ is assumed to be linear in w , $\lambda(w) \geq 0$ for all w between w_L and w_R , and expression (5.69) recovers g_L ; similarly, if $\lambda_L \leq 0$ and $\lambda_R \leq 0$, it recovers g_R . For a fan that spreads past $x = 0$, it yields

$$g_U = \frac{1}{2}(g_L + g_R) + \frac{1}{2} \int_{w_L}^{w_0} \lambda dw - \frac{1}{2} \int_{w_0}^{w_R} \lambda dw = g_0. \quad (5.70)$$

For these three cases, the result by (5.69) is the same as that by the flux-difference splitting with admissibility condition in the previous subsection.

However, if we have a shock with left and right speeds spreading past $x = 0$, i.e.,

$$\lambda_L > 0 \quad \text{and} \quad \lambda_R < 0, \quad (5.71)$$

then (5.69) yields

$$g_U = \frac{1}{2}(g_L + g_R) - \frac{1}{2} \int_{w_L}^{w_0} \lambda dw + \frac{1}{2} \int_{w_0}^{w_R} \lambda dw = g_L + g_R - g_0. \quad (5.72)$$

Next, in addition to (5.71), we assume that $\tilde{\lambda} \geq 0$. Since (5.59) yields g_L , we must add to its right hand side the quantity $g_R - g_0$ to obtain (5.69). But this quantity is given by (5.62b), and since $\Delta\lambda < 0$,

$$g_R - g_0 \approx -\frac{\lambda_R^2}{2|\Delta\lambda|}\Delta w. \quad (5.73)$$

Thus, with

$$\eta = \frac{\lambda_R^2}{|\Delta\lambda|}. \quad (5.74a)$$

the upwind flux resulting from (5.69) is given by (5.64). The case $\tilde{\lambda} < 0$ is similar except

$$\eta = \frac{\lambda_L^2}{|\Delta\lambda|}. \quad (5.74b)$$

An expression that works for all cases can be derived by assuming that we always have a fan in (5.65): $\Delta\lambda/2$ is replaced by its absolute value,

$$\lambda_N = \{|\tilde{\lambda}| - \tfrac{1}{2}|\Delta\lambda|\}^-. \quad (5.75)$$

Clearly, for a fan, the above expression reduces to (5.65) since $\Delta\lambda \geq 0$. For a shock that does not spread past $x = 0$, it yields 0. For a shock spreading past $x = 0$, there are two subcases: if $\tilde{\lambda} \geq 0$, it yields λ_R ; if $\tilde{\lambda} < 0$, it yields $-\lambda_L$.

Thus, with λ_N defined by (5.75) instead of (5.65), the algorithm for the flux-difference splitting with admissibility correction in the previous subsection yields an approximation to Osher's approximate Riemann solver. This method is named the simple-wave upwind flux.

Observe that the argument (5.72–74) can also be carried out by switching the limits of integration in (5.72) so that the variable λ corresponding to a shock is transformed into that of a fan.

Note that the flux-difference splitting turns a fan into a shock, whereas the algorithm in this subsection does the exact opposite by turning a shock into a fan. Also notice that to save computing time, one can replace the tilde state by the average state (5.5).

6. Flow with area variation. The formulation presented in the previous sections extends easily to the case of flows with source terms. As an example, consider the

unsteady one-dimensional flow in a duct of variable cross-sectional area $A(x)$, which is described by

$$(AU)_t + (AF)_x = \mathbf{B}, \quad (6.1a)$$

where \mathbf{U} and \mathbf{F} are defined in (2.1), and the source term is

$$\mathbf{B} = (0, pdA/dx, 0)^T. \quad (6.1b)$$

This problem has been studied by several authors; see for example Roe (1986) and the references given there. In the same paper, Roe presented a modified Riemann solver that takes into account the area variation. Using Hancock's half-step formulation presented in section 3, the area variation is accounted for simply in the partial differential equations for the primitive variables derived from (6.1). First, (6.1) implies

$$\mathbf{U}_t + \mathbf{F}_x + \frac{A_x}{A} \mathbf{C} = 0, \quad (6.2a)$$

where

$$\mathbf{C} = \mathbf{F} - (0, p, 0)^T = (\rho u, \rho u^2, (e + p)u)^T. \quad (6.2b)$$

Next, with \mathbf{V} and \mathbf{A}_p defined by (2.3), the above leads to

$$\mathbf{V}_t + \mathbf{A}_p \mathbf{V}_x + \frac{A_x}{A} \mathbf{K} = 0, \quad (6.3a)$$

where

$$\mathbf{K} = (\rho u, 0, \gamma p u)^T. \quad (6.3b)$$

The first component of (6.3a) follows easily from that of (6.2a); a similar statement holds for the second component, which employs the first; and the derivation of the third component employs both the first and the second.

For each j th cell, the calculation of \mathbf{S}_j , which approximates $(\mathbf{V}_x)_j$, is the same as that of section 4; however, the quantity \mathbf{T}_j , which approximates $(\mathbf{V}_t)_j$ and was estimated by (3.5), is now obtained by using (6.3):

$$\mathbf{T}_j = -(\mathbf{A}_p)_j \mathbf{S}_j - \left(\frac{A_x}{A} \right)_j \mathbf{K}_j. \quad (6.4)$$

This is where the source term plays its role. The rest of the calculations are the same as those for the Euler equations (2.1): the solution at time $t^{n+1/2}$ is given by (3.6),

and the flux at each interface is obtained by one of the methods in §5 without any modification because time is frozen at $t^{n+1/2}$, and the source term contributes nothing here. Finally, using the midpoint rule in a manner similar to the derivation of (3.3b), expression (6.1) implies

$$A_j \mathbf{U}_j^{n+1} = A_j \mathbf{U}_j^n + \frac{\tau}{h} (A_{j-1/2} \mathbf{F}_{j-1/2}^{n+1/2} - A_{j+1/2} \mathbf{F}_{j+1/2}^{n+1/2} + \mathbf{Z}_j^{n+1/2}), \quad (6.5a)$$

where

$$\mathbf{Z}_j^{n+1/2} = (0, p_j^{n+1/2} [A_{j+1/2} - A_{j-1/2}], 0)^T, \quad (6.5b)$$

$$p_j^{n+1/2} = p_j + \frac{1}{2} \tau (p_t)_j, \quad (6.5c)$$

and $(p_t)_j$ is given by the third component of (6.4). The above expression completes the algorithm for the interior points. Note that in the steady state case, one can set $p_j^{n+1/2} = p_j$.

Our boundary conditions are the standard ones. At the inflow boundary, if the flow is supersonic, we specify all three variables ρ , u , and p ; if it is subsonic, we specify two and extrapolate the third from the interior. For the outflow boundary, if the flow is supersonic, we extrapolate all three variables; if it is subsonic, we extrapolate two and specify one, say, pressure. For simplicity, and since the flow fields in our numerical examples are essentially constant near the boundary, we employ a piecewise constant reconstruction for the boundary cells. For the cell right next to the boundary ones, the slope is estimated by the parabolic (three-point) formula. One can also employ appropriate one-sided interpolations; the results are essentially the same.

7. Numerical experiments. Unless otherwise stated, in all of our numerical examples, $\gamma = 1.4$; the slope steepening coefficient κ is set equal to 5; at a ‘smooth’ interface where $k_{j+1/2} = 2$, the one-intermediate-state model is used; otherwise ($k_{j+1/2} < 2$), the flux-difference splitting with admissibility condition is employed; and finally, the CFL number is 0.8. Note that at a ‘smooth’ interface, the computing effort is only slightly more than that of the central difference scheme with artificial viscosity (5.13b) [see also Jameson, Schmidt, and Turkel (1981), and Harabetian and Pego (1993)]. Generally, most of the interfaces are ‘smooth’. Near a discontinuity, i.e., at a ‘nonsmooth’ interface, the computing effort is slightly more than that of the UNO scheme. Also note that for convenience, the one-intermediate-state model is said to be employed

everywhere if at a ‘nonsmooth’ interface ($k_{j+1/2} < 2$), this model with the modification (5.14) is used.

For shocktube problems, the initial time steps corresponding to a fixed CFL condition are often too big because the flow has not developed; e.g., for the Riemann problem used by Sod (1978), with CFL number 0.8, the first three time steps are 0.0068, 0.0040, and 0.0039, while the final time step is 0.0036. Because the solution is self-similar, all time steps are equal to 0.0036 in the exact solution. The initial time steps that are too big can occasionally cause some oscillations, especially for CFL number 0.9 and above. To avoid this problem, one can make a test run to find the appropriate time step, and then fix the time step rather than the CFL condition. This modification, however, is not appropriate in the case of Woodward and Colella’s blast wave problem, where the time steps vary as time goes by. The following simple modification works well for all test problems below: let the k th time step corresponding to a fixed CFL condition be denoted by Δt ; for the first 9 time steps, ($1 \leq k \leq 9$), we reset Δt by

$$\Delta t \leftarrow \Delta t \left[1 - \frac{(10 - k)^2}{100} \right]. \quad (7.1)$$

Generally, to arrive at the same final time, a test case with the above modification takes three time steps more than one without. For the example mentioned above, the first three time steps are 0.0013, 0.0020, and 0.0023. Expression (7.1) is employed in all of the following problems.

The first test, used by Sod (1978), is the Riemann problem (5.2) with the initial data

$$(\rho_L, u_L, p_L) = (1, 0, 1), \quad (\rho_R, u_R, p_R) = (0.125, 0, 0.1) \quad (7.2)$$

where the subscript L represents the initial condition for $0 \leq x < 0.5$, and R , for $0.5 < x \leq 1$. Figure 7.1 shows the results with 100 cells, $h = 0.01$, $x_j = (j - 1/2)h$. The final time is $t = 0.2$ (after 57 time steps). The solid line represents the exact solution; the dots, the numerical one. Note that we do not need any boundary conditions because the final time is chosen so that the waves have not arrived at the boundaries. At the end of the calculation, there are a total of fourteen ‘nonsmooth’ cells: eight around the contact discontinuity, and six around the shock; the two discontinuities at the two ends of the expansion fan are in the ‘smooth’ region because the changes in density are

within the noise level by this time. The internal energy (per unit mass) $e/\rho - u^2/2$, which magnifies any error for this problem, is shown in Fig. 7.1(d).

In the next two figures, the results for the density, velocity, and pressure are omitted since they are practically identical to those of Fig. 7.1. Figure 7.2 shows the internal energy of the same problem except for: (a) CFL number 0.9 (51 time steps), (b) CFL number 0.01 (4329 time steps), (c) $\kappa = 0$ (57 time steps), and (d) $\kappa = 10$ (also 57 time steps). As shown by Fig 7.2(a), in spite of the rather large CFL number, the solution is highly accurate thanks to the modification of the first ten time steps via expression (7.1). In Fig. 7.2(b), after a large number of time steps, the contact discontinuity still remains sharp due to the slope-steepening technique. Figure 7.2(c) shows that for Sod's problem, with CFL number 0.8, the final time is not large enough to smear the contact discontinuity because even without slope steepening ($\kappa = 0$), the solution is adequate. Comparing Fig. 7.2(d) to Fig. 7.1(d) we notice that once the contact discontinuity is resolved, increasing κ does not yield a significant improvement.

Figure 7.3 shows again the internal energy for the same problem except now we test the upwind fluxes: (a) central difference with artificial viscosity (5.13b) at 'smooth' regions, (b) one-intermediate-state (HLV) model everywhere, (c) primitive-variable splitting (5.6) at 'smooth' regions, and (d) simple-wave upwind flux (5.75) everywhere. As shown in Fig. 7.3(a), the result by the central difference (5.13b) is as good as that via a more sophisticated model. However, for the Lax problem below, the solution by this model cannot match that via an upwind one. Figure 7.3(b) shows that the modification (5.14) for the one-intermediate-state model is capable of resolving a contact discontinuity. Comparing Fig. 7.1(d) to Fig. 7.3(c), we notice that the solution via the primitive-variable splitting (at 'smooth' regions) is essentially identical to that via the one-intermediate-state model. Similarly, Figs. 7.1(d) and 7.3(d) show that at 'smooth' regions, a more sophisticated (simple-wave) model does not improve the result; in addition, for a moving shock, the simple-wave model produces a result essentially identical to that via the shock model (flux-difference splitting).

The next test, used by Lax (1954), is the Riemann problem

$$(\rho_L, u_L, p_L) = (0.445, 0.698, 3.528), \quad (\rho_R, u_R, p_R) = (0.5, 0, 0.571) \quad (7.3)$$

with the final time 0.16 (after 97 time steps). This problem is more difficult than that

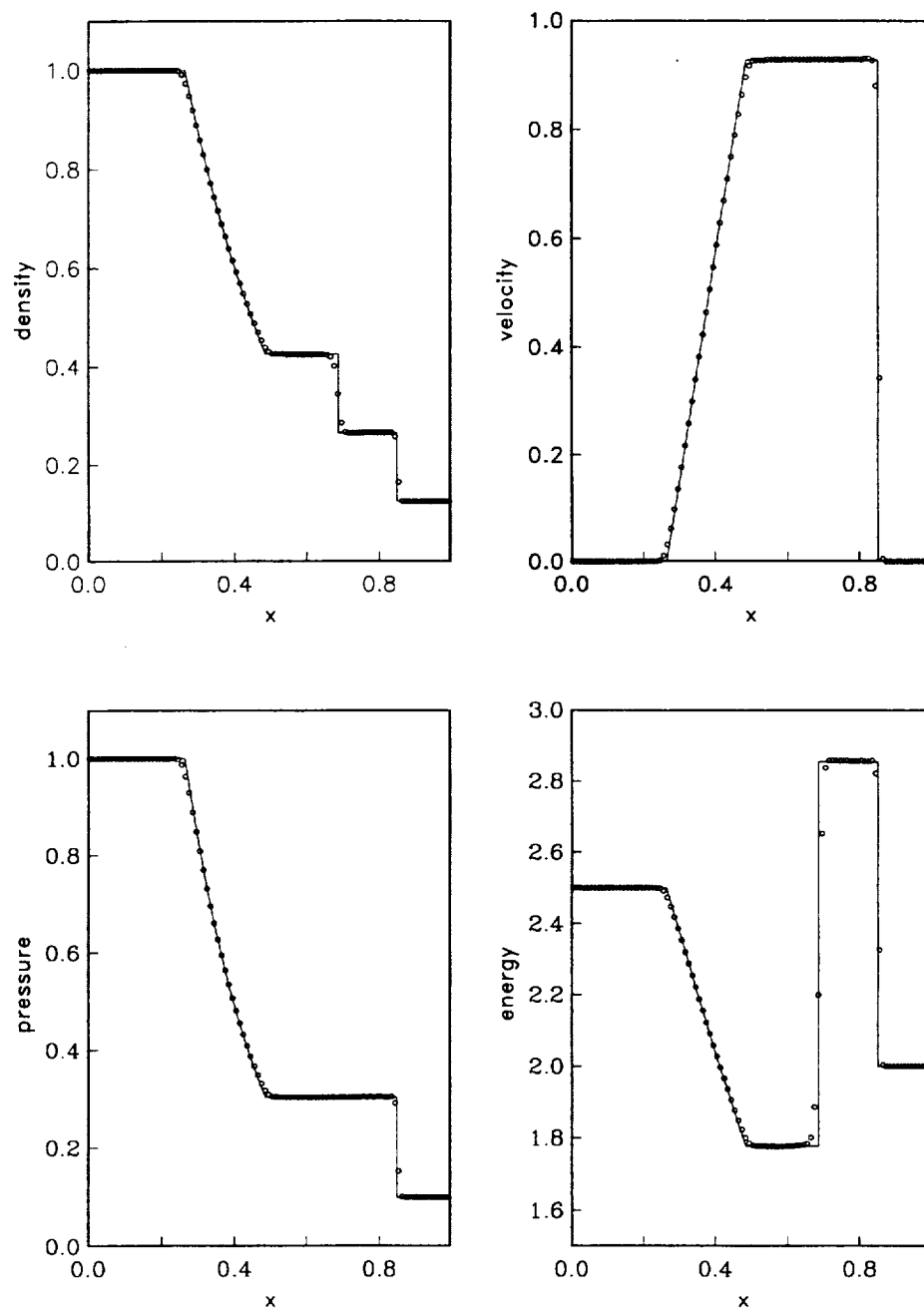


Figure 7.1. *Sod's problem.*

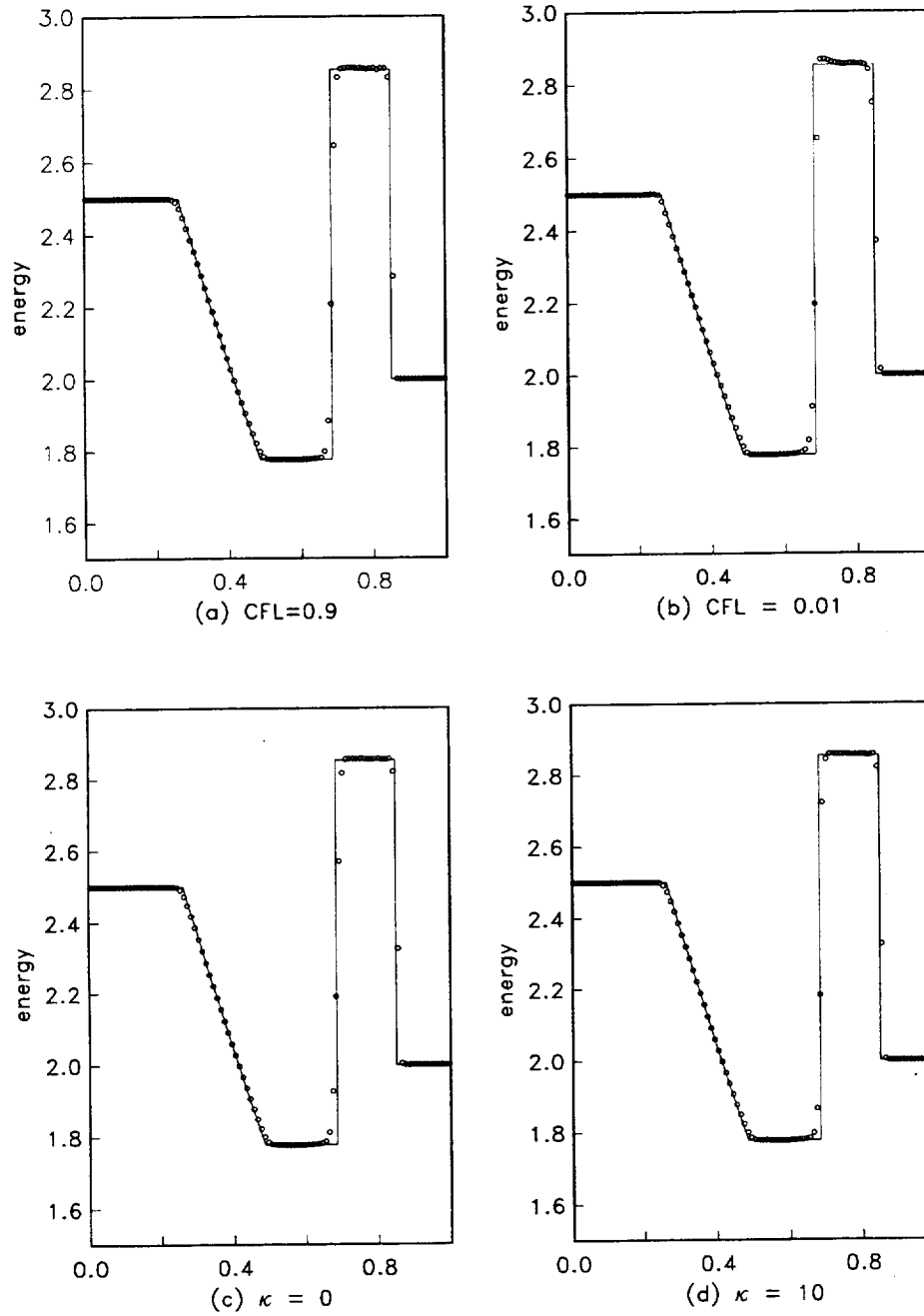


Figure 7.2. Sod's problem with different CFL numbers and κ 's.

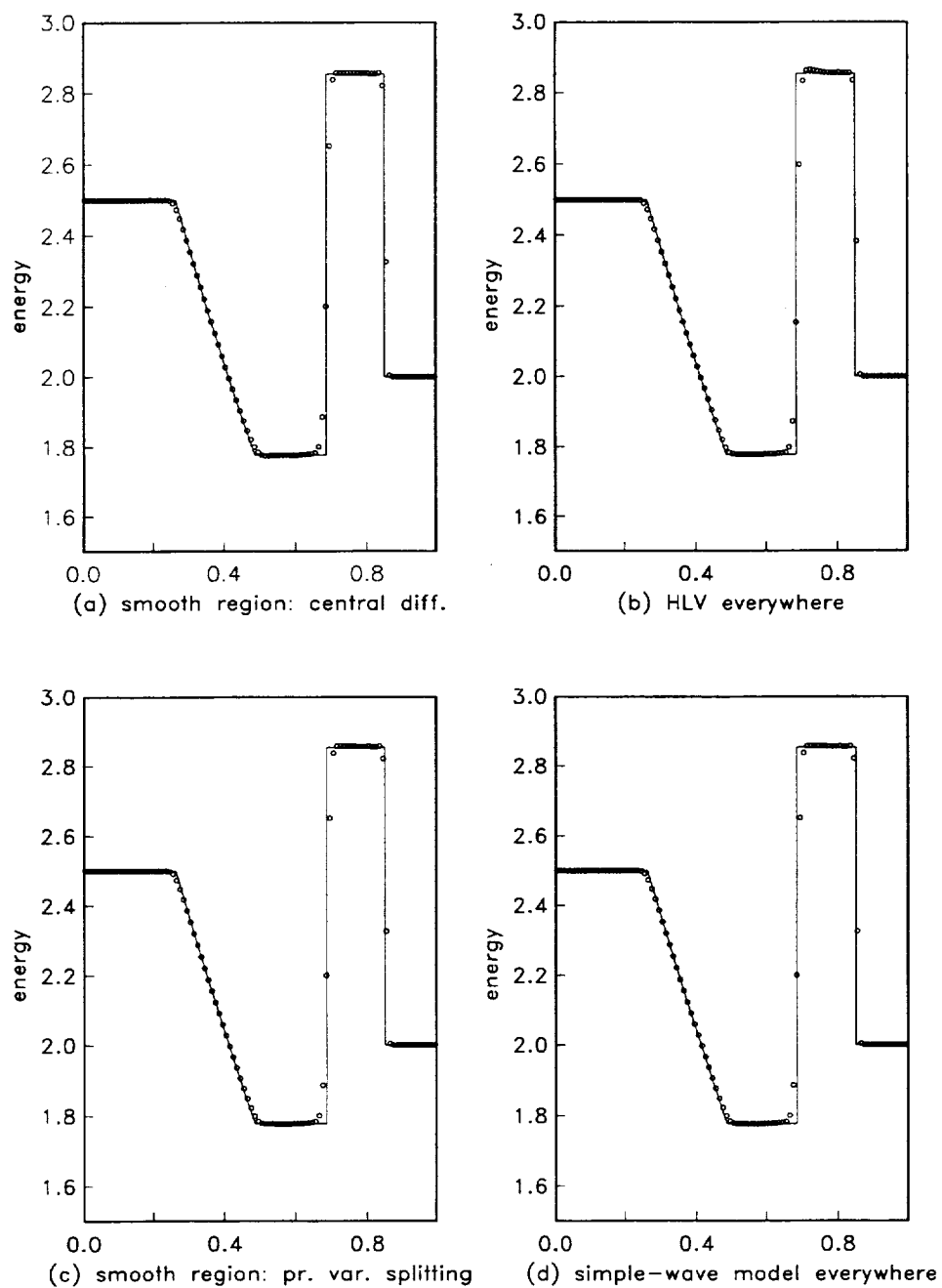


Figure 7.3. *Sod's problem with different upwind fluxes.*

of Sod because the density field has to be ‘built-up’. The results are shown in Fig. 7.4.

Figure 7.5 shows the density and velocity fields of the same problem except again (as in Fig. 7.3) we test the upwind fluxes: (a) ‘central difference’ (5.13b) at ‘smooth’ regions, and here note the bump at the tail of the expansion fan in the velocity field; (b) one-intermediate-model everywhere, and for this model, the bump is considerably smaller; (c) primitive-variable splitting at ‘smooth’ regions; and (d) simple-wave upwind flux everywhere.

The above two (and also the following) problems show that at ‘smooth’ interfaces, the one-intermediate-state model is as accurate as a more sophisticated model. It is also very economical. The flux-difference splitting with admissibility correction (or the simple-wave upwind flux) is needed only near a discontinuity.

The third problem is designed to test the admissibility condition. It is the Riemann problem corresponding to a steady Mach 3 shock, except the left and right states are reversed:

$$(\rho_L, u_L, p_L) = (3.857, 0.920, 10.333), \quad (\rho_R, u_R, p_R) = (1, 3.550, 1). \quad (7.4)$$

The solution for the final time $t = 0.09$ (after 59 time steps) is shown in Fig. 7.6.

Note that for this problem, the shock for the numerical solution is about two mesh points behind that of the exact one. To explain this phenomenon, we plot the solutions in Fig. 7.7 for the density field at time $t = 0.003$ (after 4 time steps, and the expansion fan is well captured even here), $t = 0.01$ (9 time steps), $t = 0.03$ (23 time steps), and $t = 0.04$ (27 time steps). Because the numerical solution resolves the fan first, and then the contact discontinuity, by the time the shock is formed, it is already two cells behind.

Also notice that if the flux-difference splitting without any correction is employed, then theoretically, the solution does not change with time because the fluxes are balanced. Eventually however, the solution is broken up by numerical error, but is no longer time-accurate. Next, a remark concerning the various admissibility conditions is in order. The correction for the flux-vector splitting in §5.5, that for the flux-difference splitting in §5.6, and the simple-wave upwind flux in §5.7 all produce essentially identical results for this test.

The fourth problem, due to Shu and Osher (1989), has several extrema in the

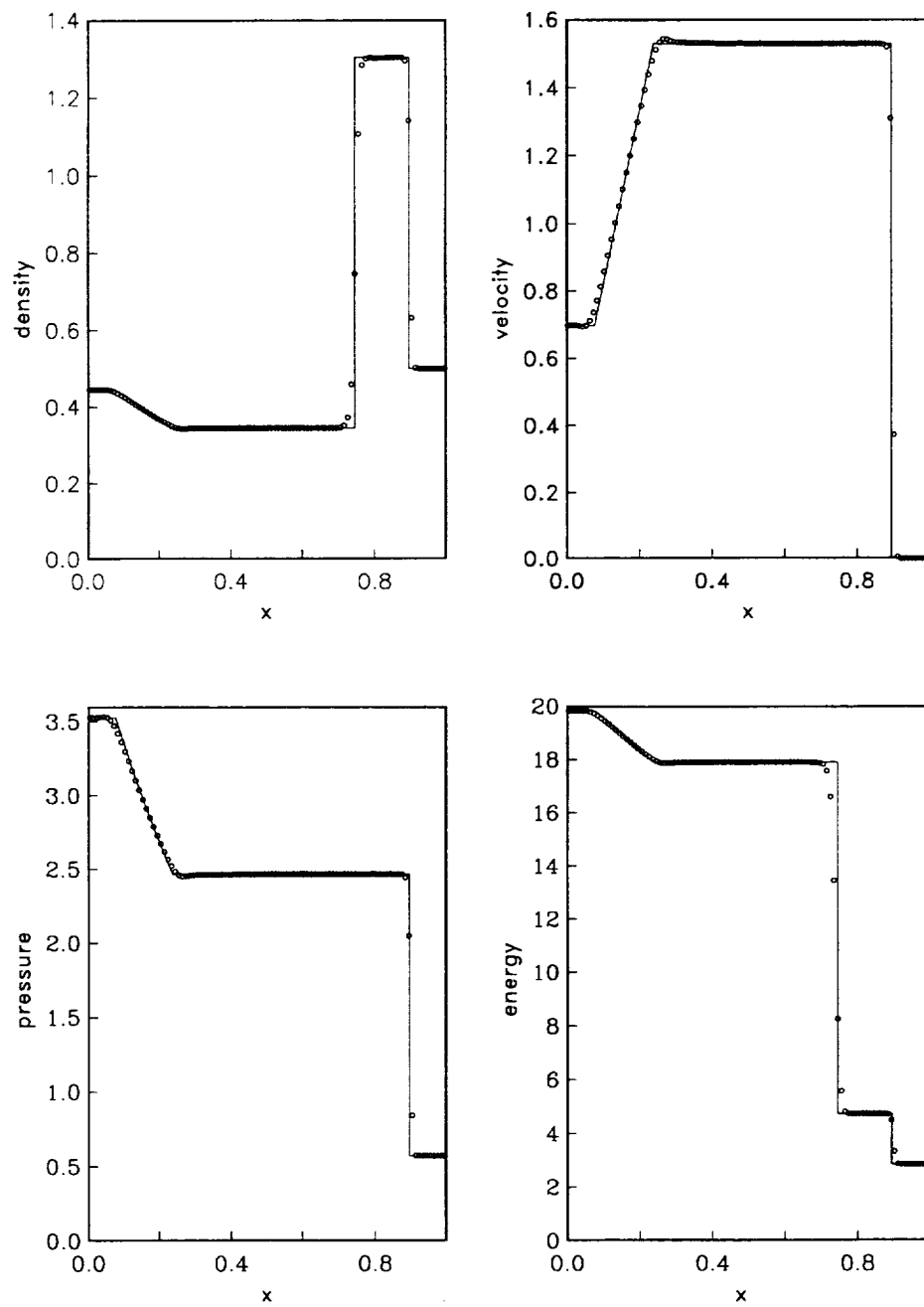


Figure 7.4. *Lax's problem.*

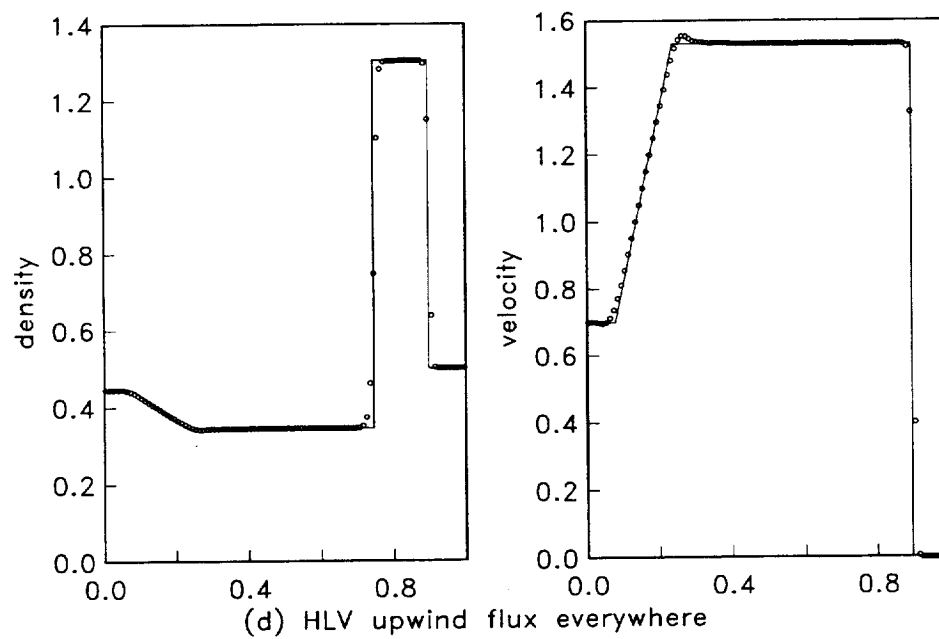
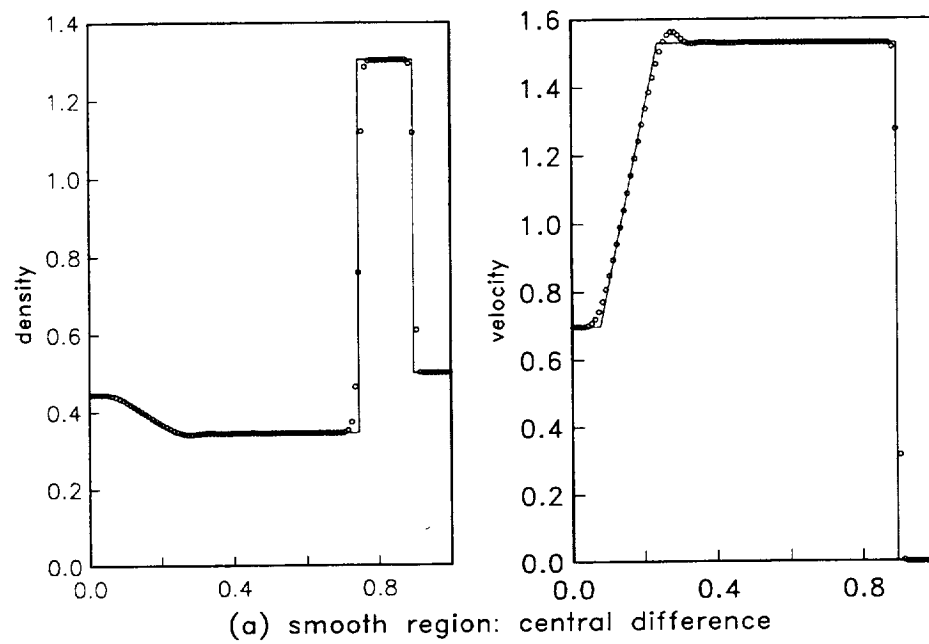


Figure 7.5. *Lax's problem with different upwind fluxes.*

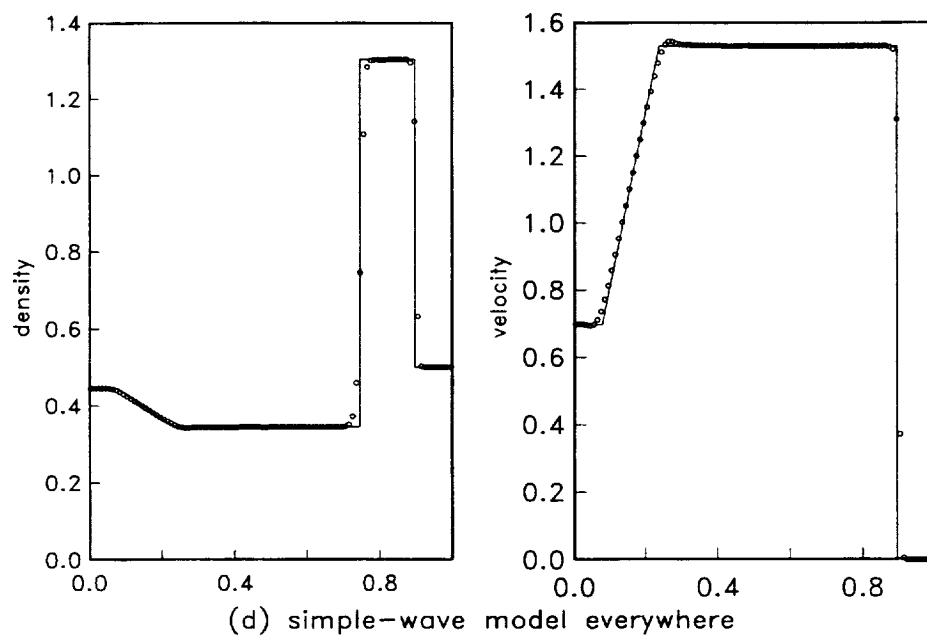
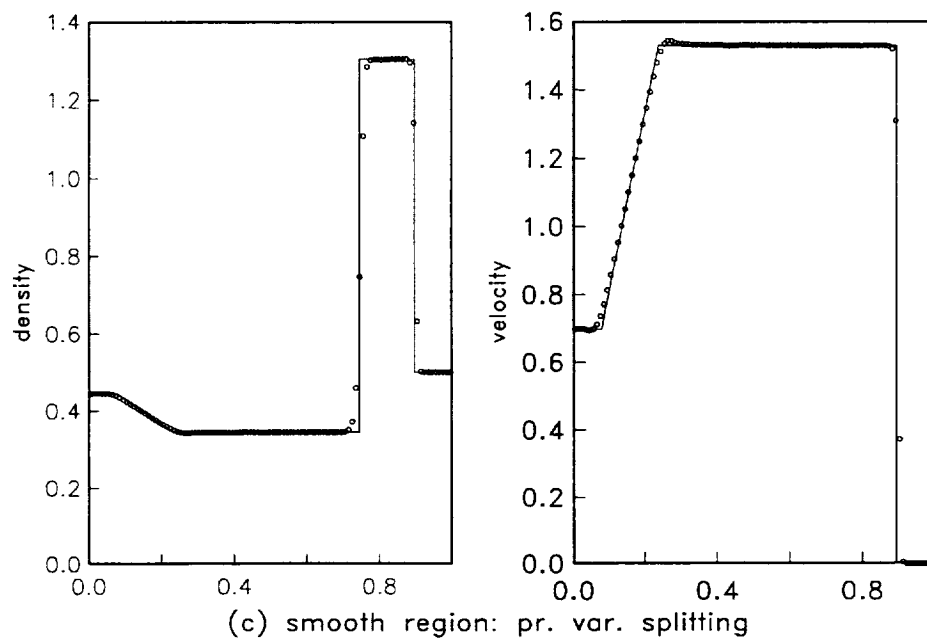


Figure 7.5 (continued). *Lax's problem with different upwind fluxes.*

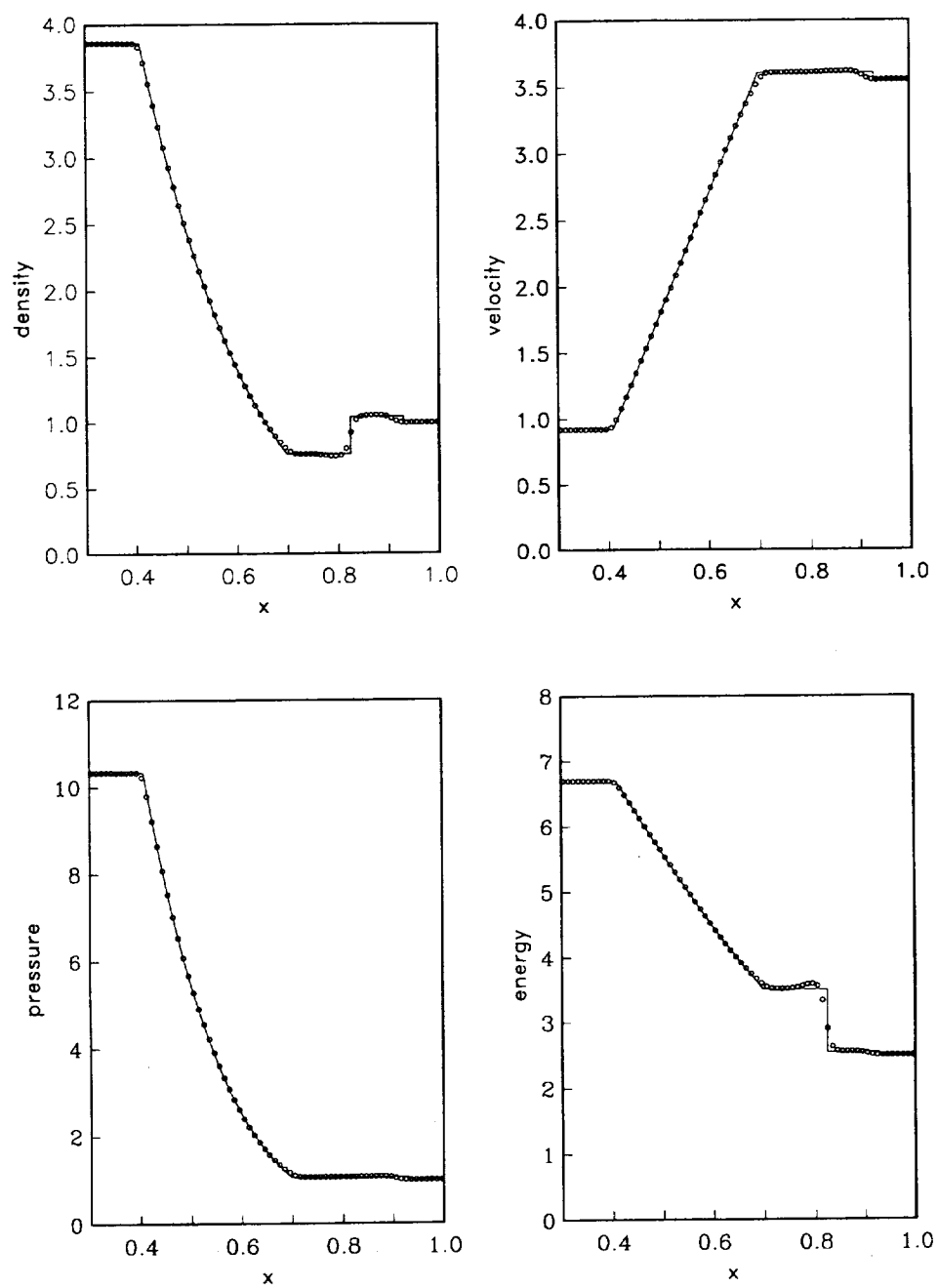


Figure 7.6. *Mach 3 expansion.*

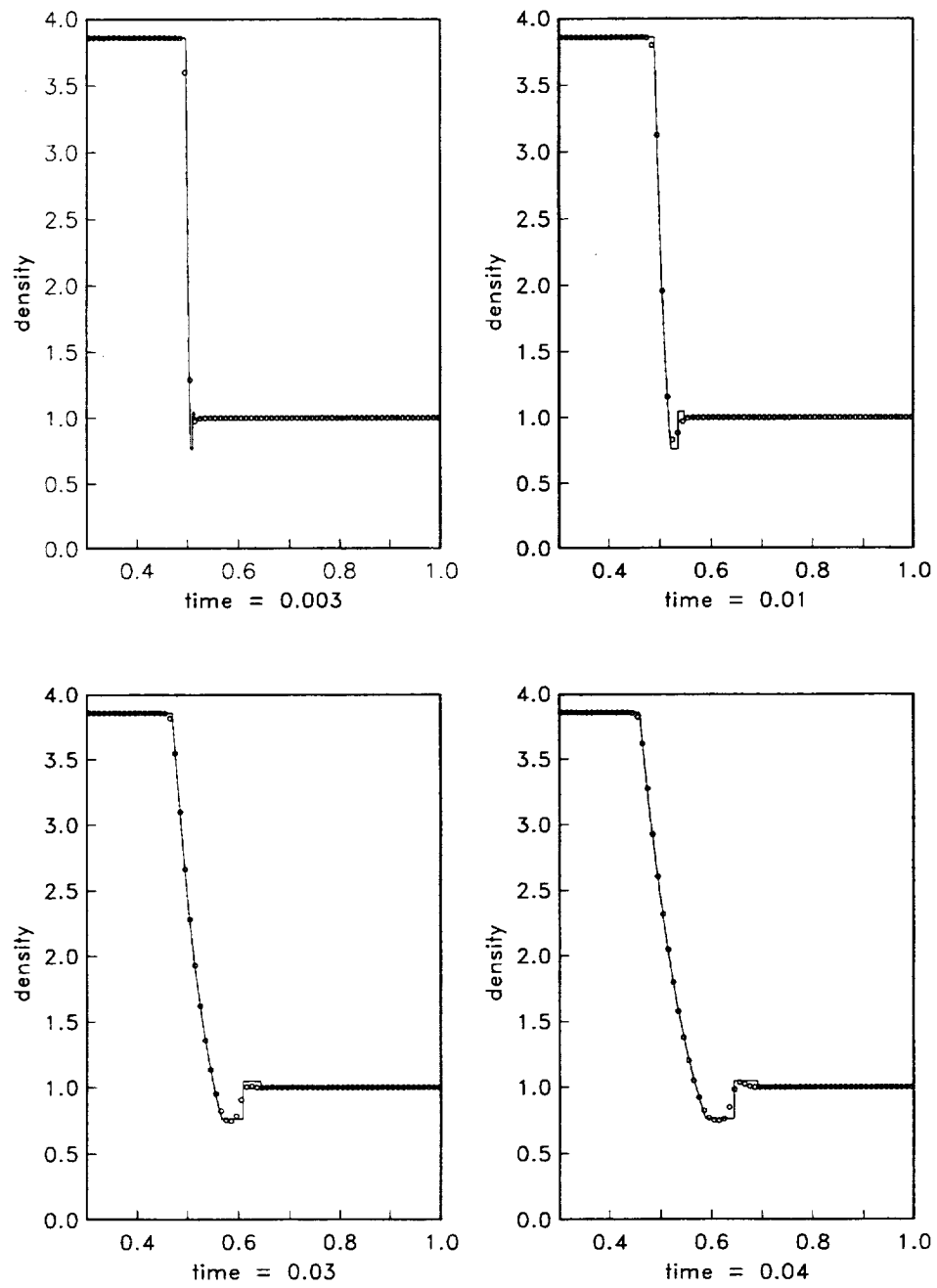


Figure 7.7. *Mach 3 expansion at various time.*

smooth regions. It is used here to test the accuracy of the schemes. In the interval $-5 \leq x \leq 5$, a moving Mach 3 shock interacts with sine waves in density as described by the following initial conditions:

$$(\rho, u, p) = \begin{cases} (3.857, 2.629, 10.333) & \text{if } x < -4, \\ (1 + 0.2 \sin 5x, 0, 1) & \text{otherwise.} \end{cases} \quad (7.5)$$

The final time is $t = 1.8$. Since the exact solution for this problem is unknown, the solid line is the solution with 1600 cells and $\kappa = 0$. Figure 7.8 shows the solutions with 400 and 800 cells, respectively ($\kappa = 5$). Note that the solution for 800 cells has essentially converged (in the sense of mesh refinement). The solutions compare favorably with those of the third-order ENO scheme (Shu and Osher 1989).

Our fifth test is the problem of two interacting blast waves (Woodward and Colella 1984). The initial condition for $0 \leq x \leq 1$ is

$$\mathbf{V} = \begin{cases} \mathbf{V}_L & \text{if } x < 0.1, \\ \mathbf{V}_M & \text{if } 0.1 \leq x < 0.9, \\ \mathbf{V}_R & \text{if } 0.9 \leq x, \end{cases} \quad (7.6)$$

where

$$\rho_L = \rho_M = \rho_R = 1, \quad u_L = u_M = u_R = 0, \quad p_L = 10^3, \quad p_M = 10^{-2}, \quad p_R = 10^2. \quad (7.7)$$

The boundaries at $x = 0$ (corresponding to the index $1/2$) and $x = 1$ (index $N + 1/2$ where N is the total number of cells) are solid walls with the following reflecting boundary conditions: at time level n (the superscript n is omitted below), we define auxiliary states $\mathbf{V}_0, \mathbf{V}_{-1}, \mathbf{V}_{-2}$ for the left boundary, and $\mathbf{V}_{N+1}, \mathbf{V}_{N+2}, \mathbf{V}_{N+3}$ for the right one by

$$\rho_{-l+1} = \rho_l, \quad u_{-l+1} = -u_l, \quad p_{-l+1} = p_l, \quad l = 1, 2, 3, \quad (7.8a)$$

$$\rho_{N+l} = \rho_{N-l+1}, \quad u_{N+l} = -u_{N-l+1}, \quad p_{N+l} = p_{N-l+1}, \quad l = 1, 2, 3. \quad (7.8b)$$

At the final time $t = 0.038$, the flow field has three contact discontinuities. The middle one, which is created by the two shocks interacting with each other, is very difficult to resolve. Even very compressive schemes such as Roe's 'Superbee' smear it (Roe and Pyke 1984).

Notice that due to the small pressure in the middle region, the reconstruction step may result in negative values for pressure and density. This problem is solved simply

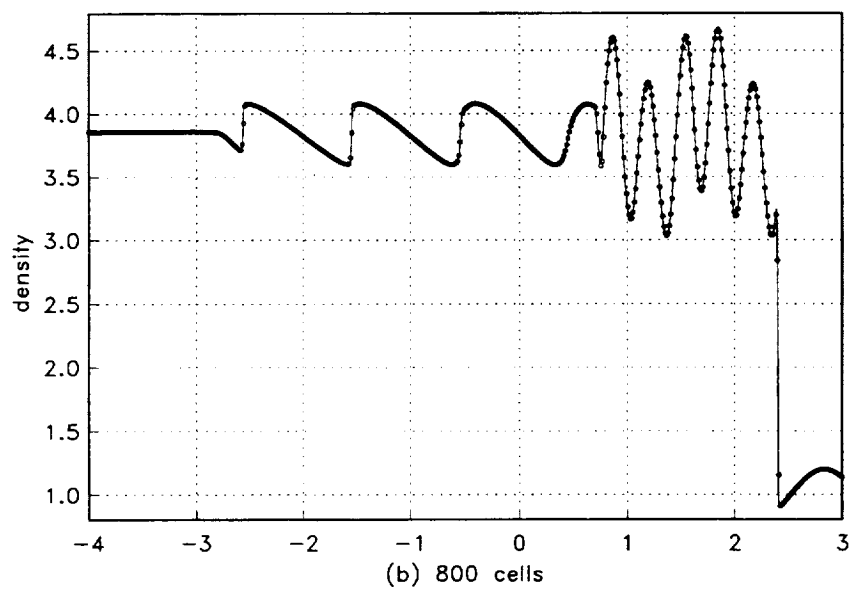
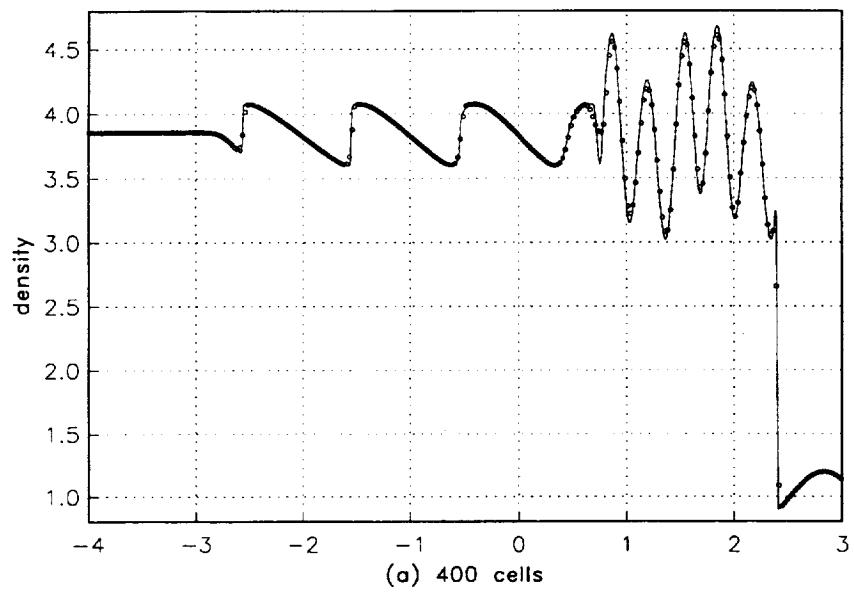


Figure 7.8. *Shu and Osher's problem.*

by replacing p by $\max(10^{-5}, p)$ (similarly for ρ) each time a new p or ρ is calculated (for ‘nonsmooth’ cells and interfaces).

Figure 7.9 shows the solution with 1200 cells, $\kappa = 12$, after 2411 time steps. In Figs. 7.9–10, this solution is represented by the continuous line since the exact solution is unknown. At the end of the calculation, there are nine ‘nonsmooth’ cells around each of the three contact discontinuities, and seven or eight around each of the two shocks. Note that again the solution compares favorably with those of higher-order schemes (Woodward and Colella 1984, Harten, Enquist, Osher, and Chakravarty 1987, Harten 1989). We also tested several values of κ between 10 and 99; the results are practically identical for this problem.

Concerning computing time, on an Iris Indigo R4000 workstation (16 Mflops, compiled with option -O2), this case takes 32 seconds. At ‘smooth’ regions, if the central difference with artificial viscosity (5.13b) is employed, there is roughly a 6 percent savings; if Roe’s flux-difference splitting (with no admissibility correction) is used, the cost is about 35 percent more. The three results are practically identical.

Figure 7.10 show the density field of the solution with (a) 600 cells and $\kappa = 16$; (b) 400 cells and $\kappa = 20$; (c) 400 cells and $\kappa = 7$, where the middle contact discontinuity is slightly smeared, but the other two still remain sharp; and (d) 400 cells and $\kappa = 5$, where the three contact discontinuities are smeared.

The last example is (essentially) the quasi one-dimensional nozzle flow problem employed by Shubin, Stephen, and Glaz (1981): the domain is $-2 \leq x \leq 10$; the area, $A(x) = 1.398 + .347 \tanh(.8x - 4)$; the supersonic inflow conditions are $\rho = 1$, $u = 1.30$, $p = 1$; the subsonic outflow, $p = 1.4$. The initial conditions are identical to the inflow conditions everywhere. Figure 7.11 shows the solution with 80 cells after 2000 iterations when machine accuracy is reached. Note that at the shock, the exact value for entropy calculated from the exact averages of density, momentum, and energy is plotted with the cross symbol. As a result, the entropy value at this point is meaningless. Also notice that the entropy field for this problem is difficult to resolve accurately; see, e.g., Hirsch (1990).

Finally, we remark that for the case of a very slowly moving shock (Roberts 1990, Lin 1991, Quirk 1993), similar to the solution by the exact Riemann solver, all upwind fluxes presented here yield oscillatory results. We omit the details. As stated by Roe

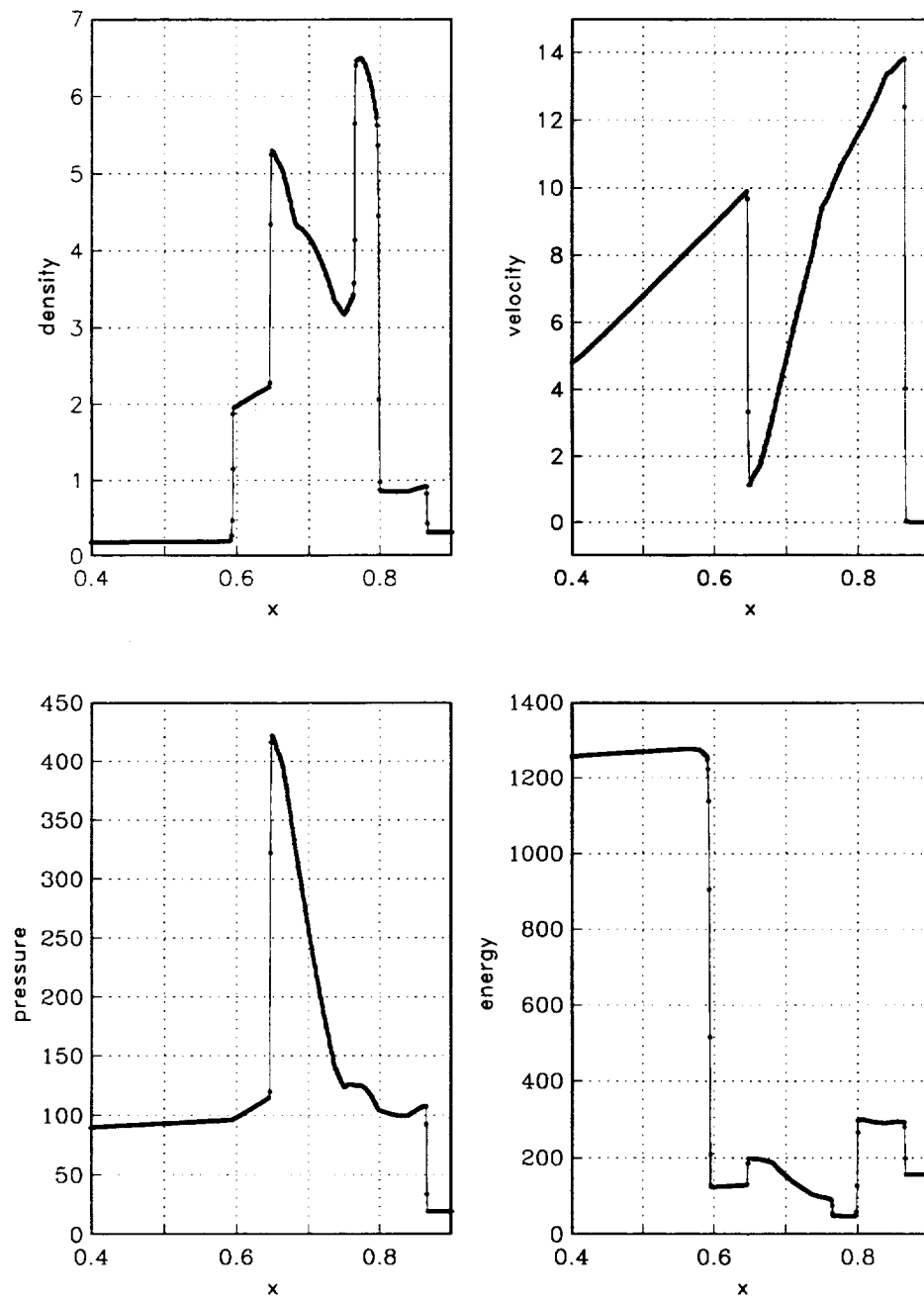


Figure 7.9. *Woodward and Colella's interacting-blast-wave problem.*

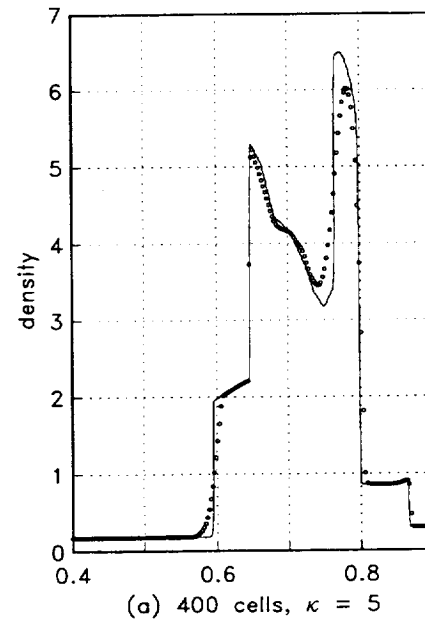
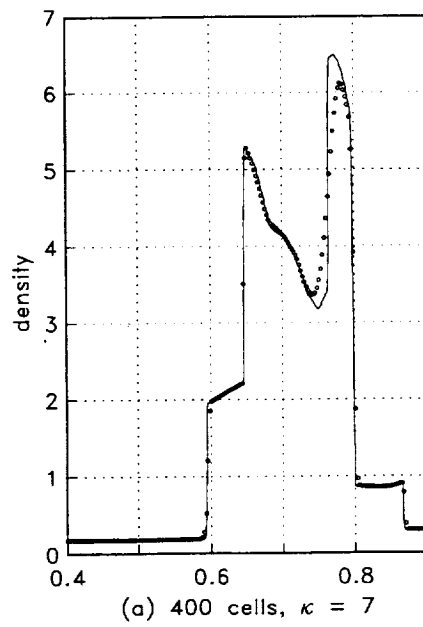
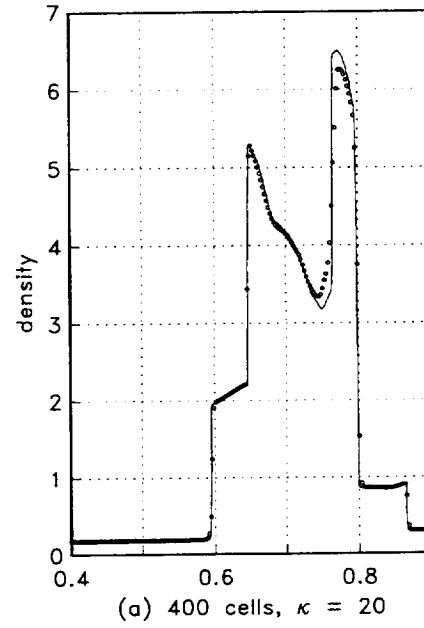
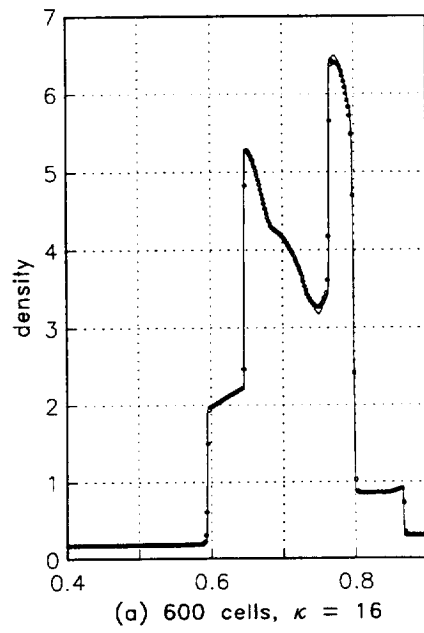


Figure 7.10. *Interacting-blast-wave problem with different mesh sizes and κ 's.*

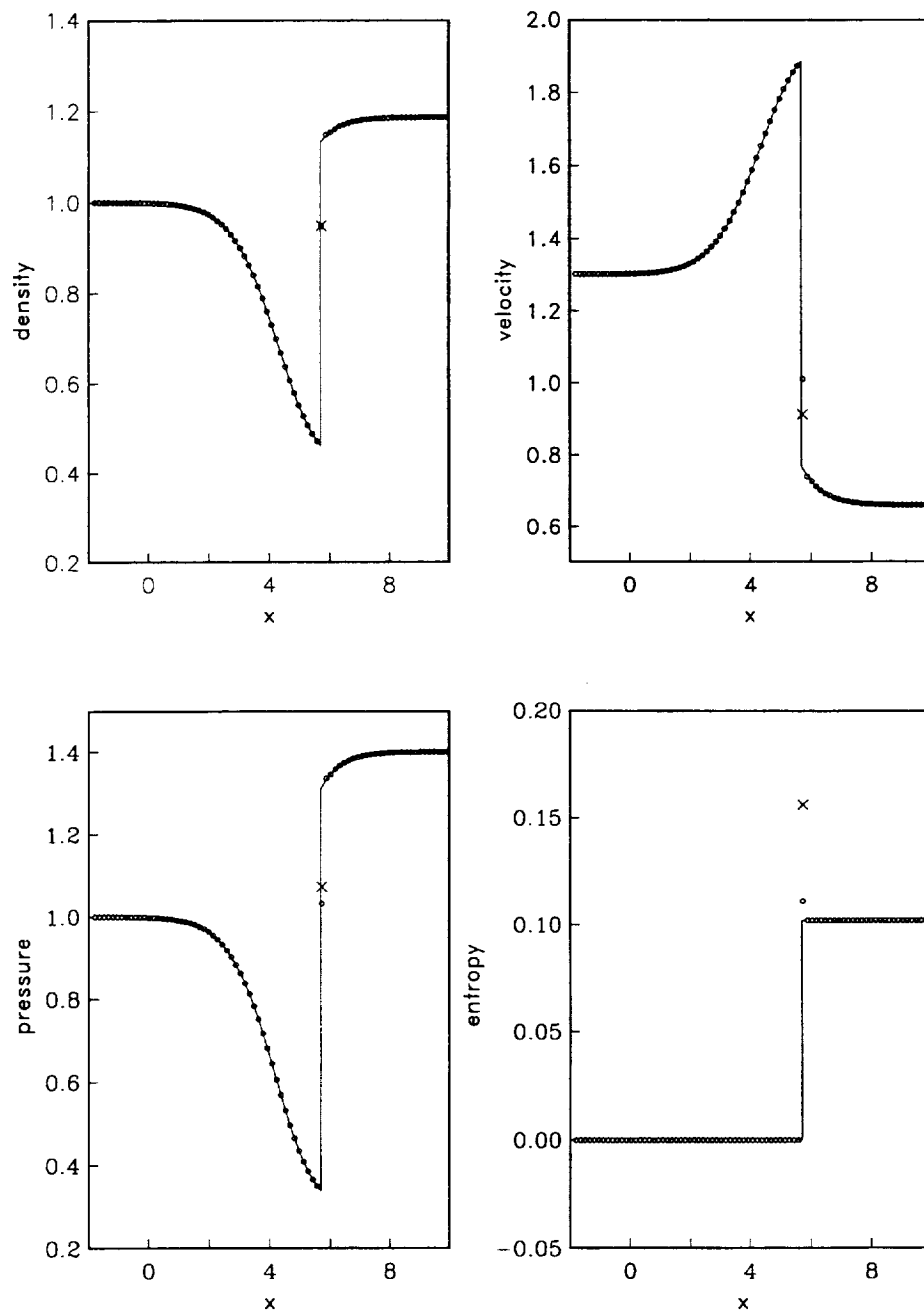


Figure 7.11. *Quasi one-dimensional nozzle flow problem.*

(1989), "The perfect Riemann solver for generating numerical schemes probably does not yet exist."

8. Discussion and conclusions. The interpolation and constraints in this paper can be extended to the piecewise-parabolic, dual-variable (scheme V in Van Leer 1977), and dual-mesh methods (Nessyahu and Tadmor 1990, Sanders and Weiser 1992). The author hopes to report on some of the results in forthcoming papers.

In summary, concerning the reconstruction step, a piecewise linear interpolation that preserves monotonicity and uniform second-order accuracy was introduced. The concept and coding of the monotonicity constraints were simplified by the use of the median function. Computational efficiency was enhanced by devising a criterion that detects the 'smooth' part of the data where the constraint is redundant. A slope-steepening technique, which can generally resolve a contact discontinuity in four cells, was presented. As for the upwind step, upwind fluxes were employed in a manner slightly different from those in the literature. They were rederived from a simple and unified point of view by approximating the Euler equations in conservation form using linearization and diagonalization. At smooth regions, the one-intermediate-state model was employed for economic reasons. (The central difference with artificial viscosity (5.13b), which is even more economical, can also be used, but for the Riemann problem of Lax, the result is not as accurate.) Near a discontinuity, either the flux-difference splitting with admissibility condition or the simple-wave upwind flux was employed for robustness. The resulting scheme is efficient and accurate.

For slow moving shocks, difficulties remain. Extensions to the multi-dimensional case, which may involve more theoretical work on constraints as well as upwind fluxes, are to be explored.

Acknowledgments. The author wishes to thank Mr. Christopher J. Steffen, Jr. and Dr. Ambady Suresh for reviewing the manuscript and for several interesting discussions.

References.

W. K. Anderson, J. L. Thomas, and B. Van Leer, *A comparison of finite volume flux vector splittings for the Euler equations*, AIAA Paper 85-0122.

- J. P. Boris and D. L. Book, *Flux-corrected transport, I. SHASTA, a fluid transport algorithm that works*, J. Comp. Phys., 11 (1973), pp. 38–69.
- P. Colella and P. Woodward, *The piecewise parabolic method (PPM) for gas-dynamical simulations*, J. Comp. Phys., 54 (1984), pp. 174–201.
- R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2 John Wiley & Sons, New York, 1962, 830 pp.
- S. F. Davis, *Simplified second-order Godunov-type methods*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 445–473.
- B. Einfeldt, *On Godunov-type methods for gas dynamics*, SIAM J. Numer. Anal., 25 (1988), pp. 294–318.
- B. Engquist and S. Osher, *One-sided difference approximations for non-linear conservation laws*, Math. Comput., 36 (1981) pp. 321–351.
- S. K. Godunov, *A finite difference method for the numerical computation of discontinuous solutions of the equations of fluid dynamics*, Mat. Sb., 47 (1959), pp. 357–393.
- J. J. Gottlieb and C. P. T. Groth, *Assessment of Riemann solvers for unsteady one-dimensional inviscid flows of perfect gas*, J. Comp. Phys., 78 (1988), pp. 437–458.
- E. Harabetian and R. Pego, *Nonconservative hybrid shock capturing schemes*, J. Comp. Phys., 105 (1993), pp. 1–13.
- A. Harten, *High resolution schemes for hyperbolic conservation laws*, J. Comp. Phys., 49 (1983), pp. 357–393.
- A. Harten, *ENO schemes with subcell resolution*, J. Comp. Phys., 83 (1989), pp. 148–184.
- A. Harten, B. Engquist, S. Osher, and S. R. Chakravarty, *Uniformly high-order accurate essentially nonoscillatory schemes. III*, J. Comp. Phys., 71 (1987), pp. 231–303.
- A. Harten, J. M. Hyman, *Self-adjusting grid methods for one-dimensional hyperbolic conservation laws*, J. Comp. Phys., 50 (1985), pp. 235–269.
- A. Harten, P. D. Lax, and B. van Leer, *On upstream differencing and Godunov-type schemes for hyperbolic conservation laws*, SIAM Rev., 25 (1983), pp. 35–61.

- A. Harten and S. Osher, *Uniformly high-order accurate nonoscillatory schemes. I*, SIAM J. Numer. Anal., 24 (1987), pp. 279–309.
- C. Hirsch, *Numerical Computation of Internal and External Flows*, Vol. 2, John Wiley & Sons, New York, 1990, 691 pp.
- M. Holt, *Numerical Methods in Fluid Dynamics*, Springer, Berlin, 1984, 273 pp. 2nd ed.
- H. T. Huynh, *Second-order accurate nonoscillatory schemes for scalar conservation laws*, Proceedings of the Sixth International Conference on Numerical Methods in Laminar and Turbulent Flows, 1989, Pineridge Press, Swansea, U.K., 1989, pp. 25–38.
- H. T. Huynh, *Accurate monotone cubic interpolation*, SIAM J. Numer. Anal., 30 (1993), pp. 57–100.
- A. Jameson, W. Schmidt, and E. Turkel, *Numerical solutions of the Euler equations by finite-volume methods using Runge-Kutta time-stepping*, AIAA Paper 81-1259
- P. D. Lax, *Weak solutions of nonlinear hyperbolic equations and their numerical computation*, Commun. Pure Appl. Math., 7 (1954) 159–193.
- P. D. Lax and B. Wendroff, *Difference schemes for hyperbolic equations with high order of accuracy*, Commun. Pure Appl. Math., 17 (1964) 381–398.
- B. P. Leonard and H. S. Niknafs, *Sharp monotonic resolution of discontinuities without clipping of narrow extrema*, Computers & Fluids, 19 (1991) 141–154.
- M.-S. Liou and C. J. Steffen, Jr., *A new flux splitting scheme*, J. Comp. Phys., 107 (1993), pp. 23–39.
- H.-C. Lin, *Dissipation additions to flux-difference splitting*, AIAA paper 91-1544, pp. 190–198.
- D-K. Mao, *A treatment of discontinuities in shock-capturing finite difference methods*, J. Comp. Phys., 92 (1991), pp. 422–455.
- H. Nessyahu and E. Tadmor, *Non-oscillatory central differencing for hyperbolic conservation laws*, J. Comp. Phys., 87 (1990), pp. 408–463.
- S. Osher, *Riemann solvers, the entropy condition, and difference approximations*, SIAM J. Numer. Anal., 21 (1984), pp. 217–235.

- S. Osher and S. R. Chakravarthy, *High resolution schemes and the entropy condition*, SIAM J. Numer. Anal., 21 (1984), pp. 955–984.
- S. Osher and F. Solomon, *Upwind difference schemes for hyperbolic systems of conservation laws*, Math. Comput., 38 (1982) pp. 339–374.
- J. J. Quirk, *A contribution to the great Riemann solver debate*, ICASE Report No. 92-64.
- T. W. Roberts, *The behavior of flux difference splitting schemes near slowly moving shock waves*, J. Comp. Phys., 90 (1990), pp. 141–160.
- P. L. Roe, *Approximate Riemann solvers, parameter vectors, and difference schemes*, J. Comp. Phys., 43 (1981), pp. 357–372.
- P. L. Roe, *Some contributions to the modelling of discontinuous flows, Large-Scale Computations in Fluid Mechanics*, Lectures on Appl. Math., 22, part 2, American Mathematical Society, Providence, RI, 1985, pp. 163–194.
- P. L. Roe, *Characteristic-based schemes for the Euler equations*, Ann. Rev. Fluid Mech., 18 (1986), pp. 337–365.
- P. L. Roe, *A survey of upwind differencing techniques*, Proceedings of the Eleventh International Conference on Numerical Methods in Fluid Dynamics, 1988, Lecture Notes in Phys., 323, Springer-Verlag, Berlin, New York, 1989, pp. 69–78.
- P. L. Roe, *Sonic flux formulae*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 611–630.
- P. L. Roe and M. J. Baines, *Assymtotic behaviour of some non-linear schemes for linear convection*, Notes on Numerical Fluid Mechanics, 7 (1984), pp 283–290. Braunschweig, Vieweg.
- P. L. Roe and J. Pyke, *Efficient construction and utilisation of approximate Riemann solutions*, Computing Methods in Applied Sciences and Engineering, Elsevier Science Publishers, Amsterdam, North-Holland, 1984, pp. 499–518.
- A. M. Rogerson and E. Meiburg, *A numerical study of the convergence properties of ENO schemes*, J. Sci. Comput., 5 (1990), pp. 151–167.
- V. V. Rusanov, *Calculation of interaction of non-steady shock waves with obstacles*, Zhur. Vychislitel'noi Matematicheskoi Fiziki, 1 (1961) pp. 267–279.

- R. H. Sanders and K. H. Prendergast, *The possible relation of the 3-kiloparsec arm to explosions*, *Astrophys. J.*, 188 (1974) pp. 489–500.
- R. Sanders and A. Weiser, *High resolution staggered mesh approach for nonlinear hyperbolic systems of conservation laws*, *J. Comp. Phys.*, 101 (1992), pp. 314–329.
- C-W. Shu, *Numerical experiments on the accuracy of ENO and modified ENO schemes*, *J. Sci. Comput.*, 5 (1990), pp. 127–149.
- C-W. Shu and S. Osher, *Efficient implementation of essentially non-oscillatory shock-capturing schemes*, II, *J. Comp. Phys.*, 83 (1989), pp. 32–78.
- G. R. Shubin, A. B. Stephens, and H. M. Glaz, *Steady shock tracking and Newton's method applied to one-dimensional duct flow*, *J. Comp. Phys.*, 39 (1981), pp. 364–374.
- G. A. Sod, *A survey of several finite difference methods for systems of non-linear hyperbolic conservation laws*, *J. Comp. Phys.*, 27 (1978), pp. 1–31.
- J. L. Steger and R. F. Warming, *Flux vector splitting of the inviscid gasdynamic equations with application to finite-difference methods*, *J. Comp. Phys.*, 40 (1981), pp. 263–293.
- A. Suresh and M.-S. Liou, *The Osher scheme for non-equilibrium reacting flows*, *International Journal for Numerical Methods in Fluids*, 15 (1992), pp. 219–232.
- P. K. Sweby, *High resolution schemes using flux limiters for hyperbolic conservation laws*, *SIAM J. Numer. Anal.*, 21 (1984), pp. 995–1011.
- E. F. Toro, *A linearised Riemann solver for the time-dependent Euler equations of gas dynamics*, *Proc. Roy. Soc. London, A* 434 (1991), pp. 683–693.
- G. D. van Albada, B. van Leer, and W. W. Roberts, Jr., *A comparative study of computational methods in cosmic gas dynamics*, *Astronom. and Astrophys.*, 108 (1982), pp. 76–84.
- B. van Leer, *Towards the ultimate conservative difference scheme. II. Monotonicity and conservation combined in a second-order scheme*, *J. Comp. Phys.*, 14 (1974), pp. 361–370.
- B. van Leer, *Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection*, *J. Comp. Phys.*, 23 (1977), pp. 276–298.

- B. van Leer, *Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method*, J. Comp. Phys., 32 (1979), pp. 101–136.
- B. van Leer, *Flux-vector splitting for the Euler equations*, Lect. Notes Phys., 170 (1982) pp. 507–512.
- B. van Leer, *Upwind-difference methods for aerodynamic problems governed by the Euler equations*, *Large-Scale Computations in Fluid Mechanics*, Lectures on Appl. Math., 22, part 2, American Mathematical Society, Providence, RI, 1985, pp. 327–336.
- R. F. Warming, R. M. Beam, and B.J. Hyett, *Diagonalization and simultaneous symmetrization of the gas-dynamic matrices*, Math. Comput., 29 (1975) pp. 1037–1045.
- P. Woodward and P. Colella, *The numerical simulation of two-dimensional fluid flow with strong shocks*, J. Comp. Phys., 54 (1984), pp. 115–173.
- H. Yang, *An artificial compression method for ENO schemes: the slope modification method*, J. Comp. Phys., 89 (1990), pp. 125–160.
- H. C. Yee, *Construction of explicit and implicit symmetric TVD schemes and their applications*, J. Comp. Phys., 68 (1987), pp. 151–179.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE November 1993		3. REPORT TYPE AND DATES COVERED Technical Memorandum
4. TITLE AND SUBTITLE Accurate Upwind Methods for the Euler Equations			5. FUNDING NUMBERS WU-505-62-52	
6. AUTHOR(S) Hung T. Huynh				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Lewis Research Center Cleveland, Ohio 44135-3191			8. PERFORMING ORGANIZATION REPORT NUMBER E-8210	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, D.C. 20546-0001			10. SPONSORING/MONITORING AGENCY REPORT NUMBER NASA TM-106388	
11. SUPPLEMENTARY NOTES Responsible person, Hung T. Huynh, (216) 433-5852.				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified - Unlimited Subject Category 64			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) A new class of piecewise linear methods for the numerical solution of the one-dimensional Euler equations of gas dynamics is presented. These methods are uniformly second-order accurate, and can be considered as extensions of Godunov's scheme. With an appropriate definition of monotonicity preservation for the case of linear convection, it can be shown that they preserve monotonicity. Similar to Van Leer's MUSCL scheme, they consist of two key steps: a reconstruction step followed by an upwind step. For the reconstruction step, a monotonicity constraint that preserves uniform second-order accuracy is introduced. Computational efficiency is enhanced by devising a criterion that detects the 'smooth' part of the data where the constraint is redundant. The concept and coding of the constraint are simplified by the use of the median function. A slope-steepening technique, which has no effect at smooth regions and can resolve a contact discontinuity in four cells, is described. As for the upwind step, existing and new methods are applied in a manner slightly different from those in the literature. These methods are derived by approximating the Euler equations via linearization and diagonalization. At a 'smooth' interface, Harten, Lax, and Van Leer's one-intermediate-state model is employed. A modification for this model that can resolve contact discontinuities is presented. Near a discontinuity, either this modified model or a more accurate one, namely, Roe's flux-difference splitting, is used. The current presentation of Roe's method, via the conceptually simple flux-vector splitting, not only establishes a connection between the two splittings, but also leads to an admissibility correction with no conditional statement, and an efficient approximation to Osher's approximate Riemann solver. These reconstruction and upwind steps result in schemes that are uniformly second-order accurate and economical at smooth regions, and yield high resolution at discontinuities.				
14. SUBJECT TERMS Upwind methods; Euler equations			15. NUMBER OF PAGES 75	
			16. PRICE CODE A04	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT	