SIAM J. SCI. STAT. COMPUT.
Vol. 5, No. 1, March 1984

0196-5204/84/0501-0001 $01.25/0

ON THE RELATION BETWEEN THE UPWIND-DIFFERENCING SCHEMES OF GODUNOV, ENGQUIST-OSHER AND ROE*

BRAM VAN LEER†

Abstract. The upwind-differencing first-order schemes of Godunov, Engquist–Osher and Roe are discussed on the basis of the inviscid Burgers equations. The differences between the schemes are interpreted as differences between the approximate Riemann solutions on which their numerical flux-functions are based. Special attention is given to the proper formulation of these schemes when a source term is present. Second-order two-step schemes, based on the numerical flux-functions of the first-order schemes are also described. The schemes are compared in a numerical experiment and recommendations on their use are included.

Key words. upwind differencing, approximate Riemann solution, conservation laws

1. Introduction. Upwind differencing, while trivial for a diagonalized hyperbolic system, is difficult to achieve when the difference scheme has to be written in conservation form. The oldest and most complicated version is due to Godunov [1], [2]; the increasing popularity of upwind differencing recently has led to a variety of simpler implementation techniques. A review of these is given by Harten, Lax and van Leer [3].

Among the recent additions to the family of upwind conservative schemes the method of Engquist and Osher [4], [5] is closest to the original Godunov scheme, while the method of Roe [6], [7] offers the greatest simplification. In the present paper the differences between these schemes are discussed on the basis of the inviscid Burgers equation. Part of the discussion covers known, but not well-known, aspects of the schemes, thus rendering the paper to some extent a review paper. For maximum clarity the presentation leans heavily on geometrical insights.

The first-order accurate schemes are explained in §§ 2, 3 and 4 for the homogeneous equation; § 6 describes how to include a source term and § 7 how to achieve second-order accuracy in a two-step format. Their rendition of a stationary shock and a transonic expansion is discussed in § 5 and illustrated in the numerical experiments of § 8. Section 9 rounds off with recommendations regarding the application of these schemes to single conservation laws and systems of conservation laws.

2. Godunov's method. Godunov's [1], [2] method for integrating a hyperbolic system of conservation laws

$$u_t + [f(u)]_x = 0 \tag{1}$$

is a scheme in conservation form:

$$(u_i^{n+1} - u_i^n)/\Delta t + \{F(u_i^n, u_{i+1}^n) - F(u_{i-1}^n, u_i^n)\}/\Delta x = 0; \tag{2}$$

here u_i^n represents the average value at time $t^n = n\Delta t$ in the computational zone centered on $x_i = i\Delta x$. The numerical flux-function $F_G(u_i^n, u_{i+1}^n)$ in the Godunov scheme

* Received by the editors June 5, 1981, and in revised form June 2, 1982. This work was supported under NASA contract NAS1-14472 while the author was in residence at ICASE, NASA Langley Research Center, Hampton, Virginia 23665.

† Leiden State University, Leiden, The Netherlands. Present address: Department of Mathematics and Computer Science, Delft University of Technology, 2600 AJ Delft, the Netherlands.

is taken to be the flux value arising at $x_{i+1/2}$ in the exact solution of the initial-value problem with a piecewise uniform initial distribution

$$u^n(x) = u_i^n, \qquad x_i - \Delta x/2 < x < x_i + \Delta x/2. \tag{3.1}$$

That is, if

$$u(x, t) = v(x/t; u_L, u_R) \tag{3.2}$$

is the (weak) similarity solution of the Riemann problem with initial values

$$u = \begin{cases} u_L, & x < 0, \\ u_R, & x > 0, \end{cases} \tag{3.3}$$

then

$$F_G(u_i^n, u_{i+1}^n) = f[v(0; u_i^n, u_{i+1}^n)]. \tag{4}$$

For Burgers' equation in the inviscid limit,

$$u_t + (\tfrac{1}{2}u^2)_x = 0, \tag{5}$$

the similarity solution to the initial-value problem (3.2) is

$$u_L \leqq u_R \text{ (expansion)} \begin{cases} v = u_L, & x/t \leqq u_L, \\ v = x/t, & u_L < x/t < u_R, \\ v = u_R, & x/t \geqq u_R, \end{cases} \tag{6.1}$$

$$u_L > u_R \text{ (shock)} \begin{cases} v = u_L, & x/t < u_S = \tfrac{1}{2}(u_L + u_R), \\ v = u_R, & x/t > u_S. \end{cases} \tag{6.2}$$

For numerical reasons it is useful to distinguish in the formula for $F_G(u_L, u_R)$ three cases:

(i) *fully supersonic/subsonic*

$$F_G(u_L, u_R) = \begin{cases} \tfrac{1}{2}u_L^2, & u_L, u_R > 0, \\ \tfrac{1}{2}u_R^2, & u_L, u_R < 0; \end{cases} \tag{7.1}$$

(ii) *transonic expansion*

$$F_G(u_L, u_R) = 0, \qquad u_L \leqq 0 \leqq u_R; \tag{7.2}$$

(iii) *transonic shock*

$$F_G(u_L, u_R) = \begin{cases} \tfrac{1}{2}u_L^2, & u_L > u_S \geqq 0 \geqq u_R, \\ \tfrac{1}{2}u_R^2, & u_L \geqq 0 > u_S > u_R. \end{cases} \tag{7.3}$$

These cases are illustrated in Figs. 1, 2 and 3. The reference to a sonic speed ($=0$) arises from the use of (5) in transonic aerodynamics.

In order to combine the formulas (7) in one compact algorithm, we follow Engquist and Osher by introducing

$$u^+ \equiv \max(u, 0), \tag{8.1}$$

$$u^- \equiv \min(u, 0), \tag{8.2}$$

$$u = u^+ + u^-, \tag{8.3}$$

$$|u| = u^+ - u^-; \tag{8.4}$$

we then have

$$F_G(u_L, u_R) = \max[\tfrac{1}{2}(u_L^+)^2, \tfrac{1}{2}(u_R^-)^2]. \tag{9}$$

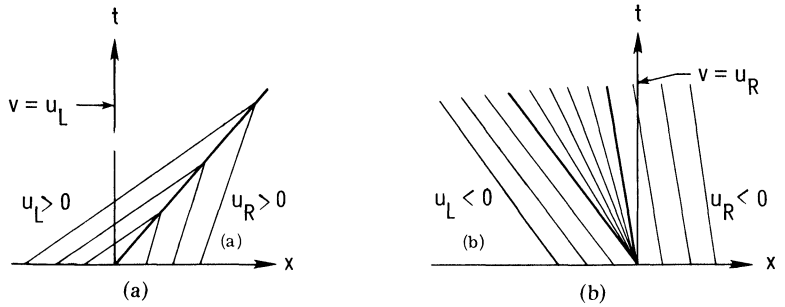

FIG. 1. *Riemann solution*: (x, t) *diagrams for case* (i), *with* $u_L > u_R > 0$ (a); $u_L < u_R < 0$ (b).

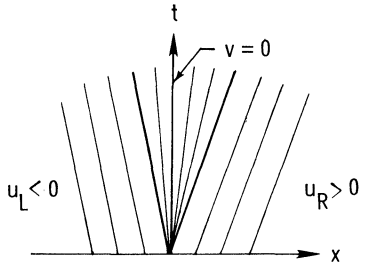

FIG. 2. (x, t) *diagram for case* (ii).

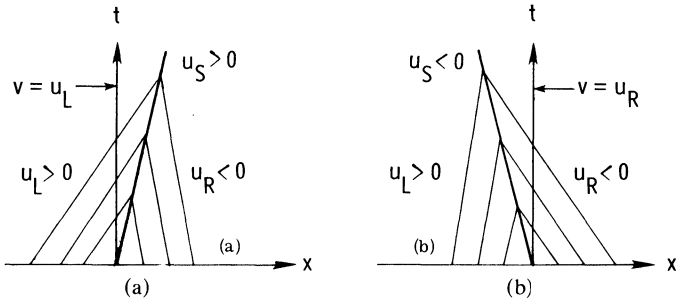

FIG. 3. (x, t) *diagram for case* (iii), *with* $u_S > 0$ (a); $u_S < 0$ (b).

3. The Engquist–Osher scheme. The Engquist–Osher [4], [5] scheme for integrating (1) also has the conservation form (2) and tacitly assumes the initial-value distribution to be (3.1); its numerical flux-function is

$$F_{EO}(u_L, u_R) = f(u_L) + \int_{u_L}^{u_R} A^-(u)\,du = f(u_R) - \int_{u_L}^{u_R} A^+(u)\,du$$
$$= \frac{1}{2}[f(u_L) + f(u_R)] - \frac{1}{2}\int_{u_L}^{u_R} |A(u)|\,du, \tag{10}$$

where the matrices $A^+(u)$, $A^-(u)$ and $|A(u)|$ are related to

$$A(u) \equiv df/du \tag{11}$$

by an extension of (8). For precise definitions and a description of the integration path used for the integrals in (10), see [5] or the review [3].

For Burgers' equation (5) the above recipe boils down to

$$F_{EO}(u_L, u_R) = \frac{1}{2}u_L^2 + \int_{u_L}^{u_R} u^- \, du = \frac{1}{2}u_R^2 - \int_{u_L}^{u_R} u^+ \, du$$
$$= \frac{1}{2}\left(\frac{1}{2}u_L^2 + \frac{1}{2}u_R^2\right) - \frac{1}{2}\int_{u_L}^{u_R} |u| \, du. \tag{12}$$

The integrals over u in (12) are defined in phase space, without reference to any mapping onto the (x, t)-plane. We may, however, introduce a mapping in the style of (3.2), namely,

$$u(x, t) = w(x/t; u_L, u_R), \qquad \min(u_L, u_R) \leqq u \leqq \max(u_L, u_R) \tag{13}$$

and consider $w(x/t; u_L, u_R)$ an approximation to the exact solution $v(x/t; u_L, u_R)$ of the Riemann problem (3.3). It follows that w is the following function of x/t:

$$w = \begin{cases} u_L, & x/t \leqq u_L, \\ x/t, & \min(u_L, u_R) < x/t < \max(u_L, u_R), \\ u_R, & x/t \geqq u_R. \end{cases} \tag{14}$$

In relating $F_{EO}(u_L, u_R)$ to $w(0; u_L, u_R)$ after the example of (4), caution is required, since, for $u_L > u_R$, w becomes multivalued in the domain $u_L \geqq x/t \geqq u_R$. Specifically, we have three branches

$$w^{(1)} = u_L, \quad w^{(2)} = x/t, \quad w^{(3)} = u_R, \qquad u_L \geqq x/t \geqq u_R. \tag{15}$$

The picture associated with this case is that of an overturned centered compression wave or folded characteristic field (see Figs. 4 and 5); in the exact Rieman solution (6) such a wave would be replaced by a shock discontinuity.

The proper formula for $F_{EO}(u_L, u_R)$ in terms of $w(x/t; u_L, u_R)$, equivalent to (12), is

$$F_{EO}(u_L, u_R) = \sum_k (-1)^{k-1} \frac{1}{2}\{w^{(k)}(0; u_L, u_R)\}^2, \tag{16}$$

where the sum is taken over all branches present at $x/t = 0$.

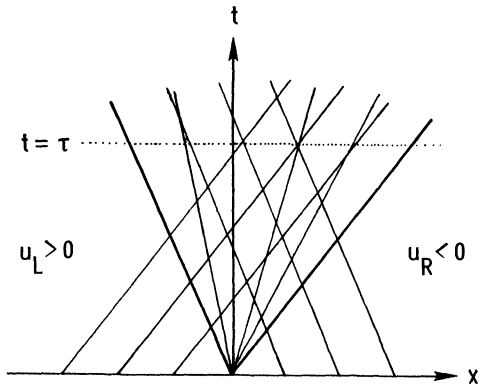

FIG. 4. *Approximate Riemann solution in the Engquist–Osher scheme*: (x, t) *diagram for case* (iii). *The multivalued solution at* $t = \tau$ *is displayed in Fig.* 5.

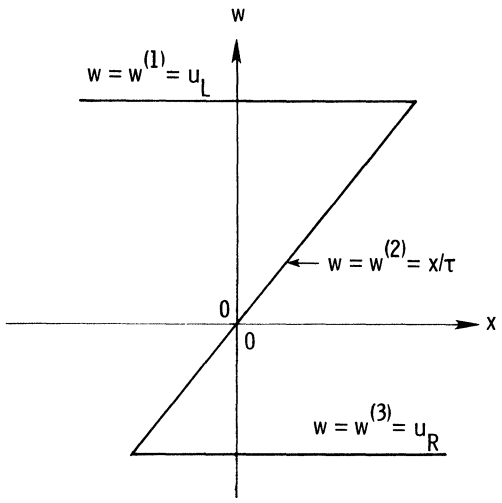

FIG. 5. *Approximate Riemann solution at* $t = \tau$.

In the cases (i), (ii) and (iii), distinguished earlier in (7) for $F_G(u_L, u_R)$, (12) or (16) yields

$$(17)\quad \begin{aligned} &\text{(i)} && F_{EO}(u_L, u_R) = F_G(u_L, u_R) = \begin{cases} \tfrac{1}{2}u_L^2, & u_L, u_R > 0, \\ \tfrac{1}{2}u_R^2, & u_L, u_R < 0, \end{cases} \\ &\text{(ii)} && F_{EO}(u_L, u_R) = F_G(u_L, u_R) = 0, && u_L \leqq 0 \leqq u_R, \\ &\text{(iii)} && F_{EO}(u_L, u_R) = \tfrac{1}{2}u_L^2 + \tfrac{1}{2}u_R^2 \neq F_G(u_L, u_R), && u_L \geqq 0 \geqq u_R. \end{aligned}$$

As indicated in [4], these formulas combine into

$$(18)\qquad F_{EO}(u_L, u_R) = \tfrac{1}{2}(u_L^+)^2 + \tfrac{1}{2}(u_R^-)^2.$$

The difference between the Godunov and Engquist–Osher schemes lies entirely in the treatment of a transonic compression (iii). The latter scheme is the simpler one, since it does away with one test, namely, a test for the sign of u_S or for the maximum of $\tfrac{1}{2}(u_L^+)^2$ and $\tfrac{1}{2}(u_R^-)^2$.

4. Roe's method. Roe's [6], [7] method for integrating (1) again has the conservation form (2) and employs the initial-value distribution (3.1). Its numerical flux-function is defined as

$$(19)\qquad F_R(u_i^n, u_{i+1}^n) = f[w(0; u_i^n, u_{i+1}^n)],$$

where $w(x/t; u_L, u_R)$ is the exact solution of the Riemann problem (3.3) for the locally linearized system

$$(20.1)\qquad w_t + A(u_L, u_R)w_x = 0.$$

The approximate Jacobian $A(u_L, u_R)$ is constructed such as to satisfy the discrete version of (11)

$$(20.2)\qquad f(u_R) - f(u_L) = A(u_L, u_R)(u_R - u_L).$$

For a scalar equation like (5), (20.2) uniquely determines $A(u_L, u_R)$:

$$A(u_L, u_R) = \tfrac{1}{2}(u_L + u_R) = u_S; \tag{21}$$

thus, the Riemann problem for the linear equation (20.1) with (21) has the similarity solution

$$w(x/t, u_L, u_R) = \begin{cases} u_L, & x/t < u_S, \\ u_R, & x/t > u_S. \end{cases} \tag{22}$$

In all cases (i), (ii) and (iii) this yields

$$F_R(u_L, u_R) = \begin{cases} \tfrac{1}{2}u_L^2, & u_S > 0, \\ \tfrac{1}{2}u_R^2, & u_S < 0. \end{cases} \tag{23}$$

The difference with Godunov's method lies in the treatment of an expansion: where the exact Riemann solution, used in Godunov's method, would include an expansion fan, Roe's method puts in a so-called expansion shock (see Fig. 6).

The numerical flux-function (23) deviates from the Godunov flux (7) only in case (ii), when the expansion is transonic. The computational simplification is, again, the elimination of one test.

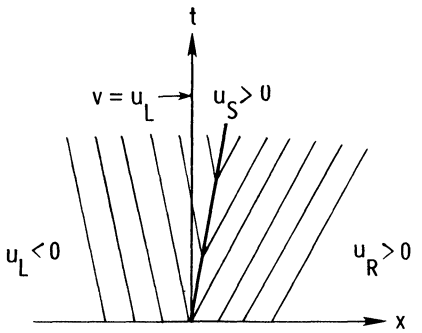

FIG. 6. *Approximate Riemann solution in Roe's scheme*: (x, t) *diagram for case* (iii), *with* $u_S > 0$, *showing the expansion shock.*

Rewriting (23) as an approximation to the Engquist–Osher flux (12),

$$F_R(u_L, u_R) = \tfrac{1}{2}(\tfrac{1}{2}u_L^2 + \tfrac{1}{2}u_R^2) - \tfrac{1}{2}|\tfrac{1}{2}(u_L + u_R)|(u_R - u_L), \tag{24}$$

we see that, for Burgers' equation, Roe's scheme is identical to the scheme of Murman and Cole [8], used in transonic aerodynamics. The latter is known to occasionally yield numerical results that include a (physically inadmissible) expansion shock. This is a direct consequence of the admission of an expansion shock in the underlying approximate Riemann solution.

Roe [14] has proposed a modification in the numerical flux-function for a transonic expansion that not only does away with expansion shocks, but may be an improvement over the flux-function in the Godunov and Engquist–Osher schemes. This is discussed at the end of the next section.

The problem of preventing inadmissible discontinuities in the use of upwind schemes has been addressed in general by Harten and Lax [15] and for gas dynamics by Colella [16].

5. Steady-shock and sonic-point representation. The fluxes in the schemes of Godunov and Engquist–Osher differ only on meshes where the data constitute a transonic compression; therefore, numerical results from these schemes differ only if a transonic shock, in particular, a stationary shock, is present. Likewise, the results of Roe's scheme will differ from those of Godunov's scheme only if a transonic expansion is present.

For Burgers' equation (5), Godunov's scheme admits the following stationary discrete representation of a stationary shock connecting the states $u_L > 0$ and $u_R = -u_L < 0$:

$$\begin{aligned} &u_i = u_L, && i \leqq -1, \\ &u_0 = u_M, && u_L \geqq u_M \geqq u_R, \\ &u_i = -u_L, && i \geqq 1. \end{aligned} \tag{25}$$

This is the only stationary distribution admitted by the scheme.

The interior value u_M indicates the subgrid position x_S of the shock; by conservation we have (see Fig. 7a),

$$u_L(x_S + \tfrac{1}{2}\Delta x) + u_R(\tfrac{1}{2}\Delta x - x_S) = u_M \Delta x \tag{26.1}$$

or

$$x_S = \tfrac{1}{2}\Delta x\, u_M / u_L. \tag{26.2}$$

A shock standing exactly on the boundary of a zone will be represented without any interior value.

The Engquist–Osher scheme admits discrete steady shock-profiles with, in general, two interior states:

$$\begin{aligned} &u_i = u_L, && i \leqq -1, \\ &u_0 = u_M, && u_L \geqq u_M \geqq 0, \\ &u_i = u_N, && 0 \geqq u_N \geqq -u_L, \\ &u_i = -u_L, && i \geqq 2, \end{aligned} \tag{27.1}$$

where u_M and u_N are constrained by

$$\tfrac{1}{2}u_M^2 + \tfrac{1}{2}u_N^2 = \tfrac{1}{2}u_L^2. \tag{27.2}$$

Again, this is the only stationary distribution admitted by the scheme.

The pair of interior states indicates the subgrid position of the shock; by conservation we have (see Fig. 7b):

$$u_L(x_S + \tfrac{1}{2}\Delta x) + u_R(\tfrac{3}{2}\Delta x - x_S) = (u_M + u_N)\Delta x \tag{28.1}$$

or

$$x_S = \tfrac{1}{2}\Delta x(u_M + u_N + u_L)/u_L. \tag{28.2}$$

A shock standing exactly in the middle of a zone will be represented with only one interior value.

The representation of a stationary shock by Roe's scheme is the same as by Godunov's scheme, at least for Burgers' equation. However, Roe's scheme also admits as a stationary solution the expansion shock

$$\begin{aligned} &u_i = u_L < 0, && i \leqq -1, \\ &u_i = -u_L > 0, && i \geqq 0. \end{aligned} \tag{29.1}$$

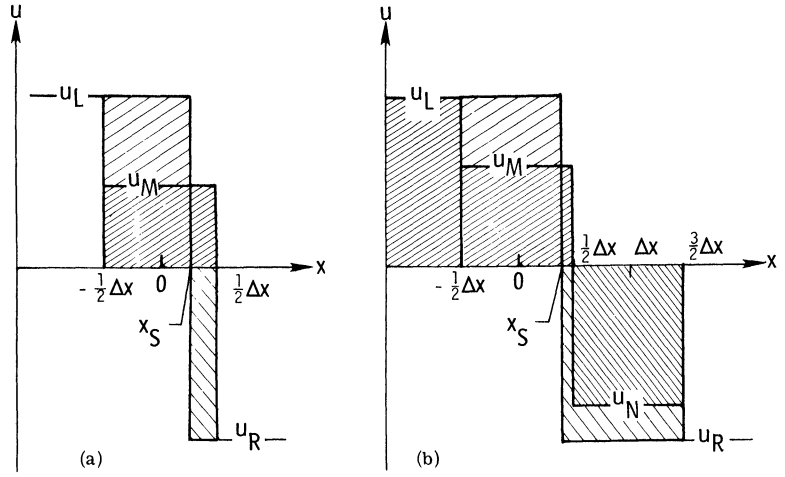

FIG. 7. *Fitting a shock discontinuity to stationary discrete shocks yielded by the schemes of Godunov or Roe* (a) *and Engquist–Osher* (b). *The numerical values* u_M *and* u_N *represent the average value of* u *in the intervals* $(-\frac{1}{2}\Delta x, \frac{1}{2}\Delta x)$ *and* $(\frac{1}{2}\Delta x, \frac{3}{2}\Delta x)$. *A subgrid distribution representing a shock transition from* u_L *to* u_R *at* x_S *must have the same integral as the numerical solution. In case* (a) *the integrals from* $-\frac{1}{2}\Delta x$ *to* $\frac{1}{2}\Delta x$ *are compared, in case* (b) *from* $-\frac{1}{2}\Delta x$ *to* $\frac{3}{2}\Delta x$.

A transonic shock profile, steady or not, obtained with any of the upwind schemes, has the property that the interior zones cannot influence the exterior solution (see Fig. 8a). Inversely, in a transonic expansion computed with the Roe–Murman–Cole scheme, the zone with the value closest to the sonic value cannot be influenced by the rest of the grid (see Fig. 8b). Once established by transient waves, the value in this zone changes no more, and it fixes the amplitude of the steady expansion shock that will remain.

Colella [16] has pointed out that, in the approximate solution to the Riemann problem, two steps can be distinguished. The first step is to determine the speeds and amplitudes of the finite-amplitude waves; the second step is to compute the full solution as a function of x/t. It is only in the second step that the method of Roe errs. A rarefaction wave must always be given a finite spread; for Burgers' equation this boils down to replacing Roe's scheme by Godunov's.

The remedy that Roe [14] proposes cannot be formulated as a change in the approximate Riemann solution. It is based on regarding the initial values as nodal values of a smooth distribution, rather than zone averages of a piecewise uniform distribution. For neighboring zones L and R enclosing a sonic point we may write down the following upwind formulas:

$$\left(\frac{\partial u}{\partial t}\right)_L = -u_L(u_R - u_L)/\Delta x, \qquad u_L \leqq 0 \leqq u_R, \tag{29.2}$$
$$\left(\frac{\partial u}{\partial t}\right)_R = -u_R(u_R - u_L)/\Delta x,$$

from which follows

$$(\partial u/\partial t)_L/(\partial u/\partial t)_R = u_L/u_R. \tag{29.3}$$

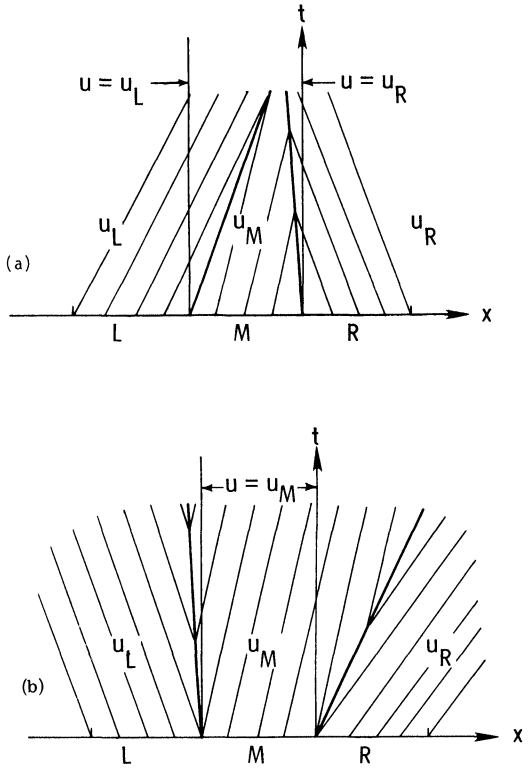

FIG. 8. *Two* (x, t) *diagrams showing the influence, through Godunov's scheme, of the adjacent zones on a zone inside a shock* (a) *and the influence, through Roe's scheme, of an almost sonic zone on the adjacent zones* (b). *In* (a), *the fluxes at the boundaries of zone M do not depend on* u_M; *in* (b) *they depend only on* u_M.

The latter relation can be forced onto the numerical flux $F_{tx}(u_L, u_R)$ for a transonic expansion:

$$\{\tfrac{1}{2}u_R^2 - F_{tx}(u_L, u_R)\}/\{F_{tx}(u_L, u_R) - \tfrac{1}{2}u_L^2\} = u_L/u_R, \tag{29.4}$$

leading to

$$F_{tx}(u_L, u_R) = \tfrac{1}{2}u_L u_R, \qquad u_L \leqq 0 \leqq u_R. \tag{29.5}$$

This unphysical flux (it is negative!) breaks down an expansion shock *faster* than the zero flux dictated by the exact Riemann solution. This is a direct consequence of the underlying assumption of smooth initial values.

While not fitting into the general framework of approximate Riemann solutions, the above approach has its appeal. The piecewise uniform values associated with a Riemann problem are not necessarily a good representation of the solution near a sonic point. In particular, the errors in the signal speed $df(u)/du$ $(= u)$ locally are of the order of the speed itself. The sonic point is always moved to a zone boundary and, in consequence, the flux gradient computed for each bordering zone becomes independent of the true gradient of the solution. In numerical solutions obtained with first-order schemes this shows up as a transonic expansion shock with amplitude $O(\Delta x)$. For a second-order scheme the effect disappears or reduces to $O[(\Delta x)^3]$.

In contrast, first-order accurate solutions obtained with the transonic flux (29.5) do not show these weak expansion shocks.

6. Inclusion of a source term. When approximating the equation

$$u_t + [f(u)]_x = s(x) \tag{30}$$

with any upwind scheme, it is not sufficient to add the source term to the right-hand side of (2); we must also change the initial-value distribution implied in our numerical model. Instead of assuming that it is piecewise uniform, as in (3.1), we rather take it to be *piecewise stationary*, that is,

$$[f(u^n(x))]_x = s(x), \qquad x_i - \Delta x/2 < x < x_i + \Delta x/2, \tag{31.1}$$

with

$$\frac{1}{\Delta x}\int_{x_i-\Delta x/2}^{x_i+\Delta x/2} u^n(x)\,dx = u_i^n. \tag{31.2}$$

This was first proposed by Liu [17] in constructing a random-choice method for the inhomogeneous conservation law (30).

The extension preserves a fundamental property of the homogeneous scheme, namely, that a zone-average can change only through the finite-amplitude waves entering from the zone boundaries. Moreover, if the numerical solution globally tends to a steady state, it will be the zone-averaged exact steady state almost everywhere (that is, provided that the scheme, like the three considered here, can render a shock transition in a finite number of zones). We shall come back to this further below.

As before, the initial values on either side of a zone boundary become the arguments of the numerical flux-function. Transient effects caused by a shock wave returning to the zone boundary under the influence of the source term, or by a curved transonic rarefaction fan, are ignored in order to keep the flux-function independent of Δt. The full scheme reads

$$(u_i^{n+1} - u_i^n)/\Delta t + \{F(u^n_{(i+1/2)-}, u^n_{(i+1/2)+}) - F(u^n_{(i-1/2)-}, u^n_{(i-1/2)+})\}/\Delta x = s_i^n, \tag{32.1}$$

with

$$s_i \equiv \frac{1}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}} s(x)\,dx. \tag{32.2}$$

For the inhomogeneous Burgers equation

$$u_t + (\tfrac{1}{2}u^2)_x = s(x) \tag{33}$$

the initial-value representation must satisfy

$$\frac{1}{2}[u^n(x)]^2 = \frac{1}{2}[u^n(x_{(i-1/2)+})]^2 + \int_{x_{i-1/2}}^{x} s(x')\,dx', \qquad x_i - \Delta/2 < x < x_i + \Delta/2, \tag{34}$$

under the constraint (31.2). This constraint, however, is not strong enough to define a unique distribution in each zone, since the distribution may (and, in some zones, must) contain a discontinuity. Uniqueness can be achieved by selecting the distribution with, say, the weakest possible shock.

If we take Δx small enough to ensure that $s(x)$ does not change sign more than once in any zone, the following algorithm for the piecewise construction of $u^n(x)$ is adequate. First, determine for the zone under consideration the smallest continuous solution (see Fig. 9). Let ω_i^n be the average value on the positive branch of this solution; then the stationary distribution inside zone i will be continuous if

$$|u_i^n| \geqq \omega_i^n. \tag{35}$$

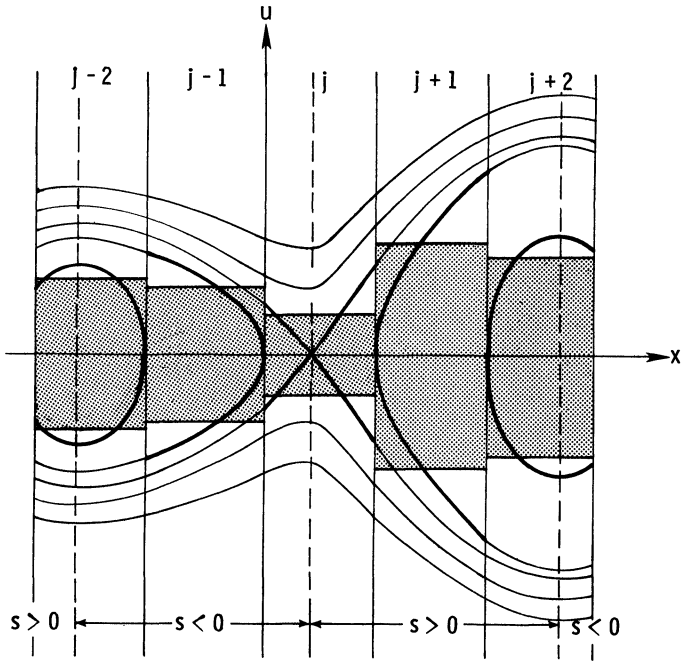

FIG. 9. *Stationary solutions of* (30) *for some source term* $s(x)$. *In each zone the heavy line traces the positive and the negative branch of the smallest continuous steady solution. The average value on these branches is indicated by the upper and lower boundary of the shaded areas. If the average value in a zone falls in the range of the shaded area, a continuous stationary solution cannot be realized.*

If this condition is satisfied, we search, iteratively, for the continuous distribution with the proper zone average. If it is violated, the stationary distribution will include parts of the upper and lower branch of the smallest continuous distribution, connected by a shock positioned so as to achieve the proper average value (see Fig. 10).

When using the upwind scheme to approach a globally stationary solution of (33), any zone containing a sonic point must be treated with special care, since the chance of numerically realizing the exact transonic structure without a shock is zero.

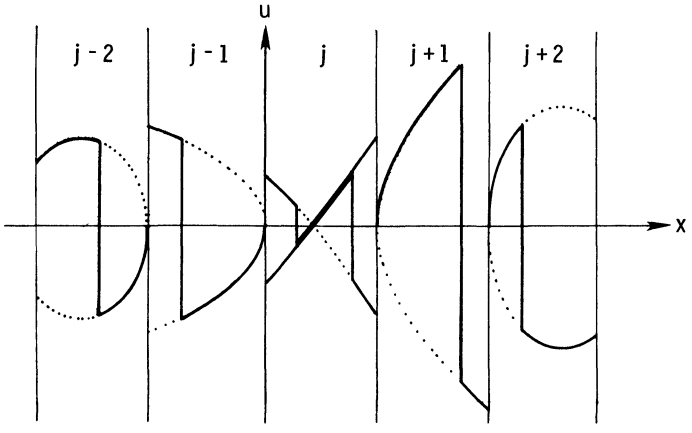

FIG. 10. *Examples of steady structure with a discontinuity, in the zones shown in Fig.* 9. *A (steady) shock connects the upper and lower branch of the smallest continuous steady solution for the zone considered. In zone j two different possibilities are shown.*

Specifically, in order to prevent the zone-boundary values from flipping sign, making global convergence impossible, smoothing may be introduced (see Fig. 11).

In a zone containing a stationary shock, no particular interior structure is needed to ensure global convergence, since the zone cannot influence its neighbors (see earlier Fig. 8a). After convergence one may insert the proper structure by enforcing continuity across the zone boundaries (see Fig. 12).

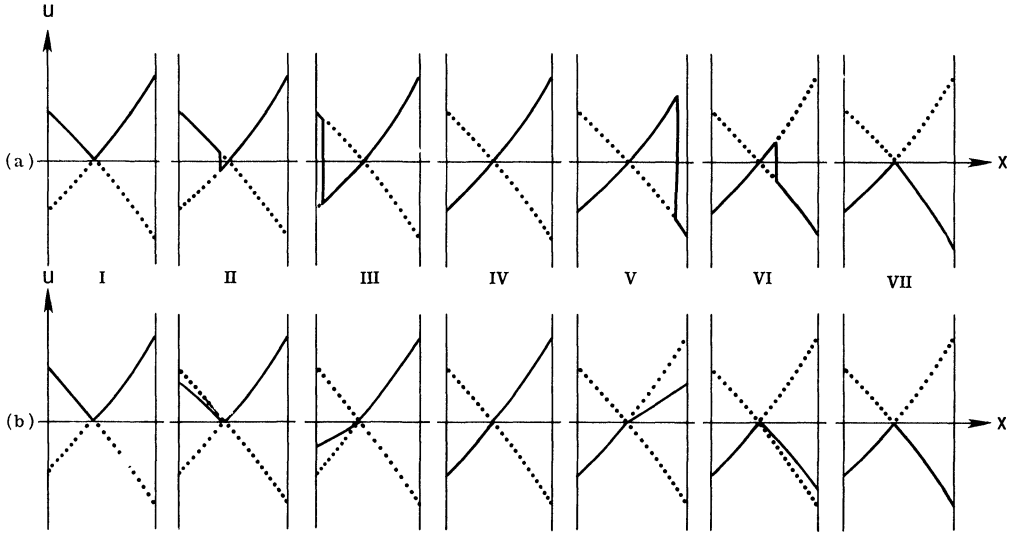

FIG. 11. (a) *A sequence of stationary structures close to the transonic expansion (graph* IV), *ordered by decreasing zone-average. Going from* III *to* IV, *and from* IV *to* V, *the slight change in zone-average causes one boundary value to flip its sign.* (b) *The solutions have been smoothed on the side where the shock occurs, so that the boundary values now vary continuously with the zone-average. The structures are not stationary but will allow convergence to the smooth transonic solution* IV.

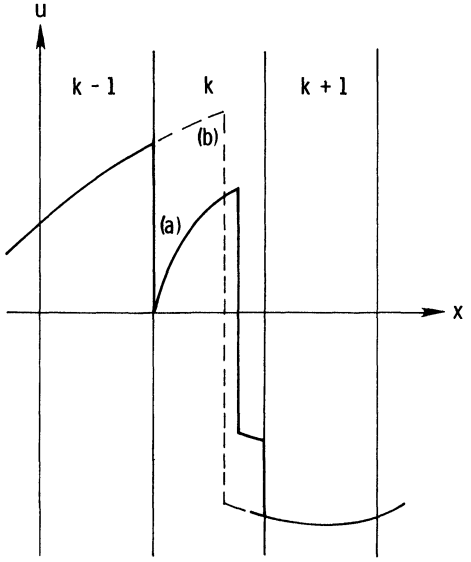

FIG. 12. *Shock structure in a globally stationary solution. Global convergence of the zone averages to a steady solution containing a shock has been obtained, say, with Godunov's scheme. For zone k containing the shock the scheme had adopted the structure* (a), *based on the smallest continuous steady solution for that zone. This may now be replaced by the stationary solution* (b) *that smoothly connects to the solution in zones* $k-1$ *and* $k+1$.

The scheme described above must be considered a research tool from which more practical schemes may be derived. In the numerical experiment of § 8 we simplified the procedure by allowing only s_i to enter the zone-boundary values. Away from sonic points or shocks the boundary values $u_{(i\pm1/2)\mp}$ can be approximated by

$$u^n_{(i\pm1/2)\mp} = u^n_i \pm \tfrac{1}{2}s_i\Delta x/u^n_i + O\{(\Delta x)^2\} \tag{36.1}$$

or

$$u^n_{(i\pm1/2)\mp} = \sqrt{(u^n_i)^2 \pm s_i\Delta x} + O\{(\Delta x)^2\}. \tag{36.2}$$

We used (36.2), extending it, for use in sonic and shock zones, to

$$u^n(x_{(i\pm1/2)\mp}) = \operatorname{sgn}(u^n_i)\sqrt{(u^n_i)^2 \pm \min\{|s_i|\Delta x, (u^n_i)^2\}\operatorname{sgn}(s_i)}. \tag{36.3}$$

Note that (34), with $x = x_{i+1/2}$, is satisfied everywhere, but transonic structure in a zone is avoided. For zones that should include a sonic point the error in the boundary values is $O(\Delta x)$.

The technique of computing the boundary values in a zone from a locally stationary solution can easily be extended to systems of conservation laws of the form

$$u_t + [f(u, x)]_x = s(x). \tag{37}$$

7. Second-order upwind schemes. Any numerical flux-function used in a first-order upwind scheme for (30) can be used in a second-order upwind scheme. For second-order accuracy in space we must introduce structure inside the zones by interpolation, while second-order accuracy in time may be achieved by advancing the cell-boundary values, to be used in the flux-function, to the intermediate time level $t^{n+1/2} = t^n + \frac{1}{2}\Delta t$. In predicting these time-centered values, *the interaction between cells can be fully ignored.* This observation, due to Hancock [9], has led to a drastic simplification of second-order upwind schemes since these first were formulated for systems of conservation laws by van Leer [10]. As in [10] we assume that the initial values form a *piecewise linear* distribution

$$u^n(x) = u^n_i + (x - x_i)\frac{(\delta u)^n_i}{\Delta x}, \qquad x_{i-1/2} < x < x_{i+1/2}, \tag{38}$$

with

$$(\delta u)^n_i = \operatorname{ave}(u^n_{i+1} - u^n_i, u^n_i - u^n_{i-1}); \tag{39}$$

ave (a, b) is an averaging procedure to be specified later. We particularly need the initial boundary values inside cell i

$$u^n_{(i\pm1/2)\mp} = u^n_i \pm \tfrac{1}{2}(\delta u)^n_i. \tag{40.1}$$

These boundary values are now advanced to $t^{n+1/2}$ by

$$u^{n+1/2}_{(i\pm1/2)\mp} = u^n_{(i\pm1/2)\mp} - \frac{\Delta t}{2\Delta x}\{f(u^n_{(i+1/2)-}) - f(u^n_{(i-1/2)+})\} + \frac{\Delta t}{2}s_i. \tag{40.2}$$

The full scheme becomes

$$(u^{n+1}_i - u^n_i)/\Delta t + \{F(u^{n+1/2}_{(i+1/2)-}, u^{n+1/2}_{(i+1/2)+}) - F(u^{n+1/2}_{(i-1/2)-}, u^{n+1/2}_{(i-1/2)+})\}/\Delta x = s_i. \tag{41}$$

The function ave (a, b) is chosen such that it tends to $\frac{1}{2}(a+b)$ if a and b are subsequent finite differences of a smooth solution, but tends to the smallest value where the solution is not smooth. Examples can be found in [11], [10]; we chose a refined formula due to Van Albada [12]

$$\text{ave}\,(a, b) = \frac{(b^2+c^2)a + (a^2+c^2)b}{a^2+b^2+2c^2}, \tag{42.1}$$

where c^2 is a small bias of the order $O((\Delta x)^3)$.

The weighted averaging prevents central differencing across a discontinuity in the solution or in its first derivative which would lead to numerical oscillations. It is an effective way of administering artificial dissipation wherever needed and nowhere else. By rewriting (42.1) as

$$\text{ave}\,(a, b) = \frac{a+b}{2}\left\{1 - \frac{(a-b)^2}{a^2+b^2+2c^2}\right\}, \tag{42.2}$$

we see that, wherever the solution is smooth, the artificial-viscosity coefficient generally is of the order $O((\Delta x)^2)$. In a smooth extremum the coefficient grows to $O(\Delta x)$; the bias $2c^2$ in the denominator prevents a further increase of the viscosity that could lead to an undesirable clipping phenomenon (see [13, § 3(e)]).

With regard to Burgers' equation, an acceptable expression for the bias is

$$c^2 = (u_{\max} - u_{\min})^2 (\Delta x)^3 / (x_{\max} - x_{\min})^3, \tag{43}$$

where $u_{\max}$ and $u_{\min}$ are certain upper and lower bounds of the numerical solution, fixed a priori or determined at each time level and $x_{\max} - x_{\min}$ is the length-scale of the problem.

8. Numerical comparison. The performance of the three schemes was tested on the basis of the periodic intial-value problem

$$u_t + (\tfrac{1}{2}u^2)_x = (\pi/2)\sin[2\pi(x-\xi)], \qquad 0 \leqq x \leqq 1, \quad 0 \leqq \xi \ll 1, \tag{44.1}$$

$$u(x, 0) = 0, \tag{44.2}$$

$$u(0, t) = u(1, t). \tag{44.3}$$

The solution tends to the steady state

$$u(x, \infty) = \begin{cases} +\sin \pi(x-\xi), & 0 \leqq x < \xi + \frac{1}{2}, \\ -\sin \pi(x-\xi), & \xi + \frac{1}{2} < x \leqq 1, \end{cases} \tag{45}$$

including a sonic point at $x = \xi$ and a shock at $x = x_S = \xi + \frac{1}{2}$. We tested the ability of the schemes to approximate this steady state, and the accuracy of the approximation. For Δx we chose a value of $\frac{1}{16}$; the zones were centered on $x_i = (i - \frac{1}{2})\Delta x$, $i = 1, \cdots, 16$. For ξ we chose 0, $\frac{1}{4}\Delta x$ and $\frac{1}{2}\Delta x$, in order to achieve different subgrid shock positions. The Courant–Friedrichs–Lewy condition on the timestep, based on the maximum characteristic speed (= 1) in the steady solution is

$$\frac{\Delta t}{\Delta x} \leqq 1; \tag{46}$$

accordingly we used $\Delta t = \frac{1}{2}\Delta x$ or $\Delta t = \Delta x$.

Table 1 shows the number of steps N_ε it took the schemes to converge according to the L_1-criterion

$$\sum_{i=1}^{16} |u_i^n - u_i^{n-1}| < \varepsilon, \tag{47}$$

with $\varepsilon = 1\times 10^{-3}$ or 1×10^{-6} and $\Delta t = \frac{1}{2}\Delta x$ or (for Godunov's scheme only) $\Delta t = \Delta x$. When the time-step is doubled, N_ε appears to be halved. The L_1-error E_ε was evaluated with respect to the zone-averaged exact solution, for the smallest value of ε. The subgrid shock positions in Table 1 were calculated with aid of (26.1) and (28.1), approximately valid because the source term is small in the shock region.

TABLE 1

Number of steps N_ε till convergence, L_1-error E_ε and steady-shock position x_S, for the initial-value problem (44), *solved with the first-order schemes of Godunov* (G), *Engquist–Osher* (EO) *and Roe* (R). *Mesh*: $\Delta x = \frac{1}{16}$, $\Delta t = \frac{1}{2}\Delta x$ *or* Δx (G *only, numbers between brackets*).

$\xi/\Delta x$	N_ε				E_ε		$(x_S - \frac{1}{2})/\Delta x$	
	$\varepsilon = 1\times 10^{-3}$		$\varepsilon = 1\times 10^{-6}$		$\varepsilon = 1\times 10^{-6}$		$\varepsilon = 1\times 10^{-6}$	
	G, R	EO	G, R	EO	G, R	EO	G, R	EO
0	62	61	112 [55]	111	8.8×10^{-3}	4.6×10^{-2}	0	0
$\frac{1}{4}$	68	66	138 [70]	136	9.6×10^{-3}	1.8×10^{-2}	.23	.26
$\frac{1}{2}$	52	52	88 [42]	88	4.6×10^{-3}	4.6×10^{-3}	$\frac{1}{2}$	$\frac{1}{2}$

The converged solutions for $\xi = \frac{1}{4}\Delta x$, with $\varepsilon = 1\times 10^{-6}$, are listed in Table 5; these are independent of Δt. The zone-averaged exact solution is given for comparison. The error in the zone with the sonic point is comparable to the value in the zone, as anticipated with the use of (36.3).

The results of Roe's scheme happen to be identical to those of Godunov's scheme; in particular, no expansion shock shows up. The reason is that, with the use of the boundary values (36.3), a sonic point in a zone is always moved to a boundary; in this case Roe's flux function (23) yields the same value (= 0) as Godunov's (7.2). Furthermore, the initial average value in the zone ultimately containing the sonic point was close to the steady value to begin with.

The results of the Engquist–Osher scheme are identical to the Godunov–Roe results except for the expected spread in the shock profile for $\xi = 0$ and $\frac{1}{4}\Delta x$. The subgrid shock positions are slightly, but not significantly, more accurate than for the other schemes. Likewise, convergence with the Engquist–Osher scheme is faster, but not significantly so.

Better results were obtained with the second-order versions of the schemes. Convergence was achieved in 10–35% fewer steps than with the first-order schemes (Table 2) and the local error near the sonic point is now of the same order as elsewhere outside the shock (Table 5).

TABLE 2

The quantities N_ε, E_ε and x_S ($\varepsilon = 1\times10^{-6}$) for the initial-value problem (44), *solved with the second-order two-step schemes based on the numerical flux-functions of Godunov* (G2, G2a), *Engquist–Osher* (EO2) *and Roe* (R2). *In* G2a *the algebraic average* (48) *is used instead of* (42.1). *Mesh*: $\Delta x = \frac{1}{16}$, $\Delta t = \frac{1}{2}\Delta x$.

$\xi/\Delta x$	N_ε			E_ε			$(x_S - \frac{1}{2})/\Delta x$		
	G2, R2	G2a	EO2	G2, R2	G2a	EO2	G2, R2	G2a	EO2
0	75	76	75	1.4×10^{-3}	3.1×10^{-2}	2.9×10^{-2}	0	0	0
$\frac{1}{4}$	89	92	89	1.3×10^{-3}	2.2×10^{-3}	4.7×10^{-3}	.24	.19	.26
$\frac{1}{2}$	79	79	79	1.3×10^{-3}	1.7×10^{-3}	1.3×10^{-3}	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

The Engquist–Osher scheme still yields a shock structure with two interior zones. For the second-order Godunov scheme we checked that the use of algebraic averaging of differences,

$$\text{ave}\,(a, b) = \tfrac{1}{2}(a+b), \tag{48}$$

causes the numerical shock to overshoot and undershoot (Table 5). The accuracy of the shock position also suffers (Table 2).

The results of Roe's scheme are practically, but not exactly, identical to the Godunov results, indicating a slightly different transient behavior. Again, the internal structure (36.3) in the zones, in combination with the particular choice of initial values, provides a safeguard against the occurrence of an expansion shock.

In order to make the problem more challenging, we changed the initial values (44.2), for $\xi = 0$, into

$$u_i^0 = \begin{cases} 1, & i = 1, \cdots, 8, \\ -1, & i = 9, \cdots, 16, \end{cases} \tag{49}$$

including an expansion shock at $x = 0$. The results are listed in Tables 3 and 6. Among the second-order schemes Roe's scheme now requires 30% more steps than the other schemes: apparently, the expansion shock is not so easily dissipated by Roe's scheme. Among the first-order schemes the discrepancy gets worse: the results of Roe's scheme quickly converge to the wrong solution, including the full initial expansion shock. This

TABLE 3

The quantities N_ε and E_ε ($\varepsilon = 1\times10^{-6}$) for the initial-value problem (44.1), (49), (44.3), *solved with both first- and second-order upwind schemes. Scheme* R *converges to the wrong weak solution. The schemes* R* *and* R2* *incorporate the flux-function* (29.5) *for a transonic expansion. Mesh*: $\Delta x = \frac{1}{16}$, $\Delta t = \frac{1}{2}\Delta x$.

$\xi/\Delta x$	N_ε								E_ε	
	G	R	EO	G2	R2	EO2	R*	R2*	R	R2
0	170	30	169	77	99	76	103	71	5.7×10^{-1}	1.4×10^{-3}

is because the zones on either side of the expansion shock cannot be influenced by their neighbors (see Fig. 8b), and they cannot change by themselves because of (36c).

We then replaced the transonic value of the flux-function in Roe's scheme by Roe's new formula (29.5), with dramatic results. Not only was the proper solution obtained, but in considerably fewer steps than required by the other schemes (Table 3). Even for the second-order scheme convergence was significantly faster.

Finally, Table 4 shows the same quantities as Table 1 for experiments in which, in the first-order schemes, the source-term dependence of zone-boundary values was dropped. It took the schemes of Engquist–Osher and Godunov 15–25% more steps than previously to converge to a much less accurate solution (Table 5). Roe's scheme is now unstable: it yields an expansion shock where a smooth transonic transition should be; the amplitude of this shock grows linearly with time. The reason is that

TABLE 4

The quantities N_ε, E_ε and x_S ($\varepsilon = 1\times10^{-6}$) for the initial-value problem (44), *solved with the first-order upwind schemes under the assumption of piecewise uniform initial values* (3.1). (*The label* "u" *in* Gu, Ru, EOu *stands for* "*uniform*".) *Mesh*: $\Delta x = \frac{1}{16}$, $\Delta t = \frac{1}{2}\Delta x$.

$\xi/\Delta x$	N_ε			E_ε			$(x_S - \frac{1}{2})/\Delta x$		
	Gu	Ru	EOu	Gu	Ru	EOu	Gu	Ru	EOu
0	135	-	135	6.0×10^{-2}	unstable	9.5×10^{-2}	0	-	0
$\frac{1}{4}$	174	-	172	6.1×10^{-2}	unstable	6.7×10^{-2}	.22	-	.24
$\frac{1}{2}$	103	103	103	4.7×10^{-2}	4.7×10^{-2}	4.7×10^{-2}	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

TABLE 5

Converged numerical solutions and zone-averaged asymptotic solutions for the experiments of Tables 1, 2 *and* 4, *with* $\xi = \frac{1}{4}\Delta x$, $\varepsilon = 1\times10^{-6}$, $\Delta t = \frac{1}{2}\Delta x$.

i	$u_i^{N_\varepsilon}$, u_i^{∞}							
	$\xi/\Delta x = \frac{1}{4}$							
	G, R	EO	G2, R2	G2a	EO2	Gu	EOu	exact
1	.09778	.09778	.04916	.04917	.04916	.13828	.13828	.04899
2	.25516	.25516	.24249	.24244	.24248	.33330	.33330	.24259
3	.43185	.43184	.42635	.42610	.42639	.51176	.51175	.42687
4	.59602	.59602	.59379	.59344	.59364	.66976	.66976	.59474
5	.73869	.73869	.73871	.73726	.73933	.80171	.80171	.73976
6	.85367	.85367	.85462	.85711	.85231	.90266	.90266	.85635
7	.93631	.93631	.94000	.91445	.95106	.96879	.96879	.94003
8	.98329	.98033	.98172	1.15231	.95622	.99759	.89488	.98759
9	-.52996	-.42699	-.49587	-.60164	-.47889	-.54359	-.44088	-.49739
10	-.96441	-.96441	-.96449	-1.01358	-.96440	-.98796	-.98796	-.96847
11	-.89927	-.89927	-.90232	-.89318	-.90235	-.94026	-.94026	-.90254
12	-.79997	-.79997	-.80042	-.80067	-.80041	-.85632	-.85632	-.80192
13	-.67047	-.67047	-.66948	-.66861	-.66948	-.73932	-.73932	-.67048
14	-.51616	-.51616	-.51252	-.51223	-.51252	-.59368	-.59368	-.51328
15	-.34426	-.34426	-.33603	-.33589	-.33603	-.42473	-.42473	-.33635
16	-.16827	-.16827	-.14650	-.14649	-.14650	-.23797	-.23797	-.14649

TABLE 6

Converged numerical solutions and zone-averaged asymptotic solution for the experiments of Table 3 with schemes R *and* R2. *Note the expansion shock in the results of* R. *Parameters*: $\xi = 0$, $\varepsilon = 1 \times 10^{-6}$, $\Delta t = \frac{1}{2}\Delta x$.

i	$u_i^{N_\varepsilon}$, u_i^∞		
	R	R2	exact
1	1.00000	.09815	.09786
2	1.03596	.28976	.28982
3	1.09933	.47012	.47064
4	1.17699	.63242	.63337
5	1.25564	.77076	.77177
6	1.32417	.87880	.88051
7	1.37431	.95648	.95540
8	1.40069	.98762	.99359
9	-1.40069	-.98762	-.99359
10	-1.37431	-.95648	-.95540
11	-1.32417	-.87880	-.88051
12	-1.25564	-.77076	-.77177
13	-1.17699	-.63242	-.63337
14	-1.09933	-.47012	-.47064
15	-1.03596	-.28976	-.28982
16	-1.00000	-.09815	-.09786

the zone containing the sonic point (say, zone j), not influenced by its neighbors, no longer has a steady internal structure, so that the scheme locally reduces to

$$u_j^n = u_j^0 + n\Delta t s_j. \tag{50}$$

9. Recommendations. For a scalar conservation law like Burgers' equation there seems to be little reason to abandon Godunov's first-order scheme in favor of the first-order Engquist–Osher scheme. The slight simplification in the flux calculation is accompanied by a degradation of steady-shock representation and no significant acceleration of the convergence to a steady state. The simplification achieved in Roe's first-order scheme is too drastic: the scheme cannot be used "as is" near a sonic point. With the modification (29.5), however, the scheme surpasses Godunov's scheme.

Both Godunov and Engquist–Osher schemes become appreciably more elaborate when applied to a nonlinear hyperbolic system like the one-dimensional Euler equations. The interpretation of the Engquist–Osher scheme as a Godunov-type scheme in which any shock in the solution of a local Riemann problem is replaced by an overturned centered compression wave remains valid. This modification makes it possible to compute $F_{EO}(u_L, u_R)$ explicitly, while $F_G(u_L, u_R)$ must be determined iteratively.

The Engquist–Osher scheme, on the other hand, requires two interior points in a discrete stationary shock [5], while Godunov's scheme probably requires only one (this has not yet been proven in general).

For the Euler equations, however, Roe's more drastic simplification of Godunov's method will pay off, even though the scheme must be somewhat modified in order to reject (almost) stationary expansion shocks. This can be achieved by spreading of rarefaction waves, as in [16], or by generalizing (29.3). Numerical experience with the latter technique is still lacking.

The first-order schemes, when formulated as in § 6, have the potential of achieving any desired order of spatial accuracy in a steady numerical solution. To what degree this potential can be realized for the one-dimensional Euler equations is at present not clear. Meanwhile, the second-order two-step schemes, intended primarily for solving transient problems, seem to outperform the first-order schemes based on (36.3) in obtaining a steady solution for Burgers' equation. Experiments for the Euler equations with both kinds of schemes, conducted for [13], but not fully reported therein, indicate a similar performance.

All schemes discussed in this paper are explicit. When using their numerical flux-functions in an implicit configuration, which may be desirable in approaching a steady state, the above recommendations do not automatically carry over. As noted by Engquist and Osher [4], the nonsmooth dependence of $F_G(u_L, u_R)$ on its arguments in the case (iii) of a transonic shock, that is: $F_G(u_L, u_R) \in C^{(0)}$, invalidates the linearization needed to invert the implicit difference equations. In contrast, $F_{EO}(u_L, u_R) \in C^{(1)}$ in all cases (i), (ii) and (iii). Whether this makes a difference in practice remains to be shown.

Acknowledgments. The ideas of Steve Hancock (Physics International, San Leandro, California) and Dick van Albada (University of Groningen, Netherlands), embodied in (40) and (42), respectively, are greatly appreciated, while discussions with Stanley Osher (University of California at Los Angeles) strengthened my belief in the treatment of source terms discussed in § 6. Phil Colella (Lawrence Berkeley Laboratory) pointed out the weak transonic expansion shocks arising in the use of the exact Riemann solution, while Phil Roe (Royal Aircraft Establishment, United Kingdom) kindly gave me permission to discuss his hitherto unpublished sonic flux-function.

REFERENCES

[1] S. K. GODUNOV, *Finite-difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics*, Mat. Sbornik, 47 (1959), pp. 271–306. (In Russian.)

[2] S. K. GODUNOV, A. W. ZABRODYN AND G. P. PROKOPOV, *A difference scheme for two-dimensional unsteady problems of gas dynamics and computation of flow with a detached shock wave*, Z. Vyčisl. Matem. Mat. Fiz., 1 (1961), pp. 1020–1050. (In Russian.)

[3] A. HARTEN, P. D. LAX AND B. VAN LEER, *On upstream differencing and Godunov-type schemes for hyperbolic conservation laws*, ICASE Report No. 82–5, 1982; SIAM Rev., 25 (1983), pp. 35–62.

[4] B. ENGQUIST AND S. OSHER, *Stable and entropy satisfying approximations for transonic flow calculations*, Math. Comp., 34 (1980), pp. 45–75.

[5] S. OSHER AND F. SOLOMON, *Upwind difference schemes for hyperbolic systems of conservation laws*, Math. Comp. (1983), to appear.

[6] P. L. ROE, *The use of the Riemann problem in finite-difference schemes*, Lecture Notes in Physics, 141, Springer-Verlag, New York, 1981, pp. 354–559.

[7] ———, *Approximate Riemann solvers, parameter vectors and difference schemes*, J. Comp. Phys., 43 (1981), pp. 357–372.

[8] E. M. MURMAN, *Analysis of embedded shock waves calculated by relaxation methods*, AIAA J., 12 (1974), pp. 626–633.

[9] S. L. HANCOCK, private communication (1980).

[10] 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.

[11] ———, *Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection*, J. Comp. Phys. 23 (1977), pp. 276–299.

[12] G. D. VAN ALBADA, private communication (1981).

[13] G. D. VAN ALBADA, B. VAN LEER AND W. W. ROBERTS, JR., *A comparative study of computational methods in cosmic gas dynamics*, Astron. Astrophys., 108 (1982), pp. 76–84.

[14] P. L. ROE, private communication (1981).
[15] A. HARTEN AND P. D. LAX, *A random-choice finite-difference scheme for hyperbolic conservation laws*, SIAM J. Numer. Anal., 18 (1981), pp. 289–315.
[16] P. COLELLA, *Glimm's method for gas dynamics*, this Journal, 3 (1982), pp. 76–110.
[17] T.-P. LIU, *Quasilinear hyperbolic systems*, Comm. Math. Phys., 68 (1979), pp. 141–172.