JOURNAL OF COMPUTATIONAL PHYSICS **135,** 260–278 (1997)
ARTICLE NO. CP975713

High Resolution Schemes for Hyperbolic Conservation Laws*

Ami Harten

School of Mathematical Sciences, Tel-Aviv University, Ramat Aviv, Israel; and Courant Institute of Mathematical Sciences, New York University, New York City, New York 10012

Received February 2, 1982; revised June 23, 1982

A class of new explicit second order accurate finite difference schemes for the computation of weak solutions of hyperbolic conservation laws is presented. These highly nonlinear schemes are obtained by applying a nonoscillatory first order accurate scheme to an appropriately modified flux function. The so-derived second order accurate schemes achieve high resolution while preserving the robustness of the original nonoscillatory first order accurate scheme. Numerical experiments are presented to demonstrate the performance of these new schemes.

1. INTRODUCTION

In this paper we consider numerical approximations to weak solutions of the initial value problem (IVP) for hyperbolic systems of conservation laws

$$u_t + f(u)_x = 0, \quad u(x,0) = \phi(x), \quad -\infty < x < \infty. \tag{1.1}$$

Here $u(x, t)$ is a column vector of m unknowns, and $f(u)$, the flux, is a vector-valued function of m components. Equation (1.1) is called hyperbolic if all eigenvalues $a^1(u)$, ..., $a^m(u)$ of the Jacobian matrix $A(u)$

$$A(u) = f_u \tag{1.2a}$$

are real and the set of right eigenvectors $R^1(u)$, ..., $R^m(u)$ is complete. We assume that the eigenvalues $\{a^i(u)\}$ are arranged in a nondecreasing order

$$a^1(u) \leqslant a^2(u) \leqslant \cdots \leqslant a^m(u). \tag{1.2b}$$

We consider systems of conservation laws (1.1) that possess an entropy function $U(u)$, defined as follows:

(i) U is a convex function of u, i.e., $U_{uu} > 0$,

(ii) U satisfies

$$U_u f_u = F_u, \tag{1.3a}$$

where F is some other function called entropy flux.

Admissible weak solutions of (1.1) satisfy, in the weak sense, the inequality

$$U(u)_t + F(u)_x \leqslant 0 \tag{1.3b}$$

(see [12]). The inequality (1.3b) is called an entropy condition.

We shall discuss numerical approximations to weak solutions of (1.1) which are obtained by $(2k + 1)$-point explicit schemes in conservation form

$$v_j^{n+1} = v_j^n - \lambda(\bar{f}_{j+1/2}^n - \bar{f}_{j-1/2}^n), \tag{1.4a}$$

where

$$\bar{f}_{j+1/2}^n = \bar{f}(v_{j-k+1}^n, ..., v_{j+k}^n). \tag{1.4b}$$

Here $v_j^n = (j\Delta x, n\,\Delta t)$, and $\bar{f}$ is a numerical flux function. We require the numerical flux function to be consistent with the flux $f(u)$ in the following sense:

$$\bar{f}(u, ..., u) = f(u). \tag{1.4c}$$

We say that difference scheme (1.4) is consistent with entropy condition (1.3b) if an inequality of the following kind is satisfied:

$$U_j^{n+1} \leqslant U_j^n - \lambda(\bar{F}_{j+1/2}^n - \bar{F}_{j-1/2}^n), \tag{1.5a}$$

where $U_j^n = U(v_j^n)$, $\bar{F}_{j+1/2}^n = \bar{F}(v_{j-k+1}^n, ..., v_{j+k}^n)$; here $\bar{F}$ is a numerical entropy flux consistent with the entropy flux $F(u)$, i.e.,

$$\bar{F}(u, ..., u) = F(u). \tag{1.5b}$$

We turn now to discuss the question of convergence of the finite-difference solution of (1.4) to weak solutions of conservation laws (1.1). Since the finite-difference scheme

Reprinted from Volume 49, Number 3, March 1983, pages 357–393.
* Prepared under Interchange NCA2-OR525-101 with the NASA Ames Research Center, Moffett Field, California; and DE-AC02-76ER03077 with the U.S. Department of Energy, Division of Basic Energy Sciences, Applied Mathematical Sciences Program.

0021-9991/97 $25.00

is nonlinear and the computed solutions are certainly not smooth, L_2-stability of a consistent finite-difference scheme does not imply convergence. One can establish convergence of finite-difference solutions of (1.4) to weak solutions of (1.1) when the following conditions are satisfied:

(i) The total variation with respect to x of the finite-difference solutions is uniformly bounded with respect to t, Δt, and Δx.

(ii) Finite-difference scheme (1.4) is consistent with entropy condition (1.3b) for all entropy functions of (1.1).

(iii) Entropy condition (1.3b) implies uniqueness of the solution of the IVP (1.1).

Using compactness arguments one can deduce from condition (i) the existence of convergent subsequences. Conservation form (1.4) and condition (ii) imply that each limit solution is a weak solution which satisfies entropy condition (1.3b). When the entropy condition implies uniqueness of the IVP (condition (iii)), then all subsequences have the same limit solution, and consequently the finite-difference scheme is convergent. (See [2, 9, 10].)

It seems possible to satisfy conditions (i) and (ii) by adding a hefty amount of artificial viscosity to finite-difference scheme (1.4). The additional viscosity terms damp possible oscillations in the computed solution and make the convergence process simulate the zero-dissipation limit which is used to select the unique physically relevant weak solution. Unfortunately, viscosity represents an irretrievable loss of information and therefore the addition of artificial viscosity brings about some deterioration in resolution.

In this paper we describe a new method to design finite-difference schemes that satisfy conditions (i) and (ii), but are second order accurate and have high resolution.

2. MONOTONICITY IN THE SCALAR CASE

In this section we consider the IVP for a scalar conservation law

$$u_t + f(u)_x \equiv u_t + a(u)u_x = 0, \qquad a(u) = df/du, \quad (2.1a)$$

$$u(x, 0) = \phi(x), \quad -\infty < x < \infty, \quad (2.1b)$$

where $\phi(x)$ is assumed to be of bounded total variation. A weak solution of scalar IVP (2.1) has the following *monotonicity property* as a function of t:

(i) No new local extrema in x may be created.

(ii) The value of a local minimum is nondecreasing, the value of a local maximum is nonincreasing.

It follows from this monotonicity property that the total variation is x, denoted $\mathrm{TV}(u(t))$, of $u(x, t)$, is nonincreasing in t, i.e.,

$$\mathrm{TV}(u(t_2)) \leqslant \mathrm{TV}(u(t_1)) \quad \text{for all } t_2 \geqslant t_1. \quad (2.2)$$

We consider now explicit $(2k + 1)$-point finite-difference schemes in conservation form (1.4) approximating (2.1)

$$v_j^{n+1} = H(v_{j-k}^n, v_{j-k+1}^n, ..., v_{j+k}^n)$$
$$= v_j^n - \lambda[\bar{f}(v_{j-k+1}^n, ..., v_{j+k}^n) - \bar{f}(v_{j-k}^n, ..., v_{j+k-1}^n)] \quad (2.3a)$$

and denote (2.3a) in an operator form as

$$v^{n+1} = L \cdot v^n \quad (2.3b)$$

We say that finite-difference scheme (2.3) is *total variation nonincreasing* (TVNI) if for all v of bounded total variation

$$\mathrm{TV}(L \cdot v) \leqslant \mathrm{TV}(v), \quad (2.4a)$$

where

$$\mathrm{TV}(u) \equiv \sum_{j=-\infty}^{\infty} |\Delta_{j+1/2} u|. \quad (2.4b)$$

Here, and throughout this paper, we use the standard notation

$$\Delta_{j+1/2} u = u_{j+1} - u_j. \quad (2.5)$$

We say that finite-difference scheme (2.3) is *monotonicity preserving* if the finite difference operator L is monotonicity preserving; i.e., if v is a monotone mesh function, so is $L \cdot v$.

We say that finite-difference scheme (2.3) is a *monotone scheme* if H in (2.3a) is a monotone nondecreasing function of each of its $2k + 1$ arguments.

The following theorem states the hierarchy of these properties:

THEOREM 2.1. (i) *A monotone scheme is* TVNI.

(ii) *A* TVNI *scheme is monotonicity preserving.*

Proof. (i) It was proven in [8] that monotone schemes form an l_1-contractive semigroup, i.e.,

$$\|L \cdot v - L \cdot z\|_{l_1} \leqslant \|v - z\|_{l_1} \quad (2.6a)$$

for all l_1-summable v and z; here $\|u\|_{l_1} = \sum_{j=-\infty}^{\infty} |u_j|$. Equation (2.4) follows immediately from applying (2.6a) to v and $z = T \cdot v$ (i.e., $z_j \equiv v_{j+1}$ for all j).

(ii) Let (2.3) be a TVNI scheme, let v be a monotone mesh function of bounded total variation, and denote $w = L \cdot v$. Since L has a finite support of $2k + 1$ points it is sufficient to prove that w is monotone for all v of the form

$$\begin{aligned} v &= \text{const} = v_L, \quad \text{if } j \leqslant J_-, \\ &= \text{monotone}, \quad \text{if } J_- \leqslant j \leqslant J_+, \quad J_+ \geqslant J_-, \\ &= \text{const} = v_R, \quad \text{if } j \geqslant J_+, \\ \mathrm{TV}(v) &= |v_R - v_L|. \end{aligned} \tag{2.6b}$$

We prove (ii) by negation. Suppose w is not monotone; then it has at least one local minimum and one local maximum. Denote by v_m and v_M the values of the first two successive local extrema, then

$$\mathrm{TV}(w) \geqslant |v_R - v_L| + |v_m - v_M| > \mathrm{TV}(v),$$

which contradicts the assumption that the scheme is TVNI. This completes the proof of Theorem 2.1. ▌

Monotone schemes approximate solutions of the viscous modified equation

$$u_t + f(u)_x = \Delta t[\beta(u, \lambda)u_x]_x, \quad \lambda = \Delta t/\Delta x; \tag{2.7a}$$

$$\beta(u, \lambda) = \frac{1}{2\lambda^2}\left[\sum_{l=-k}^{k} l^2 H_l(u, u, ..., u) - \lambda^2 a^2(u)\right]; \tag{2.7b}$$

$$\beta(u, \lambda) \geqslant 0, \quad \beta(u, \lambda) \not\equiv 0 \tag{2.7c}$$

to second-order accuracy; since $\beta(u, \lambda) \not\equiv 0$, monotone schemes are necessarily first order accurate; H_l in (2.7b) denotes $(\partial H/\partial w_l)(w_{-k}, w_{-k+1}, ..., w_k)$. (See [8].)

Since monotone schemes are TVNI, there exist convergent subsequences for all initial data of bounded total variation. Each limit is a weak solution of (2.1) that satisfies Oleinik's entropy condition [8]. Since Oleinik's entropy condition implies uniqueness of IVP (2.1), we conclude that all subsequences converge to the same limit, and therefore the scheme is convergent [2].

Let us consider now the scalar constant coefficient case $a(u) = \text{const}$ in (2.1). A linear finite-difference approximation

$$v_j^{n+1} = \sum_{l=-k}^{k} c_l v_{j+l}^n, \quad c_l = \text{const} \tag{2.8a}$$

is monotonicity preserving if and only if

$$c_l \geqslant 0, \quad -k \leqslant l \leqslant k. \tag{2.8b}$$

See [4]. Hence any *linear* monotonicity-preserving scheme, and therefore any TVNI linear scheme, is a monotone scheme and consequently first order accurate.

We remark that the previous statement does not exclude the possibility of having *nonlinear* monotonicity-preserving and TVNI schemes that are second order accurate (and consequently are not monotone schemes). In fact the schemes presented in [6, 17] are monotonicity preserving (at least in the constant coefficient case) and second order accurate.

It is the purpose of this paper to present new high-resolution second order accurate TVNI schemes. These new schemes are generated by converting known 3-point first order accurate TVNI schemes into new 5-point second order accurate TVNI schemes. Both the 3-point schemes and the new 5-point schemes can be rewritten in the form

$$v^{n+1} = L \cdot v^n, \tag{2.9a}$$

$$(L \cdot v)_j = v_j + C_{+,j+1/2}\Delta_{j+1/2}v - C_{-,j-1/2}\Delta_{j-1/2}v, \tag{2.9b}$$

where $\Delta_{j+1/2}$ is defined in (2.5) and

$$\begin{aligned} C_{+,j+1/2} &= C_+(v_{j-1}, v_j, v_{j+1}, v_{j+2}), \\ C_{-,j-1/2} &= C_-(v_{j-2}, v_{j-1}, v_j, v_{j+1}). \end{aligned} \tag{2.9c}$$

The following lemma states conditions on coefficients (2.9c) which are sufficient to ensure that scheme (2.9b) is TVNI.

LEMMA 2.2. *Let the coefficients* $C_\pm$ *in* (2.9c) *satisfy the inequalities*

$$C_{-,j+1/2} \geqslant 0, \quad C_{+,j+1/2} \geqslant 0, \tag{2.10a}$$

$$C_{-,j+1/2} + C_{+,j+1/2} \leqslant 1; \tag{2.10b}$$

then scheme (2.9) *is* TVNI.

Proof. Denote $w = L \cdot v$ and subtract (2.9b) at $j = i$ from (2.9b) at $j = i + 1$ to obtain

$$\begin{aligned} \Delta_{i+1/2}w = C_{-,i-1/2}\Delta_{i-1/2}v + (1 - C_{-,i+1/2} - C_{+,i+1/2})\Delta_{i+1/2}v \\ + C_{+,i+3/2}\Delta_{i+3/2}v. \end{aligned} \tag{2.11}$$

By (2.10) all the coefficients in (2.11) are nonnegative; therefore

$$\begin{aligned} |\Delta_{i+1/2}w| \leqslant (1 - C_{-,i+1/2} - C_{+i+1/2})|\Delta_{i+1/2}v| + C_{-,i-1/2}|\Delta_{i-1/2}v| \\ + C_{+,i+3/2}|\Delta_{i+3/2}v|. \end{aligned} \tag{2.12}$$

Summing (2.12) for $-\infty < i < \infty$ we obtain

$$\begin{aligned}\mathrm{TV}(w) \equiv \sum_{i=-\infty}^{\infty} |\Delta_{i+1/2} w| &\leqslant \sum_{i=-\infty}^{\infty} (1 - C_{-,i+1/2} - C_{+,i+1/2}) |\Delta_{i+1/2} v| \\ &+ \sum_{i=-\infty}^{\infty} C_{-,i-1/2} |\Delta_{i-1/2} v| + \sum_{i=-\infty}^{\infty} C_{+,i+3/2} |\Delta_{i+3/2} v| \\ &= \sum_{i=-\infty}^{\infty} |\Delta_{i+1/2} v| \equiv \mathrm{TV}(v),\end{aligned}$$

which shows that (2.4) is satisfied. The equality is obtained by changing the summation index in the last two sums in the RHS of the inequality.

We remark that any 3-point finite difference scheme in conservation form with a differentiable numerical flux can be rewritten as (2.9) in the following way: It follows from the mean value theorem that there exist C_+ and C_- such that

$$\lambda[\bar{f}(v_j, v_{j+1}) - \bar{f}(v_j, v_j)] = -C_+(v_j, v_{j+1}) \Delta_{j+1/2} v, \quad (2.13a)$$

$$\lambda[\bar{f}(v_{j-1}, v_j) - \bar{f}(v_j, v_j)] = -C_-(v_{j-1}, v_j) \Delta_{j-1/2} v. \quad (2.13b)$$

Expressing the numerical flux values in (2.3a) with $K = 1$ by (2.13) results in form (2.9).

In the next section we shall use Lemma 2.2 to design second order accurate TVNI schemes in the following way: We start with a 3-point first order accurate TVNI scheme which can be written in form (2.9) that satisfies the assumptions of Lemma 2.2. This scheme approximates solutions of modified equation (2.7a)

$$u_t + (f - (1/\lambda) g)_x = 0, \quad (2.14a)$$

$$g = \Delta x \beta(u, \lambda) u_x \quad (2.14b)$$

to second order accuracy.

Consider now the application of the original first order scheme to the equation

$$u_t + (f + (1/\lambda) g)_x = 0. \quad (2.15)$$

This in turn is a new scheme that is a second order approximation to its own modified equation. Since $g = O(\Delta x)$, this modified equation satisfies

$$u_t + f_x = O((\Delta x)^2). \quad (2.16)$$

Thus applying the original first order scheme to the flux $f + (1/\lambda) g$ results in a second order accurate approximation to the original equation $u_t + f_x = 0$.

To apply the scheme to the flux $f + (1/\lambda) g$ we have to be able to consider g as a differentiable function of u. To do so we smooth the point values of g in (2.14b) by the recipe devised in [5, 6]. This smoothing process enlarges the support of the scheme to five points.

The resulting scheme is TVNI due to the fact that *for any flux function* the original scheme has the form (2.9) that satisfies (2.10) under an appropriate CFL-like restriction.

3. SECOND ORDER ACCURATE TVNI SCHEMES

Let us consider a general 3-point finite-difference scheme in conservation form (1.4) with a numerical flux $\bar{f}$ of the form

$$\bar{f}(v_j, v_{j+1}) = \tfrac{1}{2}[f(v_j) + f(v_{j+1}) - (1/\lambda) Q(\lambda \bar{a}_{j+1/2}) \Delta_{j+1/2} v], \quad (3.1a)$$

where

$$\begin{aligned}\bar{a}_{j+1/2} &= [f(v_{j+1}) - f(v_j)]/\Delta_{j+1/2} v, \quad \text{when } \Delta_{j+1/2} v \neq 0, \\ &= a(v_j), \quad \text{when } \Delta_{j+1/2} v = 0.\end{aligned} \quad (3.1b)$$

Here $Q(x)$ is some function, which is often referred to as the coefficient of numerical viscosity.

LEMMA 3.1. *Let* $Q(x)$ *in* (3.1a) *satisfy the inequalities*

$$|x| \leqslant Q(x) \leqslant 1 \quad \textit{for } 0 \leqslant |x| \leqslant \mu \leqslant 1; \quad (3.2)$$

then finite-difference scheme (1.4) *with* (3.1) *is* TVNI *under the* CFL*-like restriction*

$$\lambda \max_j |\bar{a}^n_{j+1/2}| \leqslant \mu. \quad (3.3)$$

Proof. Using the notation

$$\bar{\nu}_{j+1/2} = \lambda \bar{a}_{j+1/2}, \quad (3.4a)$$

where $\bar{a}_{j+1/2}$ is (3.1b), we rewrite (3.1a) as

$$\lambda \bar{f}_{j+1/2} = \lambda f(v_j, v_{j+1}) = \lambda f(v_j) - \tfrac{1}{2}[-\bar{\nu}_{j+1/2} + Q(\bar{\nu}_{j+1/2})] \Delta_{j+1/2} v \quad (3.4b)$$

and similarly

$$\lambda \bar{f}_{j-1/2} = \lambda f(v_{j-1}, v_j) = \lambda f(v_j) - \tfrac{1}{2}[\bar{\nu}_{j-1/2} + Q(\bar{\nu}_{j-1/2})] \Delta_{j-1/2} v. \quad (3.4c)$$

Substituting (3.4) for numerical flux values in (1.4) we get form (2.9)

$$\begin{aligned} v_j^{n+1} &= v_j^n - \lambda(\bar{f}_{j+1/2} - \bar{f}_{j-1/2}) \\ &= v_j^n + \tfrac{1}{2}[Q(\bar{\nu}_{j+1/2}) - \bar{\nu}_{j+1/2}]\Delta_{j+1/2}v^n \\ &\quad - \tfrac{1}{2}[Q(\bar{\nu}_{j-1/2}) + \bar{\nu}_{j-1/2}]\Delta_{j-1/2}v^n \\ &\equiv v_j^n + C_{+,j+1/2}\Delta_{j+1/2}v^n - C_{-,j-1/2}\Delta_{j-1/2}v^n, \end{aligned} \tag{3.5a}$$

where

$$C_{\pm,j+1/2} = \tfrac{1}{2}[Q(\bar{\nu}_{j+1/2}) \mp \bar{\nu}_{j+1/2}]. \tag{3.5b}$$

Since

$$C_{+,j+1/2} + C_{-,j+1/2} = Q(\bar{\nu}_{j+1/2}), \tag{3.5c}$$

it follows from (3.2) and (3.5) that conditions (2.10) of Lemma 2.2 are satisfied under CFL restriction (3.3) and therefore finite-difference scheme (3.1) is TVNI.

The second order accurate Lax–Wendroff scheme has numerical flux (3.1) with $Q(x) = x^2$, i.e.,

$$\bar{f}^{\mathrm{LW}}_{j+1/2} \equiv \tfrac{1}{2}[f(v_j) + f(v_{j+1}) - (1/\lambda)(\bar{\nu}_{j+1/2})^2\Delta_{j+1/2}v]. \tag{3.6a}$$

Clearly a numerical flux of a second order accurate scheme $\bar{f}_{j+1/2}$ has to satisfy

$$\bar{f}_{j+1/2} - \bar{f}^{\mathrm{LW}}_{j+1/2} = O(\Delta^2), \tag{3.6b}$$

for all smooth solutions of (1.1); here Δ is the discretization parameter. When $Q(x)$ is constrained by (3.2), then the 3-point scheme (3.1) is only first order accurate, for

$$\bar{f}_{j+1/2} = \bar{f}^{\mathrm{LW}}_{j+1/2} - (1/(2\lambda))[-(\bar{\nu}_{j+1/2})^2 + Q(\bar{\nu}_{j+1/2})]\Delta_{j+1/2}v^n \tag{3.7}$$

and therefore

$$\begin{aligned} |f^{\mathrm{LW}}_{j+1/2} - \bar{f}_{j+1/2}| &\geqslant (1/(2\lambda))[|\bar{\nu}_{j+1/2}| - (\bar{\nu}_{j+1/2})^2] \cdot |\Delta_{j+1/2}v^n| \\ &= O(\Delta). \end{aligned}$$

We describe now how to convert a 3-point first order accurate TVNI scheme to a 5-point second order accurate TVNI scheme. Consider the application of a 3-point first order accurate TVNI scheme (3.1) to modified mesh values f_i^M of the original flux $f(u)$:

$$f_i^M = f(v_i) + (1/\lambda)\, g_i, \quad g_i = g(v_{i-1}, v_i, v_{i+1}), \tag{3.8a}$$

$$\bar{\nu}^M_{i+1/2} = \bar{\nu}_{i+1/2} + \gamma_{i+1/2}, \quad \gamma_{i+1/2} = (g_{i+1} - g_i)/\Delta_{i+1/2}v. \tag{3.8b}$$

The modified numerical flux $\bar{f}^M_{j+1/2} = \bar{f}^M(v_{j-1}, v_j, v_{j+1}, v_{j+2})$ is

$$\bar{f}^M_{j+1/2} = \tfrac{1}{2}[f^M_j + f^M_{j+1} - (1/\lambda)\, Q(\bar{\nu}^M_{j+1/2})\, \Delta_{j+1/2}v] \tag{3.8c}$$

or

$$\begin{aligned} \bar{f}^M_{j+1/2} &= \tfrac{1}{2}[f(v_j) + f(v_{j+1})] \\ &\quad + (1/(2\lambda))[g_j + g_{j+1} - Q(\bar{\nu}_{j+1/2} + \gamma_{j+1/2})\Delta_{j+1/2}v]. \end{aligned} \tag{3.8d}$$

LEMMA 3.2. *Suppose $Q(x)$ is Lipschitz continuous and g_j satisfies*

$$g_j + g_{j+1} = [Q(\bar{\nu}_{j+1/2}) - (\bar{\nu}_{j+1/2})^2]\, \Delta_{j+1/2}v + O(\Delta^2), \tag{3.9a}$$

$$\gamma_{j+1/2} \cdot \Delta_{j+1/2}v \equiv g_{j+1} - g_j = O(\Delta^2); \tag{3.9b}$$

then the numerical flux of (3.8) *satisfies* (3.6).

Proof. The modified numerical flux $\bar{f}^M_{j+1/2}$ of (3.8c) differs from the original flux $\bar{f}_{j+1/2}$ of (3.1) in the following way:

$$\begin{aligned} \bar{f}^M_{j+1/2} &= \bar{f}_{j+1/2} + (1/(2\lambda))\{g_j + g_{j+1} \\ &\quad + [Q(\bar{\nu}_{j+1/2}) - Q(\bar{\nu}_{j+1/2} + \gamma_{j+1/2})]\, \Delta_{j+1/2}v\}. \end{aligned} \tag{3.10a}$$

Substituting (3.7) for $\bar{f}_{j+1/2}$ in (3.10a) we get that (3.6) holds if the relation

$$\begin{aligned} &(g_j + g_{j+1}) + [Q(\bar{\nu}_{j+1/2}) - Q(\bar{\nu}_{j+1/2} + \gamma_{j+1/2})]\, \Delta_{j+1/2}v \\ &\quad = [Q(\bar{\nu}_{j+1/2}) - (\bar{\nu}_{j+1/2})^2]\, \Delta_{j+1/2}v + O(\Delta^2) \end{aligned} \tag{3.10b}$$

is satisfied.

Since $Q(x)$ is Lipschitz continuous

$$|Q(\bar{\nu}_{j+1/2}) - Q(\bar{\nu}_{j+1/2} + \gamma_{j+1/2})| \leqslant \text{const}\, |\gamma_{j+1/2}|; \tag{3.10c}$$

therefore it follows from (3.9b) that the second term on the LHS of (3.10b) is itself $O(\Delta^2)$; consequently, (3.9a) implies (3.10b). This completes the proof of Lemma 3.2. ∎

We construct $g_i = g(v_{i-1}, v_i, v_{i+1})$ that satisfies (3.9) in the following way:

$$\begin{aligned} g_i &= s_{i+1/2} \max[0, \min(|\tilde{g}_{i+1/2}|, \tilde{g}_{i-1/2} \cdot s_{i+1/2})] \\ &= s_{i+1/2} \min(|\tilde{g}_{i+1/2}|, |\tilde{g}_{i-1/2}|), \quad \text{when } \tilde{g}_{i+1/2} \cdot \tilde{g}_{i-1/2} \geqslant 0, \\ &= 0, \qquad\qquad\qquad\qquad\quad \text{when } \tilde{g}_{i+1/2} \cdot \tilde{g}_{i-1/2} \leqslant 0, \end{aligned} \tag{3.11a}$$

where

$$\tilde{g}_{i+1/2} = \tfrac{1}{2}[Q(\bar{\nu}_{i+1/2}) - (\bar{\nu}_{i+1/2})^2]\, \Delta_{i+1/2}v, \tag{3.11b}$$

$$s_{i+1/2} = \mathrm{sgn}(\tilde{g}_{i+1/2}). \tag{3.11c}$$

LEMMA 3.3. *Let g_i be defined by* (3.11); *then relations* (3.9a) *and* (3.9b) *are satisfied, and*

$$|\gamma_{j+1/2}| = |g_{j+1} - g_j|/|\Delta_{j+1/2}v| \leqslant \tfrac{1}{2}|Q(\bar{\nu}_{j+1/2}) - (\bar{\nu}_{j+1/2})^2|. \tag{3.12}$$

Proof. First let us assume that $\tilde{g}_{j+1/2}\tilde{g}_{j-1/2} \geqslant 0$; then using definition (3.11a) and the relation $\min(a, b) = \frac{1}{2}[(a + b) - |a - b|]$ we get

$$\begin{aligned} g_j &= \tfrac{1}{2}[\tilde{g}_{j-1/2} + \tilde{g}_{j+1/2} - s_{j+1/2}|\tilde{g}_{j+1/2} - \tilde{g}_{j-1/2}|] \\ &= \tilde{g}_{j\pm1/2} + \tfrac{1}{2}[\mp(\tilde{g}_{j+1/2} - \tilde{g}_{j-1/2}) - s_{j+1/2}|\tilde{g}_{j+1/2} - \tilde{g}_{j-1/2}|]. \end{aligned} \tag{3.13a}$$

From (3.11b) we conclude that if v is smooth and $Q(x)$ at least Lipschitz continuous, then

$$\tilde{g}_{j+1/2} - \tilde{g}_{j-1/2} = O(\Delta^2). \tag{3.13b}$$

Thus (3.13a) and (3.13b) imply that

$$g_j = \tilde{g}_{j\pm1/2} + O(\Delta^2). \tag{3.13c}$$

It is easy to see that (3.13c) holds even if $\tilde{g}_{j-1/2}\tilde{g}_{j+1/2} \leqslant 0$, for then $g_j = 0$ but $\tilde{g}_{j\pm1/2} = O(\Delta^2)$ itself (since $\Delta_{j\pm1/2}v = O(\Delta^2)$).

Relations (3.9a) and (3.9b) follow immediately by rewriting (3.13c) as

$$g_j = \tilde{g}_{j+1/2} + O(\Delta^2), \quad g_{j+1} = \tilde{g}_{j+1/2} + O(\Delta^2).$$

We turn now to prove (3.12). We observe from definition (3.11a) that g_j and g_{j+1} cannot be of different sign, hence

$$\begin{aligned} |g_{j+1} - g_j| &\leqslant \max(|g_j|, |g_{j+1}|) \\ &\leqslant \max[\min(|\tilde{g}_{j-1/2}|, |\tilde{g}_{j+1/2}|), \min(|\tilde{g}_{j+1/2}|, |\tilde{g}_{j+3/2}|)] \\ &\leqslant |\tilde{g}_{j+1/2}|. \end{aligned}$$

Thus it follows from (3.11b) that

$$\begin{aligned} |\gamma_{j+1/2}| = |g_{j+1} - g_j|/|\Delta_{j+1/2}v| &\leqslant |\tilde{g}_{j+1/2}|/|\Delta_{j+1/2}v| \\ &\leqslant \tfrac{1}{2}|Q(\bar{\nu}_{j+1/2}| - (\bar{\nu}_{j+1/2})^2|; \end{aligned}$$

this completes the proof of (3.12).

We show now that the 5-point second order accurate scheme (3.8) with (3.11) is TVNI under the same CFL restriction of the original 3-point first order accurate TVNI scheme (3.1).

LEMMA 3.4. *Suppose $Q(x)$ satisfies* (3.2) *and g_j is defined by* (3.11); *then finite-difference scheme* (1.4) *with numerical flux* (3.8) *is* TVNI *under* CFL *restriction* (3.3).

Proof. Since (3.8) is (3.1) applied to a modified flux f_i^M (3.8a), it can also be rewritten as (3.5) with a modified CFL number $\bar{\nu}_{j+1/2}^M$ (3.8b). Since Lemma 2.2 is valid also for 5 point schemes of form (2.9), we conclude from Lemma 3.1 that scheme (3.8) is TVNI under the modified CFL restriction

$$\max_j|\bar{\nu}_{j+1/2}^M| \leqslant \mu. \tag{3.14}$$

To complete the proof of Lemma 3.4 we show that (3.14) is implied by the original CFL condition (3.3). Using (3.12) and (3.2) we get

$$\begin{aligned} |\bar{\nu}_{j+1/2}^M| = |\bar{\nu}_{j+1/2} + \gamma_{j+1/2}| &\leqslant |\bar{\nu}_{j+1/2}| + |\gamma_{j+1/2}| \\ &\leqslant |\bar{\nu}_{j+1/2}| + \tfrac{1}{2}|\,Q(\bar{\nu}_{j+1/2}) - (\bar{\nu}_{j+1/2})^2| \\ &\leqslant |\bar{\nu}_{j+1/2} + \tfrac{1}{2}[1 - (\bar{\nu}_{j+1/2})^2] \\ &= 1 - \tfrac{1}{2}(|\bar{\nu}_{j+1/2}| - 1)^2 \leqslant 1 \end{aligned}$$

whenever $|\bar{\nu}_{j+1/2}| \leqslant Q(\bar{\nu}_{j+1/2}) \leqslant 1$; thus the original CFL restriction (3.3) guarantees that inequalities (2.10) of Lemma 2.1 are satisfied and therefore the scheme is TVNI.

Remark 1. If $\Delta_{j+1/2}v = 0$, then it follows immediately from definition (3.11) that $g_j = g_{j+1} = 0$. This shows that the modified numerical flux (3.8c) is consistent with the physical flux $f(u)$ in the sense of (1.4c). Scheme (3.8) + (3.11) is TVNI and therefore it has convergent subsequences for all initial data of bounded total variation; the limits of these subsequences are weak solutions of scalar conservation law (2.1). To complete the convergence proof one has to show that all these limits are the same. In the constant coefficient case the solution to IVP (2.1) is unique and therefore the scheme is convergent (note that the scheme is nonlinear even in the constant coefficient case!). In the nonlinear case, convergence will follow if one shows that the scheme is consistent with Oleinik's entropy condition in the sense of (1.5). We shall discuss consistency with the entropy condition in Section 5.

Remark 2. Condition (3.6) is only a necessary condition for second-order accuracy. It becomes a sufficient condition if the coefficient in the $O(\Delta^2)$ term in (3.6) is differentiable, except possibly at a finite number of points $N(t, \Delta t)$, such that $N\,\Delta t \to 0$ as $\Delta t \to 0$ for all t. It is clear from (3.13) and (3.9) that the troublesome points where the scheme (3.8) + (3.11) may degenerate locally to $O(\Delta^2)$ truncation error are those where $s_{j+1/2}$ in (3.11c) is discontinuous, i.e., where $Q(\nu) - \nu^2 = 0$ or $u_x = 0$. The fact that the scheme is TVNI controls the possible increase of the number of local extremum points in the computed solution. The schemes that we consider in Section 5 all have the monotonicity property (see Section 2), i.e., the number of local extremum points in the computed solution is nonin-

creasing in time, and thus bounded by that of the initial data.

Remark 3. The modified equation of scheme (3.1), i.e., the equation which it approximates to second-order accuracy, is (2.6) with

$$\beta(u, \lambda) = \tfrac{1}{2}[Q(\nu) - \nu^2], \quad \nu = \lambda a(u). \tag{3.15a}$$

We rewrite the modified equation as

$$u_t + \{f - [\Delta t/(2\lambda^2)][Q(\nu) - \nu^2]u_x\}_x = 0, \tag{3.15b}$$

and observe that $\tilde{g}_{i+1/2}$ in (3.11b), and consequently g_i in (3.11a), is an approximation to the term

$$g \approx [\Delta t/(2\lambda)][Q(\nu) - \nu^2]u_x = \tfrac{1}{2}[Q(\nu) - \nu^2](\Delta x u_x). \tag{3.15c}$$

Our method to convert a first order accurate TVNI scheme into a second order accurate TVNI scheme is based on the following heuristic argument described in the end of Section 2: first order scheme (3.1) approximates

$$u_t + [f - (1/\lambda)g]_x = 0 \tag{3.16a}$$

to second-order accuracy. Relation (3.7) shows that applying the same scheme to

$$u_t + [f + (1/\lambda)g]_x = 0 \tag{3.16b}$$

results in a second order accurate approximation to $u_t + f_x = 0$. To be able to apply the scheme to (3.16b) we have to define $g(u)$ with a bounded derivative dg/du; therefore we use the particular form (3.11) rather than a direct discretization of (3.15c) (see [5]). Lemma 3.3 shows that the so-defined g satisfies these requirements.

Remark 4. We observe that if the term $Q(\nu + \gamma)$ in (3.8d) is replaced by $Q(\nu) + |\gamma|$, then the resulting scheme remains second order accurate and TVNI under the same CFL restriction. The corresponding numerical flux takes the somewhat simpler form

$$\lambda \bar{f}_{j+1/2} = (\lambda/2)(f_j + f_{j+1}) - \tfrac{1}{2}Q(\bar{\nu}_{j+1/2})\Delta_{j+1/2}v + G_{j+1/2}, \tag{3.17a}$$

$$G_{j+1/2} = S \max[0, \min(\tilde{g}_{j-1/2}S, |\tilde{g}_{j+1/2}|, \tilde{g}_{j+3/2}S)], \tag{3.17b}$$

where, as before,

$$\tilde{g}_{j+1/2} = \tfrac{1}{2}[Q(\bar{\nu}_{j+1/2}) - (\bar{\nu}_{j+1/2})^2]\Delta_{j+1/2}v, \quad S = \operatorname{sgn}(\tilde{g}_{j+1/2}). \tag{3.17c}$$

To derive this expression we use the definition of $\gamma_{j+1/2}$ in (3.8b) and note that

$$\begin{aligned} G_{j+1/2} &= \tfrac{1}{2}[g_j + g_{j+1} - |g_j + g_{j+1}| \operatorname{sgn}(\Delta_{j+1/2}v)] \\ &= S \min(g_j S, g_{j+1} S). \end{aligned}$$

Form (3.17b) follows from the definition of g_j in (3.11a). Note the resemblance of (3.17b) with $Q(x) = x^2 + 1/4$ to the "anti-diffusion" term in [1].

4. SYSTEMS OF CONSERVATION LAWS

In this section we describe how to extend our new scalar scheme of Section 3 to systems of conservation laws. Our extension technique is a somewhat generalized version of the procedure suggested in [14]. The basic idea is to extend the scalar scheme to the system case by applying it "scalarly" to each of the (appropriately linearized) characteristic variables.

Let

$$S(u) = (R^1(u), R^2(u), ..., R^m(u)) \tag{4.1a}$$

be a matrix, the columns of which are the right eigenvectors of the Jacobian matrix $A(u)$ in (1.2a). Then

$$S^{-1}AS = \Lambda, \quad \Lambda_{ij} = a^i(u)\,\delta_{ij}. \tag{4.1b}$$

The rows $L^1(u)$, $L^2(u)$, ..., $L^m(u)$ of $S^{-1}(u)$ constitute a complete system of left eigenvectors of $A(u)$ which is biorthonormal to the system of right eigenvectors, i.e.,

$$L^i R^j = \delta_{ij}. \tag{4.1c}$$

In the constant coefficient case $A(u) \equiv A = \text{const}$

$$u_t + Au_x = 0, \quad u(x, 0) = \phi(x), \quad -\infty < x < \infty. \tag{4.2}$$

One defines characteristic variables $w = (w^k)$ by

$$w^k = L^k u, \quad w = S^{-1}u. \tag{4.3}$$

It follows from (4.1) that (4.2) decouples into m scalar characteristic equations, $1 \leqslant k \leqslant m$

$$w_t^k + a^k w_x^k = 0, \quad w^k(x, 0) = L^k\phi(x), \quad -\infty < x < \infty. \tag{4.4}$$

This offers a natural way of extending a scalar scheme to a constant coefficient system of equations (4.2) by applying it "scalarly" to each of the m scalar characteristic equations (4.4).

The characteristic variables w^k in (4.3) can also be viewed as the components of u in the coordinate system $\{R^k\}$, i.e.,

$$u = \sum_{k=1}^{m} w^k R^k. \tag{4.5}$$

We use this interpretation of characteristic variables to extend the scalar scheme to general nonlinear systems of conservation laws.

Let $v_{j+1/2} = V(v_j, v_{j+1})$ be an average of v_j and v_{j+1}, i.e., a smooth function $V(u, v)$ such that

$$V(u, v) = V(v, u), \tag{4.6a}$$

$$V(u, u) = u; \tag{4.6b}$$

and let $\alpha^k_{j+1/2}$ denote the component of $\Delta_{j+1/2}v = v_{j+1} - v_j$ in the coordinate system $\{R^k(v_{j+1/2})\}$

$$\Delta_{j+1/2}v = \sum_{k=1}^{m} \alpha^k_{j+1/2} R^k_{j+1/2}, \tag{4.7a}$$

$$\alpha^k_{j+1/2} = L^k_{j+1/2} \Delta_{j+1/2} v. \tag{4.7b}$$

Here we use the notation convention $b_{j+1/2} = b(v_{j+1/2}) = b(V(v_j, v_{j+1}))$.

We now extend the scalar (3.8) + (3.11) to general systems of conservation laws as follows:

$$v^{n+1}_j = v^n_j - \lambda(\bar f^n_{j+1/2} - \bar f^n_{j-1/2}), \tag{4.8a}$$

$$\bar f^n_{j+1/2} = \frac{1}{2}[f(v_j) + f(v_{j+1})] + \frac{1}{2\lambda}\sum_{k=1}^{m} R^k_{j+1/2}[g^k_j + g^k_{j+1} - Q^k(\nu^k_{j+1/2} + \gamma^k_{j+1/2})\alpha^k_{j+1/2}], \tag{4.8b}$$

where $\nu^k_{j+1/2} = \lambda a^k(v_{j+1/2})$ and

$$g^k_i = s^k_{i+1/2} \max[0, \min(|\tilde g^k_{i+1/2}|, \tilde g^k_{i-1/2} s^k_{i+1/2})],$$

$$s^k_{i+1/2} = \operatorname{sgn}(\tilde g^k_{i+1/2}), \tag{4.8c}$$

$$\tilde g^k_{i+1/2} = \tfrac{1}{2}[Q^k(\nu^k_{i+1/2}) - (\nu^k_{i+1/2})^2]\,\alpha^k_{i+1/2}, \tag{4.8d}$$

$$\gamma^k_{i+1/2} = (g^k_{i+1} - g^k_i)/\alpha^k_{i+1/2}, \quad \text{when } \alpha^k_{i+1/2} \neq 0, \tag{4.8e}$$

$$= 0, \quad \text{when } \alpha^k_{i+1/2} = 0.$$

The second order accurate one-step Lax–Wendroff scheme can be represented as

$$\bar f^{\mathrm{LW}}_{j+1/2} = \frac{1}{2}[f(v_j) + f(v_{j+1})] - \frac{1}{2\lambda}\sum_{k=1}^{m} (\nu^k_{j+1/2})^2 \alpha^k_{j+1/2} R^k_{j+1/2}$$

$$\equiv \frac{1}{2}[f(v_j) + f(v_{j+1})] - \frac{1}{2}\lambda[A(v_{j+1/2})]^2 \Delta_{j+1/2} v. \tag{4.9}$$

LEMMA 4.1. *Suppose $\{Q^k(x)\}$ are Lipschitz continuous; then* (4.8) *satisfies*

$$\bar f_{j+1/2} = \bar f^{\mathrm{LW}}_{j+1/2} + O(\Delta^2). \tag{4.10}$$

Proof. Rewrite (4.8b) as

$$\bar f_{j+1/2} = \bar f^{\mathrm{LW}}_{j+1/2} + \sum_{k=1}^{m} R^k_{j+1/2}\tau^k_{j+1/2}, \tag{4.11a}$$

where

$$2\lambda\tau^k_{j+1/2} = g^k_j + g^k_{j+1} - [Q^k(\nu^k_{j+1/2}) - (\nu^k_{j+1/2})^2]\,\alpha^k_{j+1/2} - [Q^k(\nu^k_{j+1/2} + \gamma^k_{j+1/2}) - Q^k(\nu^k_{j+1/2})]\,\alpha^k_{j+1/2}, \tag{4.11b}$$

and then use (3.10c) and conclude (3.9) from Lemma 3.3.

We define the total variation $\mathrm{TV}(v)$ of the vector mesh function v to be

$$\mathrm{TV}(v) = \sum_{j=-\infty}^{\infty}\sum_{k=1}^{m} |\alpha^k_{j+1/2}|, \tag{4.12}$$

where $\alpha^k_{j+1/2}$ is defined by (4.7), and show

LEMMA 4.2. *Suppose $Q^k(x)$ satisfies* (3.2) *for all k, and that $A(u) \equiv A = \text{const}$. Then scheme* (4.8) *is* TVNI *under the* CFL *restriction*

$$\lambda \max|a^k| \leqslant \mu = \min \mu^k \leqslant 1, \tag{4.13}$$

where μ^k are the restrictions in (3.2).

Proof. Because of the assumption $A(u) \equiv \text{const}$, $\{R^k\}$, $\{L^k\}$, and $\{a^k\}$ are all constant. Multiplying (4.8b) from the left by L^k, we obtain (3.8d) for the characteristic variable w^k in (4.3); g^k_j and $\gamma^k_{j+1/2}$ in (4.8) become identical with (3.11). Thus by Lemma 3.4 we conclude that under condition (4.3) the total variation of each of the characteristic variables is nonincreasing, and therefore the total variation (4.12) is nonincreasing as well. ∎

COROLLARY 4.3. *Scheme* (4.8) *in the constant coefficient case is convergent under restriction* (4.13) *for all initial data of bounded total variation, and is second order accurate.*

We remark that this corollary is not trivial since the scheme is highly nonlinear even in the constant coefficient case.

Our technique to extend scalar schemes to the system case does not require any particular form of averaging $V(u, v)$ (4.6). Roe in [15] uses a specific form of averaging that on top of being mathematically pleasing, also enables the computational advantage of perfectly resolving stationary discontinuities.

In [7] we show that if the system of conservation laws (1.1) possesses an entropy function (1.3), then it is symmetrizable, and there exists a mean value Jacobian $A(u, v)$ such that

$$f(v) - f(u) = A(u, v)(v - u), \tag{4.14a}$$

$$A(u, u) = A(u), \tag{4.14b}$$

$A(u, v)$ has real eigenvalues $\{a^k(u, v)\}_{k=1}^m$ and a complete set of right eigenvectors $\{R^k(u, v)\}_{k=1}^m$.

In the context of scheme (4.8), Roe's extension technique is expressed by taking $a_{j+1/2}^k$ and $R_{j+1/2}^k$ in (4.7) and (4.8) to be the eigenvalues and the right eigenvectors of the mean value Jacobian $A(v_j, v_{j+1})$ (4.14a), respectively. Thus if the $\alpha_{j+1/2}^k$ are defined by (4.7a)

$$v_{j+1} - v_j = \sum_{k=1}^m \alpha_{j+1/2}^k R_{j+1/2}^k, \tag{4.15a}$$

then it follows from (4.14a) that

$$f(v_{j+1}) - f(v_j) = \sum_{k=1}^m \alpha_{j+1/2}^k a_{j+1/2}^k R_{j+1/2}^k. \tag{4.15b}$$

Relation (4.15) makes scheme (4.8) a more faithful extension of (3.8) in the sense that (4.8) for $m = 1$ is identical with scalar scheme (3.8).

In the case of the Euler equations of gasdynamics, where the flux $f(u)$ is homogeneous function of u of degree 1, it is possible to express $A(u, v)$ in (4.14) as

$$A(u, v) = A(V(u, v)). \tag{4.16}$$

This relatively simple function $V(u, v)$ (see [15]) will be described in Section 6.

Remark 1. Note that we use $Q^k(x)$, thus allowing different functions (3.2) for different characteristic fields. As observed by Roe [14], the extension technique of this section permits even the use of completely different scalar schemes for different characteristic fields.

Remark 2. Version (3.17) extends to systems in a similar way:

$$\lambda \bar{f}_{j+1/2} = \frac{1}{2}\lambda(f_j + f_{j+1}) + \sum_{k=1}^m \left[G_{j+1/2}^k - \frac{1}{2} Q(\nu_{j+1/2}^k)\alpha_{j+1/2}^k \right] R_{j+1/2}^k, \tag{4.17a}$$

$$G_{j+1/2}^k = S \max[0, \min(\tilde{g}_{j+1/2}^k S, |\tilde{g}_{j+1/2}^k|, \tilde{g}_{j+1/2}^k S)], \tag{4.17b}$$

where

$$\tilde{g}_{j+1/2}^k = \frac{1}{2}[Q(\nu_{j+1/2}^k) - (\nu_{j+1/2}^k)^2]\,\alpha_{j+1/2}^k, \quad S = \operatorname{sgn}(\tilde{g}_{j+1/2}^k). \tag{4.17c}$$

Remark 3. The particular definition (4.12) of total variation is motivated by the definition of Glimm's functional in [3]. When applied to a piecewise smooth solution $u(x, t)$ of (1.1)

$$\lim_{\Delta x \to 0} \mathrm{TV}(u) = \int_{-\infty}^{\infty} \sum_{k=1}^m |L^k(u) u_x|\, dx + \sum_j \sum_{k=1}^m |\alpha^k(x_j)|, \tag{4.18}$$

where x_j are points of discontinuity, and $\alpha^k(x_j)$ denotes the value of $\alpha_{j+1/2}^k$ in (4.7) evaluated with respect to $v_j = u((x_j)_-, t)$, $v_{j+1} = u((x_j)_+, t)$.

There is no reason to expect that functional (4.18), and consequently (4.12), is generally nonincreasing with t. Based on Glimm's results [3] we do, however, believe that this functional (under certain conditions) is bounded in t. At this time we do not have estimates of the possible increase in total variation in solutions of scheme (4.8), and therefore cannot prove convergence in the nonlinear system case.

5. ON ENTROPY, RESOLUTION, AND $Q(x)$

The proposed scheme of this paper is in conservation form. Hence, if the scheme is convergent, then its limit is a weak solution of (1.1), i.e., it satisfies the differential equation pointwise wherever it is smooth, and at points of discontinuity it satisfies the Rankine–Hugoniot jump relations. It is well known that the fact that a scheme is in conservation form does not by itself guarantee that all the discontinuities in its limit solution are physically relevant, i.e., they do not necessarily satisfy the entropy jump inequality associated with (1.3) (see [8]).

In practice, however, we are not concerned with the limit of an infinite sequence of computations, but rather with a single computation on a relatively crude mesh. Thus, to have a viable numerical method it is not sufficient to ensure that discontinuities in its *limit* solution are admissible. We have to insist upon having a good approximation to the exact solution on a finite mesh.

To formulate such criteria let us consider the Riemann IVP for (1.1):

$$u(x,0) = \varphi(x) = u_L, \quad \text{if } x < 0, \qquad = u_R, \quad \text{if } x > 0, \tag{5.1}$$

where u_L and u_R satisfy the Rankine–Hugoniot relations with some speed of propagation s.

If $u(x, t) = \varphi(x - st)$ satisfies the entropy inequality, then we require that the numerical scheme possess a steady progressing profile with a narrow transition from u_L to u_R (see [5]); we refer to this property as *resolution.*

If, on the other hand, the solution $u(x, t) = \varphi(x - st)$ is inadmissible, then the physical solution is a fan of waves (i.e., a function of x/t) which consists of a large rarefaction wave in the same field as the initial discontinuity, and possibly weaker waves in the other fields. Since the breakup of the initial discontinuity in the exact solution occurs spontaneously, we require that the numerical scheme also break it up and at a *fast* rate; we refer to this property as *entropy enforcement.* (To demonstrate the need for this requirement, we would like to point out that the Godunov scheme, which is so attractive because of it being "physical," actually exhibits a surprisingly poor entropy enforcement; see [11].)

We consider here systems of conservation laws where the characteristic fields are either *genuinely nonlinear* ($a_u^k R^k \neq 0$) or *linearly degenerate* ($a_u^k R^k \equiv 0$; see [12]). The waves of a genuinely nonlinear field are either shocks or rarefaction waves, depending whether the characteristics are convergent or divergent. The waves of a linearly degenerate field are exclusively contact discontinuities.

The extension technique of Section 4 consists of applying scalar schemes, not necessarily the same, to each of the characteristic fields. Therefore, it makes sense to approach the questions of entropy enforcement and resolution for the system by examining the corresponding ones for the scalar scheme in each of the fields separately. We note that with this extension technique we have the possibility to custom fit the scheme to the special computational needs of each of the different characteristic fields. First let us consider the 3-point first order accurate scheme (3.1). The effective numerical viscosity coefficient (2.7a) of this scheme is

$$\beta(u, \lambda) = \tfrac{1}{2}[Q(\nu) - \nu^2], \quad \nu = \lambda a. \tag{5.2}$$

Hence a natural choice of $Q(x)$ under restriction (3.2) is $Q(x) = |x|$, as it gives the least dissipative TVNI scheme of form (3.1). This scheme with $Q(x) = |x|$ can be rewritten in form (3.5) as

$$v_j^{n+1} = v_j^n - (\overline{\nu}_{j+1/2})^- \Delta_{j+1/2} v^n - (\overline{\nu}_{j-1/2})^+ \Delta_{j-1/2} v^n, \tag{5.3a}$$

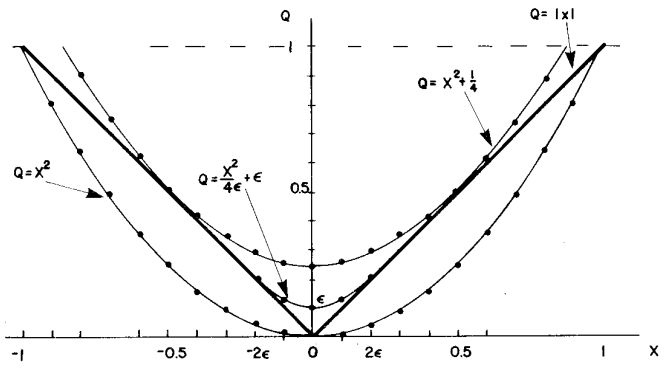

FIGURE 1

where

$$\nu^- = \min(\nu, 0) = \tfrac{1}{2}(\nu - |\nu|), \quad \nu^+ = \max(\nu, 0) = \tfrac{1}{2}(\nu + |\nu|). \tag{5.3}$$

Scheme (5.3) is a generalization of the well-known upstream differencing scheme of Courant, Isaacson, and Rees, and it is well investigated in the literature (see [5, 10]).

Consider now the application of this scheme to Riemann problem (5.1), where u_L and u_R satisfy the Rankine–Hugoniot relations with a zero speed of propagation, i.e., $f(u_L) = f(u_R)$. It is easy to see that the initial discontinuity is a stationary solution of (5.3), regardless of whether or not it satisfies the entropy condition. (The same is true for its system extension with Roe's linearization; see [10, 15]). It is clear, therefore, that this scheme admits nonphysical solutions. We note, however, that this statement applies only to discontinuities in a genuinely nonlinear field, since there are no entropy considerations in a linearly degenerate one.

This type of entropy violation in scheme (5.3) is related to the fact that numerical viscosity (5.2) vanishes for $\nu = 0$. Therefore it seems possible to eliminate this sort of entropy violation by simply modifying $Q(x) = |x|$ near $x = 0$ to be positive. To make $Q(x)$ smoother at the same time we define for $0 < \varepsilon \leqslant \frac{1}{2}$

$$Q(x) = (x^2/(4\varepsilon)) + \varepsilon, \quad \text{for } |x| < 2\varepsilon, \qquad = |x|, \quad \text{for } |x| \geqslant 2\varepsilon, \tag{5.4}$$

with, say, $\varepsilon = 0.1$ (see Fig. 1).

This change increases the amount of numerical viscosity for $|x| < 2\varepsilon$ so that now $\beta(u, \lambda) > 0$ for $|\nu| < 1$. Then $\beta(u, \lambda)$ vanishes only for $\nu = 1$, which can be handled by taking $\mu < 1$, say, $\mu = 0.95$, in CFL restriction (3.3).

We turn now to a discussion of the entropy enforcement and resolution of the proposed second order accurate

TVNI scheme of this paper. Being unable to rigorously analyze these properties, we have conducted a series of numerical experiments on Riemann problem (5.1) for the Euler equations of gasdynamics, part of which is presented in Section 7. In these experiments we have used scheme (4.8) with $Q^k(x)$ defined by (5.4) for all k, with $\varepsilon = 0.05$, 0.1, and 0.25; we have also experimented with $\varepsilon = 0$ for the linearly degenerate characteristic field. In all our experiments the scheme has demonstrated high resolution of shocks and strong entropy enforcement. Contact discontinuities, although much better resolved than those of the corresponding first order accurate scheme, were rather smeared.

As is to be expected, a larger ε brings about improvement in entropy enforcement at the expense of some deterioration in resolution. The dependence on ε is rather slight, however, and the performance of the scheme for $\varepsilon = 0.05$ is about the same as for $\varepsilon = 0.25$. We have selected $\varepsilon = 0.1$ for genuinely nonlinear field and $\varepsilon = 0$ for linearly degenerate fields as our parameters for future computations.

Next we present a heuristic analysis which may explain our numerical results and give some guidance for possible improvements. We observe that when applied to a jump discontinuity of form (5.1), g_j in (4.8) is identically zero. Hence the opening move in this entropy versus resolution game is up to the "viscous" 3-point first order accurate scheme. To understand the development of the numerical solution at later stages, we examine the operation of scalar scheme (3.8) with (3.11) on data of the form

$$\begin{aligned} u(x,0) = \varphi(x) &= u_{\mathrm{L}}, \qquad \text{if } x \leqslant x_{\mathrm{L}}, \\ &= u(x), \quad \text{if } x_{\mathrm{L}} \leqslant x \leqslant x_{\mathrm{R}}, \\ &= u_{\mathrm{R}}, \qquad \text{if } x_{\mathrm{R}} \leqslant x \end{aligned} \tag{5.5}$$

for some $x_{\mathrm{L}} \leqslant x_{\mathrm{R}}$, where $u(x)$ is some smooth monotone transition from u_{L} to u_{R}. To simplify things further, let us assume that $f(u)$ is convex, so that the asymptotic solution to the initial data (5.5) is either a shock or a centered rarefaction wave.

When operating with a monotonicity-preserving scheme on data of form (5.5), we obtain functions in the same class, i.e., monotone transitions from u_{L} to u_{R}. Hence we can identify the speed at which each value of u (excluding u_{L} and u_{R}) propagates due to the operation of the finite difference schemes, and thus introduce the notion of a *numerical characteristic speed.*

Our heuristic analysis relies on the representation of the second order accurate scheme as the original first order accurate TVNI scheme operating on a modified flux.

Let us denote the numerical characteristic speed of the first order accurate scheme by $\bar{a}$, and measure its deviation from the exact characteristic speed by

$$\bar{a} = a + (1/\lambda)\delta, \quad \delta = O(\Delta); \tag{5.6a}$$

that $\delta = O(\Delta)$ follows from the first order accuracy of the scheme. Similarly, let us denote the numerical characteristic speed of our second order accurate scheme by $\bar{\bar{a}}$, thus

$$\bar{\bar{a}} = a + O(\Delta^2). \tag{5.6b}$$

On the other hand, this second order accurate scheme is obtained by applying the original first order scheme to the modified characteristic field $a + (1/\lambda)\gamma$, where γ is defined by (3.8b) and $\gamma = O(\Delta)$ by (3.9b). Therefore, it follows from (5.6a) that

$$\bar{\bar{a}} = (a + (1/\lambda)\gamma) + (1/\lambda)\delta + O(\Delta^2). \tag{5.6c}$$

Comparing (5.6c) with (5.6b), we conclude that

$$\gamma = -\delta + O(\Delta^2). \tag{5.6d}$$

The observed excessive smearing of the first order accurate scheme indicates that $\bar{a} < a$ near u_{L} and $\bar{a} > a$ near u_{R}. Hence, it follows from (5.6a) that $\delta < 0$ near u_{L} and $\delta > 0$ near u_{R}, i.e., the δ characteristic field is divergent. From (5.6d) we conclude that the γ characteristic field is convergent.

Our second-order TVNI scheme can be described in these terms as the operation of a scheme with an overdivergent numerical characteristic field on a suitably modified overconvergent characteristic field that results in a cancellation to $O(\Delta^2)$ of the error. This rather unusual mechanism allows the scheme to achieve both high resolution and strong entropy enforcement, a task that seems self-contradictory when considered within the context of artificial viscosity methods. This model also serves to explain the rather uniform high resolution of shocks with respect to their strength and speed of propagation, and with respect to the values of ε in (5.4) and the CFL number used in the computation.

We turn now to discuss the questions of entropy and resolution in a linearly degenerate field. The only possible waves in this field are contact discontinuities, for which the entropy jump inequality turns out to be an equality, i.e., a Rankine–Hugoniot relation for the entropy function. Consequently, there are no entropy considerations in a linearly degenerate field.

The Riemann invariants of this field, one of which is the characteristic speed, are the same on both sides of the contact discontinuity; the latter travels passively with the speed of the characteristic field. Unlike shocks, contact discontinuities do not form spontaneously, and may be present in the solution either by being present in the initial data or as a result of shock interaction (see [12]). Because of all these facts, the computation of a contact discontinuity

is very much like that of (5.1) for the scalar constant coefficient case (see [5]).

Since the characteristic field in a contact discontinuity region is parallel, even a slight divergence of the numerical characteristic field ($\bar{\bar{a}} - a$) in (5.6c) brings about a rather severe loss of resolution. We may prevent excessive smearing by modifying the numerical characteristic field to be slightly convergent in this region (see [5, 6]). To do so we increase the convergence of the γ field by increasing g in (4.8) as follows: We replace (4.8c) in a linearly degenerate k-field by

$$g_i = \bar{g}_i + \theta_i \bar{\bar{g}}_i, \tag{5.7a}$$

where $\bar{g}_i$ is the quantity defined by the RHS of (4.8c), θ_i is such that

$$\theta_i = O(\Delta), \quad 0 \leqslant \theta_i \leqslant 1, \tag{5.7b}$$

and $\bar{\bar{g}}_i$ is

$$\bar{\bar{g}}_i = S \max[0, \min(S\sigma_{j-1/2}\alpha_{j-1/2}, \sigma_{j+1/2}|\alpha_{j+1/2}|)], \tag{5.7c}$$

where $S = \operatorname{sgn}(\alpha_{j+1/2})$ and $\sigma_{j+1/2} = \sigma(\nu_{j+1/2})$ is restricted by

$$0 \leqslant \sigma(\nu) \leqslant 1 - |\nu| - \tfrac{1}{2}[Q(\nu) - \nu^2] \quad \text{for } |\nu| \leqslant 1. \tag{5.7d}$$

It is easy to verify that the resulting scheme retains its second-order accuracy and remains TVNI under the same CFL restriction. A particular choice of θ and $\sigma(\nu)$ that satisfies (5.7b) and (5.7d), respectively, is

$$\theta_i = |\alpha_{i+1/2} - \alpha_{i-1/2}|/(|\alpha_{1+1/2}| + |\alpha_{i-1/2}|), \tag{5.8a}$$

$$\sigma(\nu) = \tfrac{1}{2}[1 - Q(\nu)]. \tag{5.8b}$$

We have experimented with the so-modified scheme on a variety of Riemann problems for the Euler equations of gasdynamics. In all these computations we have found high resolution of contact discontuities. Although the decomposition of $(v_{j+1} - v_j)$ into the different characteristic field (4.7) is only approximate, we find that the modification in the linearly degenerate field (5.7) does not significantly change the waves in the other fields.

Remark 1. We may enhance resolution and entropy enforcement in the genuinely nonlinear fields by increasing γ where the k-characteristic field is convergent and decreasing it where the field is locally divergent. To accomplish this, we let θ_i in (5.7a) satisfy

$$\theta_i = O(\Delta), \quad -1 \leqslant \theta_i \leqslant 1 \tag{5.9a}$$

and construct it so that

$$\operatorname{sgn} \theta_i = -\operatorname{sgn}(a_{i+1} - a_i). \tag{5.9b}$$

Our preliminary experiments in two-dimensional calculations using dimensional splitting indicate a need for such an enhancement mechanism. We find no need for improvement in our one-dimensional computations.

Remark 2. Earlier in our investigation we pursued a different avenue to enforce entropy inequality. This approach is motivated by the representation of first-order scheme (5.3) as a Godunov-type, and by Liu's theory [13] on generalizing the Oleinik entropy condition to systems of conservation laws. In this approach we add a nonnegative entropy viscosity term to $Q(x) = |x|$ in each of the characteristic fields; this term is a measure of the violation of the Oleinik entropy condition in the particular characteristic field. Thus it vanishes for admissible discontinuities, but if large enough, it enforces entropy inequality through viscosity whenever the approximate solution is inadmissible (see [9, 10, 11]).

This technique is elegant and amenable to rigorous analysis. However, it complicates the programming, increases the CPU time, and above all it does not perform as well as the simple correction to $Q(x)$ in (5.4). Though with reluctance, we give precedence to computational efficiency over neat analysis, and pursue it no further.

6. APPLICATION TO EULER EQUATIONS OF GASDYNAMICS

In this section we describe how to apply our new scheme (4.8) to the Euler equations of gasdynamics,

$$w_t + f(w)_x = 0, \tag{6.1a}$$

$$w = \begin{pmatrix} \rho \\ m \\ E \end{pmatrix}, \quad f(w) = uw + \begin{pmatrix} 0 \\ p \\ pu \end{pmatrix}, \tag{6.1b}$$

$$p = (\gamma - 1)(E - \tfrac{1}{2}\rho u^2). \tag{6.1c}$$

Here ρ, u, p, and E are the density, velocity, pressure, and total energy, respectively; $m = \rho u$ is the momentum and we take $\gamma = 1.4$.

The eigenvalues of the Jacobian matrix $A(w) = f_w$ are

$$a_1(w) = u - c, \quad a_2(w) = u, \quad a_3(w) = u + c, \tag{6.2a}$$

where c is the sound speed, $c = (\gamma p/\rho)^{1/2}$.

The corresponding right eigenvectors are

$$R_1(w) = \begin{pmatrix} 1 \\ u - c \\ H - uc \end{pmatrix}, R_2(w) = \begin{pmatrix} 1 \\ u \\ \frac{1}{2}u^2 \end{pmatrix}, R_3(w) = \begin{pmatrix} 1 \\ u + c \\ H + uc \end{pmatrix}, \tag{6.2b}$$

where $H = (E + p)/\rho = c^2/(\gamma - 1) + \frac{1}{2}u^2$ is the enthalpy.

Let $\alpha^k(w_L, w_R)$, $k = 1, 2, 3$, be the solution of the following system of linear equations (4.7),

$$w_R - w_L = \sum_{k=1}^{3} \alpha^k R^k(V(w_L, w_R)), \tag{6.3a}$$

where $V(w_L, w_R)$ (4.6) is some average state; denote its velocity and sound speed by $\hat{u}$ and $\hat{c}$, respectively. To calculate α^k in (6.3a) we first evaluate

$$C_1 = (\gamma - 1)\{[E] + \tfrac{1}{2}\hat{u}^2[\rho] - \hat{u}[m]\}/\hat{c}^2, \tag{6.3b}$$

$$C_2 = \{[m] - \hat{u}[\rho]\}/\hat{c}, \tag{6.3c}$$

where $[b] = b_R - b_L$; then the α^k in (6.3a) are obtained by

$$\alpha^1 = \tfrac{1}{2}(C_1 - C_2), \quad \alpha^2 = [\rho] - C_1, \quad \alpha^3 = \tfrac{1}{2}(C_1 + C_2). \tag{6.3d}$$

The second characteristic field corresponding to the eigenvalue u is linearly degenerate, i.e., $a_w^2 R^2 \equiv 0$; the other characteristic fields corresponding to the eigenvalues $u \pm c$ are genuinely nonlinear.

We note that in computing $\bar{f}_{j+1/2}$ in (4.8c), one could take advantage of the simple form of the R^k in (6.2b).

Next we show how to implement Roe's linearization technique (4.14)–(4.15) in the above algorithm. Roe presents a particular form of averaging $V(w_L, w_R)$ such that for the Euler equations of gasdynamics, the mean value Jacobian $A(w_L, w_R)$ in (4.14a) can be expressed by (4.16). This averaging takes the form

$$\begin{aligned} &\hat{u}_{j+1/2} = \langle \rho^{1/2} u \rangle / \langle \rho^{1/2} \rangle, \quad \hat{H}_{j+1/2} = \langle \rho^{1/2} H \rangle / \langle \rho^{1/2} \rangle, \\ &\hat{c}_{j+1/2} = \{(\gamma - 1)(\hat{H}_{j+1/2} - \tfrac{1}{2}\hat{u}^2_{j+1/2})\}^{1/2}, \end{aligned} \tag{6.4a}$$

where $\langle b \rangle$ denotes the arithmetic mean

$$\langle b \rangle = \tfrac{1}{2}(b_j + b_{j+1}). \tag{6.4b}$$

Therefore to use Roe's linearization in our scheme all one has to do is to compute $\hat{u}_{j+1/2}$ and $\hat{c}_{j+1/2}$ in (6.2)–(6.3) by (6.4).

We remark that the averaging in (6.4) is rather expensive. Our experiments indicate that in many applications the simple arithmetic average (6.4b) will do just as well.

7. NUMERICAL EXPERIMENTS

In this section we present some numerical experiments that demonstrate the performance of the proposed second order accurate scheme. We consider here three versions of it;

$$v_j^{n+1} = v_j^n - \lambda(\bar{f}_{j+1/2} - \bar{f}_{j-1/2}), \tag{7.1a}$$

$$\bar{f}_{j+1/2} = \frac{1}{2}\left[f(v_j) + f(v_{j+1}) - \frac{1}{\lambda}\sum_{k=1}^{m} \beta^k_{j+1/2} R^k_{j+1/2} \right]. \tag{7.1b}$$

By ULT 1 we denote scheme (7.1) with

$$\beta^k_{j+1/2} = Q^k(\nu^k_{j+1/2} + \gamma^k_{j+1/2})\alpha^k_{j+1/2} - (g^k_j + g^k_{j+1}), \tag{7.2}$$

where g_j and $\gamma_{j+1/2}$ are defined by (4.8). By ULT 1C we denote scheme (7.1)–(7.2), with modification (5.7)–(5.8) in the linearly degenerate field.

We also experiment with the simplified version

$$\beta_{j+1/2} = Q(\nu_{j+1/2})\alpha_{j+1/2} - 2G_{j+1/2}, \tag{7.3}$$

where $G_{j+1/2}$ is defined by (4.17); we denote this version by ULT 2.

For comparison sake we also present calculations with the following two schemes:

(i) The second order accurate Lax–Wendroff-type scheme (7.1) with

$$\beta^k_{j+1/2} = (\nu^k_{j+1/2})^2 \alpha^k_{j+1/2} \tag{7.4}$$

which is referred to as the LW scheme;

(ii) the first order accurate Godunov-type scheme of Roe (see [10, 15]), which is defined by (7.1) with

$$\beta^k_{j+1/2} = |\nu^k_{j+1/2}|\alpha^k_{j+1/2} \tag{7.5}$$

and is referred to as the ROE scheme.

In all the schemes and experiments reported herein we use the Roe linearization (6.3)–(6.4).

I. *The Shock Tube Problem*

We consider now a Riemann problem

$$\begin{aligned} w(x, 0) &= W_L, \quad x < 0, \\ &= W_R, \quad x > 0, \end{aligned} \tag{7.6}$$

for the Euler equations of a polytropic gas (6.1). Our first set of data is

$$W_L = \begin{pmatrix} 0.445 \\ 0.311 \\ 8.928 \end{pmatrix}, \quad W_R = \begin{pmatrix} 0.5 \\ 0 \\ 1.4275 \end{pmatrix}. \tag{7.7}$$

Other numerical experiments with this problem are reported in [6] and the references cited there.

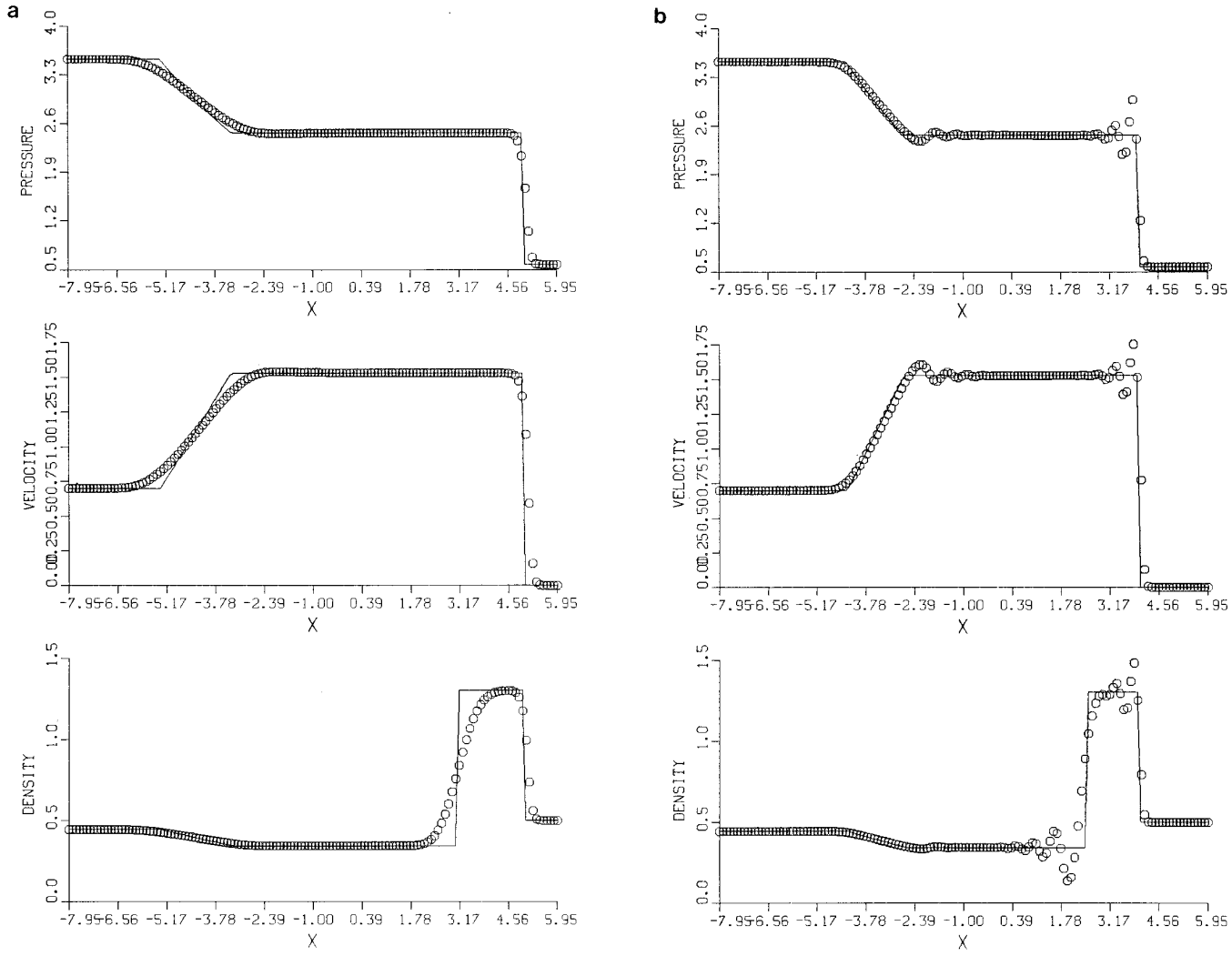

FIG. 2. (a) ROE scheme for (7.7); (b) LW scheme for (7.7); (c) ULT 1 scheme for (7.7); (d) ULT 1C scheme for (7.7).

In Figs. 2a–d we show the results obtained by the ROE, LW, ULT 1, and ULT 1C schemes. The numerical values are shown by circles; the exact solution is shown by the solid line. The calculations in Fig. 2 were performed with 100 time steps under the CFL restriction $\mu = 0.95$ in (4.13), and 140 cells.

In Fig. 3 we apply the ULT 1C scheme to a different set of data for the Riemann problem (7.6)

$$W_L = \begin{pmatrix} 1 \\ 0 \\ 2.5 \end{pmatrix}, \quad W_R = \begin{pmatrix} 0.125 \\ 0 \\ 0.25 \end{pmatrix}. \tag{7.8a}$$

Other numerical experiments with this problem are presented in [16]. The calculations in Fig. 3 were performed with 50 time steps under the CFL restriction $\mu = 0.95$ in (4.13), with 100 cells.

We would like to make the following remarks regarding the calculations in Figs. 2 and 3:

(1) Modification (5.7)–(5.8) in the linearly degenerate field improves the resolution of the contact discontinuity, but does not produce any noticeable change in the other waves.

(2) The calculations in these figures were performed with $Q^k(x) = |x|$ for all k; nevertheless, the solution has properly developed. This fact supports our assumption that entropy violation in this case is possible only if there is a sonic point in an expanding region. To test the performance of the scheme in the latter case, we have applied ULT 1 with $Q(x)$ defined by (5.4) with $\varepsilon = 0.1$ to the data.

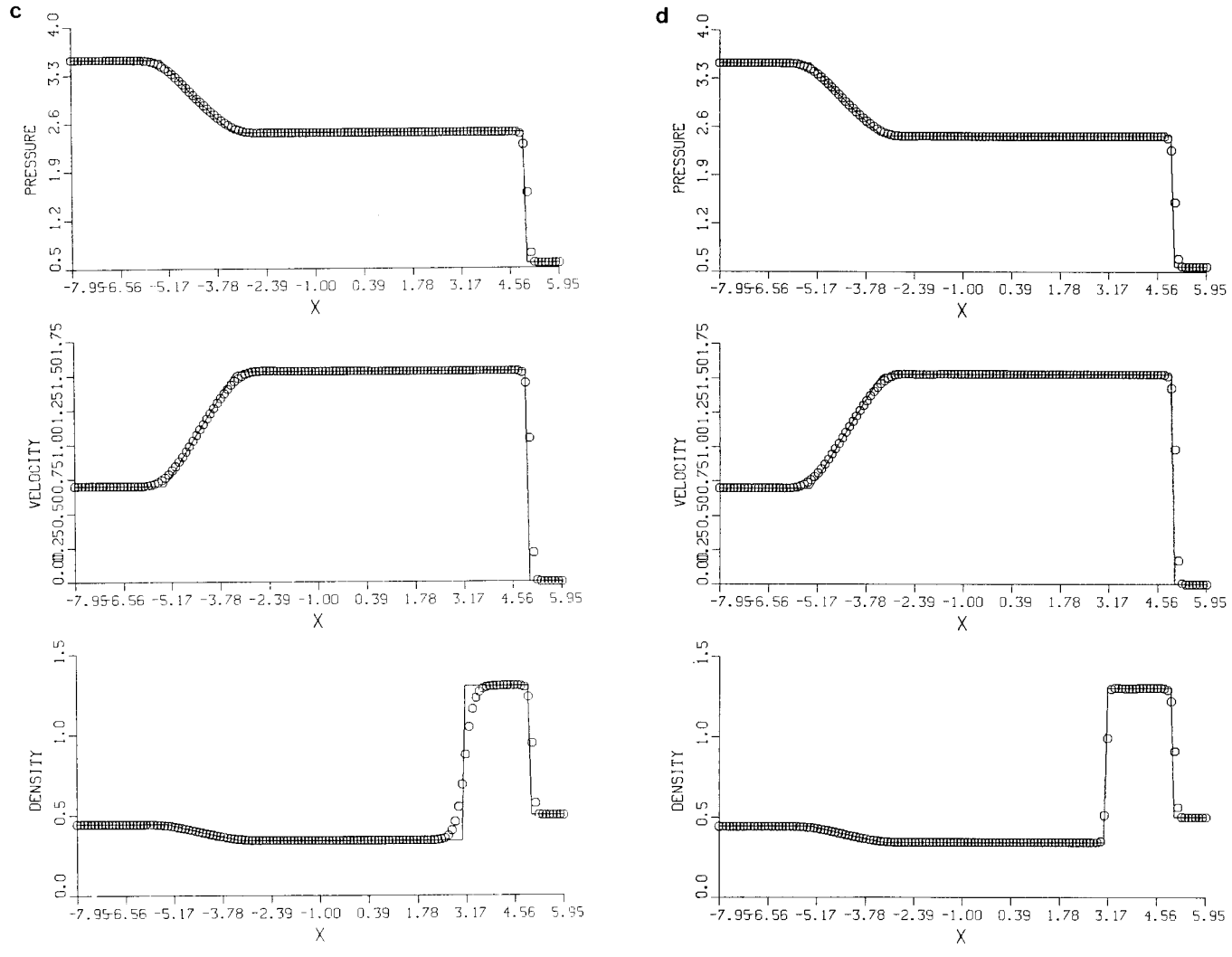

FIG. 2—*Continued*

$$W_L = \begin{pmatrix} 1 \\ 0.5 \\ 2.625 \end{pmatrix}, \quad W_R = \begin{pmatrix} 0.125 \\ 0.0625 \\ 0.265625 \end{pmatrix}; \qquad (7.8b)$$

these data are obtained by superposing a uniform translation with the speed of 0.5 on the data in (7.8a), thus resulting in a sonic point in the rarefaction wave. The numerical results (up to the constant shift in velocity) look almost identical to those of Fig. 3.

To further test the entropy enforcement and the resolution of the scheme, we consider now a Riemann problem where the two states, say v_1 and v_2, satisfy the Rankine–Hugoniot relations to the direction of the field $u - c$, with a zero speed of propagation. Fixing the state v_1 by assigning values to ρ_1, P_1, and the Mach number $M_1 = u_1/c$, we determine the state v_2 by

$$\begin{aligned} P_2/P_1 &= (2\gamma M_1^2 - \gamma + 1)/(\gamma + 1), \\ u_2/u_1 &= (2/M_1^2 + \gamma - 1)/(\gamma + 1), \\ \rho_2/\rho_1 &= u_1/u_2. \end{aligned} \qquad (7.9)$$

In Fig. 4 we test the scheme ULT 2 on stationary shock (7.6) with $W_L = v_1$ and $W_R = v_2$. First we set $\rho_1 = P_1 = 1$, $M_1 = 4$ (which gives $P_2/P_1 = 18.5$) and test the scheme for resolution. Figures 4a and b show the pressure profiles of ULT 2 with $Q(x)$ defined by (5.4) with $\varepsilon = 0.1$ and $\varepsilon = 0.25$, respectively. In Fig. 4c we test ULT 2 with $\varepsilon = 0.1$ for robustness by applying it to the very strong shock $\rho_1 = P_1 = 1$, $M_1 = 10$, i.e., a pressure ratio of $P_2/P_1 = 116.5$.

Next, in Fig. 5 we test the scheme for entropy enforcement by applying it to the inadmissible "stationary" discontinuity $W_L = v_2$, $W_R = v_1$ with $\rho_1 = P_1 = 1$, $M_1 = 4$. In Fig. 5 we show the solution of ULT 2 with $\varepsilon = 0.1$ in

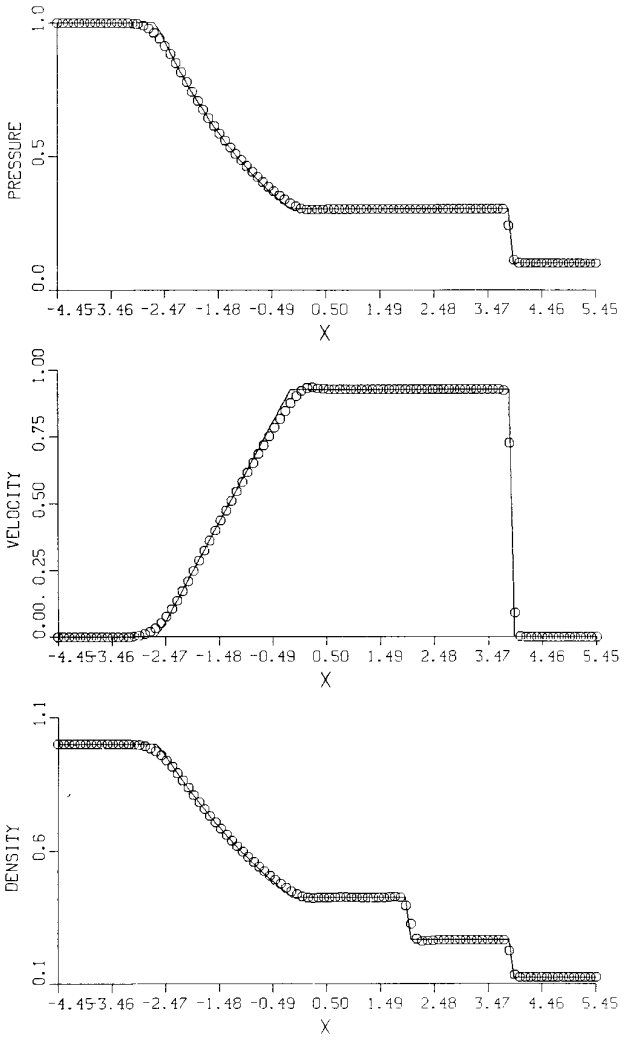

FIG. 3. ULT 1C scheme for (7.8).

(5.4). We recall that the discontinuity in Fig. 5 is an entropy-violating stationary solution for ULT 2 with $\varepsilon = 0$.

The calculations in Figs. 4 and 5 were performed with 50 time steps under the CFL restriction $M = 0.95$ in (4.13) with 100 cells.

We remark that for all the Riemann problems considered here, we found the solutions of ULT 1, ULT 1C and ULT 2 to be very much alike (except for the obvious improved resolution of contact discontinuities in ULT 1C).

II. *The Quasi* 1-*D Nozzle Problem*

We consider an axisymmetric nozzle with a cross-sectional area $A(x)$. The cross-sectional average of the flow satisfies the following one-dimensional system of equations:

$$w_t + f(w)_x = -s(w, x), \quad s(w, x) = \begin{pmatrix} 0 \\ p(dA/dx) \\ 0 \end{pmatrix}, \quad (7.10a)$$

where w, $f(w)$, and p are given in (6.1).

In Figs. 6, 7, and 8[1] we present numerical approximations to steady state solutions of (7.10a). In Figs. 6 and 7 we show solutions for a divergent nozzle with cross-sectional area

$$A(x) = 1.398 + 0.347 \tanh(0.8x - 4); \quad (7.10b)$$

the flow condition is supersonic at the entrance and subsonic at the exit. Figs. 6a and b show steady state solutions on a crude mesh of the ROE and ULT 1 schemes, respec-

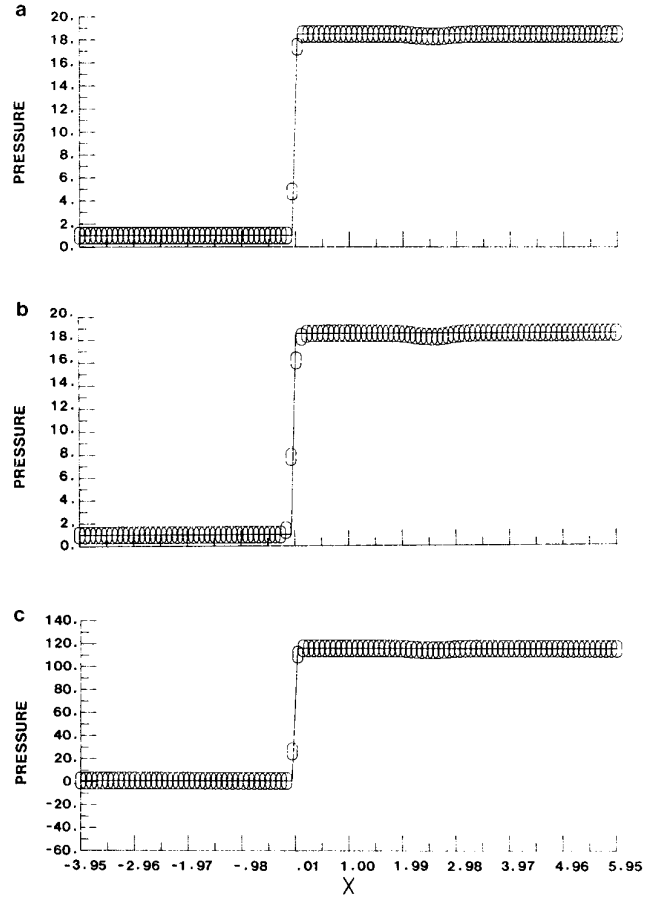

FIG. 4. ULT 2 scheme with (a) $\varepsilon = 0.1$ ($P_2/P_1 = 18.5$), (b) $\varepsilon = 0.25$ ($P_2/P_1 = 18.5$), (c) $\varepsilon = 0.1$ ($P_2/P_1 = 1.165 \times 10^2$).

[1] These figures are courtesy of Helen C. Yee of the NASA–Ames Research Center. See [19] for more details.

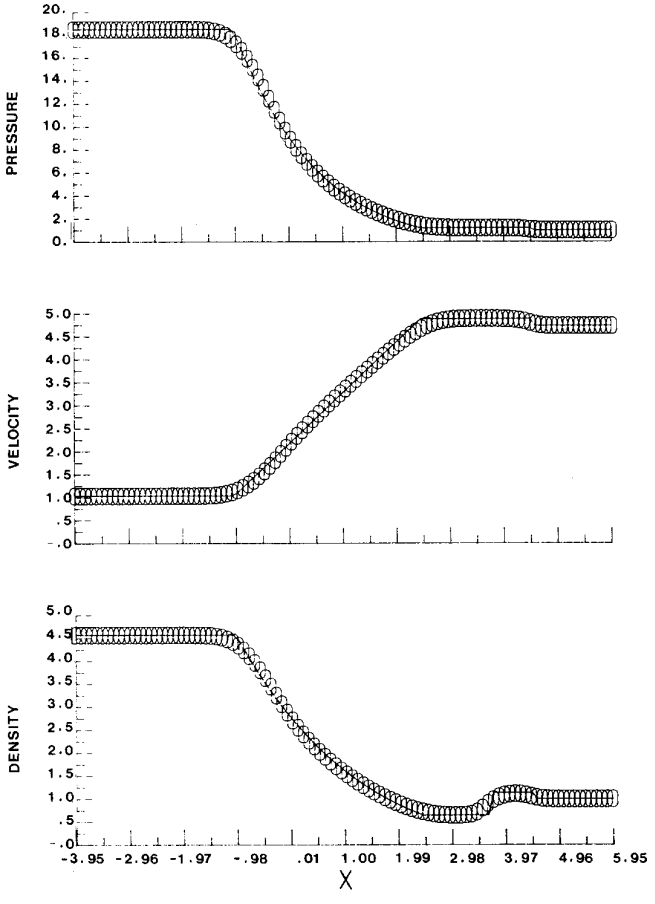

FIG. 5. ULT 2 with $\varepsilon = 0.1$ for a stationary expansion shock.

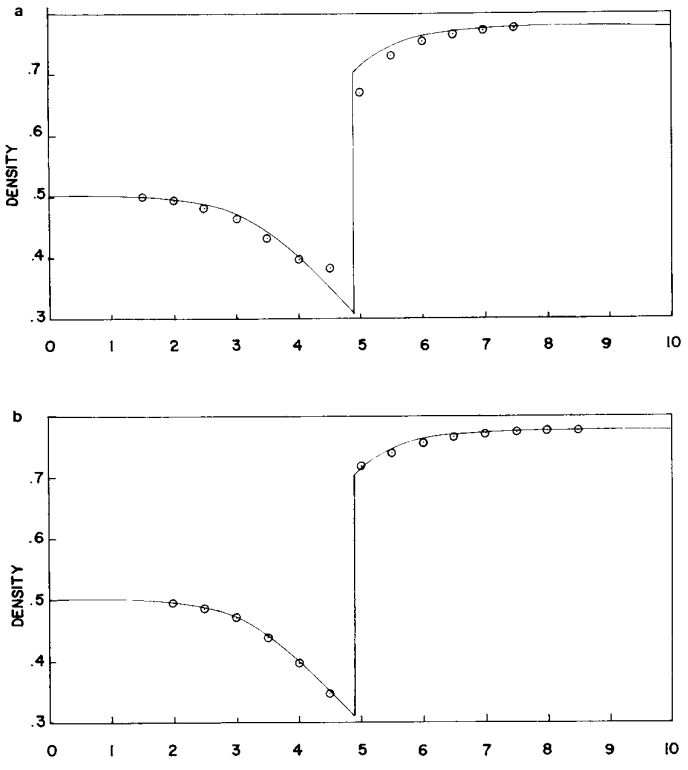

FIG. 6. (a) ROE scheme, (b) ULT 1 scheme for the divergent nozzle (7.10).

tively. Figure 7 shows the ULT 1 results for the same problem on a finer mesh.

In Fig. 8 we show a steady state solution of the ULT 1 scheme for a convergent–divergent nozzle with cross-sectional area

$$\begin{aligned} A(x) &= 1 + (A_0 - 1)(1 - x/5)^2, && x \leqslant 5, \\ &= 1 + (A_E - 1)[(x - 5)/(x_E - 5)]^2, && x > 5, \end{aligned} \tag{7.11}$$

where A_0 = entrance exit area, A_E = exit area. Here the flow is subsonic at the entrance as well as at the exit.

The exact solutions in Figs. 6–8 are shown by the solid curves; the values of the numerical solutions are indicated by a circle.

III. *2-D Flow through a Duct*

In Figs. 9a and b we show solutions to the problem of the flow of air through a duct containing a step. Initially the flow is everywhere to the right at Mach 3, with $\rho = 1.4$, $p = 1$, and $c = 1$. The duct width is 1, its length is 3, and the step of height 0.2 is located a distance 0.6 from the entrance. This problem was used by Woodward and Colella [18] to test the performance of various numerical schemes. In Fig. 9 we show the results at $t = 4$ with a crude uniform Cartesian grid with $\Delta x = \Delta y = 0.1$ (i.e., a 10×30 grid).

The solutions in Fig. 9 were obtained by a Strang-type dimensional splitting of the form

$$v^{n+2} = Lv^n, \tag{7.12a}$$

$$L = L_x L_y L_y L_x, \tag{7.12b}$$

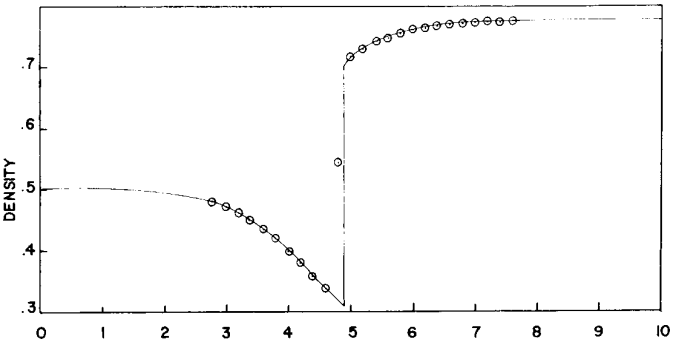

FIG. 7. ULT 1 scheme for (7.10) on a finer grid.

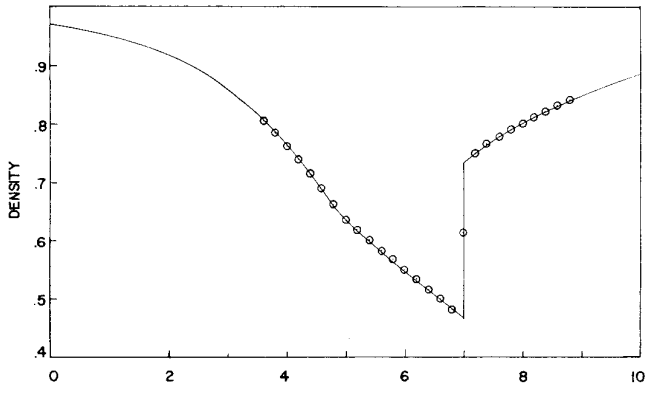

FIG. 8. ULT 1 scheme for the convergent–divergent nozzle (7.11).

where L_x and L_y are one-dimensional finite difference operators approximating

$$L_x : w_t + f(w)_x = 0, \quad L_y : w_t + g(w)_y = 0. \tag{7.12c}$$

If L_x and L_y are second order accurate approximations to the one-dimensional equations in (7.12c), then scheme (7.12a) and (7.12b) is a second order accurate approximation to the two-dimensional problem

$$w_t + f(w)_x + g(w)_y = 0. \tag{7.12d}$$

In Fig. 9a we show for comparison sake the results of second order accurate scheme (7.1) with

$$\beta^k_{j+1/2} = [(\nu^k_{j+1/2})^2 + \tfrac{1}{4}\hat{\theta}^k_{j+1/2}]\alpha^k_{j+1/2}; \tag{7.13}$$

here $\hat{\theta}_{j+1/2}$ is a switch defined by

$$\hat{\theta}_{j+1/2} = \max(\theta_j, \theta_{j+1}), \tag{7.14}$$

where θ_j is (5.8a) for all k. This scheme can be regarded as a hybrid of LW scheme (7.4) with a first order accurate scheme, which is (3.1) with $Q(x) = x^2 + \frac{1}{4}$ (see Fig. 1 and [6]).

In Fig. 9b we show the results of a version of ULT 1C which is defined by (7.1) with

$$\beta_{j+1/2} = Q(\nu_{j+1/2} + (1 + 2\hat{\theta}_{j+1/2})\gamma_{j+1/2})\alpha_{j+1/2} - (1 + 2\hat{\theta}_{j+1/2})(g_j + g_{j+1}), \tag{7.15}$$

where for all k, $Q(x) = x^2 + \frac{1}{4}$, g_j and $\gamma_{j+1/2}$ are (4.8c) and (4.8e), respectively, and $\hat{\theta}_{j+1/2}$ is (7.14).

Both calculations were performed with a CFL restriction of 0.75. The corner of the step was treated as a sharp corner without any rounding (or equivalent addition of artificial viscosity). Figures 9a and b show 30 equally spaced density contours.

Figure 9b clearly demonstrates that the high resolution of the proposed scheme in one-dimensional problems is also obtainable in two-dimensional calculations.

Altogether we find the performance of the new second order accurate scheme to be quite pleasing. We note that the scheme is simple to program and requires not much more CPU time than a Lax–Wendroff scheme with some artificial viscosity.

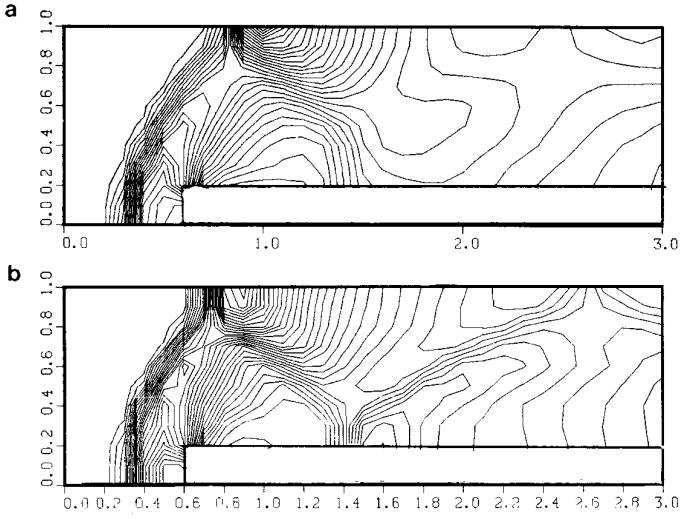

FIG. 9. Density contours with a 10×30 grid for (a) the hybrid scheme (7.13) and (b) the ULT 1C scheme.

ACKNOWLEDGMENTS

The author would like to thank the members of the CFD branch at NASA Ames Research Center, especially Helen Yee and Bob Warming, for many helpful discussions as well as for their hospitality.

REFERENCES

1. J. P. Boris and D. L. Book, *J. Comput. Phys.* **11,** 38 (1973).
2. M. G. Crandall and A. Majda, *Math. Comput.* **34,** 1 (1980).
3. J. Glimm, *Commun. Pure Appl. Math.* **18,** 697 (1965).
4. S. K. Godunov, *Math. Sb.* **47,** 271 (1959); also Cornell Aero. Lab. Transl.
5. A. Harten, *Commun. Pure Appl. Math.* **30,** 611 (1977).
6. A. Harten, *Math. Comput.* **32,** 363 (1978).
7. A. Harten, *J. Comput. Phys.* **49,** 151 (1983).
8. A. Harten, J. M. Hyman, and P. D. Lax, *Commun. Pure Appl. Math.* **29,** 297 (1976).
9. A. Harten and P. D. Lax, *SIAM J. Numer. Anal.* **18,** 289 (1981).
10. A. Harten, P. D. Lax, and B. van Leer, *SIAM Rev.* **25,** 35 (1983).
11. A. Harten and J. M. Hyman, *J. Comput. Phys.* **50,** 235 (1983).
12. P. D. Lax, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves* (SIAM, Philadelphia, 1972).

13. T. P. Liu, *J. Math. Anal. Appl.* **53,** 78 (1976).
14. P. L. Roe, in *Proceedings, 7th Int. Conf. Numer. Methods Fluid Dyn., Stanford/NASA Ames, June 1980* (Springer-Verlag, New York/Berlin 1981), p. 354.
15. P. L. Roe, *J. Comput. Phys.* **43,** 357 (1981).
16. G. A. Sod, *J. Comput. Phys.* **27,** 1 (1978).
17. B. van Leer, *J. Comput. Phys.* **14,** 361 (1974).
18. P. Woodward and P. Colella, in *Proceedings, 7th Int. Conf. Numer. Methods Fluid Dyn., Stanford/NASA Ames, June 1980* (Springer-Verlag, New York/Berlin, 1981), p. 434.
19. H. C. Yee, R. F. Warming, and A. Harten, in *Proceedings, 8th Int. Conf. Numer. Methods Fluid Dyn., Aachen, June 1982* (Springer-Verlag, New York/Berlin, 1982), p. 546.