



---

High Resolution Schemes and the Entropy Condition

Author(s): Stanley Osher and Sukumar Chakravarthy

Source: *SIAM Journal on Numerical Analysis*, Oct., 1984, Vol. 21, No. 5 (Oct., 1984), pp. 955-984

Published by: Society for Industrial and Applied Mathematics

Stable URL: <https://www.jstor.org/stable/2156937>

#### REFERENCES

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/2156937?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/2156937?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*Society for Industrial and Applied Mathematics* is collaborating with JSTOR to digitize, preserve and extend access to *SIAM Journal on Numerical Analysis*

## HIGH RESOLUTION SCHEMES AND THE ENTROPY CONDITION\*

STANLEY OSHER† AND SUKUMAR CHAKRAVARTHY‡

**Abstract.** A systematic procedure for constructing semidiscrete, second order accurate, variation diminishing, five-point band width, approximations to scalar conservation laws, is presented. These schemes are constructed to also satisfy a single discrete entropy inequality. Thus, in the convex flux case, we prove convergence to the unique physically correct solution. For hyperbolic systems of conservation laws, we formally use this construction to extend the first author's first order accurate scheme, and show (under some minor technical hypotheses) that limit solutions satisfy an entropy inequality. Results concerning discrete shocks, a maximum principle, and maximal order of accuracy are obtained. Numerical applications are also presented.

**Introduction.** Recently a number of new shock capturing finite difference approximations have been constructed and found to be very useful in shock calculations, e.g. [3], [4], [16]. In addition to conservation form, these schemes are usually constructed to have as many as possible of the following properties:

- (1) Stable and sharp discrete shock solutions.
- (2) Limit solutions which satisfy a geometric and/or analytic entropy condition.
- (3) A bound on the variation of the approximate solutions, at least in both the scalar, and linear systems, case.
- (4) Second order accuracy in regions of smoothness, (with certain isolated exceptional points, as described in § 2 below).

In this paper we shall present a general procedure for constructing schemes with a five-point band width satisfying all of the above. We shall then prove a convergence theorem for a wide class of approximations to any scalar convex conservation laws in one space dimension, using recent uniqueness results of DiPerna [5].

Our convergence proof involves the following simple steps:

- (a) We first obtain a variation bound for a wide class of second order accurate approximations. This implies that, for any sequence of mesh widths approaching zero, a subsequence of approximate solutions converges to a weak solution of the Cauchy problem.
- (b) Next, we obtain a single discrete entropy inequality for a large subclass of the approximations mentioned above. Then all limit solutions in this subclass satisfy the inequality.
- (c) We finally invoke the uniqueness results of DiPerna to prove convergence to the solution to the Cauchy problem, as the mesh width goes to zero, for the subclass mentioned in (b).

Earlier, Majda and Osher [20], modified the Lax–Wendroff scheme, keeping its essential properties, so that an entropy inequality was obtained. DiPerna (private communication) has proven  $L^2$  convergence of this modification (again approximating convex conservation laws) using the theory of compensated compactness. No variation bound has been obtained for this second order accurate approximation.

From a practical point of view, this apparent lack of variation bound means that conventional schemes such as Lax–Wendroff, when approximating hyperbolic systems

---

\* The research of this author was supported by the National Science Foundation under grant MCS 82-00788, the Army Research Office under grant DAAG 29-82-K-0090, and the National Aeronautics and Space Administration under grant NAG 1-270.

† Department of Mathematics, University of California, Los Angeles, California 90024.

‡ Rockwell Science Center, Thousand Oaks, California 91360.

of conservation laws, even with an entropy modification, suffer from a lack of robustness in computing complex flows with shock waves and steep gradients. While such schemes have been widely used in a variety of problems, (see [2] for references), that list of solved problems does not include flows with strong shocks (say Mach 5 upstream, normal to the shock), when the shocks are captured.

A main drawback of most finite-difference schemes is that discontinuities are approximated by discrete transitions, that when narrow, usually overshoot or undershoot, or when monotone, usually spread the discontinuity over many grid points.

Upwind schemes have been designed and used over the years, largely because of their success in treating this difficulty. Those based on solving the Riemann problem either exactly (Godunov's method [10]) or approximately e.g. (Osher's [24], or Roe's [28] with an entropy fix [34], [22]), have been extremely successful, especially when made second order accurate, [3], [4], or [16].

We should particularly mention the early investigations of van Leer [32], [33]. There he introduced the concepts of flux limiters and higher order Riemann solvers. Recently, Harten [16], using an argument also used in [1], [31] and [29], obtained simple sufficient conditions compatible with second order accuracy, which guarantee that a scalar one-dimensional approximation is TVD—total variation diminishing. He constructed a scheme having that property, and formally extended it to systems, using a field-by-field limiter and Roe's decomposition.

Using any of the three nonlinear decompositions (Godunov's Osher's or Roe's) one can obtain field-by-field limiters for systems, as described in § 6 below. Peter Sweby, [30], has investigated the properties of various limiters, and clarified their application considerably. The notation used in §§ 3–6 below is due to him.

We shall use the now introduced term “high resolution scheme” to mean a formal extension to systems via a field-by-field decomposition, of a scalar, second order accurate, variation diminishing scheme as in [16]. These schemes do not, in general, satisfy the entropy condition—e.g. expansion shocks exist as stable solutions of high resolution schemes based on Roe's (unmodified) scheme. In § 6, we use Osher's decomposition and certain limiters, to prove that under technically reasonable hypotheses, limit solutions of certain high resolution schemes do satisfy the entropy condition for hyperbolic systems of conservation laws.

Fundamental to our work is a formula measuring the discrete entropy in a cell for any scheme, obtained in [22]. See equations (4.2), (4.3), and (4.4) below. In [22], a class of approximations called *E* schemes was shown to be convergent, even for nonconvex, but scalar flux functions. This seems to be the widest known class of convergent schemes in this general case, but unfortunately the approximations are, at most, first order accurate.

All the results in this paper are for semi-discrete (continuous in time) approximations, and thus can serve as guidelines for a wide variety of time discretizations, both implicit and explicit. One of the principal applications of this theory has been in supersonic and transonic aerodynamics, where phenomena close to steady state are often studied. As an important example, one might mention the enormous advantage found when algorithms used to solve the transonic small disturbance equation, based on traditional Murman [21] (Roe's unfixed for scalar equations) differencing were replaced by using the *E* schemes of Engquist and Osher [12], [7], and later Godunov [13]. This minimal coding change, based on an entropy fix, increased the robustness of production codes by an order of magnitude, or more.

Goodman and LeVeque [11] have recently obtained a rather depressing result—two space dimensional, scalar approximations, cannot be TVD, and still be more than

first order accurate, if the associated flux functions are reasonably smooth. Nevertheless, two-dimensional schemes based on dimension by dimension high resolution differencing, do perform extremely well, even for complex configurations, with very strong shocks. See [4], [16], [3], and § 7 below. This phenomenon remains to be justified analytically.

The format of this paper is as follows. In § 1, we review the relevant theory of weak solutions for hyperbolic systems of conservation laws. In § 2, we do the same for the theory of approximate solutions, and rederive results about bounding the variation in Lemma 2.1. We then obtain a maximum (minimum) principle and a nonlinear saturation result—Lemmas 2.2 and 2.3. In § 3, we develop a systematic procedure used to obtain high resolution schemes, based on any three-point first order accurate TVD scheme. We use second order “upwinding,” together with flux limiters to do this. In § 4, we show how to obtain an entropy inequality, and hence convergence, which is our main result, Theorem 4.1. We use artificial compression or rarefaction operators (ACR). Harten introduced the notion of artificial compression in [15]. Here we use the compression part to firm up shocks, and contacts, and the rarefaction part to enforce an entropy condition. Rather precise admissible bounds on the ACR term are given. Section 5 is concerned with existence of steady discrete shocks for certain high resolution schemes. In § 6, we construct our high resolution scheme for systems, and then prove an entropy inequality: Theorem 6.2.

Finally, § 7 gives numerical evidence demonstrating the utility of these schemes, as well as some results illustrating a few theoretical points made throughout this paper. See also [3].

**1. Review of theory of weak solutions.** We shall consider numerical approximations to the initial value problem for nonlinear hyperbolic systems of conservation laws

$$(1.1) \quad \frac{\partial w}{\partial t} + \frac{\partial}{\partial x} f(w) = 0, \quad t > 0, \quad -1 \leq x < 1,$$

with periodic boundary conditions:

$$w(x+1, t) \equiv w(x, t),$$

and initial conditions  $w(x, 0) = w_0(x)$ .

Here  $w(x, t)$  is an  $m$ -vector of unknowns, and the flux function  $f(w)$  is vector valued, having  $m$  components. The system is hyperbolic when the Jacobian matrix has real eigenvalues.

It is well known that solutions of (1.1) may develop discontinuities in finite time, even when the initial data are smooth. Because of this, we seek a weak solution of (1.1), i.e., a bounded measurable function  $w$ , such that for all  $\varphi \in C_0^\infty(R \times R^+)$ ,

$$(1.2) \quad \begin{aligned} (a) \quad & \iint_{R \times R^+} (w \varphi_t + f(x) \varphi_x) \, dx \, dt = 0, \\ (b) \quad & \lim_{t \rightarrow 0} \|w(x, t) - w_0(x)\|_{L^1} = 0. \end{aligned}$$

Solutions of (1.2) are not necessarily unique. For physical reasons, the limit solution of the viscous equation, as viscosity tends to zero, is sought. In the scalar case, this solution must satisfy, for all  $\varphi \in C_0^\infty(R \times R^+)$ ,  $\varphi \geq 0$ , and all real constants  $c$ .

$$(1.3) \quad (a) \quad - \iint (|w - c| \varphi_t + \operatorname{sgn}(w - c)(f(w) - f(c)) \varphi_x) \, dx \, dt \leq 0.$$

This is equivalent to the statement:

$$(1.3) \quad (b) \quad \frac{\partial}{\partial t}|w - c| + \frac{\partial}{\partial x}((f(w) - f(c)) \operatorname{sgn}(w - c)) \leq 0,$$

in the sense of distributions.

Such solutions are called entropy solutions. Kruzkov has shown in [18] that two entropy solutions satisfy

$$(1.4) \quad \|w(x, t_1) - v(x, t_1)\|_{L^1} \leq \|w(x, t_0) - v(x, t_0)\|_{L^1},$$

for all  $t_1 \geq t_0$ . Hence, condition (1.3) guarantees the uniqueness of solutions to the scalar version of (1.2). Existence was also obtained in [18].

For systems of equations, Lax has defined an entropy inequality [19], with the help of an entropy function  $V(w)$  for (1.1), defined to have the following properties:

(i)  $V$  satisfies

$$(1.5) \quad V_w f_w = F_w$$

where  $F$  is some other function, called the entropy flux.

(ii)  $V$  is a convex function of  $w$ .

It follows from (1.1), upon multiplication by  $V_w$ , that every smooth solution of (1.1) also satisfies

$$(1.6) \quad V_t + F_x = 0.$$

It was also shown in [19], that if  $w$  is the bounded a.e. limit of solutions to the regularized equation, then the limit satisfies, in the weak sense, the following inequality:

$$(1.7) \quad V_t + F_x \leq 0.$$

Inequality (1.3b), for scalar equations, is just (1.7), with  $V(w) = |w - c|$ .

Inequalities (1.3) and (1.7) have important geometric consequences for piecewise continuous solutions. Suppose  $w(x, t)$  is such a solution having a jump discontinuity,  $w_L(t)$ ,  $w_R(t)$  moving with speed  $s(t)$ . Then (1.2) implies the well known jump conditions

$$(1.8) \quad f(w_L) - f(w_R) = s(w_L - w_R).$$

In the scalar case, (1.3) is equivalent to Oleinik's condition  $E$  across the shock

$$(1.9) \quad \frac{f(w) - f(w_R)}{w - w_R} \leq \frac{f(w_L) - f(w_R)}{w_L - w_R},$$

for all  $w$  between  $w_L$  and  $w_R$ .

If  $f$  is convex, then (1.9) is equivalent to the statement that characteristics flow *into* the discontinuity, as  $t$  increases. Also, for scalar convex  $f$ , inequality (1.7) for a *single fixed* convex  $V$  is equivalent to (1.3) for *all constants*  $c$ , if the solution is of bounded variation. This is a consequence of the recent results of DiPerna [5]. Thus uniqueness in this case is a consequence of Kruzkov's results [18]. This fact is crucial to our convergence proof in § 4.

For hyperbolic systems of equations, the Jacobian of  $f$ , denoted by  $\partial f$ , has real eigenvalues, which are usually assumed to be distinct:

$$\lambda_1 < \lambda_2 < \cdots < \lambda_m,$$

corresponding to right eigenvectors:  $r_1, r_2, \dots, r_m$ . Lax [19] then defines the  $k$ th field to be linearly degenerate if

$$\nabla_w \lambda_k \cdot r_k \equiv 0.$$

He also defines the field to be genuinely nonlinear if

$$\nabla_w \lambda_k \cdot r_k \neq 0.$$

For genuinely nonlinear fields, a  $k$  shock moving with speed  $s$  is defined to be a discontinuity of  $w$ , such that  $m+1$  characteristics flow into the shocks, and

$$\lambda_k(w_L) - s > 0 > \lambda_k(w_R) - s.$$

This geometric condition is equivalent to (1.7) for weak shocks [19].

**2. Preliminary theory of approximate solutions.** We consider a semi-discrete, method of lines approximation to (1.1). We break the interval  $[-1, 1]$  into subintervals

$$\mathcal{J}_j = \{x | (j - \frac{1}{2})\Delta \leq x < (j + \frac{1}{2})\Delta\},$$

$j = 0, \pm 1, \dots, \pm N$ , with  $(2N+1)\Delta = 2$ .

Let  $x_j = j\Delta$ , be the center of each interval  $\mathcal{J}_j$  with end points  $x_{j-1/2}, x_{j+1/2}$ . Define the step function for each  $t > 0$ , as

$$\mathcal{V}_\Delta(x, t) \equiv u_j(t)$$

for  $x \in \mathcal{J}_j$ .

The initial data is discretized via the averaging operator  $T_\Delta$ ,

$$(2.1) \quad T_\Delta w_0(x) = \frac{1}{\Delta} \int_{\mathcal{J}_j} w(s) ds = u_j(0) \quad \text{for } x \in \mathcal{J}_j.$$

For any step function, we define the difference operators:

$$\Delta_\pm u_j = \pm(u_{j\pm 1}, -u_j), \quad D_\pm u_j = \frac{1}{\Delta} \Delta_\pm u_j.$$

A method of lines, conservation form discretization of (1.1), is a system of differential equations

$$(2.2) \quad \begin{aligned} \frac{\partial}{\partial t} u_j + D_+ h_{j-1/2} &= 0, \quad j = 0, \pm 1, \dots, \pm N, \\ \mathcal{V}_\Delta(x, 0) &= T_\Delta w_0(x) \quad \text{for } x \in \mathcal{J}_j. \end{aligned}$$

Here the numerical flux defined by

$$(2.3) \quad h_{j-1/2} = h(u_{j+k-1}, \dots, u_{j-k}),$$

for  $k \geq 1$ , is a Lipschitz continuous function of its arguments, satisfying the consistency condition:

$$h(w, w, \dots, w) = f(w).$$

It is well known that bounded a.e. limits, as  $\Delta \rightarrow 0$ , of approximate solutions, converge to weak solutions of (1.1), i.e. (1.2a) is satisfied. However, this does not also imply that the limit solutions will satisfy any of the entropy conditions mentioned above, e.g. [8], [9]. Some restrictions on  $h$  are required.

A simple class of flux functions  $h$ , for which (2.2) converges to the unique entropy solution in  $L^\infty(L^1(R); [0, T])$ , as  $\Delta \rightarrow 0$ , for any  $T > 0$ , are scalar “ $E$ ” schemes, [22]. Such schemes satisfy the following:

DEFINITION 2.1. A consistent scheme whose numerical flux satisfies

(2.4) 
$$\operatorname{sgn}(u_j - u_{j-1})[h_{j-1/2} - f(u)] \leq 0,$$

for all  $u$  between  $u_{j-1}$  and  $u_j$ , is said to be an  $E$  scheme.

It is easy to see that this class includes the widely known class of three-point monotone schemes, i.e. those for which

$$h_{j-1/2} = h(u_j, u_{j-1})$$

with  $h_{j-1/2}$  nonincreasing in its first argument, nondecreasing in its second.

Partial derivatives of a numerical flux, when needed, will be denoted via

$$\frac{\partial}{\partial u_{j+\gamma}} h(u_{j+k}, \dots, u_j, \dots, u_{j-k+1}) = h_\gamma.$$

Thus a  $C^1$  three-point scheme is monotone iff

$$h_1 \leq 0 \leq h_0.$$

One particular three-point monotone scheme is due to Godunov [10], and has a special significance to this theory. The flux for Godunov’s scalar scheme can be defined by

(2.5) 
$$h_{j-1/2}^G = h_G(u_j, u_{j-1}) = \begin{cases} \min_{u_{j-1} \leq u \leq u_j} f(u) & \text{if } u_{j-1} \leq u_j, \\ \max_{u_{j-1} \geq u \geq u_j} f(u) & \text{if } u_{j-1} > u_j. \end{cases}$$

Thus one can characterize  $E$  schemes as precisely those for which

(2.6) 
$$\begin{aligned} \text{(a)} \quad & h_{j-1/2} \leq h_{j-1/2}^G \quad \text{if } u_{j-1} < u_j, \\ \text{(b)} \quad & h_{j-1/2} \geq h_{j-1/2}^G \quad \text{if } u_j > u_{j-1}. \end{aligned}$$

It follows from [22, Lemma 2.1] that these approximations are at most first order accurate.

Together with an entropy inequality, a key estimate involved in many convergence proofs, is a bound on the variation. We present an argument originally due to Sanders for monotone schemes [29], used to obtain this bound.

For any fixed  $t > 0$ , the  $x$  variation of  $\mathcal{V}_\Delta(x, t)$  is

$$B(\mathcal{V}_\Delta) = \sum_j |\Delta_+ u_j(t)|.$$

Let

$$\chi_{j+1/2} = \begin{cases} 1 & \text{if } \Delta_+ u_j \geq 0, \\ -1 & \text{if } \Delta_+ u_j < 0. \end{cases}$$

Then:

$$\begin{aligned}\frac{d}{dt}B(\mathcal{V}_\Delta) &= \frac{1}{\Delta} \sum_j \frac{d}{dt}(\chi_{j+1/2} \Delta_+ u_j) \\ &= \frac{1}{\Delta} \sum_j \chi_{j+1/2} \frac{d}{dt} \Delta_+ u_j = -\frac{1}{\Delta} \sum_j \chi_{j+1/2} \Delta_+ \Delta_- h_{j-1/2} \\ &= \frac{1}{\Delta} \sum_j (\chi_{j+1/2} - \chi_{j-1/2}) \Delta_+ h_{j-1/2} \leq 0,\end{aligned}$$

if we can write

$$\begin{aligned}(a) \quad & \Delta_+ h_{j-1/2} = -[C_{j+1/2} \Delta_+ u_j - D_{j-1/2} \Delta_- u_j], \\ (2.7) \quad (b) \quad & C_{j+1/2} \geq 0, \\ (c) \quad & D_{j-1/2} \geq 0.\end{aligned}$$

In the case of  $E$  (or three-point monotone) schemes, this follows by defining:

$$C_{j+1/2} = -\left(\frac{h_{j+1/2} - f(u_j)}{\Delta_+ u_j}\right), \quad D_{j-1/2} = \left(\frac{f(u_j) - h_{j-1/2}}{\Delta_- u_j}\right)$$

as in [22].

Harten in [16] pointed out for explicit methods that a variation bound could be obtained for schemes which are higher order accurate. In our present method of lines context, it involves a five-point consistent approximation

$$(2.8) \quad (a) \quad \frac{\partial u_j}{\partial t} = -\frac{1}{\Delta} \Delta_+ h_{j-1/2} = (C_{j+1/2} \Delta_+ u_j - D_{j-1/2} \Delta_- u_j) / \Delta,$$

$$\text{with } (b) \quad C_{j+1/2} = C(u_{j+2}, u_{j+1}, u_j, u_{j-1}) \geq 0,$$

$$\text{and } (c) \quad D_{j-1/2} = D(u_{j+1}, u_j, u_{j-1}, u_{j-2}) \geq 0,$$

both Lipschitz continuous functions of their arguments. (See also van Leer [32].)

We have thus proven for schemes of type (2.8):

LEMMA 2.1. For any  $t_1 > t_2 \geq 0$

$$B(\mathcal{V}_\Delta(\cdot, t_1)) \leq B(\mathcal{V}_\Delta(\cdot, t_2)).$$

We also have the maximum and minimum principles:

LEMMA 2.2. Let  $\max u_j(0) = M$ ,  $\min u_j(0) = m$ . Then, for  $0 < t$  and each  $j$ ;

$$m \leq u_j(t) \leq M.$$

Moreover, if  $C_{j+1/2} > 0$ , and  $u_j(t) = M$ , then  $u_{j+1}(t) = M$ . If  $D_{j-1/2} < 0$  and  $u_j(t) = m$ , then  $u_{j-1}(t) = m$ .

*Proof.* The proof is trivial if the initial data is constant. Otherwise, we let  $\{u_j^\varepsilon(t)\}_{j=-N}^N$ , satisfy

$$(2.9) \quad \frac{\partial}{\partial t} u_j^\varepsilon = (C_{j+1/2} + \varepsilon) \Delta_+ u_j^\varepsilon - (D_{j-1/2} + \varepsilon) \Delta_- u_j^\varepsilon$$

for  $\varepsilon > 0$ , with the initial data  $u_j^\varepsilon(0) \equiv u_j(0)$ .



Suppose the maximum of  $u_j^\varepsilon(t)$ , for  $0 \leq t \leq T$ , occurs for  $u_{j_0}^\varepsilon(t_0)$ , with  $t_0 > 0$ . Then (2.8), (2.9), imply that  $u_j^\varepsilon(t_0) = u_{j_0}^\varepsilon(t_0)$ . Hence  $u_j^\varepsilon(t) \equiv \text{constant}$ , for  $0 \leq t \leq T$ . This is a contradiction. Thus:

$$\max_{j, 0 \leq t \leq T} \{u_j^\varepsilon(t)\}_{j=-N}^N < M.$$

Let  $\varepsilon \downarrow 0$ . It follows that  $u_j(t) \leq M$  for each  $j, t$ . The remainder of the proof follows analogously.

Next, we prove a limit on the possible accuracy of approximations (2.8).

LEMMA 2.3. *Approximation (2.8) is at most first order accurate at nonsonic critical points of  $u$ . (A sonic point  $\bar{u}$  is one such that  $f'(\bar{u}) = 0$ .)*

*Proof.* For any  $C^2$  function  $u(x)$ , we let  $C_{j+1/2}(u(x+2\Delta), u(x+\Delta), u(x), u(x-\Delta))$  be denoted by  $C_{j+1/2}$ ,  $C_{j+1/2}(u(x), u(x), u(x), u(x))$  be denoted by  $C_{j+1/2}(u)$ , and similarly for  $D_{j-1/2}$ .

Then

$$\begin{aligned} & -\frac{1}{\Delta} \Delta_- h_j(u(x+2\Delta), u(x+\Delta), u(x), u(x-\Delta)) \\ (2.10) \quad & = (C_{j+1/2}(u) - D_{j-1/2}(u))u_x + \frac{\Delta}{2} u_{xx}(C_{j+1/2}(u) + D_{j-1/2}(u)) \\ & + u_x(C_{j+1/2} - C_{j+1/2}(u) - D_{j-1/2} - D_{j-1/2}(u)) + o(\Delta). \end{aligned}$$

Consistency implies

$$(2.11) \quad C_{j+1/2}(u) - D_{j-1/2}(u) = -f'(u),$$

while second order accuracy at critical points means

$$(2.12) \quad C_{j+1/2}(u) + D_{j-1/2}(u) = 0.$$

Solving (2.11), (2.12), gives us

$$C_{j+1/2}(u) = -\frac{1}{2}f'(u), \quad D_{j-1/2}(u) = \frac{1}{2}f'(u).$$

Thus one of inequalities (2.8b, c) must fail at nonsonic  $u$ .

Thus, although schemes of the type (2.8) can be made to be as high as third order accurate, Lipschitz continuity of  $C_{j+1/2}$ ,  $D_{j-1/2}$  implies a local degeneracy to first order accuracy at smooth maxima and minima. This local degeneracy, together with some results on initial boundary value problems in [14], indicate strongly that overall second order accuracy is the best possible.

**3. Total variation diminishing, second order accurate, scalar approximations.** We now describe a systematic procedure used to construct second order methods of the form (2.8) from three-point, first order methods of the same type.

We begin with a three-point scheme

$$\begin{aligned} & \frac{\partial u_j}{\partial t} = -\frac{1}{\Delta} (h(u_{j+1}, u_j) - h(u_j, u_{j-1})) \\ (3.1) \quad & = -\frac{1}{\Delta} (h(u_{j+1}, u_j) - f(u_j)) - \frac{1}{\Delta} (f(u_j) - h(u_j, u_{j-1})) \\ & = \frac{1}{\Delta} C_{j+1/2}^{(1)} \Delta_+ u_j - \frac{1}{\Delta} D_{j-1/2}^{(1)} \Delta_- u_j \end{aligned}$$

for

$$C_{j+1/2}^{(1)} = - \left( \frac{h(u_{j+1}, u_j) - f(u_j)}{\Delta_+ u_j} \right) \geq 0,$$

$$D_{j-1/2}^{(1)} = \left( \frac{f(u_j) - h(u_j, u_{j-1})}{\Delta_- u_j} \right) \geq 0.$$

These quantities are both nonnegative for  $E$  schemes. Moreover, the entropy inequality was shown to be valid in [22] for this class. Thus, using Lemmas 2.1, 2.2 and a standard argument, e.g. [29], we have convergence to the unique entropy solution, as  $\Delta \rightarrow 0$ ; for this class.

However, the nonnegativity of these functions, by itself, need not imply convergence, since the entropy condition may fail, as it does for Murman's (Roe's) scheme. See e.g. [34]. We do assume nonnegativity of  $C_{j+1/2}^{(1)}$ ,  $D_{j-1/2}^{(1)}$  in what follows, as well as Lipschitz continuity of  $h(u, v)$ .

Our first attempt at constructing a higher order accurate scheme comes from a simple upwind type of hybridization:

$$(3.2) \quad \frac{\partial u_j}{\partial t} = \frac{1}{\Delta} \left( 1 - \frac{1}{2} \Delta_+ \right) C_{j+1/2}^{(1)} \Delta_+ u_j - \frac{1}{\Delta} \left( 1 + \frac{1}{2} \Delta_- \right) D_{j-1/2}^{(1)} \Delta_- u_j = -\frac{1}{\Delta} \Delta_+ \mathcal{H}_{j-1/2},$$

with the numerical flux satisfying

$$(3.3) \quad \begin{aligned} \mathcal{H}_{j-1/2} &= h(u_j, u_{j-1}) - \frac{1}{2}(h(u_{j+1}, u_j) - f(u_j)) + \frac{1}{2}(f(u_{j-1}) - h(u_{j-1}, u_{j-2})) \\ &= \frac{1}{2}(f(u_j) + f(u_{j-1})) - \frac{1}{2} \Delta_+ \Delta_- h(u_j, u_{j-1}). \end{aligned}$$

(Throughout this work we shall often use the fact that consistency is equivalent to the statement:

$$h(u, u) = f(u). \quad )$$

This scheme is, within second order accuracy, just the central difference algorithm, and is thus always second order accurate. It cannot, by Lemma 2.3, obey (2.8).

We can define the associated quantities so that

$$-\Delta_+ \mathcal{H}_{j-1/2} = C_{j+1/2}^{(2)} \Delta_+ u_j - D_{j-1/2}^{(2)} \Delta_- u_j$$

where, in this case:

$$(3.4) \quad \begin{aligned} C_{j+1/2}^{(2)} &= C_{j+1/2}^{(1)} \left( 1 - \frac{1}{2} \frac{(h(u_{j+2}, u_{j+1}) - h(u_{j+1}, u_{j+1}))}{(h(u_{j+1}, u_j) - h(u_j, u_j))} + \frac{1}{2} \right), \\ D_{j-1/2}^{(2)} &= D_{j-1/2}^{(1)} \left( 1 + \frac{1}{2} - \frac{1}{2} \left( \frac{h(u_{j-1}, u_{j-1}) - h(u_{j-1}, u_{j-2})}{h(u_j, u_j) - h(u_j, u_{j-1})} \right) \right). \end{aligned}$$

It is clear that the two quantities on the right above may become negative if the nontrivial ratios above become sufficiently large.

One possible remedy is to employ a flux limiter, using the notation of Sweby [30] (this issue, pp. 995–1011). (See also van Leer [32] for very original related work.)

Let:

$$(3.5) \quad R_j^+ = \frac{f(u_j) - h(u_j, u_{j-1})}{f(u_{j+1}) - h(u_{j+1}, u_j)}, \quad R_j^- = \frac{h(u_{j+1}, u_j) - f(u_j)}{h(u_j, u_{j-1}) - f(u_{j-1})}.$$

Our TVD approximation will be of the form

$$(3.6) \quad \frac{\partial u_j}{\partial t} = -\frac{1}{\Delta} \left[ h(u_{j+1}, u_j) - h(u_p, u_j) - \frac{1}{2} \Delta_- (\psi(R_{j+1}^-) (h(u_{j+1}, u_j) - h(u_p, u_j))) \right. \\ \left. + h(u_p, u_j) - h(u_p, u_{j-1}) + \frac{1}{2} \Delta_+ (\psi(R_{j-1}^+) (h(u_p, u_j) - h(u_p, u_{j-1}))) \right].$$

This scheme is easily seen to be second order accurate away from critical points, (i.e. points where the denominators in (3.5) approach zero), if the Lipschitz continuous function,  $\psi(R)$ , satisfies  $\psi(1) = 1$ .

The resulting quantities  $C_{j+1/2}$ ,  $D_{j-1/2}$  are required to be positive:

$$(3.7) \quad C_{j+1/2} = -\left( \frac{h(u_{j+1}, u_j) - h(u_p, u_j)}{\Delta_+ u_j} \right) \left[ 1 + \frac{1}{2} \left[ \frac{\psi(R_j^-)}{R_j^-} - \psi(R_{j+1}^-) \right] \right] \geq 0, \\ D_{j-1/2} = \left( \frac{h(u_p, u_j) - h(u_p, u_{j-1})}{\Delta_- u_j} \right) \left[ 1 + \frac{1}{2} \left[ \frac{\psi(R_j^+)}{R_j^+} - \psi(R_{j-1}^+) \right] \right] \geq 0.$$

We thus have the required inequalities for  $\psi(R^\pm)$ :

$$(3.8) \quad (a) \quad 1 + \frac{1}{2} \left[ \frac{\psi(R_j^-)}{R_j^-} - \psi(R_{j+1}^-) \right] \geq 0, \\ (b) \quad 1 + \frac{1}{2} \left[ \frac{\psi(R_j^+)}{R_j^+} - \psi(R_{j-1}^+) \right] \geq 0.$$

Various slope limiters have been developed. See Sweby [30], for a numerical and theoretical analysis of their properties. As an example, we may take

$$(3.9) \quad \psi(R) = \begin{cases} 0 & \text{if } R < 0, \\ R, & \text{if } 0 \leq R \leq k \text{ for } 1 \leq k \leq 2, \\ k & \text{if } R > k. \end{cases}$$

Let the numerical flux defined from (3.1)–(3.8) be called

$$\tilde{\mathcal{H}}_{j-1/2} = \tilde{\mathcal{H}}(u_{j+1}, u_p, u_{j-1}, u_{j-2}).$$

Its precise definition is

$$(3.10) \quad \tilde{\mathcal{H}}(u_{j+1}, u_p, u_{j-1}, u_{j-2}) = h(u_p, u_{j-1}) - \frac{1}{2} \psi(R_j^-) (h(u_p, u_{j-1}) - h(u_{j-1}, u_{j-1})) \\ + \frac{1}{2} \psi(R_{j-1}^+) (h(u_p, u_j) - h(u_p, u_{j-1})).$$

We thus have the following:

**THEOREM 3.1.** *The approximation defined through*

$$(3.11) \quad \frac{\partial u_j}{\partial t} = -\frac{1}{\Delta} \Delta_+ \tilde{H}_{j-1/2}, \quad j = 0, \pm 1, \dots, \pm N, \\ u_j(0) = \frac{1}{\Delta} \int_{\mathcal{I}_j} w_0(s) ds,$$

with  $w_0 \in L^1 \cap L^\infty \cap BV$ , defines  $u_\Delta(x, t)$  having the property that as  $\Delta \rightarrow 0$ , there exists a converging subsequence  $\mathcal{V}_\Delta(x, t)$  which converges in  $L^\infty(L^1(R), [0, T])$ , to a weak solution of (1.1).

The proof is a routine consequence of Lemmas 2.1 and (2.2). See, e.g. [29].

THEOREM (3.2). *The approximation (3.10) is second order accurate, except at isolated zeros of  $h_1(w, w)w_x$ , or  $h_0(w, w)w_x$ .*

See Sweby [30] for a related result.

#### 4. High resolution schemes and the entropy condition for scalar approximations.

Following recent tradition, we christen the schemes constructed in the last section high resolution schemes—namely, they are second order accurate (with the usual exceptional points), variation diminishing, and have a five-point bandwidth. The question remains—is it true that limit solutions satisfy the entropy condition? For first order non- $E$  schemes, variation diminishing is not enough. The perennial example is Murman's (Roe's) scheme with  $f''(u) > 0$  and

$$(4.1) \quad h^M(u_{j+1}, u_j) = \frac{1}{2}(f(u_{j+1}) + f(u_j)) - \frac{1}{2} \left| \frac{\Delta_+ f(u_j)}{\Delta_+ u_j} \right| \Delta_+ u_j.$$

Given an entropy violating shock  $u^L < u^R$ , with  $f(u^L) = f(u^R)$ , it is well known that

$$u_j = \begin{cases} u^L & \text{for } j \leq 0, \\ u^R & \text{for } j > 0 \end{cases}$$

is a steady solution to (3.1). Moreover, since  $h(u_{j+1}, u_j) = f(u^L) = f(u^R) = h(u_j, u_{j-1})$ , in this case, then any of our high resolution schemes (3.11) will have entropy violating solutions, if  $h(u_{j+1}, u_j) = h^M(u_{j+1}, u_j)$ .

Let  $V(w)$  be any convex function. In [22, § 3] it was shown for any solution of (2.2) that

$$(4.2) \quad \Delta \left( \frac{d}{dt} V(u_j) + D_+ F_A(u_j) \right) = \int_{u_j}^{u_{j+1}} dw V''(w) [h_{j+1/2} - f(w)],$$

where the approximate entropy flux is defined through

$$(4.3) \quad F_A(u_j) = F(u_j) + V'(u_j)[h_{j-1/2} - f(u_j)].$$

Thus a sufficient condition that any limit solution satisfy (1.7), for a fixed  $V$ , is that

$$(4.4) \quad \int_{u_j}^{u_{j+1}} V''(w) [h_{j+1/2} - f(w)] dw \leq 0.$$

In order that the above inequality be valid for all convex  $V$ , it is necessary and sufficient that  $h_{j+1/2}$  correspond to an  $E$  scheme, and hence that the approximation be at most first order accurate, [22]. Thus we shall only obtain our entropy inequality for a single  $V(w)$ , say  $V(w) = \frac{1}{2}w^2$ . This is sufficient for convergence, if  $f$  is convex, as will be shown in § 4.

We now proceed to modify the flux difference quantities, to ensure that (4.4) is valid, and, moreover, to sharpen discrete shock profiles. We shall do this by using the notion of artificial compression introduced by Harten [15], cf. also [16]. We also add negative artificial compression (=artificial rarefaction), if  $f'(u_j) < f'(u_{j+1})$ , and a certain amount of (positive) artificial compression if  $f'(u_j) > f'(u_{j+1})$ . The high resolution properties are preserved, and (4.4) is shown to be valid.

It is known that too much artificial compression can cause expansion shocks to develop. We shall obtain fairly precise bounds on the amount allowed.

Once again, we define

$$h_1(u, v) = \frac{\partial}{\partial u} h(u, v), \quad h_0(u, v) = \frac{\partial}{\partial v} h(u, v).$$

We shall also use two Lipschitz continuous functions of  $u_{j+1}$ ,  $u_j$  denoted by  $a_{j+1/2}^+$ ,  $a_{j-1/2}^-$  and defined below. Our scheme is an artificially compressed or rarefied version of (3.6).

We first modify our flux differences via the following:

$$\begin{aligned}
 (a) \quad & (h(u_{j+1}, u_j) - h(u_j, u_j))^m \\
 &= (h(u_{j+1}, u_j) - h(u_j, u_j)) \\
 &\quad \cdot \left[ 1 + a_{j+1/2}^- \frac{(h_1(u_{j+1}, u_{j+1}) - h_1(u_j, u_j)) \Delta_+ u_j}{(h(u_{j+1}, u_j) - h(u_j, u_j))} \right] \\
 (4.5) \quad (b) \quad & (h(u_{j+1}, u_{j+1}) - h(u_{j+1}, u_j))^m \\
 &= (h(u_{j+1}, u_{j+1}) - h(u_{j+1}, u_j)) \\
 &\quad \cdot \left[ 1 - a_{j+1/2}^+ \frac{(h_0(u_{j+1}, u_{j+1}) - h_0(u_j, u_j)) \Delta_+ u_j}{(h(u_{j+1}, u_{j+1}) - h(u_{j+1}, u_j))} \right].
 \end{aligned}$$

Here  $a_{j+1/2}^\mp$  are both positive numbers, chosen first so the quantities in brackets in (4.5a, b) are both always between 0 and 2, i.e., we must have:

$$\begin{aligned}
 (a) \quad & |a_{j+1/2}^- (h_1(u_{j+1}, u_{j+1}) - h_1(u_j, u_j))| \leq - \frac{(h(u_{j+1}, u_j) - h(u_j, u_j))}{\Delta_+ u_j}, \\
 (4.6) \quad (b) \quad & |a_{j+1/2}^+ (h_0(u_{j+1}, u_{j+1}) - h_0(u_j, u_j))| \leq \frac{h(u_{j+1}, u_{j+1}) - h(u_{j+1}, u_j)}{\Delta_+ u_j}.
 \end{aligned}$$

We now let

$$\begin{aligned}
 R_j^{m+} &= \frac{(h(u_j, u_j) - h(u_j, u_{j-1}))^m}{(h(u_{j+1}, u_{j+1}) - h(u_{j+1}, u_j))^m}, \\
 R_j^{m-} &= \frac{(h(u_{j+1}, u_j) - h(u_j, u_j))^m}{(h(u_j, u_{j-1}) - h(u_{j-1}, u_{j-1}))^m}.
 \end{aligned}$$

Our scheme, which uses ACR (artificial compression-rarefaction) is

$$\begin{aligned}
 (4.7) \quad \frac{\partial u_j}{\partial t} &= -\frac{1}{\Delta} \left[ h(u_{j+1}, u_j) - h(u_j, u_j) \right. \\
 &\quad - \frac{1}{2} \Delta_- (\psi(R_{j+1}^m) (h(u_{j+1}, u_j) - h(u_j, u_j))^m) \\
 &\quad + h(u_j, u_j) - h(u_j, u_{j-1}) \\
 &\quad \left. + \frac{1}{2} \Delta_+ (\psi(R_{j-1}^m) (h(u_j, u_j) - h(u_j, u_{j-1}))^m) \right]
 \end{aligned}$$

where we take  $\psi(R)$  to be as defined in (3.4), with  $k=1$ .

We let the numerical flux defined through (4.5), (4.6), (4.7) be denoted by  $\tilde{H}_{j-1/2}^{ac}$ . It is precisely defined by (3.10), with the superscript  $m$  attached to both  $R_j^\pm$  and the two flux differences.

Now, for TVD, we must have

$$\begin{aligned}
 0 &\leq - \left( \frac{h(u_{j+1}, u_j) - h(u_j, u_j)}{\Delta_+ u_j} \right) \\
 &\quad \cdot \left[ 1 + \frac{1}{2} \frac{(h(u_{j+1}, u_j) - h(u_j, u_j))^m}{(h(u_{j+1}, u_j) - h(u_j, u_j))} \left[ \frac{\psi(R_j^{m-})}{R_j^{m-}} - \psi(R_{j+1}^{m-}) \right] \right], \\
 0 &\leq \left( \frac{h(u_j, u_j) - h(u_j, u_{j-1})}{\Delta_- u_j} \right) \\
 &\quad \cdot \left[ 1 + \frac{1}{2} \frac{(h(u_j, u_j) - h(u_j, u_{j-1}))^m}{(h(u_{j+1}, u_j) - h(u_j, u_j))} \left[ \frac{\psi(R_j^{m+})}{R_j^{m+}} - \psi(R_{j-1}^{m+}) \right] \right]
 \end{aligned}$$

which is valid, because  $h$  satisfies (2.7), and because of (4.5), (4.6) and (3.10). Theorems (3.1) and (3.2) are also valid for this ACR version of our scheme, but our real goal is to choose the  $a_{j+1/2}^\pm$  so that the inequality

$$(4.8) \quad \int_{u_j}^{u_{j+1}} [\mathcal{H}_{j+1/2}^{ac} - f(w)] dw \leq 0$$

is valid, i.e. to enforce inequality (4.4), with  $V(w) = \frac{1}{2}w^2$ .

It is easy to see that

$$\begin{aligned}
 (4.9) \quad - \int_{u_j}^{u_{j+1}} f(w) dw &= -\frac{1}{2} \Delta_+ u_j [f(u_{j+1}) + f(u_j)] \\
 &\quad + \frac{1}{2} \int_{u_j}^{u_{j+1}} \left[ \left( \frac{1}{2} (\Delta_+ u_j) \right)^2 - \left( w - \frac{1}{2} (u_{j+1} + u_j) \right)^2 \right] f''(w) dw.
 \end{aligned}$$

Substituting this into (4.8) gives us the equivalent, desired inequality

$$\begin{aligned}
 (4.10) \quad &\frac{1}{2} \int_{u_j}^{u_{j+1}} \left[ \left( \frac{1}{2} \Delta_+ u_j \right)^2 - \left( w - \frac{1}{2} (u_{j+1} + u_j) \right)^2 \right] f''(w) dw \\
 &+ \frac{1}{2} \Delta_+ u_j [(h(u_{j+1}, u_j) - h(u_j, u_j))^m (1 - \psi(R_{j+1}^{m-})) \\
 &\quad - a_{j+1/2}^- (h_1(u_{j+1}, u_{j+2}) - h_1(u_j, u_j)) \Delta_+ u_j] \\
 &+ \frac{1}{2} (\Delta_+ u_j) [(h(u_{j+1}, u_j) - h(u_{j+1}, u_{j+1}))^m (1 - \psi(R_j^{m+})) \\
 &\quad - a_{j+1/2}^+ (h_0(u_{j+1}, u_{j+1}) - h_0(u_j, u_j)) \Delta_+ u_j] \leq 0.
 \end{aligned}$$

From (4.5), (4.6), and the fact that  $h$  is an  $E$  flux, it suffices to prove (after cancelling the factor  $\frac{1}{2}$ )

$$\begin{aligned}
 (4.11) \quad &\int_{u_j}^{u_{j+1}} f''(w) dw \left[ \left( \frac{1}{2} \Delta_+ u_j \right)^2 - \left( w - \frac{1}{2} (u_{j+1} + u_j) \right)^2 \right] \\
 &- a_{j+1/2}^- (\Delta_+ u_j)^2 [h_1(u_{j+1}, u_{j+1}) - h_1(u_j, u_j)] \\
 &- a_{j+1/2}^+ (\Delta_+ u_j)^2 [h_0(u_{j+1}, u_{j+1}) - h_0(u_j, u_j)] \leq 0.
 \end{aligned}$$

We thus have:

LEMMA 4.1. *Any solution to (1.1) which is the limit of a subsequence of approximate solutions  $\mathcal{V}_{\Delta'}$  as  $\Delta' \rightarrow 0$  of*

$$(4.12) \quad \begin{aligned} \frac{\partial u_j}{\partial t} &= -\frac{1}{\Delta} \Delta_+ \mathcal{H}_{j-1/2}^{ac}, \quad j = 0, \pm 1, \dots, \pm N, \\ u_j(0) &= \frac{1}{\Delta} \int_{\mathcal{J}_j} w_0(s) \, ds, \end{aligned}$$

*satisfies the entropy inequality (1.6), provided that inequalities (4.6) and (4.11) are valid.*

As mentioned in the introduction, for scalar *convex* conservation laws where solutions lie in  $BV$ , a single entropy inequality implies uniqueness of solutions. This follows from the recent results of DiPerna [5]. We know that if  $w(x, 0) \in BV$ , then the function  $\mathcal{V}_{\Delta}(x, t)$  has  $x$  variation bounded by that of  $w(x, 0)$ , for any fixed  $t$ . In particular

$$(4.13) \quad \frac{1}{h} \int_{-1}^1 |\mathcal{V}_{\Delta}(x+h, t) - \mathcal{V}_{\Delta}(x, t)| \, dx \leq \text{var}(w(x, 0))$$

for any  $t > 0$ ,  $\Delta > 0$ ,  $h > 0$ .

The dominated convergence theorem then guarantees that any limit solution  $w(x, t)$  will have  $x$  variation bounded by  $w(x, 0)$ , uniformly in  $t$ . Consider the interval  $0 \leq t \leq T$ , for any  $T > 0$ . Then

$$(4.14) \quad \begin{aligned} &\frac{1}{h} \int_{-1}^1 dx \int_0^T dt |\mathcal{V}_{\Delta}(x, t+h) - \mathcal{V}_{\Delta}(x, t)| \\ &= \frac{1}{h} \int_{-1}^1 dx \int_0^T dt \left| \int_t^{t+h} \frac{\partial \mathcal{V}_{\Delta}}{\partial s}(x, s) \, ds \right| \\ &= \frac{1}{h} \int_0^T dt \left| \int_t^{t+h} \Delta_- \mathcal{H}_{j+1/2}(s) \, ds \right| \leq 4MT \text{var } w(x, 0), \end{aligned}$$

where  $M$  is the Lipschitz constant in

$$|\mathcal{H}(u_{j+1}, u_j, u_{j-1}, u_{j-2}) - \mathcal{H}(v_{j+1}, v_j, v_{j-1}, v_{j-2})| \leq M \sum_{r=-2}^1 |u_{j+r} - v_{j+r}|$$

for any pair of vectors  $\{u_j\}, \{v_j\}$ .

Thus the dominated convergence theorem implies that any limit solution  $w(x, t)$  will have bounded variation in  $x$  and  $t$ . This together with Lemma 4.1 and the above mentioned uniqueness result gives our

THEOREM 4.1 (convergence). *The sequence of approximate solutions  $\mathcal{V}_{\Delta}$  converges a.e. as  $\Delta \rightarrow 0$  to the unique solution of the scalar convex conservation law (1.1) provided that the initial data is in  $BV$  and that inequalities (4.6) and (4.11) are valid.*

For simplicity, we exemplify this theory by considering semi-discrete versions of the following three monotone schemes.

(a) Lax-Friedrichs:

$$h^{LF}(u_{j+1}, u_j) = \frac{1}{2}f(u_{j+1}) + \frac{1}{2}f(u_j) - \frac{1}{2}k(u_{j+1} - u_j)$$

with  $0 < k$  chosen so that  $|f'(u)| \leq k\theta$ , for some  $\theta$  with  $0 < \theta < 1$ .

(b) Engquist–Osher:

$$(4.15) \quad h^{EO}(u_{j+1}, u_j) = \int_0^{u_{j+1}} \min(f'(s), 0) ds + \int_0^{u_j} \max(f'(s), 0) ds - f(0) \\ = f_-(u_{j+1}) + f_+(u_j) - f(0).$$

(c) Godunov:

$$h^G(u_{j+1}, u_j) = \begin{cases} \min_{u_j \leq u \leq u_{j+1}} f(u) & \text{if } u_j < u_{j+1}, \\ \max_{u_j \geq u \geq u_{j+1}} f(u) & \text{if } u_j > u_{j+1}. \end{cases}$$

For Lax–Friedrichs, (4.6a, b) become

$$(4.16) \quad |a_{j+1/2}^\pm(f'(u_{j+1}) - f'(u_j))| \leq k \pm \frac{(f(u_{j+1}) - f(u_j))}{\Delta_+ u_j},$$

while (4.11) is valid if

$$(4.17) \quad \int_{u_j}^{u_{j+1}} f''(w) dw \left[ \left( \frac{\Delta_+ u_j}{2} \right)^2 - \left( w - \frac{1}{2}(u_{j+1} + u_j) \right)^2 \right] \\ \leq \min_{+,-} a_{j+1/2}^\mp (\Delta_+ u_j)^2 (f'(u_{j+1}) - f'(u_j)).$$

Clearly these inequalities are compatible if  $\theta$  is sufficiently small. Moreover, if  $f$  is convex, then (4.16) becomes(a) if  $u_j < u_{j+1}$  (rarefaction)

$$(4.18) \quad a_{j+1/2}^\mp \geq \frac{\int_{u_j}^{u_{j+1}} f''(w) dw \left[ \left( \frac{\Delta_+ u_j}{2} \right)^2 - \left( w - \frac{1}{2}(u_{j+1} + u_j) \right)^2 \right]}{(\Delta_+ u_j)^2 (f'(u_{j+1}) - f'(u_j))}$$

(b) if  $u_j > u_{j+1}$  (shock)

$$a_{j+1/2}^\mp \leq \frac{\int_{u_j}^{u_{j+1}} f''(w) dw \left[ \left( \frac{\Delta_+ u_j}{2} \right)^2 - \left( w - \frac{1}{2}(u_{j+1} + u_j) \right)^2 \right]}{(\Delta_+ u_j)^2 (f'(u_{j+1}) - f'(u_j))}.$$

The right side of (4.18) is always positive for convex  $f$ . For example, for  $f(w) = \frac{1}{2}w^2$ , the right side is  $\frac{1}{6}$ .

The precise restrictions for this case become for Burgers' equation:

$$(4.19) \quad (a) \quad |a_{j+1/2}^\pm \Delta_+ u_j| + \frac{1}{2}|u_{j+1} + u_j| \leq k$$

and

$$(b) \quad \operatorname{sgn}(u_{j+1} - u_j)(a_{j+1/2}^\pm - \frac{1}{6}) \geq 0.$$

Thus the high resolution scheme based on  $LF$  is a convergent approximation to Burgers' equation if  $\theta$  is small enough and the  $a_{j+1/2}^\pm$  are then chosen as above.

For the Engquist–Osher scheme, (4.6a,b) become

$$(4.20) \quad |a_{j+1/2}^\mp (f'_\mp(u_{j+1}) - f'_\mp(u_j))| \leq \frac{\mp (f_\mp(u_{j+1}) - f_\mp(u_j))}{\Delta_+ u_j},$$



while (4.11) is valid if

$$(4.21) \quad \int_{u_j}^{u_{j+1}} f''_{\mp}(w) dw \left[ \left( \frac{\Delta_+ u_j}{2} \right)^2 - \left( w - \frac{1}{2}(u_{j+1} + u_j) \right)^2 \right] \\ \leq a_{j+1/2}^{\pm} (\Delta_+ u_j)^2 [f'_{\mp}(u_{j+1}) - f'_{\mp}(u_j)].$$

Inequalities (4.20) and (4.21) are obviously compatible for convex  $f$ , for certain bounded  $a_{j+1/2}^{\pm}$ , if  $|\Delta_+ u_j|$  is sufficiently small (i.e. in a region of smoothness). They are also compatible for  $|\Delta_+ u_j|$  bounded away from zero. This follows since we may take in (4.20)

$$(4.22) \quad a_{j+1/2}^{\pm} = \frac{\mp (\Delta_+ f_{\mp} u_j)}{(\Delta_+ u_j)(\Delta_+ f'_{\mp}(u_j))}.$$

It now remains to prove that (4.21) is valid, i.e.,

$$(4.23) \quad \int_{u_j}^{u_{j+1}} f''(w) dw \left[ \left( \frac{\Delta_+ u_j}{2} \right)^2 - \left( w - \frac{1}{2}(u_{j+1} + u_j) \right)^2 \right] \pm (\Delta_+ f_{\mp}(u_j)) \Delta_+ u_j \leq 0,$$

or equivalently

$$(4.24) \quad 2 \int_{u_j}^{u_{j+1}} (f_{-}(u_{j+1}) - f_{-}(w)) dw \leq 0, \\ 2 \int_{u_j}^{u_{j+1}} (f_{+}(u_j) - f_{+}(w)) dw \leq 0.$$

These last are obvious facts.

In the case  $f(w) = \frac{1}{2}w^2$ , then (4.21) becomes

(Burgers' equation)

(a) if  $u_j < u_{j+1}$  (rarefaction), then:

$$(i) \quad \begin{aligned} a_{j+1/2}^{-} &\geq \frac{1}{6} \\ a_{j+1/2}^{+} &\geq 0 \end{aligned} \quad \text{if } u_{j+1} \leq 0,$$

$$(ii) \quad \begin{aligned} a_{j+1/2}^{-} &\geq 0 \\ a_{j+1/2}^{+} &\geq \frac{1}{6} \end{aligned} \quad \text{if } u_j \geq 0,$$

$$(iii) \quad \begin{aligned} a_{j+1/2}^{-} &\geq \frac{1}{4} - \frac{1}{12} \left( \frac{3u_{j+1}^2 + u_j^2}{(\Delta_+ u_j)^2} \right) \\ a_{j+1/2}^{+} &\geq -\frac{1}{12} + \frac{1}{12} \left( \frac{3u_{j+1}^2 + u_j^2}{(\Delta_+ u_j)^2} \right) \end{aligned} \quad \text{if } u_{j+1} > 0 > u_j;$$

(4.25)

(b) if  $u_j \geq u_{j+1}$  (shock), then

$$(i) \quad \begin{aligned} a_{j+1/2}^{-} &\leq \frac{1}{6} \\ a_{j+1/2}^{+} &\leq 0 \end{aligned} \quad \text{if } u_j \leq 0,$$

$$(ii) \quad \begin{aligned} a_{j+1/2}^{-} &\leq 0 \\ a_{j+1/2}^{+} &\leq \frac{1}{6} \end{aligned} \quad \text{if } u_{j+1} \geq 0,$$

$$(iii) \quad \begin{aligned} a_{j+1/2}^{-} &\leq \frac{1}{4} - \frac{1}{12} \left( \frac{3u_j^2 + u_{j+1}^2}{(\Delta_+ u_j)^2} \right) \\ a_{j+1/2}^{+} &\leq -\frac{1}{12} + \frac{1}{12} \left( \frac{3u_j^2 + u_{j+1}^2}{(\Delta_+ u_j)^2} \right) \end{aligned} \quad \text{if } u_{j+1} < 0 < u_j$$

Moreover, if  $|a_{j+1/2}^\mp| \leq \frac{1}{2}$  (which is admissible from (4.25)), then (4.20) is no restriction at all.

Thus, the E–O scheme incorporated into  $H_{j+1/2}^{ac}$  yields a convergent high resolution scheme for any convex  $f$ , and any initial datum having bounded variation, provided the  $a_{j+1/2}^\pm$  are judiciously chosen.

A tedious calculation yields the analogous results for Godunov's scheme.

**5. Steady discrete shocks.** We now check for the existence of discrete, steady, shock solutions to the high resolution, entropy condition satisfying schemes constructed in the previous section, based on the Engquist–Osher flux. Since the scheme satisfies Lemma 2.1, the profiles must be monotone. We shall also show that they are sharp.

Let  $u^L, u^R$ , be the left and right states for a physically correct steady shock, i.e.

$$\begin{aligned} f(u^L) &= f(u^R), \\ f(u) &< f(u^L) \quad \text{for all } u \text{ between } u^L \text{ and } u^R. \end{aligned}$$

A steady, discrete shock  $\{u_j\}$  will satisfy

$$(5.1) \quad \lim_{j \rightarrow -\infty} u_j = u^L, \quad \lim_{j \rightarrow \infty} u_j = u^R, \\ \tilde{H}_{j+1/2}^{ac} \equiv f(u^L) \quad \text{for all } j.$$

For simplicity, we assume  $f''(u) > 0$ , with  $f'(0) = 0$  and  $f(0) = 0$ . The numerical flux becomes

$$(5.2) \quad \mathcal{H}_{j+1/2}^{ac} = f_-(u_{j+1}) + f_+(u_j) - \frac{1}{2} \psi(R_{j+1}^{M-})(f_-(u_{j+1}) - f_-(u_j))^M \\ + \frac{1}{2} \psi(R_j^{M+})(f_+(u_{j+1}) - f_+(u_j))^M.$$

If we take  $k = 1$  in (3.9), we have

$$(5.3) \quad \begin{aligned} (a) \quad & \psi(R_{j+1}^{M-})(f_-(u_{j+1}) - f_-(u_j))^M \\ &= -\Delta_+ u_j \max \left[ 0, \min_{r=0,1} ([\Delta_+ f_-(u_{j+r}) \right. \\ & \quad \left. + a_{j+1/2+r}^-(\Delta_+ f'_-(u_{j+r}) \Delta_+ u_{j+r})] (-\Delta_+ u_j)^{-1}) \right], \\ (b) \quad & \psi(R_j^{M+})(f_+(u_{j+1}) - f_+(u_j))^M \\ &= \Delta_+ u_j \max \left[ 0, \min_{r=0,1} ([\Delta_+ f_+(u_{j+r-1}) \right. \\ & \quad \left. - a_{j-1/2+r}^+(\Delta_+ f'_+(u_{j+r-1}) \Delta_+ u_{j+r-1})] (\Delta_+ u_j)^{-1}) \right]. \end{aligned}$$

We shall seek steady discrete shocks of the same general form obtained in [8] for the E–O scheme.

These are

$$(5.4) \quad u_j \equiv \begin{cases} u^L, & j \leq -1, \\ u^R, & j \geq 2, \end{cases} \quad u^L > u_0 \geq 0 > u_1 \geq u^R.$$

For the first order scheme,  $u_0$  can be viewed as a smooth function of  $u_1$ , satisfying the above and

$$f(u_0) + f(u_1) = f(u^L).$$

For the present scheme, we shall get a different one-parameter family of intermediate states.

It follows for all  $j \neq 0$ , that (5.3) implies:  $\tilde{H}_{j+1/2}^{ac} \equiv f(u^L)$ .

For  $j=0$ , we have the following equation:

$$\begin{aligned}
 0 &= f(u_1) + f(u_0) - f(u^L) \\
 &+ \frac{1}{2}(u_1 - u_0) \min \left[ \frac{f(u_1) + a_{1/2}^- f'(u_1)(u_1 - u_0)}{u_0 - u_1}, \right. \\
 (5.5) \quad &\left. \frac{f(u^R) - f(u_1) + a_{3/2}^- (f'(u^R) - f'(u_1))(u^R - u_1)}{u_0 - u_1} \right] \\
 &+ \frac{1}{2}(u_1 - u_0) \min \left[ \frac{f(u_0) - f(u^L) - a_{-1/2}^+ (f'(u_0) - f'(u^L))(u - u^L)}{u_1 - u_0}, \right. \\
 &\left. \frac{-f(u_0) - a_{-1/2}^+ (-f'(u_0))(u_1 - u_0)}{u_1 - u_0} \right] = G(u_1, u_0).
 \end{aligned}$$

One special solution is, again,  $u_1 = u^R$ ,  $u_0 = 0$ . We have

$$\begin{aligned}
 (5.6) \quad &\frac{\partial G}{\partial u_1}(u^R, 0) = \frac{3}{2} f'(u^R) < 0, \\
 &\frac{\partial G}{\partial u_0}(u^R, 0) = 0, \\
 &\frac{\partial^2 G}{\partial u_1^2}(u^R, 0) = \left( \frac{3}{2} - \frac{1}{2} a_{3/2}^- \right) f''(u^R) > 0, \\
 &\frac{\partial^2 G}{\partial u_0 \partial u_1}(u^R, 0) = 0, \\
 &\frac{\partial^2 G}{\partial u_0^2} = \frac{1}{2} f''(0) > 0.
 \end{aligned}$$

A simple application of the implicit function theorem gives us:

**THEOREM 5.1.** *There exists a family of sharp, discrete, shock solutions to (4.12) of the type (5.3), with  $u_1$  a smooth function of  $u_0$  and  $0 < u_0$  small enough. By symmetry, the same is true with  $u_0$  a smooth function of  $u_1$ , and  $-u_1$  small enough.*

**6. Systems of conservation laws.** We shall now build a second order accurate scheme approximating (1.1) for systems:  $m > 1$  and verify some of the desirable properties mentioned above. Our basic three point first order scheme will be the one devised by the first author [24], with Solomon [26], and analyzed jointly by us in [2], [25].

The numerical flux for this scheme is constructed as follows. First, we define a piecewise smooth, continuous path in phase space, connecting  $w_j$  to  $w_{j+1}$ , made up of  $m$  subpaths. Along subpath  $k$ , we have

$$(6.1) \quad \frac{dw}{ds} = r_k(w(s)),$$

with  $r_k$ , the  $k$ th eigenvalue of  $\partial f(w)$  corresponding to eigenvalue  $\lambda_k$ . (Here  $\lambda_1 < \lambda_2 < \dots < \lambda_m$ .)

On such a subpath, the  $m-1$  independent Riemann invariants, corresponding to field  $k$ , are constants. Our ordering is as follows. We begin at  $w$  with  $k=m$ , stop at some endpoint, use  $k=m-1$ , etc., arriving at  $w_{j+1}$  with the subpath corresponding to  $k=1$ . We call each such subpath  $\Gamma_{j+1/2}^\gamma$ ,  $\gamma=m, m-1, \dots, 1$ . There exists exactly one such path  $\Gamma_{j+1/2} = \bigcup_{\gamma=m}^1 \Gamma_{j+1/2}^\gamma$ , for  $|w_{j+1} - w_j|$  sufficiently small (actually, for physical systems such as Euler's equations, the construction works in the large, as long as cavitation does not take place [26]).

We define the numerical flux to be

$$(6.2) \quad h_{j+1/2}^0 = \frac{1}{2} \left[ f(w_{j+1}) + f(w_j) - \int_{\Gamma_{j+1/2}} |\partial f(w)| dw \right].$$

Here  $|\cdot|$  denotes the absolute value of a diagonalizable matrix.

This algorithm yields a closed form expression for  $h_{j+1/2}^0$  for various physical systems, including Euler's equations for compressible, inviscid, gas dynamics, [26], [3], because their Riemann invariants can be easily tabulated, and because all fields are either linearly degenerate, or genuinely nonlinear. We assume this property throughout the remainder of this paper, for (1.1).

On such a  $k$  subpath, the integral in (6.2) is given by a simple expression involving  $\pm f(w)$  at endpoints of the subinterval, and, in the genuinely nonlinear case, it may also involve  $\pm 2f(\bar{w})$ , where  $\bar{w}$  is the unique sonic point for which  $\lambda_k(\bar{w}) = 0$ .

The full algorithm can be interpreted, geometrically, as solving the incoming Riemann problem in phase space, using only rarefaction, compression, or contact waves, then averaging the resulting multivalued solution, in physical space, as in Godunov's method [34].

The first order scheme may be written so as to yield a field by field decomposition:

$$(6.3) \quad \frac{\partial}{\partial t} w_j = -\frac{1}{\Delta} (h^0(w_{j+1}, w_j) - h^0(w_j, w_j)) - \frac{1}{\Delta} (h^0(w_j, w_j) - h^0(w_j, w_{j-1})),$$

with

$$(6.4) \quad \begin{aligned} (a) \quad h^0(w_{j+1}, w_j) - h^0(w_j, w_j) &= \int_{\Gamma_{j+1/2}} (\partial f(w))^+ dw, \\ (b) \quad h^0(w_j, w_j) - h^0(w_j, w_{j-1}) &= \int_{\Gamma_{j-1/2}} (\partial f(w))^- dw. \end{aligned}$$

Here

$$(\partial f)^+ = \frac{1}{2}((\partial f) + |\partial f|), \quad (\partial f)^- = \frac{1}{2}((\partial f) - |\partial f|).$$

For a linearly degenerate  $k$  field, we have  $\lambda_k$  constant on  $\Gamma_{j+1/2}^k$ , and hence:

$$(6.5) \quad \begin{aligned} \int_{\Gamma_{j+1/2}^k} (\partial f)^+ dw &= \chi(\lambda_k^{j+1/2}) [f(w^u) - f(w^L)], \\ \int_{\Gamma_{j+1/2}^k} (\partial f)^- dw &= (1 - \chi(\lambda_k^{j+1/2})) [f(w^u) - f(w^L)] \end{aligned}$$

where  $\chi(x) \equiv 1$  if  $x > 0$ ,  $\chi(x) \equiv 0$  if  $x \leq 0$ , and  $w^u, w^L$  are the upper and lower limits of integration.

For a genuinely nonlinear  $k$  field, we have  $\lambda_k$  monotone on  $\Gamma_{j+1/2}^k$ , and hence:

$$(6.6) \quad \left. \begin{aligned} (a) \quad \int_{\Gamma_{j+1/2}^k} (\partial f(w))^+ dw &= f(w) \\ (b) \quad \int_{\Gamma_{j+1/2}^k} (\partial f(w))^- dw &= f(w) \end{aligned} \right\} \begin{aligned} & \left. \begin{aligned} w^u, & \text{ if } \lambda_k(w^u) > \lambda_k(\bar{w}), \\ \bar{w}, & \text{ if } \lambda_k(\bar{w}) \geq \lambda_k(w^u), \\ w^L, & \text{ if } \lambda_k(w^L) > \lambda_k(\bar{w}), \\ \bar{w}, & \text{ if } \lambda_k(\bar{w}) \geq \lambda_k(w^L); \end{aligned} \right\} \\ & \left. \begin{aligned} w^u, & \text{ if } \lambda_k(w^u) \leq \lambda_k(\bar{w}), \\ \bar{w}, & \text{ if } \lambda_k(\bar{w}) < \lambda_k(w^u), \\ w^L, & \text{ if } \lambda_k(w^L) \leq \lambda_k(\bar{w}), \\ \bar{w}, & \text{ if } \lambda_k(\bar{w}) < \lambda_k(w^L). \end{aligned} \right\} \end{aligned}$$

As mentioned above,  $\bar{w}$  is the unique sonic point (if it exists) for which  $\lambda_k(\bar{w}) = 0$  on  $\Gamma_{j+1/2}^k$ .

This nonlinear field by field decomposition enables us to construct second order accurate schemes while working separately in each field. See, [16] for a similar treatment (without an entropy inequality), using Roe's nonlinear decomposition [28].

Let the end points of  $\Gamma_{j+1/2}^k$  be denoted by  $w^L = w_{j+1-k/m}$ ,  $w^U = w_{j+1-(k-1)/m}$ , and let the associated

$$(6.7) \quad \begin{aligned} (a) \quad \int_{\Gamma_{j+1/2}^k} (\partial f)^- dw &= N_{k,j+1/2}^-, \\ (b) \quad \int_{\Gamma_{j-1/2}^k} (\partial f)^+ dw &= N_{k,j-1/2}^+. \end{aligned}$$

As in the scalar case, a first attempt at a second order accurate extension comes via

$$(6.8) \quad \frac{\partial w}{\partial t} = -\frac{1}{\Delta} \left( \sum_{k=m}^1 \left( 1 - \frac{1}{2} \Delta_+ \right) N_{k,j+1/2}^- + \sum_{k=m}^1 \left( 1 + \frac{1}{2} \Delta_- \right) N_{k,j-1/2}^+ \right).$$

This differencing, although formally second order accurate, again leads to overshoot, undershoot, and occasionally nonlinear instability.

To remedy this, we limit the flux differencing, as in the scalar case, except that we do this separately in each field. We shall use the entropy function, discussed in the first section, in the construction of the scheme. We remark, that although the existence of a convex  $V(x)$  follows only if the overdetermined system

$$(\nabla_w V)^T \left( \frac{\partial f}{\partial w} \right) = \nabla_w F$$

has a convex solution, the fact is that most of the equations of physics are endowed with such an entropy. See [6] for further details.

The entropy gradient will be used to construct a linear functional applied to vectors  $w$ . This is correct dimensionally, as pointed out by Harten-Lax in [17].

We first define

$$(6.9) \quad V_w(w_{j+1-(k-1)/m}) - V_w(w_{j+1-(k/m)}) = \Delta_+^{(j,k)} V_w.$$

The algorithm in each field will involve the quantities

$$(6.10) \quad R_{k,j}^+ = \frac{(N_{k,j-1/2}^+, \Delta_+^{(j,k)} V_w)}{(N_{k,j+1/2}^+, \Delta_+^{(j,k)} V_w)}, \quad R_{k,j}^- = \frac{(N_{k,j+1/2}^-, \Delta_+^{(j,k)} V_w)}{(N_{k,j-1/2}^-, \Delta_+^{(j,k)} V_w)}$$

for each field  $k = 1, \dots, m$ .

Here  $(\cdot, \cdot)$  denotes the usual inner product in  $R^m$ .

Using any flux limiter of the type described in § 3 (i.e.  $\psi(R)$  satisfies (3.8), with  $\psi(1) = 1$ ), we define a second order accurate, TVD type discretization:

$$(6.11) \quad \frac{\partial w_j}{\partial t} = -\frac{1}{\Delta} \left( \sum_{k=1}^m \left( N_{k,j+1/2}^- - \frac{1}{2} \Delta_- (\psi(R_{k,j+1}^-) N_{k,j+1/2}^-) \right) + \sum_{k=1}^m \left( N_{k,j-1/2}^+ + \frac{1}{2} \Delta_+ (\psi(R_{k,j-1}^+) N_{k,j-1/2}^+) \right) \right).$$

We now show that for linear systems, this nonlinear scheme decouples to a second order accurate variation diminishing algorithm, and hence approximate solutions converge, as  $\Delta \rightarrow 0$ , to the unique solution of (1.1). (See also [16], for the first proof of such a result.)

We have

$$\partial f(w) = A = R \Lambda L,$$

with:

$$R = [r_1, r_2, \dots, r_m] \quad \text{and} \quad L = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_m \end{bmatrix},$$

the constant matrices made up of left and right eigenvalues of  $A$ , suitably normalized so that

$$(l_i, r_j) = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j, \end{cases}$$

and

$$\Lambda = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_m \end{bmatrix}.$$

Then

$$N_{k,j-1/2}^+ = (l_k, \Delta_- w_j) \max(\lambda_k, 0) r_k, \\ N_{k,j+1/2}^- = (l_k, \Delta_+ w_j) \min(\lambda_k, 0) r_k.$$

Moreover, since the entropy is just defined by  $V(w) = \frac{1}{2}(w, w)$  for linear systems, then

$$\Delta_+^{(j,k)} V_w = (l_k, \Delta_+ w) r_k.$$

Thus, if we take the inner product of (6.11) with  $l_k$ , and denote  $(l_k, w_j)$  by  $w_j^{(k)}$ , we obtain

$$\begin{aligned} \frac{\partial}{\partial t} w_j^{(k)} = & -\frac{1}{\Delta} \left[ \min(\lambda_k, 0) \left( \Delta_+ w_j^{(k)} - \frac{1}{2} \Delta_- (\psi(R_{k,j+1}^-) (\Delta_+ w_j^{(k)})) \right) \right. \\ & \left. + \max(\lambda_k, 0) \left( \Delta_- w_j^{(k)} + \frac{1}{2} \Delta_+ (\psi(R_{k,j-1}^+) (\Delta_- w_j^{(k)})) \right) \right]. \end{aligned}$$

This scheme is of the type (3.6)–(3.8), and hence is TVD. Thus we have the variation bound for each  $w_j^{(k)}$ , as in § 2, as well as the maximum (minimum) principles stated there. We thus have the simple:

**THEOREM 6.1.** *The approximation to the linear hyperbolic system (1.1), defined by (6.11), with initial data in  $L^1 \cap L^\infty \cap BV$ , converges as  $\Delta \rightarrow 0$  in  $L^\infty(L^1(R), [0, T])$ , to the unique solution of (1.1).*

Next, in order to sharpen shocks and contacts, and to enforce a discrete entropy condition for systems, we use ACR, as in the scalar case. We again (as in the scalar case) require that  $\psi(R)$  be defined by (3.4), with  $k = 1$ .

In order to add ACR, we use the  $k$  eigenvalue difference along a  $k$  subpath. This is defined by

$$\begin{aligned} (6.12) \quad & \max(\lambda_k(w_{j+1-(k-1/m)}), 0) - \max(\lambda_k(w_{j+1-(k/m)}), 0) = \Delta_+^{(j,k)} \lambda_+, \\ & \min(\lambda_k(w_{j+1-(k-1/m)}), 0) - \min(\lambda_k(w_{j+1-(k/m)}), 0) = \Delta_+^{(j,k)} \lambda_-. \end{aligned}$$

Next we define

$$(6.13) \quad (a) \quad N_{k,j+1/2}^{M_+} = N_{k,j+1/2}^+ \left( 1 - a_{k,j+1/2}^+ (\Delta_+^{(j,k)} \lambda_+) \frac{(\Delta_+^{(j,k)} w, \Delta_+^{(j,k)} V_w)}{(N_{k,j+1/2}^+, \Delta_+^{(j,k)} V_w)} \right)$$

for genuinely nonlinear  $k$  fields, and

$$(6.13) \quad (b) \quad N_{k,j+1/2}^{M_+} = N_{k,j+1/2}^+ (1 + a_{k,j+1/2}^+ (\Delta_+^{(j,k)} w, \Delta_+^{(j,k)} V_w)^{1/2})$$

for linearly degenerate fields.

Analogously, we define

$$(6.14) \quad (a) \quad N_{k,j+1/2}^{M_-} = N_{k,j+1/2}^- \left( 1 + a_{k,j+1/2}^- (\Delta_+^{(j,k)} \lambda_-) \frac{(\Delta_+^{(j,k)} w, \Delta_+^{(j,k)} V_w)}{(N_{k,j+1/2}^-, \Delta_+^{(j,k)} V_w)} \right)$$

for genuinely nonlinear fields, and

$$(6.14) \quad (b) \quad N_{k,j+1/2}^{M_-} = N_{k,j+1/2}^- (1 + a_{k,j+1/2}^- (\Delta_+^{(j,k)} w, \Delta_+^{(j,k)} V_w)^{1/2})$$

for linearly degenerate fields.

We make the following restriction on the Lipschitz continuous functions  $a_{k,j+1/2}^\mp$ :

**Restriction 6.1.** The values of the coefficients multiplying  $N_{k,j+1/2}^\mp$  in (6.13) and (6.14) are always strictly between 0 and 2.

We now define  $R_{k,j}^{M_{\pm}}$  as in (6.10), replacing  $N_{k,j+1/2}^{\pm}$  by these compressed or rarefied flux differences. The TVD and ACR scheme is constructed as follows:

$$\begin{aligned}
 \frac{\partial w_j}{\partial t} = & -\frac{1}{\Delta} \left( \sum_{k=1}^m \left( N_{k,j+1/2}^- - \frac{1}{2} \Delta_- (\psi(R_{k,j+1}^M) N_{k,j+1/2}^M) \right) \right. \\
 (6.15) \quad & \left. + \sum_{k=1}^m \left( N_{k,j-1/2}^+ + \frac{1}{2} \Delta_+ (\psi(R_{k,j-1}^M) N_{k,j-1/2}^M) \right) \right) \\
 = & -\frac{1}{\Delta} \Delta_- \mathcal{H}_{j+1/2}^{ac}.
 \end{aligned}$$

We shall now prove our last result.

**THEOREM 6.2.** *There exist values of  $a_{k,j+1/2}^{\pm}$ , for each  $j, k$  such that if the solutions to (6.15) with the usual initial data have, for  $\Delta$  small enough, (a) sufficiently small oscillation, and (b) the property that eigenvalues  $\lambda_k$  corresponding to linearly degenerate fields stay bounded away from zero, then bounded a.e. limits of the approximate solution satisfy the entropy inequality (1.7).*

*Proof.* It was shown in [22] that it suffices to obtain the inequality

$$(6.16) \quad \int_{\Gamma_{j+1/2}} (dw)^T V_{ww}(w) [\mathcal{H}_{j+1/2}^{ac} - f(w)] \leq 0,$$

for all  $j$ , under the given hypotheses.

An integration by parts and rearrangement of terms gives us an equivalent problem:

$$\begin{aligned}
 \int_{\Gamma_{j+1/2}} (dw)^T (\partial f(w)) \left[ V_w(w) - \frac{1}{2} (V_w(w_{j+1}) + V_w(w_j)) \right] \\
 - \frac{1}{2} \left( (\Delta_+ V_w(w_j)), \sum_{k=1}^m N_{k,j+1/2}^M (1 - \psi(R_{k,j}^M)) \right) \\
 + \frac{1}{2} \left( (\Delta_+ V_w(w_j)), \sum_{k=1}^m N_{k,j+1/2}^M (1 - \psi(R_{k,j+1}^M)) \right) \\
 (6.17) \quad - \frac{1}{2} \left( (\Delta_+ V_w(w_j)), \sum_{k=1}^m (N_{k,j+1/2}^+ - N_{k,j+1/2}^M) \right) \\
 + \frac{1}{2} \left( (\Delta_+ V_w(w_j)), \sum_{k=1}^m (N_{k,j+1/2}^- - N_{k,j+1/2}^M) \right) \leq 0.
 \end{aligned}$$

Let  $k'$  denote indices corresponding to genuinely nonlinear fields, and  $k''$  denote the remaining fields (which are all linearly degenerate).

Then

$$\begin{aligned}
 (a) \quad N_{k',j+1/2}^M - N_{k',j+1/2}^{\pm} &= \mp a_{k',j+1/2}^{\pm} (\Delta_+^{(j,k')} \lambda_{\pm}) \frac{(\Delta_+^{(j,k')} w, \Delta_+^{(j,k')} V_w)}{(N_{k',j+1/2}^{\pm}, \Delta_+^{(j,k')} V_w)} N_{k',j+1/2}^{\pm}, \\
 (6.18) \quad (b) \quad N_{k'',j+1/2}^M - N_{k'',j+1/2}^{\pm} &= a_{k'',j+1/2}^{\pm} (\Delta_+^{(j,k'')} w, \Delta_+^{(j,k'')} V_w)^{1/2} N_{k'',j+1/2}^{\pm}.
 \end{aligned}$$

(Notice that  $a_{k',j+1/2}^{\pm}$  are not dimensionless quantities—they have dimension inverse to that of  $(V)^{1/2}$ . The  $a_{k'',j+1/2}^{\pm}$  are normalized to be dimensionless.)



Using the argument (and notation) in [22, § 3], it is easy to show that

$$\begin{aligned}
 & \frac{1}{2} \left( \Delta_+ V_w(w_j), \int_{\Gamma_{j+1/2}^k} (\partial f(w)) dw \right) \\
 (6.19) \quad &= \frac{1}{2} s_k \int_0^{s_k} dt (r_k(w(t))^T V_{ww} | \lambda_k(w(t)) | r_k(w(t))) \\
 &+ \frac{1}{2} \sum_{\nu=1}^m \int_0^{s_\nu} ds \int_0^{s_k} dt (V_{ww}(w(s)) r_\nu(w(s)) - V_{ww}(w(t)) r_\nu(w(t)))^T \\
 & \quad | \lambda_k(w(t)) | r_\nu(w(t)).
 \end{aligned}$$

Let  $\varepsilon_j = |\Delta_+ w_j|$  be sufficiently small. Then our assumption about the fields, the argument in [22], together with (6.20) and Restriction 6.1 guarantees that the second and third terms in (6.17) are nonpositive. Moreover, regarding the last two terms, we have, using (6.18), and letting  $C > 0$  be a universal constant:

$$\begin{aligned}
 & (\Delta_+ V_w(w_j))^T \left( \sum_{k=1}^m (N_{k,j+1/2}^+ - N_{k,j-1/2}^M) + \sum_{k=1}^m (N_{k,j+1/2}^M - N_{k,j+1/2}^-) \right) \\
 (6.20) \quad & \geq (1 - C\varepsilon_j) \left( \sum_{k'} (a_{k',j+1/2}^+ (\Delta_+^{(j,k')} \lambda_+) + a_{k',j+1/2}^- (\Delta_+^{(j,k')} \lambda_-)) (\Delta_+^{(j,k')} w, \Delta_+^{(j,k')} V_w) \right) \\
 & + (1 - C\varepsilon_j) \left( \sum_{k''} (-a_{k'',j+1/2}^+ (\Delta_+^{(j,k'')} V_w, N_{k'',j+1/2}^+) \right. \\
 & \quad \left. + a_{k'',j+1/2}^- (\Delta_+^{(j,k'')} V_w, N_{k'',j+1/2}^-)) (\Delta_+^{(j,k'')} w, \Delta_+^{(j,k'')} V_w)^{1/2} \right),
 \end{aligned}$$

recalling again, that the  $k'$  are the genuinely nonlinear, and the  $k''$  are the linearly degenerate, fields.

Here we have taken

$$\begin{aligned}
 (6.21) \quad (a) \quad & a_{k',j+1/2}^\pm (\Delta_+^{(j,k')} \lambda_\pm) \geq 0, \\
 (b) \quad & a_{k'',j+1/2}^\pm \leq 0.
 \end{aligned}$$

This means that we have added only expansion, not compression, to our scheme. The freedom to add compression at shocks or contacts is lacking, only because of technical points in our proof.

The integral in (6.17) can be rewritten as

$$\begin{aligned}
 (6.22) \quad & -\frac{1}{2} \int_{\Gamma_{j+1/2}} \left( \int_{\Gamma_{j+1/2,w,U}} ((dv)^T V_{ww}(v)) \partial f(w) \right) dw \\
 & + \frac{1}{2} \int_{\Gamma_{j+1/2}} \left( \int_{\Gamma_{j+1/2,w,L}} ((dv)^T V_{ww}(v)) \partial f(w) \right) dw.
 \end{aligned}$$

Here,  $\Gamma_{j+1/2,w,U}$  is that part of  $\Gamma_{j+1/2}$  which starts at  $w_j$  and ends at  $w_{j+1}$ ,  $\Gamma_{j+1/2,w,L}$  starts at  $w$ , and ends at  $w$ . Thus  $\Gamma_{j+1/2} = \Gamma_{j+1/2,w,L} \cup \Gamma_{j+1/2,w,U}$ . Using the arguments of [22, § 5], we can show that the sum of all the integrals along subsegments is

$$O \left( \sum_{k=1}^m |s_k|^2 \left| \int_0^{s_k} | \lambda_k(w(s)) | ds \right| \right),$$

except, perhaps for a term

$$(6.23) \quad \frac{1}{2} \sum_k \int_0^{s_k} ds \left( \int_0^s dt - \int_s^{s_k} dt \right) ((r_k(w(t)))^T V_{ww}(w(t)) r_k(w(t)) \lambda_k(w(t))).$$

Let  $t = s_k - t'$  in the second integral. Using our hypotheses about the fields, easily gives us an upper bound

$$(6.24) \quad \sum_{k=1}^M D_k (s_k)^2 \left| \int_0^{s_k} |\lambda_k(w(s))| ds \right|$$

for this integral, where the  $D_k$  are certain universal positive constants.

Comparing (6.23), (6.24) to the right side of (6.20) shows that we may choose each constant  $a_{k,j+1/2}^\pm$  subject to Restriction 6.1, and sufficiently large in magnitude, (if  $\varepsilon_j$  is sufficiently small) so that (6.17) is valid.

**7. Results of numerical experiments.** Since our preceding theory concerned only semidiscrete, continuous in time, approximations, we begin by describing the time

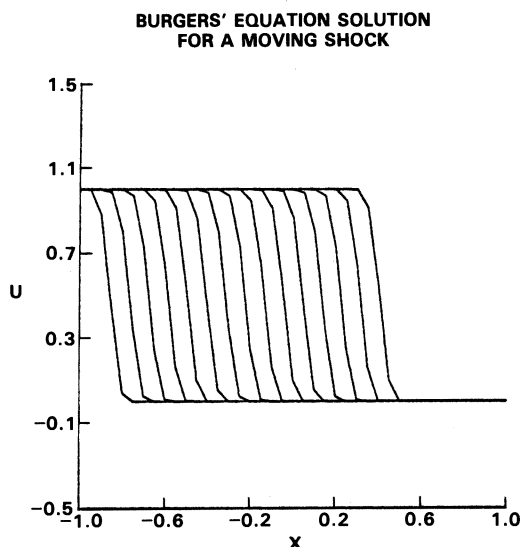


FIG. 1.

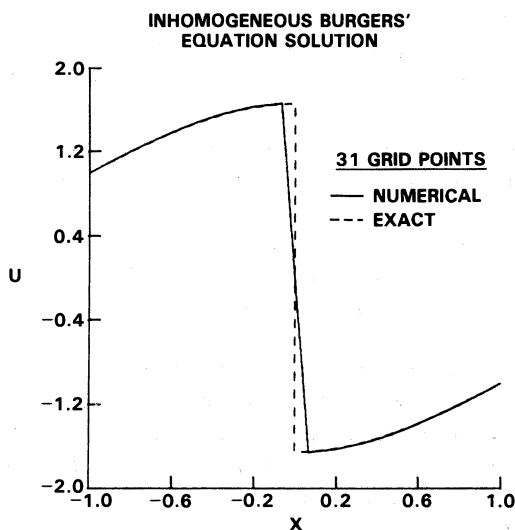


FIG. 2.

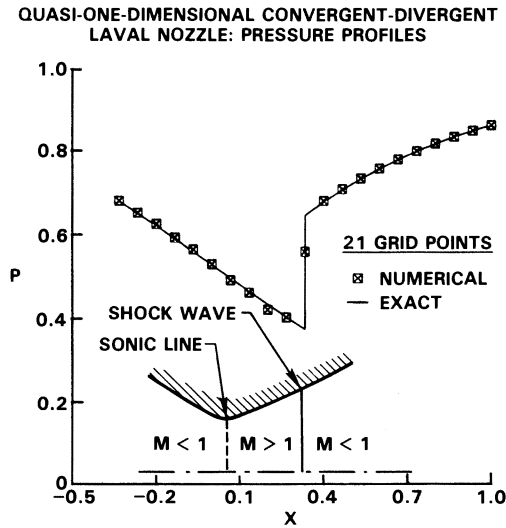


FIG. 3.

discretizations used for all calculations presented here. That was a two-step time differencing found in Richtmyer and Morton [27]. In one space dimension it can be written as

$$\begin{aligned}
 (7.1) \quad (a) \quad w^{n+1/2} &= w^n - \frac{\Delta t}{2} \{\text{first order approximation to } f_x^n\}, \\
 (b) \quad w^{n+1} &= w^n - \Delta t \{\text{first order approximation to } f_x^{n+1/2}\} \\
 &\quad - \frac{\Delta t \Delta x}{2} \{\text{“upwind” approximation to } \pm f_{xx}^n \text{ using flux limiters}\}.
 \end{aligned}$$

The resulting explicit algorithm is not in general, truly variation diminishing, but still works well, even in multidimensional calculations.

We begin by computing a solution to inviscid Burger’s equation which is a shock moving with speed  $\frac{1}{2}$ . The results are shown in Fig. 1. Next we compute a solution to the steady state Burger’s equation having an inhomogeneous right-hand side, and compare the result with the known exact solution—Fig. 2.

Both of these scalar experiments were done using the flux limiter in (3.9), with  $k=1$ , and with the high resolution scheme based on Engquist–Osher first order differencing.

Next we compute solutions to Euler’s equations of compressible gas dynamics, using the schemes constructed in the previous section. See [2], [3], for the precise equations, and algorithm used here. Figure 3 shows the results of quasi-one-dimensional Laval nozzle flow. Reference [2] has the equation with precise source term.

We next consider the case of an oblique shock reflecting from a flat plate, in two space dimensions. Figure 4a shows a cartesian grid with  $61 \times 21$  points. Figure 4b shows a first order accurate solution, 4c was obtained using a non-TVD second order extension, and 4d shows the contours obtained using the TVD second order accurate scheme. Figure 4b has highly smeared incident, and reflected shocks, 4c has thick profiles because of undershoots and overshoots near the shock, and 4d has relatively sharp, nonoscillatory profiles.

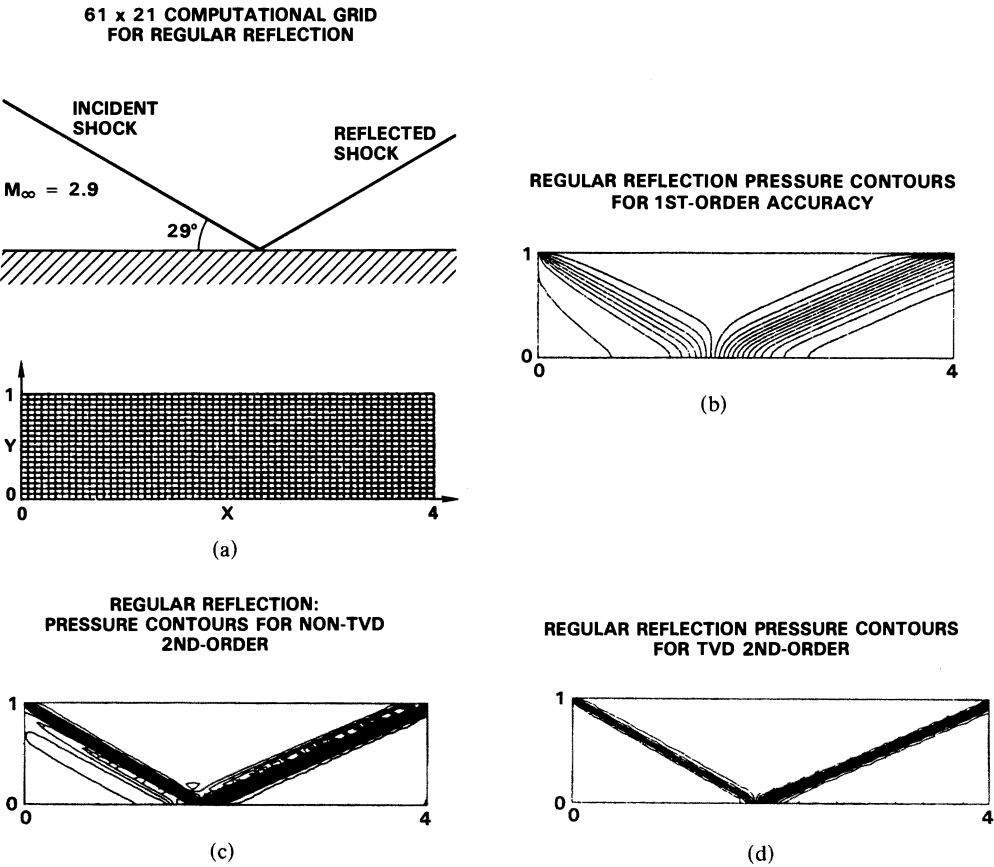


FIG. 4.

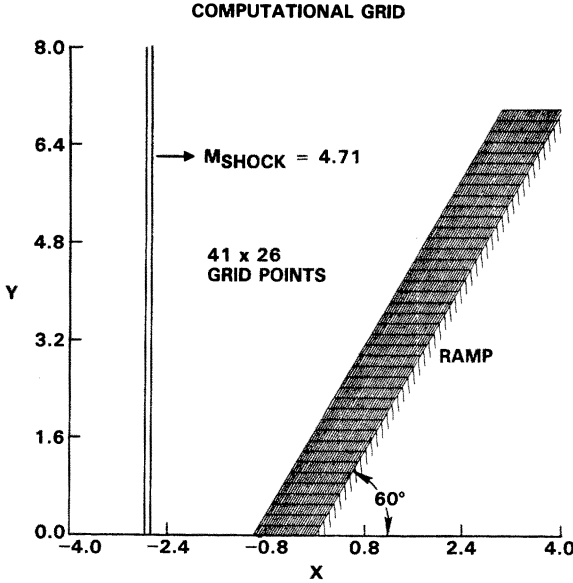


FIG. 5.

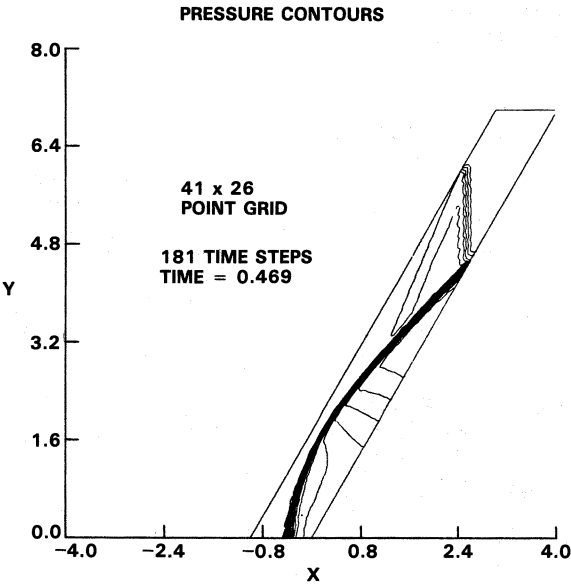


FIG. 6.

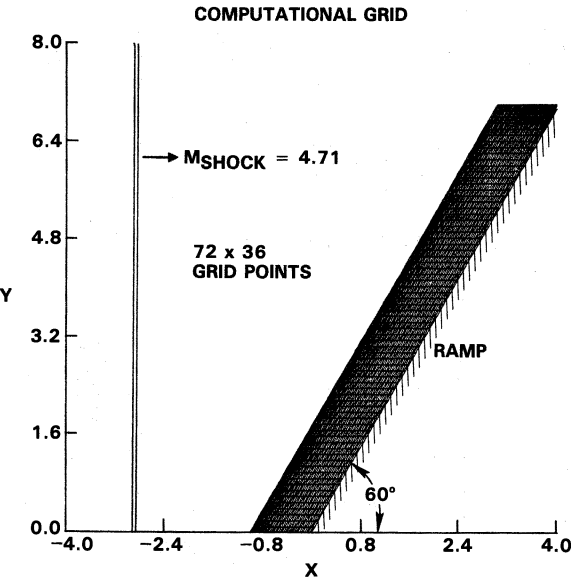


FIG. 7.

In Fig. 5, we present the computational grid to be used for a blast wave problem solved as an unsteady problem. In Fig. 6, we give the computed pressure contour using the TVD, second order scheme, at time  $T = 0.469$ . In Fig. 7, we show a fine computational mesh, and in Fig. 8, we present the pressure contours using the TVD second order scheme at time  $T = 0.459$ .

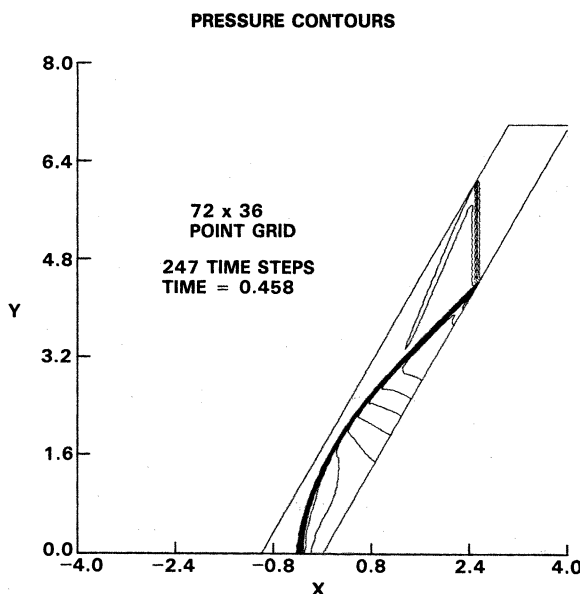


FIG. 8.

**Acknowledgment.** The authors would like to thank Peter Sweby and Bram van Leer for some very valuable conversations concerning flux limiters.

## REFERENCES

- [1] L. ABRAHAMSSON AND S. OSHER, *Monotone difference schemes for singular perturbation problems*, this Journal, 19 (1982), pp. 979–991.
- [2] S. CHAKRAVARTHY AND S. OSHER, *Numerical experiments with the Osher upwind scheme for the Euler equations*, AIAA paper # 82-0975, AIAA J., to appear.
- [3] ———, *High resolution applications of the Osher upwind scheme for the Euler equations*, AIAA paper # 83-1943, Danvers, MA, 1983.
- [4] P. COLELLA AND P. R. WOODWARD, *The piecewise-parabolic method (PPM) for gas-dynamical simulations*, LBL report # 14661, (July 1982).
- [5] R. J. DiPERNA, *Convergence of approximate solutions to conservation laws*, Arch. Rat. Mech. Anal., 82 (1983), pp. 27–70.
- [6] ———, *Decay of solutions of hyperbolic systems of conservation laws with a convex extension*, Arch. Rat. Mech. Anal., 64 (1977), pp. 1–92.
- [7] J. W. EDWARDS, R. M. BENNETT, W. WHITLOW, JR. AND D. A. SEIDEL, *Time marching transonic flutter solutions including angle-of-attack effects*, AIAA paper # 82-0685, New Orleans, LA, 1982.
- [8] B. ENGQUIST AND S. OSHER, *Stable and entropy condition satisfying approximations for transonic flow calculations*, Math. Comp., 34 (1980), pp. 45–75.
- [9] ———, *One sided difference approximations for nonlinear conservation laws*, Math. Comp., 36 (1981), pp. 321–351.
- [10] S. K. GODUNOV, *A finite difference method for the numerical computation of discontinuous solutions of the equations of fluid dynamics*, Mat. Sb., 47(1959), pp. 271–290.
- [11] J. B. GOODMAN AND R. J. LEVEQUE, *On the accuracy of stable schemes for two dimensional conservation laws*, preprint.
- [12] P. M. GOORJIAN AND R. VAN BUSKIRK, *Implicit calculations of transonic flow using monotone methods*, AIAA paper # 81-0331, St. Louis, MO, 1981.
- [13] P. M. GOORJIAN, M. MEAGHER, AND R. VAN BUSKIRK, *Monotone implicit algorithms for the small disturbance and full potential equation applied to transonic flow*, AIAA paper # 83-0371, Reno, NV, 1983.

- [14] B. GUSTAFSSON, *The convergence rate for difference approximations to mixed initial boundary value problems*, Math. Comp., 29 (1975), pp. 396–406.
- [15] A. HARTEN, *The artificial compression method for computing shocks and discontinuities. I. Single conservation laws*, Comm. Pure Appl. Math., 30 (1977), pp. 611–638.
- [16] ———, *High resolution schemes for hyperbolic conservation laws*, J. Comp. Physics, to appear.
- [17] A. HARTEN AND P. D. LAX, *A random choice finite-difference scheme for hyperbolic conservation laws*, this Journal, 18 (1981), pp. 289–315.
- [18] S. N. KRZKOV, *First order quasi-linear equations in several independent variables*, Math. USSR Sb., 10 (1970), pp. 217–243.
- [19] P. D. LAX, *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*, CBMS Regional Conference Series in Applied Mathematics 11, Society for Industrial and Applied Mathematics, Philadelphia, 1973.
- [20] A. MAJDA AND S. OSHER, *Numerical viscosity and the entropy condition*, Comm. Pure Appl. Math., 32 (1979), pp. 797–838.
- [21] E. M. MURMAN, *Analysis of embedded shock waves calculated by relaxation methods*, AIAA J., 12 (1974), pp. 626–633.
- [22] S. OSHER, *Riemann solvers, the entropy condition, and difference approximations*, this Journal, 21 (1984), pp. 217–235.
- [23] ———, *Shock modelling in transonic and supersonic flow*, Recent Advances in Numerical Methods in Fluids, Vol. 4, Advances in Computational Transonics, W. G. Habashi, ed., Pineridge Press, Swansea, UK, 1984.
- [24] ———, *Numerical solution of singular perturbation problems and hyperbolic systems of conservation laws*, Mathematical Studies 47, S. Axelsson, L. S. Frank, and A. van der Sluis, eds., North-Holland, Amsterdam, 1981, pp. 179–205.
- [25] S. OSHER AND S. CHAKRAVARTHY, *Upwind schemes and boundary conditions with applications to Euler equations in general geometries*, J. Comp. Phys., to appear.
- [26] S. OSHER AND F. SOLOMON, *Upwind schemes for hyperbolic systems of conservation laws*, Math. Comp., 38 (1982), pp. 339–377.
- [27] R. D. RICHTMYER AND K. W. MORTON, *Difference Methods for Initial-Value Problems*, 2nd ed., Interscience, New York, 1967.
- [28] P. L. ROE, *Approximate Riemann solvers, parameter vectors, and difference schemes*, J. Comp. Phys., 43 (1981), pp. 357–372.
- [29] R. SANDERS, *On convergence of monotone finite difference schemes with variable spatial differencing*, Math. Comp., 40 (1983), pp. 91–106.
- [30] P. K. SWEBY, *High resolution schemes using flux limiters for hyperbolic conservation laws*, this Journal, this issue, pp. 995–1011.
- [31] P. K. SWEBY AND M. J. BAINES, *Convergence of Roe's scheme for the general nonlinear scalar wave equation*, Numerical Analysis Report 8/81, Univ. Reading, 1981.
- [32] B. VAN LEER, *Towards the ultimate conservative difference scheme, II. Monotonicity and conservation combined in a second-order scheme*, J. Comp. Phys., 14 (1974), pp. 361–370.
- [33] ———, *Towards the ultimate conservative difference scheme, V. A second-order sequel to Godunov's method*, J. Comp. Phys., 32 (1979), pp. 101–136.
- [34] ———, *On the relation between the upwind differencing schemes of Godunov, Engquist–Osher, and Roe*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 1–20.