DISCRETE SHOCKS AND GENUINE NONLINEARITY

M. S. Mock

1. INTRODUCTION

We consider the suitability of numerical methods for problems of the form

$$(1.1) \qquad u_t + f_x = 0, \qquad x \in \mathbb{R}, t > 0, \qquad u(\cdot,0) \text{ given},$$

where $u = (u^{(1)}, u^{(2)}, \ldots, u^{(m)})^t$, $f = (f^{(1)}, f^{(2)}, \ldots, f^{(m)})^t$, and each $f^{(j)}$ is a smooth function of u. We restrict attention to problems for which $u(x,t) \in D \subset \mathbb{R}^m$ for all (x,t), and we assume throughout that (1.1) is strictly hyperbolic in D; *i.e.*, for all $u \in D$, the eigenvalues of $f_u(u)$ are real and distinct.

It is well known that smooth solutions in the large do not exist for (1.1), even for smooth initial data, and that uniqueness fails when jump discontinuities are permitted in weak solutions.

Numerical methods seek to approximate particular weak solutions of (1.1) which satisfy, in addition, an entropy condition, which may take any of several forms [8,9,10,14]. We consider below the application of a one-parameter family of schemes to such problems; our schemes are obtained from a finite element procedure, which is known to work for linear symmetric problems, and contains the Lax-Friedrichs scheme as a special case [13]. However, we use no very special properties of these schemes, and our results are somewhat more general.

We consider necessary conditions on f for discrete approximations of jump discontinuities in the solution of (1.1) (hereafter referred to as "discrete shocks") to have several properties, which appear to be essential for the success of numerical methods, even if one restricts attention to the Riemann problem for (1.1), which is thoroughly discussed in [8,10,11,15,16]. Specifically, we require that discrete shocks exist for all pairs of points in D which can be connected by a shock; *i.e.*, which satisfy the Rankine-Hugoniot relations. We require that the discrete shocks converge in L_1 to shocks as the mesh is refined, that they be stable under suitable perturbation of f, and that they exhibit a tolerance for inexact solution of the discrete equations.

Our main result is that under these conditions, the shock speed between any two points in D (which can be connected by a shock) must not be a characteristic speed at either point. This may be viewed as a statement of genuine nonlinearity of f, and of genuine coupling of the equations. For $m = 1$, it reduces to convexity of f; for $m = 2$, it is almost equivalent to the usual nonlinearity condition [8]; for $m > 2$, it is stronger. Under this stronger condition on f, the entropy conditions [8,9,10,14] are equivalent. Such discrete shocks are shown to exist only when

Received August 23, 1976.

Michigan Math. J. 25 (1978).

the entropy condition is satisfied, using primarily the requirement of stability under perturbation of f. Presumably, in practice one computes only solutions which are stable under perturbation; thus our results provide a mechanism by which discrete schemes impose an entropy selection rule on discontinuities. The requirement of discrete shock "tails" in L_1 is justified as necessary for their formation in a finite time during a computation. It is used in an apparently essential way in the proof of the main theorem. However, the requirement of tolerance for inexact discrete solution is used only to analyze a pathological case, and is not even needed for the case $m = 1$.

For $m = 1$, the existence of such discrete shocks is implied by the results of Jennings [6]. For this case, much more is known; for example, it is shown in [4] that whenever the solutions of a monotone difference scheme converge, the limit satisfies an entropy condition. For $m = 2$, we show existence using ideas and techniques similar to those of [1,3,12,2].

2. NONLINEARITY AND ENTROPY CONDITIONS

We call two distinct points $u, v \in D$ *connected* if they can be connected by a shock; *i.e.,* if for some scalar $s(u,v)$, the shock speed, the Rankine-Hugoniot relation holds:

$$(f(u) - f(v)) = s(u,v)(u - v). \tag{2.1}$$

Let $S(u)$ denote the set of points in D connected to a given point u. We assume that

$$S(u) = \bigcup_{k=1}^{m} S_k(u), \qquad S_k(u) = S_k^+(u) \cup S_k^-(u), \tag{2.2}$$

where each $S_k^{\pm}(u)$ is a smooth one-dimensional manifold with u as one endpoint and no other endpoints within D. Primes are used to denote differentiation along such manifolds, in the direction initially away from u. Let $\lambda_k(u)$, $r_k(u)$ denote the eigenvalues (arranged in increasing order) and right-eigenvectors, respectively, of $f_u(u)$; then $S_k^+(u)$ ($S_k^-(u)$) is, as usual, the manifold approaching u in the direction of $r_k(u)$, such that $\lambda_k' \geq 0$ ($\lambda_k' \leq 0$) near u. Finally, for $u \in D$, $v \in S(u)$, let $S(u,v)$ denote that segment of $S(u)$ between u and v; *i.e.,* with u and v as its endpoints.

Our condition for genuine nonlinearity may be stated in either of the following forms:

(N1) For all $u \in D$, the manifolds $S_k^{\pm}(u)$, $k = 1,2, \ldots, m$ are disjoint and $s'(u, \cdot) \neq 0$.

(N2) For all $u \in D$, $v \in S(u)$, $s(u,v)$ is not an eigenvalue of $f_u(u)$ (or of $f_u(v)$).

THEOREM 2.1. *The statements* (N1) *and* (N2) *are equivalent.*
Hereafter we call this *condition* N; it is assumed unless explicitly stated otherwise.

Proof. Differentiating (2.1), we have for all $v \in S(u)$,

$$(f_u(v) - s(u,v))v' = s'(u,v)(v - u), \tag{2.3}$$

from which it is immediate that (N1) implies (N2) and (N2) implies $s' \neq 0$.

Suppose that the $S_k(u)$ are not disjoint, that there exists $v \in S_k(u)$, $v \in S_j(u)$, $j \neq k$. Let $S_k(u,v)$, $S_j(u,v)$ denote those segments of $S_k(u)$, $S_j(u)$ between u and v.

Lax [7] showed that as $w \in S_k(u)$ approaches u,

$$\lim_{w\to u} s(u,w) = \lambda_k, \quad \text{and} \quad \lim_{w\to u} s'(u,w) = \frac{1}{2}\lim_{w\to u} \lambda'_k(w)\,. \tag{2.4}$$

Without loss of generality, we assume $\lambda_j(u) < \lambda_k(u)$. Suppose first that

$$\lambda_j(u) < s(u,v) < \lambda_k(u).$$

Then $s' < 0$ along $S_k(u,v)$ towards v; from (2.4) and (N2), it follows that for all $w \in S_k(u,v)$, $\lambda_k(w) < s(u,w)$. In particular, $\lambda_k(v) < s(u,v)$. A similar argument shows $\lambda_j(v) > s(u,v)$. But then strict hyperbolicity fails at some point between u and v. Suppose next that $\lambda_j(u) < \lambda_k(u) < s(u,v)$; then similar arguments give $s(u,v) < \lambda_j(v) < \lambda_k(v)$. But then along $S_k(u,v)$ towards v, $s(u,\cdot)$ increases from $\lambda_k(u)$ to $s(u,v)$, while λ_j increases from $\lambda_j(u)$ to $\lambda_j(v)$. Thus (N2) is violated at some point. A similar argument holds for the case $s(u,v) < \lambda_j(u)$, and completes the proof.

COROLLARY. *Suppose condition* N *holds,* $u \in D$, *and* $v \in S_k(u)$; *then* $u \in S_k(V)$.

For u near v, this is shown by Lax [8]; for u not near v, the result then follows by continuity and the assumption of disjoint manifolds $S_k(u)$.

We next establish the connection between this nonlinearity condition and that proposed in [8].

THEOREM 2.2. *Suppose condition* N *holds; then for any* $k = 1, \ldots, m$, $r_k \quad \nabla_u\lambda_k$ *cannot change sign within* D.

Remark. If $r_k \cdot \nabla_u\lambda_k$ vanishes on an open ball, $B \subset D$, then so does $s'(u,v)$, for all $u, v \in B$, $v \in S_k(u)$.

Proof. Suppose for some k, $r_k \cdot \nabla_u\lambda_k$ does change sign and does not vanish on any open ball within D; then it vanishes on a manifold q of dimension $m - 1$. Let B denote a small open ball in D, with center in q. Then

$$B = R_+ \cup (B \cap q) \cup R_-,$$

where $r_k(u) \cdot \nabla_u\lambda_k(u) > 0$ (<0) for $u \in R_+$ (R_-); $R_\pm$ are not empty. If $r_k(u)$ is tangent to q for all $u \in B \cap q$, then for any $u \in B \cap q$, a segment of $S_k(u)$ lies within $B \cap q$ and $s'(u,\cdot)$ vanishes on this segment [8]. Thus we may assume the existence of a point $w \in B \cap q$ such that $r_k(w)$ crosses q at w. Thus the two branches of $S_k(w)$ are on opposite sides of q, as shown in Figure 1. We

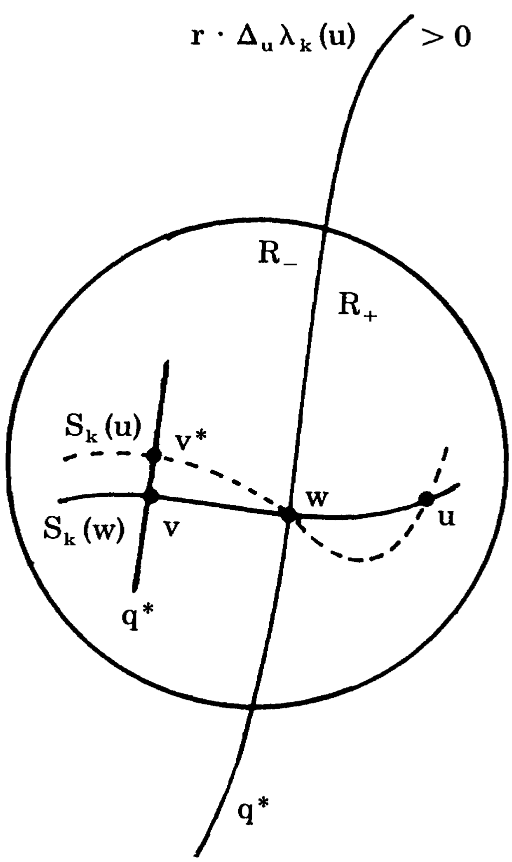

Figure 1: Proof of Theorem 2.2

first choose a point $v \in R_- \cap S_k(w)$. Then $w \in S_k^-(v)$, so $s(v,w) < \lambda_k(v)$. Let q^* be the hyperplane through v perpendicular to $r_k(v)$. We next choose a point $u \in R_+ \cap S_k(w)$. Then $w \in S_k^-(u)$, and $S_k^-(u)$ intersects q^* at a point v^* close to v. We choose u close to w so that v^* is sufficiently close to v that $s(u,v^*) < \lambda_k(v^*)$. But by condition N, since $s(u,\cdot)$ decreases along $S_k^-(u)$ between u and w, it decreases between u and v^*, and so $s(u,v^*) > \lambda_k(v^*)$, which is our desired contradiction.

THEOREM 2.3. *Suppose* $m = 2$, *and that for all* $u \in D$, *the following hold:*

$$r_k(u) \cdot \nabla_u \lambda_k(u) > 0, \qquad k = 1,2; \tag{2.5}$$

$$\frac{\partial f^{(1)}}{\partial u^{(2)}}(u) > 0, \qquad \frac{\partial f^{(2)}}{\partial u^{(1)}}(u) > 0; \tag{2.6}$$

$$(v^{(1)'} v^{(2)'}) > 0 \quad (< 0) \qquad \textit{for all}\; v \in S_2(u) \quad (S_1(u)); \tag{2.7}$$

then condition N *holds.*

Hereafter we use f_{jk} for $\partial f^{(j)}/\partial u^{(k)}$; (2.6) is a statement implying hyperbolicity and genuine coupling of the equations, adopted by several authors [1,2,10]. The hypothesis (2.7) is always true for v close to u; it may be regarded as a statement of global solvability of the Rankine-Hugoniot relations (2.1) if one component of v is given. Some such additional hypotheses are clearly necessary; it is otherwise possible for (2.5) to hold and condition N to fail, as in the case where the equations are uncoupled, f_1 dependent only on $u^{(1)}$ and f_2 dependent on $u^{(2)}$.

The proof of this theorem requires the following lemma, which is given for v close to u by Lax [8] and with different hypotheses by Liu [10].

LEMMA 2.4. *For all* $v \in S_2(u)$ ($v \in S_1(u)$), $s'(u,v) > 0$ *if and only if* $s < \lambda_2(v)$ ($s < \lambda_1(v)$).

Proof. It follows from (2.6) that for all $v \in D$,

$$\lambda_1(v) < f_{11}(v), \quad f_{22}(v) < \lambda_2(v). \tag{2.8}$$

Suppose $v \in S_2(u)$, $v^{(1)} > u^{(1)}$ and $v^{(2)} > u^{(2)}$ (the proof for $v^{(1)} < u^{(1)}$, $v^{(2)} < u^{(2)}$ is similar). We write (2.3) as

$$\begin{aligned} (f_{11} - s)v^{(1)\prime} + f_{12}v^{(2)\prime} &= s'(v^{(1)} - u^{(1)}) \\ f_{21}v^{(1)\prime} + (f_{22} - s)v^{(2)\prime} &= s'(v^{(2)} - u^{(2)}) \end{aligned} \tag{2.9}$$

For $s' > 0$, in view of (2.8), it suffices to consider the case $s > f_{11}$ and f_{22}. Then (2.9) becomes

$$(s - f_{11})v^{(1)\prime} < f_{12}v^{(2)\prime} \tag{2.10}$$

$$(s - f_{22})v^{(2)\prime} < f_{21}v^{(1)\prime}; \tag{2.11}$$

multiplying (2.10) and (2.11), we easily obtain $\lambda_1 < s < \lambda_2$.

For $s' < 0$, the inequalities are reversed in (2.10) and (2.11). Then $s > f_{11}$ or f_{22}, and so $s > \lambda_1$. We can again multiply the inequalities; solving the resulting quadratic equation for s gives $s > \lambda_2$.

Finally, suppose $v \in S_1(u)$, $v^{(1)} > u^{(1)}$, $v^{(2)} < u^{(2)}$; for $s' > 0$ we obtain

$$f_{12}v^{(2)\prime} > (s - f_{11})v^{(1)\prime} \tag{2.12}$$

$$f_{21}v^{(1)\prime} < (s - f_{22})v^{(2)\prime} \tag{2.13}$$

From $v^{(2)\prime} < 0$, we find $s < f_{11} < \lambda_2$. Then we can multiply (2.13) by the negative of (2.12) and obtain $s < \lambda_1$ as above.

For the case $s' < 0$, we wish to show $s < \lambda_1$; as above, it suffices to consider the case $s < f_{11}$ and f_{22}. The inequalities (2.12) and (2.13) are both preserved in this case, and a similar argument applies.

Proof of Theorem 2.3. Consider the case $v \in S_2^+(u)$, so that λ' and s' are both positive for v sufficiently close to u. Let v_0 be the first point on $S_2^+(u)$ such that $s'(u,v_0) = 0$. From hypothesis (2.7), it follows that $s(u, v_0) = \lambda_2(v_0)$, not $\lambda_1(v_0)$; then for all $v \in S(u,v_0)$, $s'(u,v) > 0$ and $s(u,v) < \lambda_2(v)$. Thus $\lambda_2 - s$ is a decreasing function on $S(u,v_0)$ sufficiently near v_0. But if $s = \lambda_2$ at v_0, $v' = r_2$ and from (2.5), $\lambda_2'(v_0) > 0$. Since $s'(v_0) = 0$, we obtain $\lambda_2 - s$ an increasing function on $S(u,v_0)$ near v_0. Thus no such point v_0 exists. A similar argument applies to the other cases and completes the proof.

We next consider entropy conditions. Let u_ℓ, u_r be two connected points in D. Among others, the following conditions have been proposed for the admissibility of a discontinuity in the solution of (1.1), with u_ℓ on the left and u_r on the right sides.

(E1) For some k, $\lambda_k(u_\ell) > s(u_\ell, u_r) > \lambda_k(u_r)$, $\lambda_{k-1}(u_\ell) < s(u_\ell, u_r) < \lambda_{k+1}(u_r)$;

(E2) For all $u \in S(u_r, u_\ell)$, $s(u_r, u) < s(u_r, u_\ell)$;

(E3) For all $u \in S(u_\ell, u_r)$, $s(u_\ell, u) > s(u_r, u_\ell)$;

(E4) $s[U] - [F] < 0$, where $U(u)$ is a strictly convex function of u, F is a function of u such that differentiable solutions of (1.1) satisfy $U_t + F_x = 0$, and $[\ ]$ denotes the variation of a quantity across the discontinuity.

Condition (E1) is the classical entropy condition of Lax [8]; conditions (E2) and (E3) are the Oleinik condition for $m = 1$ [14]; they have been applied to systems by Liu [10,11]. Condition (E4) has been adopted by Hopf [5], Lax [9], and Krushkov [7].

THEOREM 2.5. *Suppose condition* N *holds; then conditions* (E1)-(E4) *are all equivalent.*

Hereafter we call this *condition* E.

Proof. The equivalence of (E1),(E2) and (E3) follows from the following two lemmas:

LEMMA 2.6. *Suppose condition* N *holds; then for all* $u \in D$ *and* $v \in S_k(u)$, $s(u,v)$ *lies between* $\lambda_k(u)$ *and* $\lambda_k(v)$.

LEMMA 2.7. *Suppose condition* N *holds; then for all* $u \in D$ *and* $v \in S_k(u)$,

$$\lambda_{k-1}(u) < s(u,v) < \lambda_{k+1}(u). \tag{2.14}$$

The equivalence of (E1) and (E4) for weak shocks is proved in [9]. To extend this result to strong shocks, it suffices to show that $s[U] - [F]$ does not change sign as w moves along some branch of $S(v)$. But in [8] it is shown that

$$(s[U] - [F])' = s'(v,w)(U(v) - U(w) - \frac{\partial U(w)}{\partial u} \cdot (v - w)),$$

the first factor of which is nonvanishing by condition N and the second nonvanishing by the strict convexity of U.

Proof of Lemma 2.6. Suppose $v \in S_k(u)$, and $s(u,v) > \lambda_k(u)$. Then $s(u,\cdot)$ increases along $S(u,v)$; in particular, it increases on $S(u,v)$ near u. From (2.4), it follows that for $w \in S(u,v)$ sufficiently close to u, $\lambda_k(w) > s(u,w)$. But then this holds for all $w \in S(u,v)$, in particular, for $w = v$.

Proof of Lemma 2.7. Suppose $v \in S_k(u)$ and $\lambda_{k+1}(u) < s(u,v)$. By Lemma 2.6, $\lambda_k(v) > s(u,v)$. Furthermore, $u \in S_k(v)$. As w moves along $S(v,u)$ from v towards u, $s(u,w)$ decreases from $\lambda_k(v)$ to $s(u,v)$, while $\lambda_{k+1}(w)$ decreases from $\lambda_{k+1}(v) > \lambda_k(v)$ to $\lambda_{k+1}(u) < s(u,v)$. Thus for some $w \in S(v,u)$, $s(v,w) = \lambda_{k+1}(w)$, contradicting condition N. The case $\lambda_{k-1}(u) > s(u,v)$ is entirely analogous.

COROLLARY. *Suppose condition* N *holds; then a discontinuity between any two connected points in* D *satisfies condition* E *if and only if the points are properly oriented.*

Proof. Let $u, v \in D$, $v \in S_k(u)$. By Lemma 2.6 and (N2), it follows that $\lambda_k(u) \neq \lambda_k(v)$. Suppose $\lambda_k(u) > \lambda_k(v)$; then condition E is satisfied if and only if u is on the left and v on the right of the discontinuity.

3. DISCRETIZATION SCHEMES

We next consider the application of discretization schemes to very special problems of the form (1.1), *i.e.*, for two connected points u_ℓ, $u_r \in D$, the problem

$$u_t + f_x - s(u_r, u_\ell)u_x = 0; \tag{3.1}$$

$$u(x,0) = \begin{cases} u_\ell, & x < 0 \\ u_r, & x \geq 0 \end{cases}. \tag{3.2}$$

Thus if f is genuinely nonlinear and u_ℓ, u_r are oriented so that an entropy condition is satisfied, the solution to (3.1), (3.2) is a single stationary shock at $x = 0$; *i.e.*, $u(x,t) = u(x,0)$.

We introduce a class of discrete schemes for (3.1) as follows: let h denote the mesh spacing in x, *i.e.*, $x_j = jh$, $j = 0, \pm1, \pm2 \ldots$ Let δt be the time step, $t_n = n\delta t$, $n = 0,1,2, \ldots$, and set $\beta = \delta t/h$. Our approximation to $u(jh, n\delta t)$ is denoted as usual by u_j^n. Let X_h denote the space of piecewise linear functions (in x) over this grid. For each fixed n, $u^n \in X_h$; *i.e.*, u^n is the piecewise linear interpolation of the discrete values u_j^n. Our discrete approximation to (3.1) is then given by

$$u^{n+1} - \alpha h^2 u_{xx}^{n+1} - u^n + \gamma h^2 u_{xx}^n + h\beta(f(u^n))_x - h\beta s(u_r, u_\ell)u_x^n \perp X_h; \tag{3.3}$$
$$u^{n+1}(+\infty) = u_r, \quad u^{n+1}(-\infty) = u_\ell.$$

In (3.3), α, γ are parameters; presumably $\alpha = 1/6$, so that (3.3) gives u_j^{n+1} explicitly, and $\gamma < \alpha$, β sufficiently small, so that the scheme would be stable if f were linear and symmetric [13]. For $\gamma = -1/3$, this scheme is of the Lax-Friedrichs type; for $m = 1$, $-1/3 < \gamma < 1/6$, sufficiently small β, this scheme is monotone in the sense of Jennings [6] and Harten *et al.* [4], and their results apply.

We are interested in solutions of (3.3) which resemble stationary shock waves; *i.e.*, for fixed h, β, functions $\phi \in X_h$ satisfying

$$ah^2\phi_{xx} + h\beta s(u_r, u_\ell)\phi_x - h\beta(f(\phi))_x \perp X_h, \qquad \phi(+\infty) = u_r, \quad \phi(-\infty) = u_\ell, \tag{3.4}$$

where $a = \alpha - \gamma$ is positive and hereafter fixed; after an integration by parts, this becomes

$$a\frac{h}{\beta}\phi_x + s(\phi - \phi_r) - (f(\phi) - f_r) \perp X_h', \tag{3.5}$$

where we use $s = s(u_r,u_\ell)$ and $f_r = f(u_r)$, where no ambiguity arises; X_h' is the space of derivatives of the elements of X_h; *i.e.,* the space of piecewise constant functions in (x_j,x_{j+1}) for all j. Denoting $\phi(x_j)$ by ϕ_j, the explicit form of (3.5) is

$$\left(a\frac{h}{\beta}\right)\left(\frac{\phi_{j+1}-\phi_j}{h}\right)+s\left(\frac{\phi_{j+1}+\phi_j}{2}-u_r\right) - \int_0^1 f((1-\xi)\phi_j+\xi\phi_{j+1})d\xi + f_r = 0. \tag{3.6}$$

We note that for fixed β, (3.6) is independent of h; thus we may describe solutions of (3.4) as of the form $\phi(x/h,\beta)$. From (3.6) it follows for sufficiently small β that such solutions are uniquely specified by the values of β and $\phi(0,\beta)$.

We will be interested in solutions of (3.4) satisfying several additional requirements, related to their use in more general computations. Let ψ denote the shock at $x = 0$; *i.e.,*

$$\psi(x) = \begin{cases} u_\ell, & x < 0 \\ u_r, & x \geq 0 \end{cases}.$$

It is essential that our discrete shocks satisfy $\phi - \psi \in L_1(-\infty,\infty)$. This is because discrete schemes like (3.3) conserve $\int u\,dx$, and so cannot form shock "tails" for which this quantity is infinite. We shall require a slightly stronger condition; namely, the existence of ϕ satisfying (3.4) and

$$\|\phi(\cdot,\beta)-\psi\|_{L_1} \leq c_0 h/\beta, \tag{3.7}$$

for some c_0 and sufficiently small β. Obviously, such ϕ converge to ψ in L_1 as $h \to 0$ with β and $\phi(0,\beta)$ fixed; this is the natural limit of discretization schemes of the form (3.3). Such ϕ describe shocks located at $x = 0$; the dependence of (3.7) on h is then simply by homogeneity. The "viscous limit" is obtained as $h \to 0$ with h/β fixed. In this limit, ϕ approaches the solution of the continuous form of (3.5) (or (4.1) below). Although this limiting process is irrelevant to actual computations, the existence of such "viscous profiles" is essential to the success of this type of discretization scheme. The dependence of (3.7) on β may be expected from (3.5); however, a bound as $h \to 0$, with h/β fixed, appears to be needed for uniform approach in time of solutions of (3.3) to shocks as $\delta t \to 0$, with $n\delta t$ and h fixed.

Since discrete equations are not exactly satisfied in practice, we require the following stability property:

(S1) For any $\phi \in X_h$, satisfying (3.4) and (3.7), there exists a positive δ such that all $\eta \in X_h$ which satisfy: $\eta(x) \to u_r$ as $x \to +\infty$, $\eta(x) \to u_\ell$ as $x \to -\infty$, and

$$(3.8)\qquad |\phi(0,\beta)-\eta(0)| + \sum_{j=-\infty}^{\infty}{}' \left| \frac{a}{\beta}(\eta_{j+1}-\eta_j) + s\left(\frac{\eta_{j+1}+\eta_j}{2} - u_t\right) - \int_0^1 f((1-\xi)\eta_j + \xi\eta_{j+1})\,d\xi + f_r \right| < \delta,$$

also satisfy $\eta - \psi \in L_1(-\infty,\infty)$. In (3.8) and below, we also use $|\cdot|$ for a norm on m-dimensional vectors and matrices.

Practical discretization schemes must also tolerate appropriate perturbations of the function f. The specific condition we require is as follows:

(S2) Suppose f is replaced by $\tilde{f}$ in (3.4),(3.5),(3.6), and (3.8), with $\tilde{f}$ satisfying

$$(3.9)\qquad |\tilde{f}(v) - f(v)| \le \varepsilon \quad\text{and}\quad |\tilde{f}_u(v) - f_u(v)| \le \varepsilon \qquad \text{for all } v \in D;$$

$$(3.10)\qquad \tilde{f}(u_r) - \tilde{f}(u_\ell) = f(u_r) - f(u_\ell).$$

Then if ε is sufficiently small, (3.7) and (S1) are maintained.

From (3.10), it follows that u_ℓ, u_r remain connected under such perturbations, with unchanged shock speed $s(u_\ell,u_r)$.

4. MAIN THEOREM

THEOREM 4.1. *Suppose that for every pair of connected points (satisfying the Rankine-Hugoniot relation* (2.1)) u_ℓ, $u_r \in D$, *and for all sufficiently small* β, *there exist solutions of* (3.4), *satisfying* (3.7), *uniformly bounded* (*within* D) *with respect to x and β, and such that* (S1) *and* (S2) *are satisfied. Then condition* N *holds in* D. *Furthermore, such solutions are obtained only if* u_ℓ, u_r *are oriented so that condition* E *is satisfied.*

Remark. In the case that differentiable solutions of (1.1) satisfy an additional conservation law $U_t + F_x = 0$, with U convex, Lax [9] shows that weak limits of dissipative discrete schemes necessarily satisfy (E4), and thus condition E.

Proof. For a sequence of values of β approaching zero, we choose a sequence of solutions of (3.4), also denoted by $\phi = \phi(x/h,\beta)$, whose values at $x = 0$ converge to some point $u_0 \in D$.

For $z \in \mathbb{R}$, let θ be the solution of

$$(4.1)\qquad \frac{d\theta}{dz} + s(\theta - u_r) = f(\theta) - f_r, \qquad \theta(0) = u_0,$$

where here and below $s = s(u_\ell,u_r)$. For h/β fixed and $z = x\beta/ah$, (3.6) is a second order accurate difference approximation to (4.1). Thus as $h \to 0$ with h/β fixed,

$$\phi(x/h,\beta) \to \theta(x\beta/ah), \qquad \text{for every } x,$$

and $\phi \to \theta$ uniformly over bounded regions.

Thus by the triangle inequality and (3.7),

$$\|\theta - \psi\|_{L_1} < \infty, \tag{4.2}$$

where the L_1 – norm is with respect to z. It follows that u_0 is not u_r or u_ℓ; otherwise the solution of (4.1) would be θ = constant and (4.2) would fail. From the pointwise convergence of ϕ to θ, it follows that θ is uniformly bounded; then from (4.1), so are its derivatives. It then follows from (4.2) that $\theta(z) \to u_r$ as $z \to +\infty$ and $\theta(z) \to u_\ell$ as $z \to -\infty$.

The essential step is to show that (4.2), or at least (S1), fails if s is an eigenvalue of $f_u(u_r)$ or $f_u(u_\ell)$. Suppose $s = \lambda_k(u_r)$; we expand (4.1) in a sufficiently small neighborhood of u_r. In view of (S2), the requirement of stability under perturbation of f, we may alter the quadratic and higher order terms in such a neighborhood, and it suffices to consider the system

$$\nu_{j,z} = (\lambda_j - s)\nu_j, \qquad j \neq k, \tag{4.3}$$

$$\nu_{k,z} = \lambda_1^2 + \ldots + \lambda_{k-1}^2 - \lambda_k^2, \tag{4.4}$$

where $\nu = T(\theta - u_r)$, and T the similarity transformation diagonalizing $f_u(u_r)$.

For $j > k$, we must have $\nu_j = 0$ in (4.3), or else ν does not approach the origin as $z \to +\infty$. Since u_0 is not u_r, we may exclude the case ν identically zero. For $j < k$, the ν_j decay exponentially as $z \to \infty$. If for all $j < k$,

$$|\nu_j| = o(|\nu_k|^{1/2}) \qquad \text{as } z \to \infty,$$

then $|\nu_{k,z}| = O(\nu_k^2)$, $|\nu_k| \geq O(1/z)$ as $z \to \infty$, and (4.2) fails. However, there are special orbits for which there exists $j < k$ such that

$$|\nu_k| = O(\nu_j^2); \tag{4.5}$$

for such orbits, we show that (S1) fails.

Note that (4.5) can hold for at most one value of j, the largest $j < k$ for which ν_j is not identically zero. The other values of j are unimportant below; they are conveniently viewed as higher order terms in a two-dimensional autonomous system.

Our argument proceeds as follows: In Figure 2, no orbits leave region I, and if ν approaches the origin from region I, then $|\nu_{k,z}| = O(\nu_k^2)$ and (4.2) fails, as above. No orbits approach the origin from regions II or III; in III, $\nu_{k,z} \geq 0$ so all orbits enter II; and in II, $\mu = \nu_j / \nu_k^2$ satisfies $\mu_z = \mu(\lambda_j - s + O(\nu_k))$, and so decreases exponentially. Thus any orbit crossing the positive ν_j axis eventually enters region I.

The existence of an orbit ζ approaching the origin from region IV can be proved by continuity. We show its uniqueness; $\rho = -\nu_k / \nu_j$ satisfies

$$\rho_z = (s - \lambda_j)\rho + \nu_j(\rho^2 - 1 + o(\nu_j)). \tag{4.6}$$

To approach the origin from region IV, ρ must be founded in [0,1] for all z.

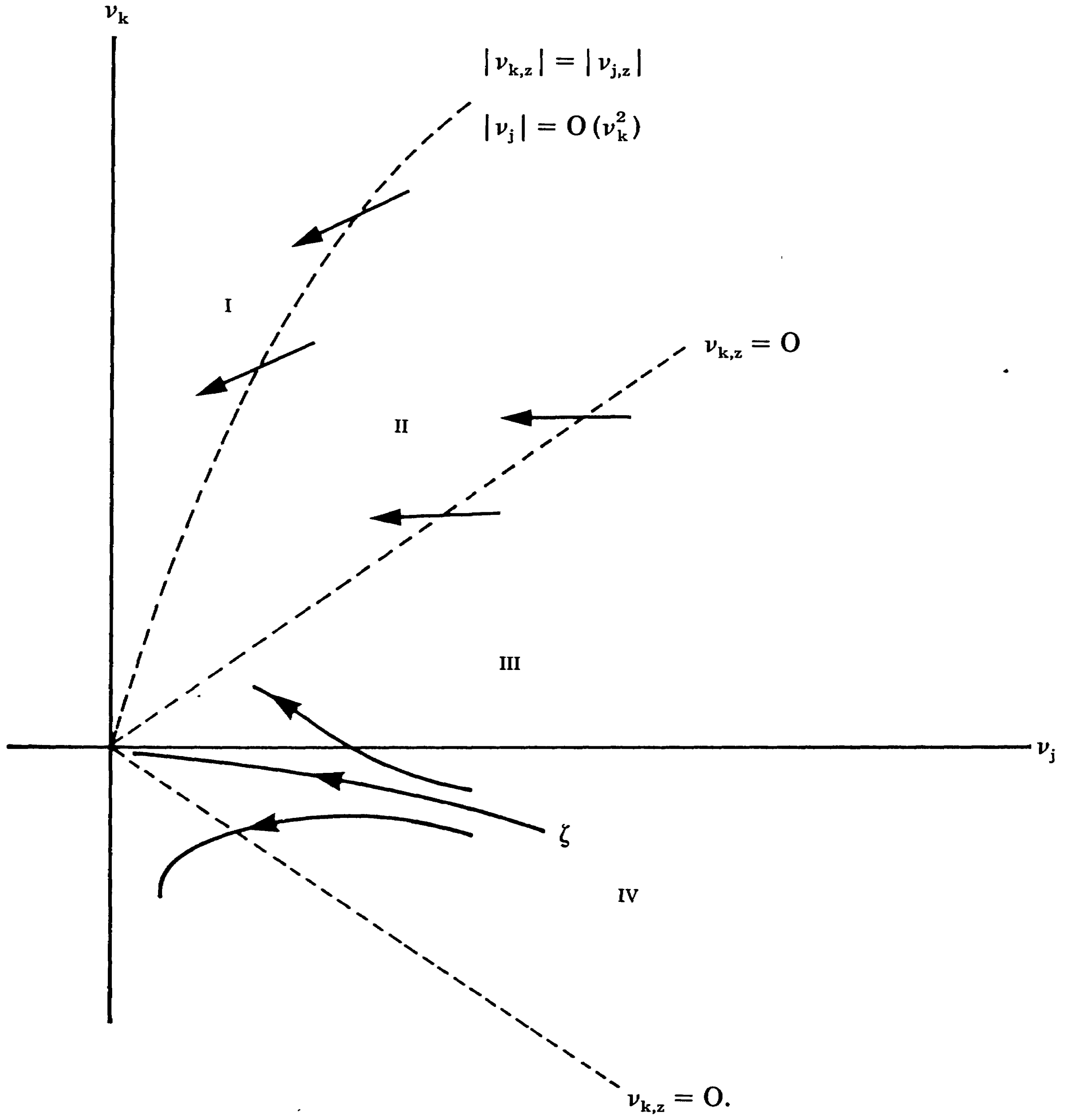

Figure 2: Proof of Theorem 4.1

But (4.6) obviously admits exponential growth of perturbations in ρ. Thus for any z there are orbits arbitrarily close to $\zeta(z)$ which approach the origin, but not from region IV; *i.e.*, from region I. If θ approaches u_r along such an orbit ζ, condition (S1) will therefore fail.

Thus condition N holds in D. To show that condition E is satisfied, let k be such that $u_r \in S_k(u_\ell)$. We claim that $\lambda_k(u_\ell) > \lambda_k(u_r)$. For the case $m = 1$, this follows immediately from Lemma 2.6 and the asymptotic conditions on θ. An intuitive argument for any m is obtained by counting the number of degrees of freedom in the orbit θ between u_ℓ and u_r in D. For stability under perturbation, we must have $m + 1$ degrees of freedom, since there are m components of f which can be perturbed, and one for the location of the shock. But if $\lambda_k(u_\ell) < \lambda_k(u_r)$, using Lemma 2.7, we have only $m - k$ degrees of freedom at u_ℓ (since

$$\lambda_k(u_\ell) < s < \lambda_{k+1}(u_\ell))$$

and $k - 1$ at u_r (since $\lambda_{k-1}(u_r) < s < \lambda_k(u_r)$), for a total of $m - 1$.

5. EXISTENCE OF DISCRETE SHOCKS

In this section we show the existence of discrete shocks satisfying the hypotheses of Theorem 4.1, for a special case. Our results are summarized by the following theorem.

THEOREM 5.1. *Suppose* $m = 2$ *and* (2.5),(2.6), and (2.7) *hold in* D. *Let* $u_\ell, u_r \in D$, $u_r \in S_k(u_\ell)$, *be such that* $\Omega \subset D$, *where* $\Omega \subset \mathbb{R}^2$ *is the closed rectangle with sides parallel to the* $u^{(1)}$, $u^{(2)}$ *axes and* u_ℓ and u_r *are at opposite corners. For any* $b \in (u_r^{(1)}, u_\ell^{(1)})$ *and sufficiently small* β, *there exists an element* $\phi \in X_h$ *satisfying* (3.4),(3.7),(S1), *and* (S2), *with* $\phi_0^{(1)} = b$, *if and only if* $\lambda_k(u_\ell) > \lambda_k(u_r)$.

Remarks. If $\Omega \subset D$ for every connected pair; *e.g.*, if D is a rectangle, half-plane, etc., then the existence of discrete shocks satisfying these conditions is equivalent to genuine nonlinearity. Since condition N holds (Theorem 2.3), necessity of the entropy condition follows from Theorem 4.1, and it suffices to prove sufficiency.

We give a proof for the case $k = 2$, $u_\ell^{(1)} > u_r^{(1)}$, $u_\ell^{(2)} > u_r^{(2)}$; the other cases are completely analogous. We first show that u_ℓ, u_r can be connected by a viscous profile [1,2,3,12].

LEMMA 5.2. *Under the hypotheses of Theorem* 5.1, *there exists a smooth function*

$$\theta: \mathbb{R} \to \Omega, \quad \theta_{(0)}^{(1)} = b, \quad \theta(+\infty) = u_r, \quad \theta(-\infty) = u_\ell,$$

satisfying

$$\frac{d\theta}{dz} + s(\theta - u_r) = f(\theta) - f_r, \tag{5.1}$$

where $s = s(u_r, u_\ell)$.

Proof. In this case, from Lemma 2.6, 2.7, it follows that with respect to the system (5.1), u_r is an attractive improper node and u_ℓ is a saddle. From (2.6) and (2.7), no orbits can leave Ω, and since condition N holds, there are no other critical points in Ω. Also from (2.6) and (2.7), an orbit enters Ω at u_ℓ; this orbit must end at u_r, and the first component must assume the value b at some point (Figure 3).

Next for fixed small positive δ, let $I = (\theta_{(0)}^{(2)} - \delta, \theta_{(0)}^{(2)} + \delta)$; we consider the solutions of (3.5) with $\phi \in X_h$, $\phi_{(0)}^{(1)} = b$, $\phi_{(0)}^{(2)} \in I$. Throughout this discussion, we consider h/β fixed. We choose h (or β) sufficiently small that all such ϕ approach u_r as $x \to +\infty$, and all such ϕ enter the region

$$B = \{u \in \Omega : |u - u_\ell| < \delta\}$$

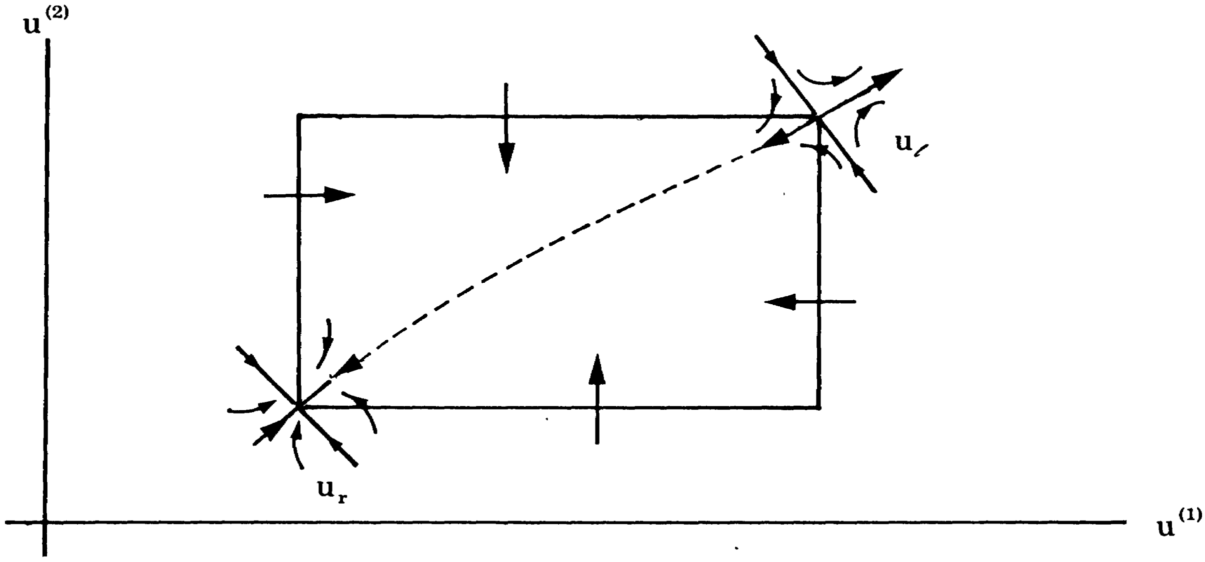

Figure 3

for sufficiently large negative x.

LEMMA 5.3. The solutions of (3.5) which enter B either cross $\partial\Omega$ at some finite value of x, or else approach u_ℓ as $x \to -\infty$.

Proof. Let $\nu = T(\phi - u_\ell)$, with T the similarity transformation diagonalizing $f_u(u_\ell)$; from (3.6), we obtain the explicit form

$$(5.2) \qquad \left(a\frac{h}{\beta}\right)\left(\frac{\nu_{j-1}^{(1)} - \nu_j^{(1)}}{h}\right) = (s - \lambda_1(u_\ell))\nu_j^{(1)} + o(|\nu_j|^2),$$

$$(5.3) \qquad \left(a\frac{h}{\beta}\right)\left(\frac{\nu_{j-1}^{(2)} - \nu_j^{(2)}}{h}\right) = (s - \lambda_2(u_\ell))\nu_j^{(2)} + o(|\nu_j|^2).$$

From (5.3), $\nu_j^{(2)} \to 0$ as $j \to -\infty$. Then from (5.2), $\nu_j^{(1)}$ either also approaches zero, or else increases exponentially. In the latter case, the orbit will eventually cross the boundary $\partial\Omega$.

We can now show the existence of a $\phi \in X_h$ satisfying (3.4) with $\phi_{(0)}^{(1)} = b$. For $y \in I$, let G(y) be the point on $\partial\Omega$ where the solution of (3.5) satisfying $\phi_{(0)}^{(1)} = b$, $\phi_{(0)}^{(2)} = y$ crosses $\partial\Omega$, for some negative x depending on y; if this ϕ never crosses $\partial\Omega$, set $G(y) = u_\ell$. From (3.6), G is obviously continuous, except possibly where $G(y) = u_\ell$. Now reducing h if necessary, the two sets

$$\{y \in I\colon G(y) \in \{u^{(1)} = u_\ell^{(1)}, u^{(2)} < u_\ell^{(2)}\}\}, \quad \text{and}$$
$$\{y \in I\colon G(y) \in \{u^{(1)} < u_\ell^{(1)}, u^{(2)} = u_\ell^{(2)}\}\}$$

are both nonempty. By the continuity of G, there exists a $y_0 \in I$ such that $G(y_0) = u_\ell$. Then the solution of (3.5) with $\phi_{(0)}^{(1)} = b$, $\phi_{(0)}^{(2)} = y_0$ satisfies (3.4). Figure 4 is an illustration.

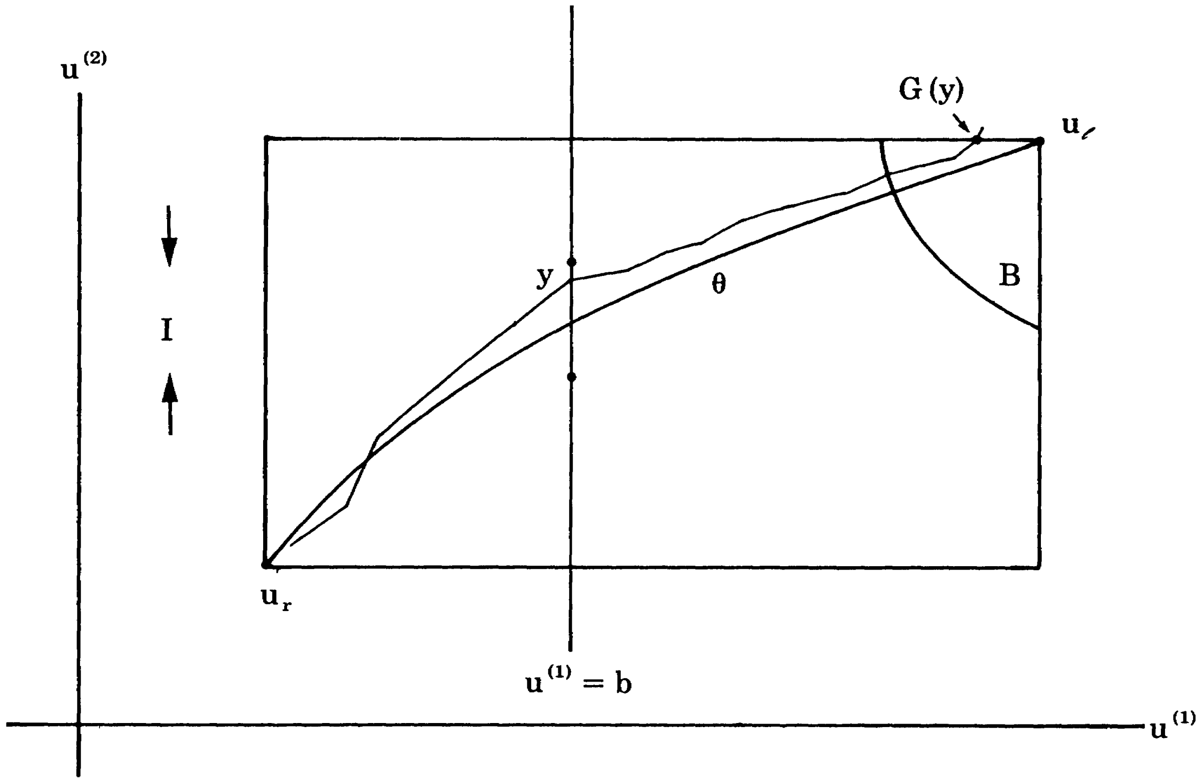

Figure 4

Denote the obtained solution of (3.4) by ϕ; we wish to show that (3.7) and (S1) are satisfied by this solution. For u outside of B, ϕ is close to θ and there is no trouble. For u inside B, both (3.7) and (S1) follow from the following:

LEMMA 5.4. *Suppose for* $k = 1,2$, $j = 0,1,2, \ldots$, $\eta_j^{(k)}$ *satisfy*

$$\eta_{j+1}^{(1)} = \left(1 + \frac{\beta}{a}(s - \lambda_1(u_\ell))\right)\eta_j^{(1)} + \varepsilon_j^{(1)} + O(\beta|\eta_j|^2) \tag{5.4}$$

$$\eta_{j+1}^{(2)} = \left(1 + \frac{\beta}{a}(s - \lambda_2(u_\ell))\right)\eta_j^{(2)} + \varepsilon_j^{(2)} + O(\beta|\eta_j|^2), \tag{5.5}$$

$$|\eta_j^{(k)}| \le \delta, \qquad k = 1,2, \quad j = 0,1,2, \ldots, \tag{5.6}$$

and

$$\eta_j^{(k)} \to 0 \quad \text{as} \quad j \to \infty, \qquad k = 1,2. \tag{5.7}$$

Then

$$\sum_{j=1}^{\infty} |\eta_j| \le O\left(\frac{1}{\beta}\right)\left(\sum_{j=1}^{\infty} (|\varepsilon_j^{(1)}| + |\varepsilon_j^{(2)}|) + \delta\right). \tag{5.8}$$

Proof. Rearranging (5.4) and (5.5), and summing with respect to j, we obtain

$$\begin{aligned}
&\sum_{j=1}^{N}\left[\left(1+\frac{\beta}{a}(s-\lambda_1(u_\ell))\right)|\eta_j^{(1)}|+|\eta_{j+1}^{(2)}|\right] \\
&\quad= \sum_{j=1}^{N} |\eta_{j+1}^{(1)} - \varepsilon_j^{(1)} + O(\beta\delta|\eta_j|)| \\
&\qquad+ \left|\left(1-\frac{\beta}{a}(s-\lambda_2(u_\ell))\right)\eta_j^{(2)}+\varepsilon_j^{(2)}+O(\beta\delta|\eta_j|)\right|
\end{aligned} \tag{5.9}$$

for any N. Since $\lambda_1(u_\ell) < s < \lambda_2(u_\ell)$, (5.9) becomes

$$\begin{aligned}
&\frac{\beta}{a}(s-\lambda_1(u_\ell))\sum_{j=1}^{N}|\eta_j^{(1)}| + \frac{\beta}{a}(\lambda_2(u_\ell)-s)\sum_{j=1}^{N}|\eta_j^{(2)}| \\
&\leq |\eta_{N+1}^{(1)}| + |\eta_1^{(2)}| + \sum_{j=1}^{N}(|\varepsilon_j^{(1)}| + |\varepsilon_j^{(2)}| + O(\beta\delta|\eta_j|)) \\
&\leq 2\delta + \sum_{j=1}^{N}(|\varepsilon_j^{(1)}| + |\varepsilon_j^{(2)}|) + O\left(\beta\delta\sum_{j=1}^{N}|\eta_j|\right),
\end{aligned}$$

using (5.6), from which (5.8) follows easily.

Perturbation of f as described in (S2) will not change the character of the critical points u_r, u_ℓ. It will also not cause condition N to fail. Thus the above argument is maintained under such perturbation, and condition (S2) is satisfied. This completes the proof of Theorem 5.1.

REFERENCES

1. C. C. Conley and J. A. Smoller, *Shock waves as limits of progressive wave solutions of higher order equations.* Comm. Pure Appl. Math. 24 (1971), 459–472.

2. ———, *Shock waves as limits of progressive wave solutions of higher order equations,* II. Comm. Pure Appl. Math. 25 (1972), 133–146.

3. R. L. Foy, *Steady state solutions of hyperbolic systems of conservation laws with viscosity terms.* Comm. Pure Appl. Math. 17 (1964), 177–188.

4. A. Harten, J. M. Hyman, P. D. Lax, and B. Keyfitz, *On finite-difference approximations and entropy conditions for shocks.* Comm. Pure Appl. Math. 29 (1976), 297–322.

5. E. Hopf, *On the right weak solution of the Cauchy problem for quasilinear equations of first order.* J. Math. and Mech. 19 (1969), 483-487.

6. G. Jennings, *Discrete shocks.* Comm. Pure Appl. Math. 27 (1974), 25–37.

7. S. N. Krushkov, *Results on the character of continuity of solutions of parabolic equations and some of their applications.* Math. Zametky 6 (1969), 97–108.

8. P. D. Lax, *Hyperbolic systems of conservation laws,* II, Comm. Pure Appl. Math. 10 (1957), 537-566.

9. ———, *Shock waves and entropy,* in Contributions to Nonlinear Functional Analysis, E. H. Zarantanello, ed., Academic Press (1971), 603-634.

10. T. P. Liu, *The Riemann problem for general* 2 × 2 *conservation laws.* Trans. Amer. Math. Soc. 199 (1974), 89–112.

11. ———, *The Riemann problem for general systems of conservation laws.* J. Diff. Eq. 18 (1975), 218–234.

12. ———, *The Entropy condition and admissibility of shocks.* J. Math. Anal. Appl. 53 (1976), 78–88.

13. M. S. Mock, *Explicit finite element schemes for first order symmetric hyperbolic systems.* Numer. Math. 26 (1976), 367–378.

14. O. A. Oleinik, *Uniqueness and stability of the generalized solution of the Cauchy problem for a quasilinear equation.* Amer. Math. Soc. Transl. Ser. 2, No. 33, 285–290.

15. J. A. Smoller, *On the solution of the Riemann problem with general step data for an extended class of hyperbolic systems.* Michigan Math. J. 16 (1969), 201–210.

16. ———, *A uniqueness theorem for Riemann problems.* Arch. Rational Mech. Anal. 33 (1969), 110–115.

Department of Mathematics
Rutgers University
New Brunswick, New Jersey 08903