

On the Accuracy of Stable Schemes for 2D Scalar Conservation Laws

By Jonathan B. Goodman and Randall J. LeVeque*

Abstract. We show that any conservative scheme for solving scalar conservation laws in two space dimensions, which is total variation diminishing, is at most first-order accurate.

1. Introduction. We study finite-difference schemes for computing weak solutions (e.g., solutions with shocks) to scalar conservation laws of the form

$$(1.1) \quad u_t + f(u)_x + g(u)_y = 0$$

in two space dimensions. In one dimension Harten [3] has given conditions which guarantee that a scheme is strongly nonlinearly stable in the sense that the total variation of the numerical solution does not increase in time. He has also shown how to construct schemes which have this property and are second-order accurate on smooth solutions.

By contrast, in two dimensions such schemes do not exist. We show that any conservative scheme which is total variation diminishing (TVD) is at most first-order accurate. Briefly, the proof goes as follows. For any two-dimensional scheme there is an associated one-dimensional scheme with the same order of accuracy. If the 2D scheme is TVD, then the corresponding 1D scheme is monotone at least on certain initial data and therefore at most first-order accurate by the results of Harten, Hyman and Lax [4].

While it is not logically necessary for a scheme to be TVD in order to be total variation stable, or to converge, we know of no method for computing weak solutions that is not TVD and yet can be shown to converge. In any case, diminishing total variation is a very desirable property which, needless to say, is shared by the exact solution (see [1]).

In Section 2 we review and slightly extend some known results on monotone schemes. These are used in Section 3 to prove our main result. We also make some remarks on its significance and limitations.

2. Monotone Schemes. A finite-difference scheme for the problem (1.1) is said to be in conservation form if it is of the form

$$(2.1) \quad U_{j,k}^{n+1} = U_{j,k}^n - \lambda_x (\bar{f}_{j+1,k}^n - \bar{f}_{j,k}^n) - \lambda_y (\bar{g}_{j,k+1}^n - \bar{g}_{j,k}^n),$$

Received October 18, 1983; revised May 1, 1984.

1980 *Mathematics Subject Classification.* Primary 65M10; Secondary 76N15.

* Prepared under Contract DE-AC02-76ER03077 with the U.S. Department of Energy. The second author was supported in part by a National Science Foundation Postdoctoral Fellowship.

©1985 American Mathematical Society
0025-5718/85 \$1.00 + \$.25 per page

where

$$\lambda_x = \Delta t / \Delta x, \quad \lambda_y = \Delta t / \Delta y$$

and the numerical fluxes \tilde{f} and \tilde{g} are functions of neighboring values of U :

$$(2.2) \quad \begin{aligned} \tilde{f}_{j,k}^n &= \tilde{f}(U_{j-p,k-q}^n, \dots, U_{j+r,k+s}^n), \\ \tilde{g}_{j,k}^n &= \tilde{g}(U_{j-p,k-q}^n, \dots, U_{j+r,k+s}^n). \end{aligned}$$

We assume \tilde{f} and \tilde{g} are consistent with f and g ,

$$(2.3) \quad \tilde{f}(u, \dots, u) = f(u), \quad \tilde{g}(u, \dots, u) = g(u),$$

and are Lipschitz continuous in all arguments.

The discrete L_1 -norm is defined by

$$(2.4a) \quad \|U\| = \Delta x \Delta y \sum_{j,k} |U_{j,k}|$$

or, by

$$(2.4b) \quad \|U\| = \Delta x \sum_j |U_j|$$

in one space dimension. The total variation of the grid function U , a discrete approximation to $\iint |\nabla u| dx dy$ with $|\nabla u| = |u_x| + |u_y|$, is given by

$$(2.5) \quad \text{TV}(U) = \sum_{j,k} \left[\Delta y |U_{j+1,k} - U_{j,k}| + \Delta x |U_{j,k+1} - U_{j,k}| \right]$$

or, in one dimension, by

$$\text{TV}(U) = \sum_j |U_{j+1} - U_j|.$$

We denote the scheme (2.1) symbolically by $U^{n+1} = F(U^n)$. Since we are studying a single timestep, we will generally suppress the index n and use the notation

$$\bar{U} = F(U).$$

We recall two definitions.

Definition 2.1. The scheme (2.1) is *monotone* if each component of \bar{U} is a monotone nondecreasing function of U , i.e., if

$$(2.6) \quad U_{j,k} \geq V_{j,k} \quad \forall j, k \Rightarrow \bar{U}_{j,k} \geq \bar{V}_{j,k} \quad \forall j, k.$$

Definition 2.2. The scheme (2.1) is L_1 -contracting if for all U and V ,

$$(2.7) \quad \|\bar{U} - \bar{V}\| \leq \|U - V\|.$$

These two properties are equivalent.

PROPOSITION 2.1. *The scheme (2.1) is monotone if and only if it is L_1 -contracting.*

Remark. This proposition, and its proof, are closely related to the lemma of Crandall and Tartar [2] which appears, for example, in [1].

Proof. Assume that $\|U - V\| < \infty$, for otherwise (2.7) is automatically satisfied. It suffices to consider the case when U and V differ at only a single point. Monotonicity on general data is equivalent to monotonicity on such data and similarly for L_1 -contraction. This follows by induction on the number of points at which U and V

differ. One must use the fact that if W differs from U at one set of mesh points and W differs from V at a disjoint set of mesh points then $\|U - V\| = \|U - W\| + \|W - V\|$.

We suppose $U_{jk} = V_{jk}$ for $(j, k) \neq (0, 0)$ and that

$$(2.8) \quad U_{00} > V_{00}.$$

From the conservation form (2.1) of the difference scheme it follows that

$$(2.9) \quad \sum_{j,k} (\bar{U}_{jk} - \bar{V}_{jk}) = \sum_{j,k} (U_{jk} - V_{jk}) = U_{00} - V_{00}.$$

Using this, we obtain

$$(2.10) \quad \frac{1}{\Delta x \Delta y} \|\bar{U} - \bar{V}\| = \sum_{j,k} |\bar{U}_{jk} - \bar{V}_{jk}| \geq \left| \sum_{j,k} (U_{jk} - V_{jk}) \right| = U_{00} - V_{00},$$

i.e.,

$$(2.11) \quad \|\bar{U} - \bar{V}\| \geq \|U - V\|$$

with equality holding if and only if $\bar{U}_{jk} - \bar{V}_{jk}$ are all of the same sign (which must then be positive by (2.8) and (2.9)). Comparing (2.7) and (2.11) we conclude that the scheme is L_1 -contracting if and only if it is monotone, completing the proof.

In Section 3 we will need to consider schemes in one dimension of the form

$$(2.12) \quad U_j^{n+1} = U_j^n - \lambda_x (\bar{f}_{j+1}^n + \bar{f}_j^n),$$

which have a weaker monotonicity property—we only require monotonicity on special data of the form

$$(2.13) \quad U_j = \begin{cases} u_0, & j < N, \\ U_N, & j = N, \\ u_\infty, & j > N, \end{cases}$$

which is monotone, i.e., $u_0 \leq U_N \leq u_\infty$ or $u_0 \geq U_N \geq u_\infty$. From the proof of Proposition 2.1 we easily obtain the following restricted form:

COROLLARY 2.1. *The scheme (2.12) is monotone on monotone data of the form (2.13) if and only if it is L_1 -contracting on pairs of data U, V , both of the form (2.13), which differ only at $j = N$.*

We have shifted the discontinuity to $j = N$ to facilitate the application in Section 3. Combining this with the results of Harten, Hyman, and Lax [4] gives

PROPOSITION 2.2. *A scheme (2.12) which is L_1 -contracting on monotone data of the form (2.13) is at most first-order accurate at any point where U_j^{n+1} is a twice differentiable function of U^n .*

Remark. There are many schemes, such as the original Godunov scheme, which have Lipschitz continuous fluxes which are not continuously differentiable everywhere. However, for all known schemes, the fluxes are infinitely differentiable except at a few points (e.g., sonic points). Consequently, we believe there are no second-order accurate schemes which are L_1 -contracting on monotone data.

There may be an alternative proof of this which does not depend on differentiability. We conjecture that monotone schemes exhibit false entropy production for smooth solutions which is $O(h)$ whenever the fluxes are Lipschitz continuous, thus implying first-order accuracy. (See Crandall and Majda [1] for entropy fluxes in 2D.) We do not pursue this approach here but rather follow the proof of Harten, Hyman and Lax [4], who show that a scheme which is L_1 -contracting on all data is at most first-order accurate (again at points of differentiability, though they do not make an issue of this).

Proof. If

$$(2.14) \quad \frac{\partial \bar{U}_j}{\partial U_N} \geq 0 \quad \text{for all } j, N$$

when U_k is constant in k , then by [4] the truncation error is of the form $h(\beta(u)u_x)_x + O(h^2)$, where β is an explicit continuous combination of partials of the form (2.14). By conservation and (2.14) it follows that $\beta \neq 0$ and, hence, the scheme is first-order accurate. To verify (2.14) for constant U we first consider data of the form (2.13) for which (2.14) holds by monotonicity and then use continuity of the derivatives as $u_0 \rightarrow U_N$ and $u_\infty \rightarrow U_N$.

3. First-Order Accuracy of TVD Schemes in 2D. We state the main result imprecisely to avoid a complicated (and unenlightening) statement.

Assertion 3.1. Except in certain unlikely cases, any TVD scheme of the form (2.1) in two dimensions is at most first-order accurate.

Proof. From any scheme of the form (2.1) for the 2D problem (1.1) we can construct a scheme of the form (2.12) for the 1D problem $u_t + f(u)_x = 0$ that has the same order of accuracy. We show that if the 2D scheme is TVD, then the resulting 1D scheme will be L_1 -contracting on monotone data (2.13). This, together with Proposition 2.2, proves the assertion.

Suppose we apply (2.1) to a grid function U which is independent of y ,

$$(3.1) \quad U_{jk} = V_j \quad \forall j, k.$$

Then (2.1) reduces to the following 1D scheme for V :

$$(3.2) \quad \bar{V}_j = V_j - \lambda_x (\bar{f}_{j+1} - \bar{f}_j),$$

where

$$\bar{f}_j = \bar{f}(V_{j-p}, \dots, V_{j+r}) \equiv \bar{f}(U_{j-p, k-q}, \dots, U_{j+r, k+s}).$$

The value of k chosen in evaluating f is arbitrary by (3.1). From (2.3) we see that \bar{f} is consistent with the flux function f ,

$$\bar{f}(u, \dots, u) = f(u).$$

Moreover, (3.2) will be a second-order accurate scheme for $u_t + f(u)_x = 0$ if (2.1) is second-order on (1.1).

The total variation as defined in (2.5) is the sum of the x and y variations,

$$(3.3a) \quad \text{TV}_x(U) = \Delta y \sum_{j, k} |U_{j+1, k} - U_{j, k}|$$

and

$$(3.3b) \quad \text{TV}_y(U) = \Delta x \sum_{j,k} |U_{j,k+1} - U_{j,k}| = \sum_k \|U_{\cdot,k+1} - U_{\cdot,k}\|,$$

where $\|\cdot\|$ is the 1D L_1 -norm (2.4b).

We construct data for which TV_x is conserved. Then by (3.3b), total variation in 2D is related to L_1 -contraction in 1D. Suppose V and W are two one-dimensional grid functions satisfying

$$(3.4) \quad \begin{cases} V_j = W_j = u_\infty & \text{for } j < -2N, \\ V_j \leq V_{j+1}, W_j \leq W_{j+1} & \text{for } -2N \leq j \leq -N, \\ V_j = W_j = u_0 & \text{for } -N < j < N, \\ V_j \geq V_{j+1}, W_j \geq W_{j+1} & \text{for } N \leq j \leq 2N, \\ V_j = W_j = u_\infty & \text{for } j > 2N, \end{cases}$$

where $u_0 > u_\infty$ and N is large relative to the stencil size determined by p , q , r and s in (2.2). Suppose that L and M are also large integers, and define U from V and W to be constant in y along each of $2M$ horizontal strips of width L , taking the values of V and W alternately

$$(3.5) \quad U_{jk} = \begin{cases} V_j & \text{if } k = 2k_1L + k_2 \\ & \text{with } 0 \leq k_1 < M, 0 \leq k_2 < L, \\ W_j & \text{if } k = (2k_1 + 1)L + k_2 \\ & \text{with } 0 \leq k_1 < M, 0 \leq k_2 < L, \\ u_\infty & \text{otherwise.} \end{cases}$$

Then $\|U - u_\infty\| < \infty$ and

$$(3.6a) \quad \text{TV}_x(U) = 4LM(u_0 - u_\infty)\Delta y,$$

$$(3.6b) \quad \text{TV}_y(U) = (2M - 1)\|V - W\| + \|V - u_\infty\| + \|W - u_\infty\|.$$

Here we mean the one-dimensional L_1 -norm (2.4b).

Since N is large compared to the stencil size determined by p , q , r , and s , and U_{jk} is constant in j and k for $|j| < N$ and $0 \leq k \leq 2LM$, it follows from the consistency of (2.1) that

$$\bar{U}_{0k} = u_0 \quad \text{for } q < k < 2LM - s.$$

Since $\bar{U}_{jk} = u_\infty$ for $|j|$ sufficiently large, it follows that the one-dimensional variation satisfies

$$\text{TV}(\bar{U}_{\cdot,k}) \geq \text{TV}(U_{\cdot,k}) \quad \text{for } q < k < 2LM - s$$

and hence,

$$(3.7) \quad \text{TV}_x(\bar{U}) \geq \text{TV}_x(U) - C,$$

where the constant C depends only on possible edge effects near $k = 0$ and $k = 2LM$ and is independent of L and M .

In the y -direction we obtain a lower bound on the total variation of \bar{U} by comparing values at the middle of successive strips where the 2D scheme reduces to

the 1D scheme (3.2) for L sufficiently large. Suppose \bar{V} and \bar{W} result from applying the 1D scheme (3.2) to V and W respectively. Then we obtain

$$(3.8) \quad \text{TV}_y(\bar{U}) \geq (2M - 1)\|\bar{V} - \bar{W}\|.$$

Now suppose the 1D scheme is not L_1 -contracting on V and W and that, consequently,

$$(3.9) \quad \|\bar{V} - \bar{W}\| = \|V - W\| + \varepsilon$$

with $\varepsilon > 0$. Then by (3.6b), (3.7), (3.8) and (3.9),

$$\begin{aligned} \text{TV}(\bar{U}) &= \text{TV}_x(\bar{U}) + \text{TV}_y(\bar{U}) \\ &\geq \text{TV}_x(U) + \text{TV}_y(U) + (2M - 1)\varepsilon - C_1, \end{aligned}$$

where $C_1 = C + \|V - u_\infty\| + \|W - u_\infty\|$. By taking M sufficiently large, we obtain

$$(3.10) \quad \text{TV}(\bar{U}) > \text{TV}(U).$$

If the 2D scheme is TVD, then (3.10) cannot occur and hence the corresponding 1D scheme must be L_1 -contracting on all data V, W of the form (3.4).

Using this construction, we can show that the 1D scheme is also L_1 -contracting on monotone data of the form (2.13). Suppose V and W are of the form (2.13). Then we simply redefine

$$\begin{aligned} V_j &= u_\infty \quad \text{for } j < -N, \\ W_j &= u_\infty \quad \text{for } j < -N. \end{aligned}$$

This does not change $\|V - W\|$ nor does it change $\|\bar{V} - \bar{W}\|$ and hence L_1 -contraction on the original data is equivalent to L_1 -contraction on the redefined data. But this new data is of the form (3.4) and hence the scheme is L_1 -contracting.

We now apply Proposition 2.2 to conclude that the 1D scheme (and therefore also the 2D scheme) is at most first-order accurate.

Remarks. (1) A. Harten has pointed out to us that our restriction to explicit schemes is probably not essential, since implicit schemes are also rather “local”. He conjectures that our proof would apply to implicit schemes as well by taking L , the strip width, to be very large.

(2) Despite our results, split schemes in 2D seem to work quite well for gas dynamics. It would therefore be interesting to prove stability and convergence for some second-order accurate 2D scheme for scalar conservation laws. This was in fact our original intention.

(3) The differences between scalars and systems are more pronounced in higher dimensions than in one dimension. For example, the Euler equations are not well-posed in $L_1(\mathbf{R}^2)$, and gas dynamical shock fronts are more stable than scalar shock fronts in two dimensions [5].

Department of Mathematics
Courant Institute of Mathematical Sciences
New York University
New York, New York 10012

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA 90024

1. M. G. CRANDALL & A. MAJDA, "Monotone difference approximations for scalar conservation laws," *Math. Comp.*, v. 34, 1980, pp. 1–21.
2. M. G. CRANDALL & L. TARTAR, "Some relations between nonexpansive and order preserving mappings," *Proc. Amer. Math. Soc.*, v. 78, 1980, pp. 385–390.
3. A. HARTEN, "High resolution schemes for hyperbolic conservation laws," *J. Comput. Phys.*, v. 49, 1983, pp. 357–393.
4. A. HARTEN, J. M. HYMAN & P. D. LAX, "On finite-difference approximations and entropy conditions for shocks," *Comm. Pure Appl. Math.*, v. 29, 1976, pp. 297–322.
5. A. MAJDA, "The existence of multi-dimensional shock fronts," *Mem. Amer. Math. Soc.*, no. 281, 1983.