

UNIFORMLY HIGH-ORDER ACCURATE NONOSCILLATORY SCHEMES. I*

AMI HARTEN† AND STANLEY OSHER‡

Abstract. We begin the construction and the analysis of nonoscillatory shock capturing methods for the approximation of hyperbolic conservation laws. These schemes share many desirable properties with total variation diminishing schemes, but TVD schemes have at most first-order accuracy, in the sense of truncation error, at extrema of the solution. In this paper we construct a uniformly second-order approximation, which is nonoscillatory in the sense that the number of extrema of the discrete solution is not increasing in time. This is achieved via a nonoscillatory piecewise-linear reconstruction of the solution from its cell averages, time evolution through an approximate solution of the resulting initial value problem and an average of this approximate solution over each cell.

Key words. conservation laws, finite difference scheme, nonoscillatory, TVD

AMS(MOS) subject classifications. Primary 65 M10; secondary 65 M05

1. Introduction. In this paper we consider numerical approximations to weak solutions of the scalar initial value problem (IVP)

$$(1.1a) \quad u_t + f(u)_x = u_t + a(u)u_x = 0,$$

$$(1.1b) \quad u(x, 0) = u_0(x).$$

The initial data $u_0(x)$ are assumed to be piecewise-smooth functions that are either periodic or of compact support.

Let $v_j^n = v_h(x_j, t_n)$, $x_j = jh$, $t_n = n\tau$, denote a numerical approximation in conservation form

$$(1.2a) \quad v_j^{n+1} = v_j^n - \lambda(\hat{f}_{j+1/2} - \hat{f}_{j-1/2}) \equiv (E_h \cdot v^n)_j.$$

Here E_h is the numerical solution operator, $\lambda = \tau/h$, and $\hat{f}_{j+1/2}$, the numerical flux is a function of $2k$ variables

$$(1.2b) \quad \hat{f}_{j+1/2} = \hat{f}(v_{j-k+1}^n, \dots, v_{j+k}^n),$$

which is consistent with (1.1a) in the sense that

$$(1.2c) \quad \hat{f}(u, u, \dots, u) = f(u).$$

We consider the numerical approximation $v_h(x, t)$ in (1.2) to be a piecewise-constant function

$$(1.3) \quad v_h(x, t) = v_j^n, \quad x_{j-1/2} < x < x_{j+1/2}, \quad n\tau < t \leq (n+1)\tau.$$

Accordingly we define its total variation in x to be

$$(1.4) \quad TV(v^n) = TV(v_h(\cdot, t_n)) = \sum_j |v_{j+1}^n - v_j^n|.$$

* Received by the editors May 1, 1985; accepted for publication (in revised form) April 12, 1986.

† School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv, Israel and Department of Mathematics, University of California, Los Angeles, California 90024. The research of this author was supported by Army Research Office contract DAAG-29-80 C-0041 while he was in residence at the Mathematics Research Center, University of Wisconsin, Madison, Wisconsin 53706, and by National Aeronautics and Space Administration Consortium Agreement NCA2-IR390-403 and Army Research Office grant DAAG 29-82-0090 while at the University of California, Los Angeles, California 90024.

‡ Department of Mathematics, University of California, Los Angeles, California 90024. The research of this author was supported by National Science Foundation grant 82-00788, Army Research Office grant DAAG 29-82-0090, National Aeronautics and Space Administration Consortium Agreement NCA2-IR390-403, and National Aeronautics and Space Administration Langley grants NAG 1-270 and NAGI-506.

If the total variation of the numerical solution is uniformly bounded in h for $0 \leq t \leq T$

$$(1.5) \quad TV(v_h(\cdot, t)) \leq C \cdot TV(u_0)$$

then any refinement sequence $h \rightarrow 0$, $\tau = O(h)$ has a subsequence $h_j \rightarrow 0$ so that

$$(1.6) \quad v_{h_j} \xrightarrow{L_1} u$$

where u is a weak solution of (1.1).

If all limit solutions (1.6) of the numerical solution (1.2) satisfy an entropy condition that implies uniqueness of the IVP (1.1), then the numerical scheme is convergent (see e.g. [3], [12]).

Recently we have introduced the notion of total variation diminishing (TVD) schemes (see [3]), where the approximate solution operator is required to diminish the total variation (1.4) of the numerical solution at each time-step

$$(1.7) \quad TV(v^{n+1}) \leq TV(v^n);$$

these schemes trivially satisfy (1.5) with $C = 1$. It follows from (1.7) that in the numerical solution of a TVD scheme the value of an isolated local maximum may only decrease in time, while that of a local minimum may only increase. Some early work along these lines was done by van Leer in [15].

We were able to construct TVD schemes that in the sense of local truncation error are high-order accurate everywhere except at local extrema where they necessarily degenerate into first-order accuracy (see [4], [13], [10], [11], [14]). The perpetual damping of local extrema determines the cumulative global error of the "high-order TVD schemes" to be $O(h)$ in the L_∞ norm, $O(h^{3/2})$ in the L_2 norm and $O(h^2)$ in the L_1 norm (see [17]).

In this paper we introduce a new class of nonoscillatory schemes, in which the solution operator is only required to diminish the *number* of local extrema in the numerical solution (as is customary we use "diminishing" loosely as short for "nonincreasing," throughout this paper). This property is satisfied by all the essentially 3-pt TVD schemes that can be described as an average of monotone Riemann solvers; most of the computationally interesting TVD schemes (with a more restrictive CFL condition) are of this type (see [12]). Unlike TVD schemes, nonoscillatory schemes are not required to damp the values of each local extremum at every single time-step, but are allowed to occasionally accentuate a local extremum.

In a sequence of papers, of which the present paper is the first, we show how to construct nonoscillatory schemes that are uniformly high-order accurate (in the sense of global error for smooth solutions of (1.1)). In this first paper we describe a second-order accurate scheme of this type.

The fact that the number of local extrema in the numerical solution may only diminish in time is sufficient by itself to guarantee that the application of the scheme to monotone data results in a monotone function. Thus nonoscillatory schemes, like TVD schemes, are monotonicity preserving. In particular, when applied to a step-function, they do not generate spurious oscillations.

We note that since the number of local extrema in the solution of nonoscillatory schemes is bounded by that of the initial data, uniform boundedness of its total variation (1.5) follows immediately if the maximum norm of the solution is shown to be uniformly bounded.

2. Design principle and overview. In this section we describe how to construct a nonoscillatory scheme that is uniformly second-order accurate.

Integrating the partial differential equation (1.1a) over the computational cell $(x_{j-1/2}, x_{j+1/2}) \times (t_n, t_{n+1})$ we get

$$(2.1a) \quad \bar{u}_j^{n+1} = \bar{u}_j^n - \lambda [\hat{f}_{j+1/2}(u) - \hat{f}_{j-1/2}(u)]$$

where

$$(2.1b) \quad \hat{f}_{j+1/2}(u) = \frac{1}{\tau} \int_{t_n}^{t_{n+1}} f(u(x_{j+1/2}, t)) dt,$$

and

$$(2.1c) \quad \bar{u}_j^n = \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} u(x, t_n) dx.$$

We observe that although (2.1a) is a relation between the cell-averages \bar{u}_j^n and \bar{u}_j^{n+1} , the evaluation of the fluxes $\hat{f}_{j+1/2}(u)$ in (2.1b) requires knowledge of the solution itself and not its cell-averages.

As in Godunov's scheme and its second-order extension by van Leer [16] and Colella and Woodward [2], we derive our scheme as a direct approximation to (2.1). We denote by ν_j^n the numerical approximation to the cell-averages \bar{u}_j^n of the exact solution in (2.1c), and set ν_j^0 to be the cell-averages of the initial data. Given $\nu^n = \{\nu_j^n\}$ we compute ν^{n+1} as follows.

First we reconstruct $u(x, t_n)$ out of its approximate cell-averages $\{\nu_j^n\}$ to the appropriate accuracy and denote the result by $L(x; \nu^n)$. Next we solve the IVP

$$(2.2) \quad \nu_t + f(\nu)_x = 0, \quad \nu(x, 0) = L(x; \nu^n),$$

and denote its solution by $\nu(x, t)$. Finally we obtain ν_j^{n+1} by taking cell-averages of $\nu(x, \tau)$

$$(2.3) \quad \nu_j^{n+1} = \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \nu(x, \tau) dx.$$

The averaging operator in (2.3) is nonoscillatory, therefore the number of local extrema in ν^{n+1} (interpreted as a mesh-function or the piecewise-constant function (1.3)) does not exceed that of $\nu(x, \tau)$. Assuming $\nu(x, t)$ to be the exact solution of (2.2) implies (since the exact solution operator is TVD) that the number of local extrema in $\nu(x, \tau)$ is less than or equal to that of $\nu(x, 0) = L(x; \nu^n)$. Therefore if the number of local extrema in $L(x; \nu^n)$ does not exceed that of ν^n , then the resulting scheme is nonoscillatory.

We conclude that the design of nonoscillatory high-order accurate schemes essentially boils down to a problem on the level of approximation of functions: Given cell-averages \bar{u}_j of a piecewise-smooth function $u(x)$, reconstruct $u(x)$ to a desired accuracy. Prior to studying this problem we tackle another related question in approximation of functions, that of constructing a nonoscillatory high-order accurate interpolation of piecewise-smooth functions.

In § 3 we construct a nonoscillatory piecewise-parabolic function $Q(x; u)$ that interpolates a piecewise-smooth function $u(x)$ at the mesh points

$$(2.4a) \quad Q(x_j; u) = u(x_j)$$

and satisfies, wherever $u(x)$ is smooth,

$$(2.4b) \quad Q(x; u) = u(x) + O(h^3),$$

$$(2.4c) \quad \frac{d}{dx} Q(x \pm 0; u) = \frac{d}{dx} u(x) + O(h^2).$$

In § 4 we make use of this nonoscillatory piecewise-parabolic interpolant to design a nonoscillatory reconstruction of a piecewise-smooth function from its cell-averages. As in [16], [2], [5] and [9] we take $L(x; \bar{u})$ to be the following piecewise-linear function:

$$(2.5a) \quad L(x; \bar{u}) = \bar{u}_j + S_j(x - x_j)/h \quad \text{for } |x - x_j| < \frac{h}{2}.$$

Unlike the above references that present “second-order accurate” TVD schemes, we compute the slopes S_j/h from $Q(x; \bar{u})$ by

$$(2.5b) \quad S_j/h = m\left(\frac{d}{dx}Q(x_j - 0; \bar{u}), \frac{d}{dx}Q(x_j + 0; \bar{u})\right).$$

Here $m(x, y)$ is the min mod function

$$(2.6) \quad m(x, y) = \begin{cases} s \cdot \min(|x|, |y|) & \text{if } \text{sgn}(x) = \text{sgn}(y) = s, \\ 0 & \text{otherwise.} \end{cases}$$

We show in § 4 that $L(x; \bar{u})$ is a proper reconstruction of $u(x)$ in the sense that whenever $u(x)$ is smooth

$$(2.7a) \quad L(x; \bar{u}) = u(x) + O(h^2)$$

and

$$(2.7b) \quad \bar{L}(x; \bar{u}) = \bar{u}(x) + O(h^3).$$

Here

$$\bar{u}(x) = h^{-1} \int_{-h/2}^{h/2} u(x+y) dy \quad \text{and} \quad \bar{L}(x; \bar{u}) = h^{-1} \int_{-h/2}^{h/2} L(x+y; \bar{u}) dy;$$

like $Q(x; \bar{u})$, the latter is also a nonoscillatory piecewise-parabolic interpolant of $\bar{u}(x)$,

$$(2.7c) \quad \bar{L}(x_j; \bar{u}) = \bar{u}(x_j).$$

We remark that the “second-order accurate” TVD schemes described in the above-mentioned references use a slope S_j/h in (2.5a) that approximates $(d/dx)u(x_j)$ to $O(h)$, and their loss of second-order accuracy at local extrema points is due to lack of smoothness of the coefficient in the $O(h)$ term at these points.¹ This problem is circumvented in the present scheme by taking S_j/h to be (2.5b) which is an $O(h^2)$ approximation to $(d/dx)u(x_j)$. Unfortunately there is a price to pay for this extra accuracy, namely the loss of the TVD property. As in TVD schemes

$$(2.8) \quad TV(\nu^{n+1}) \leq TV(L(\cdot; \nu^n));$$

however, here

$$TV(L(\cdot; \nu^n)) \geq TV(\nu^n)$$

and indeed the scheme may occasionally increase the variation of the numerical solution. Although we prove that the scheme is nonoscillatory, we have not been able as yet to complete a proof of uniform boundedness of the total variation of the numerical solution; this is due to lack of techniques to verify uniform boundedness of the maximum norm of the numerical solution.

¹ We repeat that the results of [8] and [11] imply that TVD schemes, no matter how they are constructed, must have this loss of accuracy at local extrema.

In § 5 we study the proposed scheme in the constant coefficient case. We verify that it is uniformly second-order accurate, examine its behavior at local extrema points and get estimates for the possible increase in total variation per time-step.

In this paper where we consider numerical schemes of the form (1.2) that are derived from approximating the relation (2.1), it is only natural to consider truncation error in the sense of cell-averages. That is, we say that the scheme (1.2) is second-order accurate if

$$(2.9) \quad \bar{u}^{n+1} = E_h \cdot \bar{u}^n + O(h^3)$$

where \bar{u}^n is the cell-average (2.1c) of the exact solution. Since

$$(2.10) \quad \bar{u}(x) = u(x) + O(h^2)$$

whenever $u(x)$ is smooth, (2.9) holds also for pointwise values of the solution. However, in the context of third and higher order accurate schemes, this difference in definitions of truncation error will be not only conceptual but of practical importance as well.

Up to this point we have assumed that $\nu(x, \tau)$ in (2.3) is the exact solution to (2.2). The resulting scheme

$$(2.11a) \quad \nu_j^{n+1} = \nu_j^n - \lambda [\hat{f}_{j+1/2}(\nu) - \hat{f}_{j-1/2}(\nu)],$$

where $\hat{f}_{j+1/2}(\nu)$ is (2.1b) applied to $\nu(x, t)$,

$$(2.11b) \quad \hat{f}_{j+1/2}(\nu) = \frac{1}{\tau} \int_0^\tau f(\nu(x, t)) dt,$$

is certainly second-order accurate in the sense of (2.9). Starting with the exact cell-averages $\nu_j^n = \bar{u}_j^n$ in (2.11) we get from (2.7a) that

$$(2.12a) \quad \nu(x, t) = u(x, t + t_n) + O(h^2) \quad \text{for } 0 \leq t \leq \tau,$$

and consequently,

$$(2.12b) \quad \hat{f}_{j+1/2}(\nu) = \hat{f}_{j+1/2}(u) + O(h^2),$$

which implies (2.9) due to the sufficient smoothness of the coefficient in the $O(h^2)$ term in (2.12b).

In § 6 we replace the exact solution $\nu(x, t)$ in (2.3) by an approximate one, which we denote by $\nu_n(x, t)$. This approximate solution is conservative, TVD, and second-order accurate in the sense of (2.12a). Thus replacing $\nu(x, t)$ in (2.3) by this approximate solution results in a conservative scheme that is nonoscillatory and uniformly second-order accurate.

We remark that an alternative approach to the above is to approximate $\hat{f}_{j+1/2}(\nu)$ in (2.11b) by using a midpoint rule (or trapezoidal rule) for the integral and by replacing $\nu(x, t)$ with a nonoscillating second-order accurate approximate one $\nu_n(x, t)$ (see [16] and [2]). The resulting scheme

$$(2.13a) \quad \nu_j^{n+1} = \nu_j^n - \lambda (\hat{J}_{j+1/2} - \hat{J}_{j-1/2}),$$

$$(2.13b) \quad \hat{J}_{j+1/2} = f(\nu_n(x_{j+1/2}, \tau/2)),$$

is certainly second-order accurate, and it is nonoscillatory in the constant coefficient case. Since we have not used the cell-averaging (2.3) to derive this scheme, we cannot ascertain in general that the resulting scheme is nonoscillatory. Nevertheless, our numerical experiments as well as many other experiments in the context of TVD schemes (see e.g. [1], [2]) demonstrate that the numerical results are nonoscillatory in many (if not all) applications.

In § 7 we present some numerical experiments that compare the present scheme with a typical “second-order accurate” TVD scheme.

3. Nonoscillatory interpolation. The oscillatory nature of second-order accurate Lax–Wendroff type schemes results from a Gibbs phenomenon associated with high-order interpolation across discontinuities. In this section, as a preparatory step towards designing a nonoscillatory approximation to (1.1), we construct a nonoscillatory piecewise-parabolic interpolant $Q(x; u)$ to a piecewise-smooth function $u(x)$ such that

$$(3.1a) \quad Q(x_i; u) = u(x_i),$$

$$(3.1b) \quad Q(x; u) \equiv q_{i+1/2}(x; u), \quad x_i \leq x \leq x_{i+1},$$

where $q_{i+1/2}$ is a quadratic polynomial, and

$$(3.1c) \quad Q(x; u) - u(x) = O(h^3), \quad \frac{d}{dx}Q(x \pm 0; u) - \frac{d}{dx}u(x) = O(h^2)$$

wherever $u(x)$ is smooth.

$Q(x; u)$ is nonoscillatory in the sense that the number of its local extrema does not exceed that of $u(x)$.

Since

$$q_{i+1/2}(x_i; u) = u_i, \quad q_{i+1/2}(x_{i+1}; u) = u_{i+1},$$

it can be written in the form

$$(3.2a) \quad q_{i+1/2}(x; u) = u_i + d_{i+1/2}u \cdot (x - x_i)/h + \frac{1}{2}D_{i+1/2}u \cdot (x - x_i)(x - x_{i+1})/h^2$$

where

$$(3.2b) \quad d_{i+1/2}u = u_{i+1} - u_i$$

and $D_{i+1/2}u$ is yet to be determined.

$$D_{i+1/2}u = h^2 q''_{i+1/2}(x; u), \quad x_i \leq x \leq x_{i+1}.$$

We consider as candidates for $q_{i+1/2}$ the two quadratic polynomials \bar{q}_i and \bar{q}_{i+1} , interpolating $u(x)$ at (x_{i-1}, x_i, x_{i+1}) and (x_i, x_{i+1}, x_{i+2}) , respectively, and choose $q_{i+1/2}$ to be the one that is least oscillatory in $[x_i, x_{i+1}]$. Both \bar{q}_j , $j = i$ and $j = i + 1$, can be written as (3.2a) with $D_{i+1/2}u = D_j u$ where

$$(3.2c) \quad D_j u = d_{j+1/2}u - d_{j-1/2}u = u_{j+1} - 2u_j + u_{j-1}.$$

Since the least oscillatory of \bar{q}_i and \bar{q}_{i+1} can be characterized as the one that deviates the least from the line connecting (x_i, u_i) with (x_{i+1}, u_{i+1}) we choose $D_{i+1/2}u$ in (3.2a) to be

$$(3.2d) \quad D_{i+1/2}u = m(D_i u, D_{i+1} u),$$

where $m(x, y)$ is the min mod function

$$(3.3) \quad m(x, y) = \begin{cases} s \cdot \min(|x|, |y|) & \text{if } \text{sgn}(x) = \text{sgn}(y) = s, \\ 0 & \text{otherwise.} \end{cases}$$

If $u(x)$ is smooth in $[x_{j-1}, x_{j+1}]$, then \bar{q}_j as a quadratic interpolant of u satisfies

$$(3.4) \quad \bar{q}_j(x) - u(x) = O(h^3), \quad \frac{d}{dx}\bar{q}_j(x) - \frac{d}{dx}u(x) = O(h^2), \quad x_{j-1} \leq x \leq x_{j+1}.$$

If $D_i u \cdot D_{i+1} u \geq 0$ then $q_{i+1/2}$ is either \bar{q}_i or \bar{q}_{i+1} . Otherwise we set $D_{i+1/2} u = 0$, but then smoothness of u implies that $D_j u = O(h^3)$ and consequently $q_{i+1/2} - \bar{q}_j = O(h^3)$ for $j = i, i + 1$. Thus (3.1c) follows from (3.4).

We turn now to prove that $Q(x; u)$ is a nonoscillatory interpolant of u , i.e., that the number of its local extrema does not exceed that of u . We do so by showing a one-to-one correspondence between local extrema of Q to those of the mesh function $\{u_j\}$, the number of which certainly does not exceed that of $u(x)$.

Q may have a local extremum in either the interior of some interval (x_i, x_{i+1}) or at a mesh point x_i . The first case, which will be referred to as interior-extremum, occurs when there is a point x^* , $x_i < x^* < x_{i+1}$, such that

$$\frac{d}{dx} Q(x^*; u) = 0 \quad \text{but} \quad \frac{d}{dx} q_{i+1/2} \neq 0.$$

From (3.2a) it follows that Q has an interior-extremum in (x_i, x_{i+1}) if and only if

$$(3.5) \quad |D_{i+1/2} u| > 2|d_{i+1/2} u|.$$

$q_{i+1/2}^* = q_{i+1/2}(x^*)$, the value of the interior-extremum is then

$$(3.6) \quad q_{i+1/2}^* = u_i - \frac{1}{2} D_{i+1/2} u \cdot \left(\frac{d_{i+1/2} u}{D_{i+1/2} u} - \frac{1}{2} \right)^2;$$

if $D_{i+1/2} u < 0$ it is a local maximum; if $D_{i+1/2} u > 0$ it is a local minimum. Since $D_{i+1/2} u = m(D_i u, D_{i+1} u)$, (3.5) holds if and only if

$$(3.7a) \quad D_i u \cdot D_{i+1} u > 0,$$

$$(3.7b) \quad |D_j u| > 2|d_{i+1/2} u|, \quad j = i, \quad i + 1.$$

This implies that $q_{i+1/2}$ has a local extremum in (x_i, x_{i+1}) if and only if both \bar{q}_i and \bar{q}_{i+1} also have a local extremum in (x_i, x_{i+1}) and of the same kind. Since a parabola has at most one local extremum, it follows then that \bar{q}_i does not have a local extremum in (x_{i-1}, x_i) and \bar{q}_{i+1} does not have one in (x_{i+1}, x_{i+2}) . Consequently Q is monotone in both (x_{i-1}, x_i) and (x_{i+1}, x_{i+2}) , but in an opposite sense, i.e., $d_{i-1/2} u \cdot d_{i+3/2} u < 0$; the latter implies that u has a local extremum in $[x_i, x_{i+1}]$ and that either u_i or u_{i+1} is a local extremum of the mesh function $\{u_j\}$ (for obvious reasons the case $u_i = u_{i+1}$ is counted as a single-extremum). The above analysis also shows that interior-extrema are isolated, i.e., if Q has an interior-extremum in (x_i, x_{i+1}) , then it is the only local extremum of Q in (x_{i-1}, x_{i+2}) .

We turn now to examine the case that Q has a local extremum at a mesh point x_i ; this will be referred to as a mesh-extremum. The above observation that interior-extrema are isolated excludes the possibility that Q has an interior-extremum in either (x_{i-1}, x_i) or (x_i, x_{i+1}) and consequently Q is monotone in these intervals. This implies that $d_{i-1/2} u \cdot d_{i+1/2} u < 0$ and therefore u_i is a local extremum of the mesh function $\{u_j\}$. This concludes the proof that $Q(x; u)$ is a nonoscillatory interpolant of u .

We next express the nonoscillatory nature of Q in terms of total variation. If $|D_{j+1/2} u| \leq 2|d_{j+1/2} u|$ then (2.5) implies that Q is monotone in $[x_j, x_{j+1}]$. Thus

$$(3.8a) \quad |D_{j+1/2} u| \leq 2|d_{j+1/2} u| \Rightarrow TV_{[x_j, x_{j+1}]}(Q) = |d_{j+1/2} u|.$$

If $|D_{j+1/2} u| > 2|d_{j+1/2} u|$ then Q has a local extremum in (x_j, x_{j+1}) and

$$TV_{[x_j, x_{j+1}]}(Q) = |q_{j+1/2}^* - u_j| + |u_{j+1} - q_{j+1/2}^*|.$$

Using (2.6) we get

$$(3.8b) \quad \begin{aligned} |D_{j+1/2}u| > 2|d_{j+1/2}u| &\Rightarrow TV_{[x_j, x_{j+1}]}(Q) \\ &= |d_{j+1/2}u| + |D_{j+1/2}u| \left(\left| \frac{d_{j+1/2}u}{D_{j+1/2}u} \right| - \frac{1}{2} \right)^2. \end{aligned}$$

We conclude that

$$(3.8c) \quad \begin{aligned} 0 \leq TV(Q) - \sum_j |d_{j+1/2}u| &\leq \sum_{m \in M} |D_{m+1/2}u| \left(\left| \frac{d_{m+1/2}u}{D_{m+1/2}u} \right| - \frac{1}{2} \right)^2 \\ &\leq \frac{1}{4} \sum_{m \in M} |D_{m+1/2}u|. \end{aligned}$$

The sum in the RHS of (3.8c) is taken over the set of indices M of intervals (x_m, x_{m+1}) in which $|D_{m+1/2}u| > 2|d_{m+1/2}u|$, i.e., where Q has interior-extrema.

Next we show that if $u(x)$ is a piecewise-smooth function of bounded variation, then

$$(3.9) \quad \lim_{h \rightarrow 0} TV(Q(\cdot; u)) = TV(u).$$

We observe that in this case the number of intervals in M is finite and is uniformly bounded by the number of local extrema in $u(x)$. Hence (3.9) will follow if we prove that $D_{m+1/2}u \rightarrow 0$ as $h \rightarrow 0$ for all $m \in M$. To accomplish this we show that for h sufficiently small M does not include intervals (x_m, x_{m+1}) in which $u(x)$ is discontinuous. To see that let us examine the case where $u(x)$ has a discontinuity at $\bar{x} \in (x_i, x_{i+1})$. Clearly $d_{i+1/2}u$ approaches the size of the jump in u while $d_{i-1/2}u$ approaches zero as $h \rightarrow 0$. Consequently

$$(3.10a) \quad |D_i u / d_{i+1/2} u| = |1 - d_{i-1/2} u / d_{i+1/2} u| \rightarrow 1 \quad \text{as } h \rightarrow 0.$$

Hence for h sufficiently small

$$(3.10b) \quad 2|d_{i+1/2}u| > |D_i u| \geq |D_{i+1/2}u|$$

which implies $i \notin M$.

4. Nonoscillatory reconstruction. Let $u(x)$ be a piecewise-smooth function and denote by $\bar{u}(x)$ its mean over $(x - h/2, x + h/2)$, i.e.,

$$(4.1) \quad \bar{u}(x) = \frac{1}{h} \int_{x-h/2}^{x+h/2} u(y) dy.$$

We denote $\bar{u}_j = \bar{u}(x_j)$ and refer to these values as cell-averages of $u(x)$. Given $\{\bar{u}_j\}$, the task at hand is to reconstruct $u(x)$ to $O(h^2)$ in a nonoscillatory way; denote the approximately reconstructed function by $L(x; \bar{u})$. To achieve $O(h^2)$ accuracy it is sufficient to consider $L(x; \bar{u})$ to be a piecewise-linear function. To make $L(x; \bar{u})$ a nonoscillatory approximation we use the nonoscillatory piecewise parabolic interpolation $Q(x; \bar{u})$ to compute its slopes as follows:

$$(4.2a) \quad L(x; \bar{u}) = \bar{u}_j + S_j(x - x_j)/h \quad \text{for } |x - x_j| < \frac{1}{2}h,$$

$$(4.2b) \quad S_j = h \cdot m \left(\frac{d}{dx} Q(x_j - 0; \bar{u}), \frac{d}{dx} Q(x_j + 0; \bar{u}) \right).$$

Here m is the min mod function (3.3); $d_{i+1/2}\bar{u}$ and $D_{i+1/2}\bar{u}$ are (3.2b) and (3.2d), respectively.

We note that $L(x; \bar{u})$ may be discontinuous at $\{x_{j+1/2}\}$ and that

$$(4.3a) \quad L(x_j; \bar{u}) = \bar{u}_j.$$

To see that wherever $u(x)$ is smooth

$$(4.3b) \quad L(x, \bar{u}) - u(x) = O(h^2)$$

we observe that in this case

$$(4.4a) \quad \bar{u}(x) = u(x) + O(h^2)$$

and therefore it follows from (3.1c) that

$$(4.4b) \quad \frac{1}{h} S_j = \frac{d}{dx} \bar{u}(x_j) + O(h^2) = \frac{d}{dx} u(x_j) + O(h^2).$$

Consequently the RHS of (4.2a) can be expanded as

$$(4.4c) \quad \begin{aligned} L(x; \bar{u}) &= u_j + (x - x_j) \frac{d}{dx} u(x_j) + O(h^2) \\ &= u(x) + O(h^2) \quad \text{for } |x - x_j| < \frac{1}{2}h, \end{aligned}$$

and thus (4.3b) follows.

Denote by $\bar{L}(x; \bar{u})$ the mean value of $L(x; \bar{u})$ in $(x - h/2, x + h/2)$, i.e.,

$$(4.5a) \quad \bar{L}(x; \bar{u}) = \frac{1}{h} \int_{x-h/2}^{x+h/2} L(y; \bar{u}) dy.$$

Using (4.2a) to evaluate the integral in (4.5a) we find

$$(4.5b) \quad \bar{L}(x; \bar{u}) = \bar{u}_j + d_{j+1/2} \bar{u}(x - x_j)/h + \left(\frac{1}{2}\right)(S_{j+1} - S_j)(x - x_j)(x - x_{j+1})/h^2$$

for $x_j \leq x \leq x_{j+1}$,

$$(4.5c) \quad \bar{L}(x_j; \bar{u}) = \bar{u}_j.$$

Hence $\bar{L}(x; \bar{u})$, like $Q(x; \bar{u})$, is a piecewise-parabolic interpolant of $\bar{u}(x)$. Comparing (4.5b) with (3.2) we find that for $x_j \leq x \leq x_{j+1}$

$$(4.6a) \quad \bar{L}(x; \bar{u}) - Q(x; \bar{u}) = \frac{1}{2}(S_{j+1} - S_j - D_{j+1/2} \bar{u})(x - x_j)(x - x_{j+1})/h^2.$$

From (4.4b) we see that $S_i = h(d/dx)\bar{u}(x_i) + O(h^3)$ (note that this is true also at local extrema points) and therefore

$$S_{j+1} - S_j = h^2 \frac{d^2}{dx^2} u(x_{j+1/2}) + O(h^3).$$

On the other hand, (3.2) shows that

$$D_{j+1/2} \bar{u} = h^2 \frac{d^2}{dx^2} u(x_{j+1/2}) + O(h^3).$$

Therefore

$$(4.6b) \quad S_{j+1} - S_j - D_{j+1/2} \bar{u} = O(h^3)$$

which shows that RHS of (4.6a) is $O(h^3)$. Since (2.1c) shows that

$$Q(x; \bar{u}) - \bar{u}(x) = O(h^3)$$

we conclude from (4.6a)-(4.6b) that

$$(4.6c) \quad \bar{L}(x; \bar{u}) - \bar{u}(x) = O(h^3).$$

We turn now to prove that $L(x; \bar{u})$ is a nonoscillatory approximation to $\bar{u}(x)$; this certainly implies that $L(x; \bar{u})$ is a nonoscillatory approximation to $u(x)$. We shall do so by showing that $TV_{[x_j, x_{j+1}]}(L(\cdot; \bar{u}))$, the total-variation of $L(x; \bar{u})$ in $[x_j, x_{j+1}]$, which has the value

$$(4.7a) \quad TV_{[x_j, x_{j+1}]}(L(\cdot; \bar{u})) = \frac{1}{2}(|S_j| + |S_{j+1}|) + |d_{j+1/2}\bar{u} - \frac{1}{2}(S_j + S_{j+1})|,$$

can also be expressed as

$$(4.7b) \quad TV_{[x_j, x_{j+1}]}(L(\cdot; \bar{u})) = \max(|d_{j+1/2}\bar{u}|, \frac{1}{2}|D_{j+1/2}\bar{u}|).$$

Then it follows immediately from (4.7b), (4.3a) and (3.8) that L is monotone in $[x_j, x_{j+1}]$ if and only if Q is; consequently L is a nonoscillatory approximation to $u(x)$ in exactly the same sense as Q is to the interpolated function (see § 3).

Next let us denote

$$(4.8a) \quad S_j^\pm = h \cdot \frac{d}{dx} Q(x_j \pm 0; \bar{u}),$$

i.e.,

$$(4.8b) \quad S_j^- = d_{j-1/2}\bar{u} + \frac{1}{2}D_{j-1/2}\bar{u}, \quad S_j^+ = d_{j+1/2}\bar{u} - \frac{1}{2}D_{j+1/2}\bar{u},$$

and observe that (4.2b) implies that

$$(4.8c) \quad \begin{aligned} \frac{1}{2}(|S_j| + |S_{j+1}|) &= \frac{1}{2}[|m(S_j^-, S_j^+)| + |m(S_{j+1}^-, S_{j+1}^+)|] \\ &\leq \frac{1}{2}(|S_j^+| + |S_{j+1}^-|) = \frac{1}{2}(|d_{j+1/2}\bar{u} - \frac{1}{2}D_{j+1/2}\bar{u}| + |d_{j+1/2}\bar{u} + \frac{1}{2}D_{j+1/2}\bar{u}|) \\ &= \max(|d_{j+1/2}\bar{u}|, \frac{1}{2}|D_{j+1/2}\bar{u}|). \end{aligned}$$

We note that if $|d_{j+1/2}\bar{u}| \geq \frac{1}{2}|D_{j+1/2}\bar{u}|$ then

$$\text{sgn}(d_{j+1/2}\bar{u} \pm \frac{1}{2}D_{j+1/2}\bar{u}) \cdot \text{sgn}(d_{j+1/2}\bar{u}) \geq 0$$

which in turn implies

$$\text{sgn}(S_j) \cdot \text{sgn}(d_{j+1/2}\bar{u}) \geq 0, \quad \text{sgn}(S_{j+1}) \cdot \text{sgn}(d_{j+1/2}\bar{u}) \geq 0.$$

It follows then from (4.8c) that the RHS of (4.7a) is $|d_{j+1/2}\bar{u}|$. This shows that

$$(4.9a) \quad |d_{j+1/2}\bar{u}| \geq \frac{1}{2}|D_{j+1/2}\bar{u}| \Rightarrow TV_{[x_j, x_{j+1}]}(L(\cdot; \bar{u})) = |d_{j+1/2}\bar{u}|.$$

To complete the verification of (4.7b) we still have to show that

$$(4.9b) \quad |d_{j+1/2}\bar{u}| < \frac{1}{2}|D_{j+1/2}\bar{u}| \Rightarrow TV_{[x_j, x_{j+1}]}(L(\cdot; \bar{u})) = \frac{1}{2}|D_{j+1/2}\bar{u}|.$$

First we observe that

$$(4.10) \quad \begin{aligned} S_i^+ - S_i^- &= (d_{i+1/2}\bar{u} - \frac{1}{2}D_{i+1/2}\bar{u}) - (d_{i-1/2}\bar{u} + \frac{1}{2}D_{i-1/2}\bar{u}) \\ &= D_i\bar{u} - \frac{1}{2}(D_{i+1/2}\bar{u} + D_{i-1/2}\bar{u}). \end{aligned}$$

Since (3.2d) implies

$$(4.11) \quad |D_i\bar{u}| \leq \frac{1}{2}(|D_{i-1/2}\bar{u}| + |D_{i+1/2}\bar{u}|),$$

we conclude from (4.10) that

$$(4.12) \quad (S_i^+ - S_i^-) \cdot \text{sgn}(D_i\bar{u}) \geq 0.$$

We turn now to prove (4.9b). First let us consider the case that $Q(x; \bar{u})$ has a local maximum in (x_j, x_{j+1}) , i.e., $D_j\bar{u} < 0$, $D_{j+1}\bar{u} < 0$, and $|d_{j+1/2}\bar{u}| < \frac{1}{2}|D_{j+1/2}\bar{u}|$.

It follows from (4.12) that

$$(4.13a) \quad S_j^- \cong S_j^+ = d_{j+1/2}\bar{u} - \frac{1}{2}D_{j+1/2}\bar{u} > 0,$$

$$(4.13b) \quad 0 > d_{j+1/2}\bar{u} + \frac{1}{2}D_{j+1/2}\bar{u} = S_{j+1}^- \cong S_{j+1}^+.$$

The relations (4.13) and the definitions (4.8a), (4.2b) imply that

$$(4.14a) \quad S_j = S_j^+ = d_{j+1/2}\bar{u} - \frac{1}{2}D_{j+1/2}\bar{u},$$

$$(4.14b) \quad S_{j+1} = S_{j+1}^- = d_{j+1/2}\bar{u} + \frac{1}{2}D_{j+1/2}\bar{u}.$$

The same analysis shows that (4.14) holds also for the case that $Q(x; \bar{u})$ has a local minimum in (x_j, x_{j+1}) . (4.9b) follows immediately from (4.14) and (4.7a).

We note that since $L(x; \bar{u})$ is continuous at x_j

$$(4.15) \quad \begin{aligned} TV(L(\cdot; \bar{u})) &= \sum_j TV_{[x_j, x_{j+1}]}(L(\cdot; \bar{u})) = \sum_j \max \left(|d_{j+1/2}\bar{u}|, \frac{1}{2}|D_{j+1/2}\bar{u}| \right) \\ &= \sum_j |d_{j+1/2}\bar{u}| + \sum_{m \in M} \left(\frac{1}{2}|D_{m+1/2}\bar{u}| - |d_{m+1/2}\bar{u}| \right). \end{aligned}$$

Here M is the set of indices of intervals (x_m, x_{m+1}) in the interior of which L (and also Q) has a local extremum. The number of these intervals is finite and is bounded by the number of local extrema of $\bar{u}(x)$. Comparing (4.9) with (3.8) we note that

$$(4.16) \quad TV(L(\cdot; \bar{u})) \geq TV(Q(\cdot; \bar{u})).$$

5. The constant coefficient case. In this section we study the constant coefficient case

$$(5.1) \quad u_t + au_x = 0, \quad a = \text{const.}$$

The exact solution of the IVP (2.2) is

$$(5.2) \quad v(x, t) = L(x - at; v^n).$$

Hence our scheme (2.3) is

$$(5.3a) \quad v_j^{n+1} = \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} L(x - at; v^n) dx = \bar{L}(x_j - at; v^n)$$

where \bar{L} is (4.5a). We have shown in § 3 that the number of local extrema in $L(x; v^n)$ does not exceed that of v^n . Since v_j^{n+1} in (5.3a) is a cell-average of L , it follows that the number of local extrema in v^{n+1} does not exceed that of v^n , and consequently the scheme (5.3a) is nonoscillatory.

Using (4.5b) in (5.3a) we get the following expression for the scheme

$$(5.3b) \quad \begin{aligned} v_j^{n+1} &= (E_h \cdot v^n)_j \\ &= \begin{cases} v_j^n - \mu d_{j-1/2} v^n - \frac{1}{2}\mu(1-\mu)(S_j^n - S_{j-1}^n) & \text{if } a > 0, \\ v_j^n - \mu d_{j+1/2} v^n + \frac{1}{2}\mu(1+\mu)(S_{j+1}^n - S_j^n) & \text{if } a < 0, \end{cases} \end{aligned}$$

E_h denotes the operator form of the finite difference scheme; $\mu = \lambda a$, the CFL-number, is assumed to satisfy

$$(5.3c) \quad |\mu| \leq 1.$$

We turn now to prove that (5.3) is second-order accurate in the sense of (2.9), i.e., if $\bar{u}^n(x)$ denotes the mean value (4.1) of $u(x, t_n)$ then

$$(5.4a) \quad \bar{u}_j^{n+1} - (E_h \cdot \bar{u}^n)_j = O(h^3).$$

To show that we observe that in the constant coefficient case (5.1) $\bar{u}_j^{n+1} = \bar{u}^n(x_j - a\tau)$, and by (5.3a) $(E_h \cdot \bar{u}^n)_j = \bar{L}(x_j - a\tau; \bar{u}^n)$. Hence the LHS of (5.4a) is nothing but

$$(5.4b) \quad \bar{u}^n(x_j - a\tau) - \bar{L}(x_j - a\tau; \bar{u}^n),$$

which is $O(h^3)$ as a direct consequence of (4.6c).

Next we study the time-dependence of the total variation and the maximum norm of the numerical solution (5.3). In § 2 we have pointed out that

$$(5.5a) \quad TV(\nu^{n+1}) \leq TV(L(\cdot; \nu^n)).$$

Using (4.15) and (5.5a) we get the following upper bound on the possible growth of the total variation of the numerical solution per time-step:

$$(5.5b) \quad TV(\nu^{n+1}) - TV(\nu^n) \leq \sum_{m \in M_n} \left(\frac{1}{2} |D_{m+1/2} \nu^n| - |d_{m+1/2} \nu^n| \right).$$

Here M_n is the set of indices of intervals (x_m, x_{m+1}) in the interior of which $L(x; \nu^n)$ (and also $Q(x; \nu^n)$) has a local extremum. The number of these intervals is finite and remains uniformly bounded in time by the number of local extrema in the initial data.

Clearly the upper bound (5.5b) is overly pessimistic. It estimates the possible increase in variation in the reconstruction step due to replacing the cell-averages ν_j^n by the piecewise-linear function $L(x; \nu^n)$. It does not take into account the possible decrease in variation in the averaging step (2.3), resulting from doing just the opposite, i.e., replacing the piecewise-linear function $L(x - a\tau; \nu^n)$ in (5.2) by its cell-averages (5.3a).

In the following we shall examine the temporal behavior of the local extrema of the numerical solution and its total variation by analyzing the explicit values of the cell-averages ν_j^{n+1} given by (5.3b). To simplify our presentation let us assume that $a > 0$.

First we note that (4.8b) implies

$$(5.6a) \quad |S_j - S_{j-1}| \leq |S_j| + |S_{j+1}| \leq 2 \max(|d_{j-1/2} \nu^n|, \frac{1}{2} |D_{j-1/2} \nu^n|).$$

Hence

$$(5.6b) \quad |d_{j-1/2} \nu^n| \geq \frac{1}{2} |D_{j-1/2} \nu^n| \Rightarrow |\gamma_{j-1/2}| = |S_j^n - S_{j-1}^n| / |d_{j-1/2} \nu^n| \leq 2.$$

Rewriting (5.3) in this case as

$$(5.7a) \quad \begin{aligned} \nu_j^{n+1} &= \nu_j^n - \mu d_{j-1/2} \nu^n - \frac{1}{2} \mu (1 - \mu) \gamma_{j-1/2} d_{j-1/2} \nu^n \\ &= (1 - \sigma_{j-1/2}) \nu_j^n + \sigma_{j-1/2} \nu_{j-1}^n \end{aligned}$$

where

$$(5.7b) \quad \sigma_{j-1/2} = \mu + \frac{1}{2} \mu (1 - \mu) \gamma_{j-1/2},$$

we see that the CFL condition $0 < \mu \leq 1$ and (5.6b) imply that

$$(5.7c) \quad 0 \leq \sigma_{j-1/2} \leq 1;$$

thus we conclude

$$(5.8) \quad |d_{j-1/2} \nu^n| \geq \frac{1}{2} |D_{j-1/2} \nu^n| \Rightarrow \nu_j^{n+1} \in [\nu_{j-1}^n, \nu_j^n].$$

Relation (5.8) shows that if ν^n is monotone for $J_L \leq j \leq J_R$, i.e. $\nu_{J_L} \leq \nu_{J_L+1} \leq \dots \leq \nu_{J_R}$, or $\nu_{J_L} \geq \nu_{J_L+1} \geq \dots \geq \nu_{J_R}$, then ν^{n+1} is monotone for $J_L + 1 \leq j \leq J_R$, and in the same

sense. Relation (5.8) also shows that mesh-extrema of ν^n , i.e., those for which Q has its local extremum at a mesh point, are being damped at the n th time-step. Namely,

$$(5.9a) \quad |d_{j\pm 1/2}\nu^n| \geq \frac{1}{2}|D_{j\pm 1/2}\nu^n|, \quad \nu_{j-1}^n \leq \nu_j^n \leq \nu_{j+1}^n \Rightarrow \max(\nu_j^{n+1}, \nu_{j+1}^{n+1}) \leq \nu_j^n,$$

$$(5.9b) \quad |d_{j\pm 1/2}\nu^n| \geq \frac{1}{2}|D_{j\pm 1/2}\nu^n|, \quad \nu_{j+1}^n \geq \nu_j^n \leq \nu_{j+1}^n \Rightarrow \min(\nu_j^{n+1}, \nu_{j+1}^{n+1}) \geq \nu_j^n.$$

We turn now to consider interior local extrema of ν^n , i.e. those for which Q has its local extremum in the interior of some (x_i, x_{i+1}) . We recall that such an extremum is characterized by $|d_{i+1/2}\nu^n| < \frac{1}{2}|D_{i+1/2}\nu^n|$ and that S_i^n and S_{i+1}^n in this case are given by (4.14); therefore $S_{i+1}^n - S_i^n = D_{i+1/2}\nu^n$. From (5.3a) and (4.6a) we see that in general

$$(5.10a) \quad \nu_{i+1}^{n+1} - Q(x_{i+1} - a\tau; \nu^n) = \frac{1}{2}\mu(1 - \mu)(D_{i+1/2}\nu^n - S_{i+1}^n + S_i^n).$$

Hence

$$(5.10b) \quad |d_{i+1/2}\nu^n| < \frac{1}{2}|D_{i+1/2}\nu^n| \Rightarrow \nu_{i+1}^{n+1} = Q(x_{i+1} - a\tau; \nu^n).$$

Relation (5.10b) confirms the second-order accuracy of the scheme at local extrema. Although it does not necessitate accentuation of the extremal values, as ν_{i+1}^{n+1} in (5.10b) may still be in $[\nu_i^n, \nu_{i+1}^n]$, it does allow ν_{i+1}^{n+1} to deviate from this interval by as much as

$$(5.10c) \quad \frac{1}{2}|D_{i+1/2}\nu^n|(|d_{i+1/2}\nu^n|/|D_{i+1/2}\nu^n| - \frac{1}{2})^2.$$

Thus (5.10b) is the essential difference between the present scheme and the ‘‘second-order’’ TVD schemes.

A similar analysis, which we do not present here, shows that if ν_i^n is a mesh-extremum then ν_j^{n+1} , $j = i, i + 1$, relates to $Q(x_j - a\tau; \nu^n)$ in the following way:

$$(5.11a) \quad \nu_j^{n+1} \geq Q(x_j - a\tau; \nu^n), \quad j = i, i + 1 \quad \text{if } \nu_i^n \text{ is a maximum,}$$

$$(5.11b) \quad \nu_j^{n+1} \leq Q(x_j - a\tau; \nu^n), \quad j = i, i + 1 \quad \text{if } \nu_i^n \text{ is a minimum.}$$

From (5.9)–(5.11) we deduce the following relation between the total variation of the numerical solution and that of its piecewise-parabolic interpolant Q :

$$(5.12) \quad TV(\{Q(x_j - a\tau; \nu^n)\}) \leq TV(\nu^{n+1}) \leq TV(Q(\cdot; \nu^n)).$$

The LHS of (5.12) is the total variation of the mesh function $\{Q(x_j - a\tau; \nu^n)\}$. Relation (5.12) suggests to consider an equivalent definition \overline{TV} of the total variation of the numerical approximation of the form

$$\overline{TV}(\nu^n) = TV(Q(\cdot; \nu^n))$$

with the hope that the scheme (5.3) is TVD with respect to this modified definition. Unfortunately our numerical experiments have shown that there are instances, although rather rare, that $\overline{TV}(\nu^n)$ is increasing with n ; the same is true for $\overline{TV}(\nu^n) = TV(L(\cdot; \nu^n))$.

As we have mentioned in the Introduction, because of the nonoscillatory nature of the scheme, uniform total variation boundedness of the numerical solution is implied by uniform boundedness of its maximum norm. If we follow a particular local maximum of the initial data we see from (5.9)–(5.10a) that it actually decreases most of the time, and whenever it does increase (5.10c) and (3.10) suggest that it does so by a ‘‘small amount’’ that vanishes with $h \rightarrow 0$. Since the initial data is only piecewise-smooth we have not been able as yet to rigorize these arguments.

We remark that our numerical experiments clearly indicate that in a normal computational situation the maximum norm of the numerical solution is indeed uniformly bounded. We feel that our inability to prove this fact stems only from lack of theoretical tools to analyze pointwise regularity of the numerical solution.

6. The nonlinear case. In this section we describe an approximate solution $\nu_n(x, t)$ of [5] for the IVP (2.2)

$$(6.1) \quad \nu_t + f(\nu)_x = 0, \quad \nu(x, 0) = L(x; \nu^n).$$

This approximate solution is consistent with the conservation form of the equation (6.1) in the sense that the cell-averaging (2.3) results in a scheme in conservation form, i.e.,

$$(6.2) \quad \nu_j^{n+1} = \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} \nu_n(x, \tau) dx = \nu_j^n - \lambda (\hat{f}_{j+1/2} - \hat{f}_{j-1/2})$$

where the numerical flux $\hat{f}_{j+1/2}$ is consistent with $f(u)$ in the sense of (1.2c). Furthermore, the approximate solution operator is TVD

$$(6.3) \quad TV(\nu_n(\cdot; t)) \leq TV(\nu_n(\cdot; 0)) = TV(L(\cdot; \nu^n)) \quad \text{for } 0 \leq t \leq \tau,$$

and thus by the reasoning presented in § 2, the resulting scheme (6.2) is nonoscillatory.

We turn now to outline the derivation of this approximate solution. To simplify our presentation we ignore entropy considerations and refer the reader to future papers for details of appropriate modifications. We assign to the point $x_{j+1/2}$ a characteristic speed that corresponds to the Rankine-Hugoniot speed $\bar{a}_{j+1/2}$ of the two neighboring cell-averages ν_j^n and ν_{j+1}^n

$$(6.4a) \quad \bar{a}_{j+1/2} = \begin{cases} \frac{f(\nu_{j+1}^n) - f(\nu_j^n)}{\nu_{j+1}^n - \nu_j^n} & \text{if } \nu_j^n \neq \nu_{j+1}^n, \\ a(\nu_j^n) & \text{if } \nu_j^n = \nu_{j+1}^n, \end{cases}$$

and denote by $\bar{a}(x)$ the piecewise-linear interpolant of $\{\bar{a}_{j+1/2}\}$, i.e.,

$$(6.4b) \quad \bar{a}(x) = \bar{a}_{j-1/2} + (\bar{a}_{j+1/2} - \bar{a}_{j-1/2}) \cdot (x - x_{j-1/2})/h \quad \text{for } x_{j-1/2} \leq x \leq x_{j+1/2}.$$

The approximate solution $\nu_n(x, t)$ is defined by specifying its constancy along the approximate characteristics

$$(6.5a) \quad x(t) = x_0 + \bar{a}(x_0) \cdot t,$$

i.e.,

$$(6.5b) \quad \nu_n(x(t), t) \equiv \nu_n(x_0, 0) = L(x_0; \nu^n).$$

Using (6.5a) and (6.4b) to express x_0 in terms of x and t

$$(6.5c) \quad x_0(x, t) = x_{j-1/2} + h \cdot [x - x_{j-1/2}(t)]/[x_{j+1/2}(t) - x_{j-1/2}(t)] \quad \text{for } x_{j-1/2}(t) < x < x_{j+1/2}(t)$$

we get from (6.5b) that

$$(6.6a) \quad \nu_n(x, t) = L\left(x_{j-1/2} + h \cdot \frac{x - x_{j-1/2}(t)}{x_{j+1/2}(t) - x_{j-1/2}(t)}; \nu^n\right) \quad \text{for } x_{j-1/2}(t) < x < x_{j+1/2}(t)$$

where

$$(6.6b) \quad x_{i+1/2}(t) = x_{i+1/2} + t \cdot \bar{a}_{i+1/2}.$$

Taking cell-averages of the approximate solution (6.6) we get the conservation form (6.2)

$$(6.7a) \quad \nu_j^{n+1} = \nu_j^n - \lambda (\hat{f}_{j+1/2} - \hat{f}_{j-1/2})$$

with the numerical flux

$$(6.7b) \quad \hat{f}_{j+1/2} = \begin{cases} f(\nu_j^n) + \frac{1}{2}\bar{a}_{j+1/2}(1 - \lambda\bar{a}_{j-1/2}) \cdot \hat{S}_j & \text{if } \bar{a}_{j+1/2} \geq 0, \\ f(\nu_{j+1}^n) - \frac{1}{2}\bar{a}_{j+1/2}(1 + \lambda\bar{a}_{j+3/2}) \cdot \hat{S}_{j+1} & \text{if } \bar{a}_{j+1/2} \leq 0, \end{cases}$$

where

$$(6.7c) \quad \hat{S}_j = S_j^n / [1 + \lambda(\bar{a}_{j+1/2} - \bar{a}_{j-1/2})].$$

Note that (6.7) is identical to (5.30) in the constant coefficient case.

We turn now to prove that the scheme (6.7) is uniformly second-order accurate in the sense of (2.9). We start with the exact cell-averages $\nu_j^n = \bar{u}_j^n$ in (6.7); this amounts to showing that

$$(6.8a) \quad \hat{f}_{j+1/2} = \hat{f}_{j+1/2}(u) + O(h^2)$$

with a sufficiently smooth coefficient in the $O(h^2)$ term; here $\hat{f}_{j+1/2}$ is the numerical flux (6.7b) computed with the exact cell-averages, and $\hat{f}_{j+1/2}(u)$ is (2.1b). We shall do this in two steps: first we shall show that

$$(6.8b) \quad \hat{f}_{j+1/2}(u) = \frac{1}{2}[f(L(x_{j+1/2}; \bar{u}^n)) + f(L(x_0(x_{j+1/2}, \tau); \bar{u}^n))] + O(h^2)$$

where $x_0(x_{j+1/2}, \tau)$ is (6.5c), and then we shall verify that

$$(6.8c) \quad \frac{1}{2}[f(L(x_{j+1/2}; \bar{u}^n)) + f(L(x_0(x_{j+1/2}, \tau); \bar{u}^n))] = \hat{f}_{j+1/2} + O(h^2).$$

Special attention will be given to the smoothness of the $O(h^2)$ coefficients.

To show (6.8b) we start by using the trapezoidal rule to approximate the integral in (2.1b); we get

$$(6.9a) \quad \hat{f}_{j+1/2}(u) = \frac{1}{2}[f(u(x_{j+1/2}, t_n)) + f(u(x_{j+1/2}, t_n + \tau))] + O(h^2).$$

The smoothness of the $O(h^2)$ term follows from that of $f(u)$ and $u(x, t)$. Next we observe that $\bar{a}(x)$ in (6.4b) approximates $a(u(x, t_n))$ to $O(h^2)$, and therefore we can use the approximate characteristic line (6.5c) to trace $u(x_{j+1/2}, t_n + \tau)$ to $u(x_0(x_{j+1/2}, \tau), t_n)$ with $O(h^3)$ accuracy; consequently

$$(6.9b) \quad f(u(x_{j+1/2}, t_n + \tau)) = f(u(x_0(x_{j+1/2}, \tau), t_n)) + O(h^3).$$

Finally we obtain (6.8b) by approximating $u(x, t_n)$ in (6.9a) and (6.9b) to $O(h^2)$ by $L(x; \bar{u}^n)$ (see (4.4)). The smoothness of the $O(h^2)$ term in this approximation is due to (4.4c):

$$S_j^n = h \cdot u_x(x, t_n) + O(h^3).$$

(We recall that the degeneracy to first-order accuracy at local extrema points of some "second-order accurate" TVD schemes is due to lack of smoothness there of the $O(h^2)$ term in (2.7a).)

We turn now to verify (6.8c). First let us consider the case $\bar{a}_{j+1/2} \geq 0$:

$$\begin{aligned} L(x_{j+1/2}; \bar{u}^n) &= \bar{u}_j^n + \frac{1}{2}S_j^n, \\ L(x_0(x_{j+1/2}, \tau); \bar{u}^n) &= \bar{u}_j^n + \left[\frac{1}{2} - \frac{\lambda\bar{a}_{j+1/2}}{1 + \lambda(\bar{a}_{j+1/2} - \bar{a}_{j-1/2})} \right] S_j^n, \end{aligned}$$

and expand the LHS of (6.8c) around \bar{u}_j^n . We get

$$(6.10a) \quad \begin{aligned} f(\bar{u}_j^n) + \frac{1}{2}a'(\bar{u}_j^n)(1 - \lambda\bar{a}_{j+1/2})\hat{S}_j + \frac{1}{8}a'(\bar{u}_j^n)[(1 - \lambda\bar{a}_{j-1/2})^2 + (\lambda\bar{a}_{j+1/2})^2](\hat{S}_j)^2 + O(h^3) \\ = \hat{f}_{j+1/2} + \frac{h^2}{8}(2\lambda^2 a^2 - 1) \cdot a' \cdot (u_x)^2|_{j+1/2} + O(h^3). \end{aligned}$$

Similarly in the case $\bar{a}_{j+1/2} \leq 0$,

$$\begin{aligned} L(x_{j+1/2}; \bar{u}^n) &= \bar{u}_{j+1} - \frac{1}{2} S_{j+1}^n, L(x_0(x_{j+1/2}, \tau); \bar{u}^n) \\ &= \bar{u}_{j+1}^n - \left[\frac{1}{2} + \frac{\lambda \bar{a}_{j+1/2}}{1 + \lambda(\bar{a}_{j+3/2} - \bar{a}_{j+1/2})} \right] \cdot S_{j+1}^n, \end{aligned}$$

we expand the LHS of (6.8c) around \bar{u}_{j+1}^n to get

$$\begin{aligned} (6.10b) \quad & f(\bar{u}_{j+1}^n) - \frac{1}{2} a(\bar{u}_{j+1}^n)(1 + \lambda \bar{a}_{j+3/2}) \hat{S}_{j+1} \\ & + \frac{1}{8} a'(u_{j+1}^n)[(1 + \lambda \bar{a}_{j+3/2})^2 + (\lambda \bar{a}_{j+1/2})^2] \cdot (\hat{S}_{j+1})^2 + O(h^3) \\ & = \hat{f}_{j+1/2} + \frac{h^2}{8} (2\lambda^2 a^2 - 1) \cdot a' \cdot (u_x)^2|_{j+1/2} + O(h^3). \end{aligned}$$

We see from (6.10a) and (6.10b) that independently of the sign of $\bar{a}_{j+1/2}$, the $O(h^2)$ term in (6.8c) is the same, namely

$$\frac{h^2}{8} (2\lambda^2 a^2 - 1) \cdot a' \cdot (u_x)^2|_{j+1/2}.$$

This completes the proof that the scheme (6.7) is second-order accurate in the sense of (2.9) wherever $u(x, t)$ is smooth, including local extrema and sonic ($f' = 0$) points.

Remarks. (1) The numerical flux (6.7b) can be rewritten as

$$(6.11) \quad \begin{aligned} \hat{f}_{j+1/2} &= \frac{1}{2}[f(v_j^n) + f(v_{j+1}^n)] - |\bar{a}_{j+1/2}|(v_{j+1}^n - v_j^n) + \max(0, \bar{a}_{j+1/2}) \cdot (1 - \lambda \bar{a}_{j-1/2}) \cdot \hat{S}_j \\ &\quad - [\min(0, \bar{a}_{j+1/2}) \cdot (1 + \lambda \bar{a}_{j+3/2}) \cdot \hat{S}_{j+1}]. \end{aligned}$$

(2) Our proof that the scheme (6.7) is nonoscillatory is based on the representation of (6.7) as the cell-average (6.2) of the nonoscillatory approximate solution $v_n(x, t)$ in (6.6). To ensure that $v_n(x, t)$ remains univalued for $0 \leq t \leq \tau$ we have to restrict the time-step τ so that for all j

$$(6.12a) \quad x_{j+1/2}(\tau) > x_{j-1/2}(\tau).$$

This implies the condition

$$(6.12b) \quad \tau \max_j (\bar{a}_{j-1/2} - \bar{a}_{j+1/2}) \leq h.$$

On the other hand, to derive the particular form of the numerical flux (6.7b) we have made the assumption

$$(6.13a) \quad |x_{j+1/2}(\tau) - x_{j+1/2}| < h \quad \text{for all } j,$$

which implies the condition

$$(6.13b) \quad \tau \max_j |\bar{a}_{j+1/2}| < h.$$

The two time-step restrictions (6.12b) and (6.13b) are characteristic to Godunov-type schemes. The practice of reconciling the two conditions by

$$\tau \max_j |\bar{a}_{j+1/2}| \leq h/2$$

is completely unnecessary: The scheme (6.7), as the original Godunov scheme, remains nonoscillatory (or TVD in the case analyzed in [10]) for the full CFL-restriction (6.13b). The reasoning for this observation is as follows: The approximate solution (6.6) under the CFL restriction (6.13b) may fail to be univalued in the j th cell only if $\bar{a}_{j-1/2} > 0$ and $\bar{a}_{j+1/2} < 0$. In this case the numerical fluxes $\hat{f}_{j\pm 1/2}$ as defined by cell-averaging in the neighboring cells $j \pm 1$, remain (6.7b). Thus the only thing that needs to be changed in this case is the definition of $v_n(x, \tau)$ in the j th cell.

(3) We observe that once $\nu_n(x, \tau)$ is defined globally in (6.6) there is no intrinsic need to average it on the original mesh. We may average it on different intervals and still conclude that the resulting approximation is nonoscillatory and conservative. Furthermore, the construction of the interpolant Q , the approximation L and the approximate characteristic field $\bar{a}(x)$ needed to define $\nu_n(x, t)$, does not depend on the uniformity of the mesh. Therefore the scheme (6.7) generalizes immediately to nonuniform moving meshes. Of particular computational interest are the self-adjusting moving grids of the type described in [12], which make it possible to obtain perfectly resolved shocks and contact discontinuities.

(4) We note that since the approximate solution $\nu_n(x, t)$ in (6.6) is conservative, it is possible to consider an associated random-choice method obtained by replacing the cell-averaging in (6.2) by a sampling with respect to a variable that is uniformly distributed in the cell, i.e.,

$$\nu_j^{n+1} = \nu_n(x_j + \theta_j^n h, \tau)$$

where θ_j^n is uniformly distributed in $[-\frac{1}{2}, \frac{1}{2}]$. Following the reasoning of [7] it is clear that the resulting random-choice method is nonoscillatory and that its limits are weak solutions of (1.1). Although the random-choice approach has many attractive computational features, it has been our experience that in many applications it is possible to accomplish the same computational goals with a self-adjusting moving grid. In this case the use of the latter is preferable as it offers gain in resolution without a loss in pointwise accuracy that is associated with sampling.

7. Numerical illustration. In this section we compare the new uniformly second-order nonoscillatory scheme of this paper (to be referred to as UNO2) to the typical second-order TVD scheme (to be referred to as TVD2). Both schemes can be written in the form (6.7), i.e.,

$$(7.1a) \quad \nu_j^{n+1} = \nu_j^n - \lambda(\hat{f}_{j+1/2} - \hat{f}_{j-1/2}),$$

(7.1b)

$$\hat{f}_{j+1/2} = \begin{cases} f(\nu_j^n) + \frac{1}{2}\bar{a}_{j+1/2}(1 - \lambda\bar{a}_{j-1/2})S_j^n / [1 + \lambda(\bar{a}_{j+1/2} - \bar{a}_{j-1/2})] & \text{if } \bar{a}_{j+1/2} \geq 0, \\ f(\nu_{j+1}^n) - \frac{1}{2}\bar{a}_{j+1/2}(1 + \lambda\bar{a}_{j+3/2})S_{j+1}^n / [1 + \lambda(\bar{a}_{j+3/2} - \bar{a}_{j+1/2})] & \text{if } \bar{a}_{j+1/2} \leq 0, \end{cases}$$

$$(7.1c) \quad S_j^n = m(S_j^+, S_j^-);$$

here $\bar{a}_{j+1/2}$ is (6.4a) and $m(x, y)$ is the min mod function (3.3). S_j^\pm are different for TVD2 and UNO2:

$$(7.2) \quad \text{TVD2: } S_j^\pm = d_{j\pm 1/2} \nu^n,$$

$$(7.3) \quad \text{UNO2: } S_j^\pm = d_{j\pm 1/2} \nu^n \mp \frac{1}{2} D_{j\pm 1/2} \nu^n,$$

where $d_{i+1/2}$ and $D_{i+1/2}$ are defined in (3.2).

UNO2 and TVD2 are both second-order accurate Godunov-type schemes that differ only in the reconstruction step (4.2a):

$$(7.4) \quad L(x; u) = u_j + S_j(x - x_j)/h \quad \text{for } |x - x_j| < \frac{h}{2},$$

where the slopes of the lines are calculated by (7.3) and (7.2), respectively. Therefore we start our comparison on the approximation level.

In Table 1 and Fig. 1 we present approximations to $u(x) = \sin \pi x$, $-1 \leq x \leq 1$. We divide $[-1, 1]$ into N equal intervals and define

$$(7.5) \quad x_j = -1 + j \cdot \frac{2}{N}, \quad 0 \leq j \leq N.$$

TABLE 1
Approximations to $u(x) = \sin(\pi x)$, $-1 \leq x \leq 1$, with periodic boundary conditions.

N	L_∞ -ERROR			L_1 -ERROR		
	Q	L_{UNO2}	L_{TVD2}	Q	L_{UNO2}	L_{TVD2}
10	1.545×10^{-2}	5.122×10^{-2}	1.420×10^{-1}	1.494×10^{-2}	2.467×10^{-2}	7.016×10^{-2}
20	1.971×10^{-3}	1.231×10^{-2}	3.558×10^{-2}	1.802×10^{-3}	5.576×10^{-3}	1.525×10^{-2}
40	2.476×10^{-4}	3.083×10^{-3}	9.163×10^{-3}	2.148×10^{-4}	1.355×10^{-3}	3.902×10^{-3}
80	3.104×10^{-5}	7.710×10^{-4}	2.308×10^{-3}	2.617×10^{-5}	3.351×10^{-4}	9.787×10^{-4}

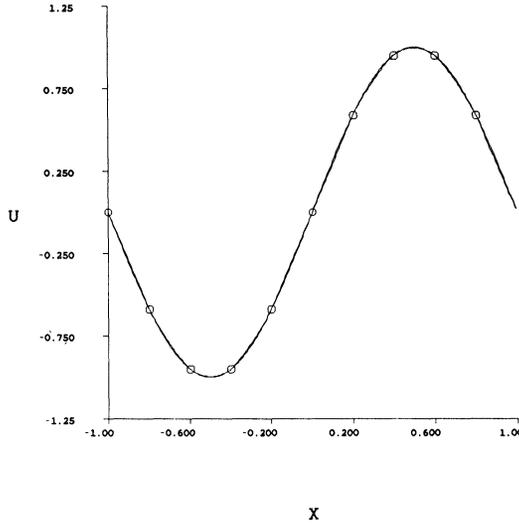


FIG. 1(a). *Approximations of $u = \sin \pi x$, $-1 \leq x \leq 1$, with $N = 10$. Piecewise-parabolic interpolant $Q(x, u)$.*

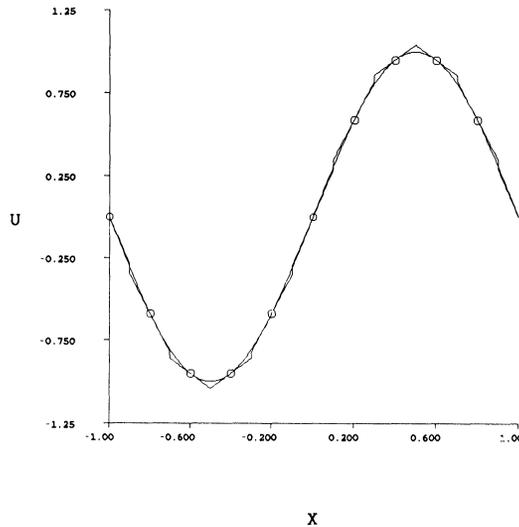


FIG. 1(b). *Approximations of $u = \sin \pi x$, $-1 \leq x \leq 1$, with $N = 10$. Piecewise-linear approximation $L^{UNO2}(x, u)$.*

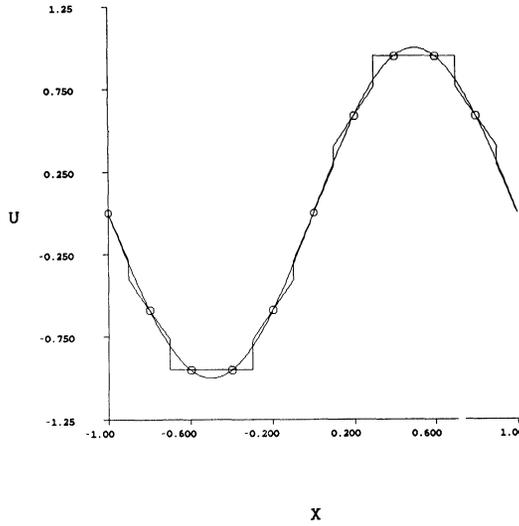


FIG. 1(c). Approximations of $u = \sin \pi x$, $-1 \leq x \leq 1$, with $N = 10$. Piecewise-linear approximation $L^{TVD2}(x; u)$.

The symbols in Fig. 1 denote values of $u_j = \sin \pi x_j$ for $N = 10$ in (7.5). In Fig. 1(a) we show the piecewise-parabolic interpolant $Q(x; u)$ (see § 3). In Fig. 1(b) we show the piecewise-linear approximation $L^{UNO2}(x; u)$ which is (7.4) with (7.1c) and (7.3). In Fig. 1(c) we show the piecewise-linear approximation $L^{TVD2}(x; u)$ which is (7.4) with (7.1c) and (7.2). We make the following observations regarding Fig. 1: (i) Q is a better approximation than L^{UNO2} ; L^{UNO2} is a better approximation than L^{TVD2} . (ii) $TV(L^{UNO2}) > TV(u) > TV(L^{TVD2})$. In Table 1 we quantify the first observation; we list the L_∞ -error and the L_1 -error of these approximations to $\sin \pi x$ for a refinement sequence of $N = 10, 20, 40, 80$ in (7.5). Clearly Q is an $O(h^3)$ approximation, while L^{UNO2} and L^{TVD2} are $O(h^2)$. The error in L^{UNO2} is about a $\frac{1}{3}$ of the error in L^{TVD2} .

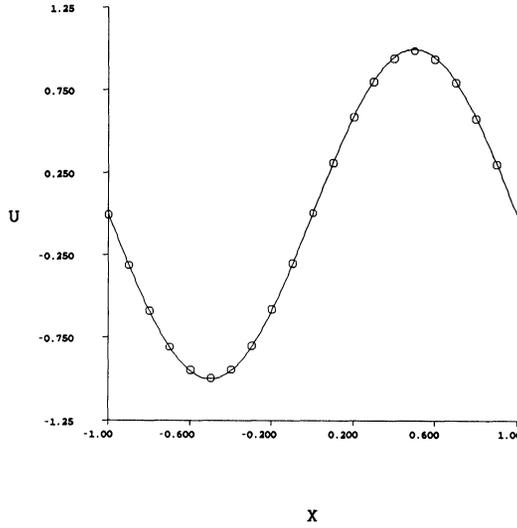
In Table 2 and Fig. 2 we present solutions of UNO2 and TVD2 for the constant coefficient case

$$(7.6) \quad u_t + u_x = 0, \quad u(x, 0) = \sin \pi x, \quad -1 \leq x \leq 1,$$

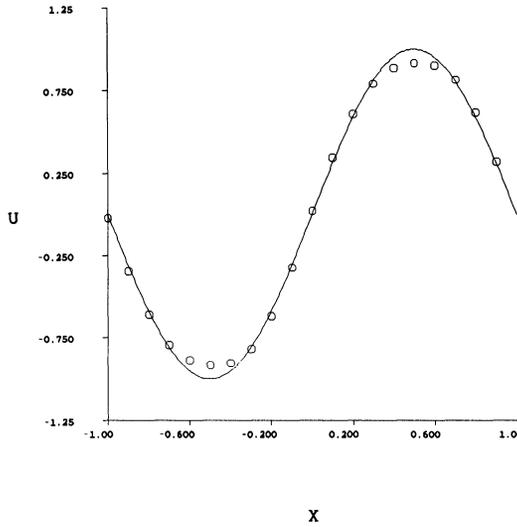
with periodic boundary conditions, at $t = 2$ with $\tau/h = 0.8$. Figures 3(a) and 3(b) show UNO2 and TVD2, respectively, with $N = 20$ in (7.5). In Table 2 we list the L_∞ -error and L_1 -error for a refinement sequence with $N = 20, 40, 80, 160$. Clearly UNO2 is second-order accurate in both L_∞ and L_1 , while TVD2 is second-order accurate in L_1 but only first-order accurate in L_∞ . Figure 3(b) demonstrates that the degeneracy to

TABLE 2
 Numerical solutions of $u_t + u_x = 0, u(x, 0) = \sin \pi x, -1 \leq x \leq 1$ at $t = 2$ with periodic boundary conditions and $\tau/h = 0.8$.

N	L_∞ -ERROR		L_1 -ERROR	
	UNO2	TVD2	UNO2	TVD2
20	7.097×10^{-3}	8.119×10^{-2}	8.944×10^{-3}	6.778×10^{-2}
40	1.607×10^{-3}	3.477×10^{-2}	2.044×10^{-3}	2.033×10^{-2}
80	3.870×10^{-4}	1.453×10^{-2}	4.926×10^{-4}	5.626×10^{-3}
160	9.201×10^{-5}	5.975×10^{-3}	1.172×10^{-4}	1.528×10^{-3}



(a)



(b)

FIG. 2. Numerical solutions of $u_t + u_x = 0$, $u(x, 0) = \sin \pi x$ at $t = 2$, with $N = 20$ and $\tau/h = 0.8$. (a) UNO2. (b) TVD2.

first-order accuracy at local extrema in the TVD scheme adversely affects the accuracy everywhere. (Because the scheme is TVD it cannot switch abruptly to second-order accuracy as this would introduce oscillations; consequently it takes quite a while to recover the second-order accuracy.)

Next we approximate the discontinuous function

$$(7.7) \quad u(x) = \begin{cases} -x \sin(3\pi x^2/2), & -1 < x < -\frac{1}{3}, \\ |\sin(2\pi x)|, & |x| < \frac{1}{3}, \\ 2x - 1 - \frac{1}{6} \sin(3\pi x), & \frac{1}{3} < x < 1, \end{cases}$$

which we extend to have period 2 outside $[-1, 1]$.

In Fig. 3 we present approximations to $\phi(x)$, using $N = 20$. Figure 3(a) shows $Q(x; \bar{u})$, Fig. 3(b) shows $L^{\text{UNO2}}(x; \bar{u})$, and Fig. 3(c) shows $L^{\text{TVD2}}(x; \bar{u})$. We again observe that Q is a better approximation than L^{UNO2} , while L^{UNO2} is a better approximation than L^{TVD2} .

In Fig. 4 we present solutions of UNO2 and TVD2 for the constant coefficient problem (7.6), initial data given by (7.2), and periodic boundary conditions. We take $t = 2$ and $\tau/h = 0.8$. Figures 4(a) and 4(b) show UNO2 and TVD2, respectively, with $N = 40$. Figure 4(b) shows the damping effect that the TVD scheme imposes due to its degeneracy to first-order accuracy at local extrema.

In Fig. 5 we solve the same problems, except we impose boundary conditions. At $x = -1$ we impose the given function (7.7) evaluated at $-1 - t$. No boundary conditions are imposed at $x = 1$. We implement this numerically using UNO2 and TVD2 except

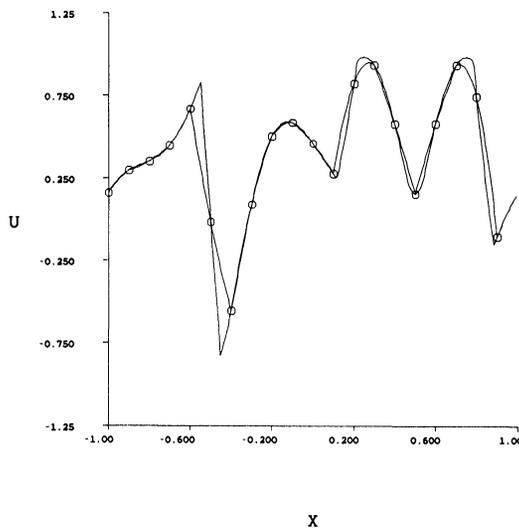


FIG. 3(a). $Q(x; \bar{u})$ for u given (7.7) with $N = 20$.

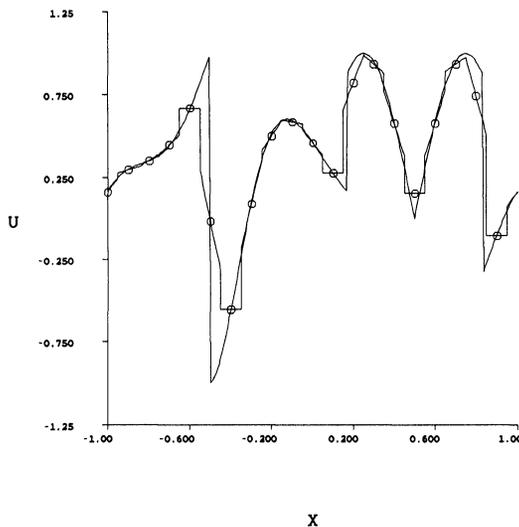


FIG. 3(b). $L^{\text{UNO2}}(x; \bar{u})$ for u given by (7.7) with $N = 20$.

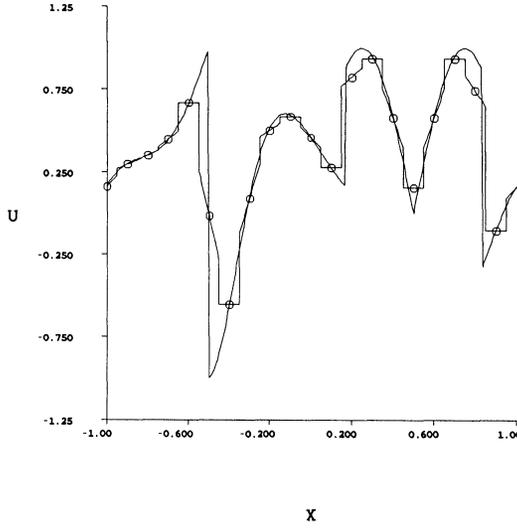


FIG. 3(c). $L^{TVD2}(x; \bar{u})$ for u given by (7.7) with $N = 20$.

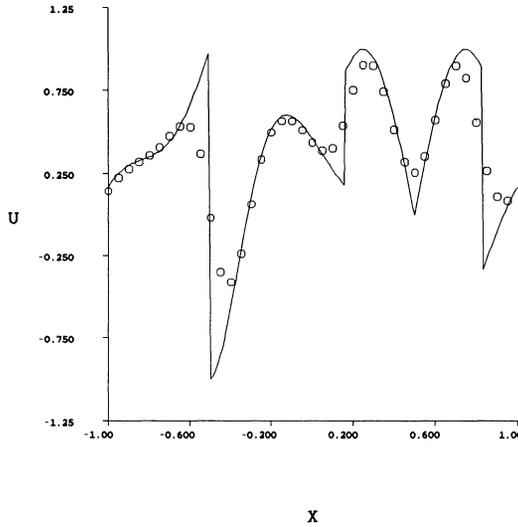


FIG. 4(a). Numerical solution of $u_t + u_x = 0$, $u(x, 0)$ defined by (7.7) at $t = 2$ with periodic boundary conditions $N = 40$ and $\tau/h = 0.8$, using UNO2.

at the boundary points. There we are, in general, unable to construct nonoscillatory piecewise-parabolic interpolants $Q(x, \bar{u})$, so we construct the only possible parabolic interpolant through x_i, x_{i+1} and the point to either the left or right which lies in the region. The analogous procedure is carried out at the reconstruction stage. Figures 5(a) and 5(b) again show the results at $t = 2$ with $\tau/h = 0.8$.

The possible introduction of oscillations through the boundary conditions does not seem to have degraded the performance of either scheme (in fact the opposite is observed). Again the TVD2 scheme shows a damping effect.

In Table 3 and Fig. 6 we present results for Burgers' equation

$$(7.8) \quad u_t + uu_x = 0, \quad u(x, 0) = \alpha + \sin \pi(x + \beta), \quad -1 \leq x \leq 1,$$

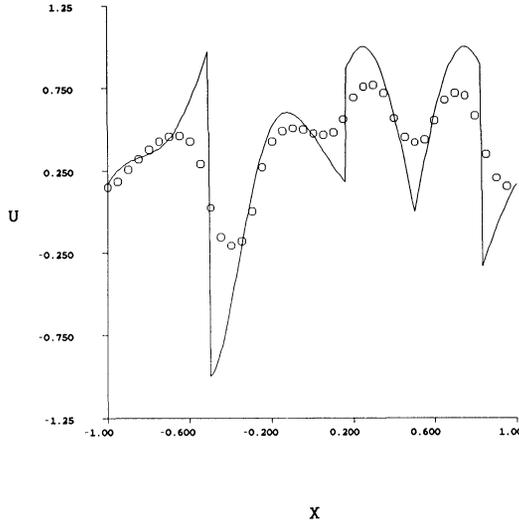


FIG. 4(b). Same as Fig. 4(a) using TVD2.

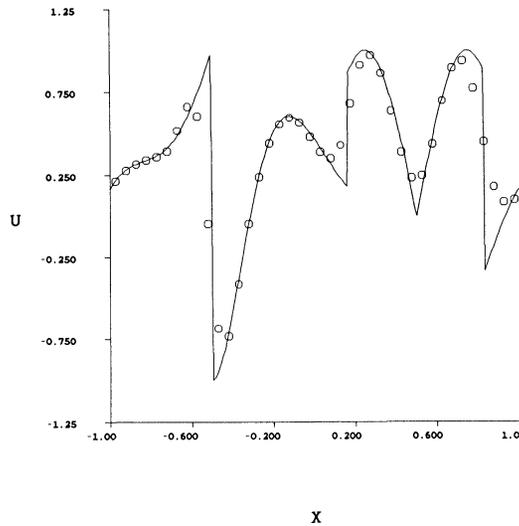


FIG. 5(a). Same as Fig. 4(a) except with given data at $x = -1$, and outflow at $x = 1$.

with periodic boundary conditions and $\tau/h(1+|\alpha|) = 0.5$. The solution to (7.7) is smooth for $t < 1/\pi$; at $t = 1/\pi$ it develops shocks. In Table 3(a) we list the L_∞ -error and L_1 -error of UNO2 and TVD2 at $t = 0.15$ for $\alpha = \beta = 0$ in (7.7). This table shows the same asymptotic behavior as Table 2, except that because of the large gradients it shows for a smaller h .

In Figs. 6(a) and 6(b) we show results of UNO2 and TVD2 for (7.8) with $\alpha = 2$ and $\beta = 1$ at $t = 0.35$ with $N = 20$. In this case the solution to (7.8) develops a shock moving with speed 2 beginning at time $t = 1/\pi \approx 0.318$.

In Table 3(b) and Figs. 6(c) and 6(d) we repeat the previous calculations for the schemes (2.13):

$$(7.9a) \quad \nu_j^{n+1} = \nu_j^n - \lambda(\hat{f}_{j+1/2} - \hat{f}_{j-1/2}),$$

$$(7.9b) \quad \hat{f}_{j+1/2} = f(\nu_n(x_{j+1/2}, \tau/2)) = f(L(x_0(x_{j+1/2}, \tau/2), \nu^n))$$

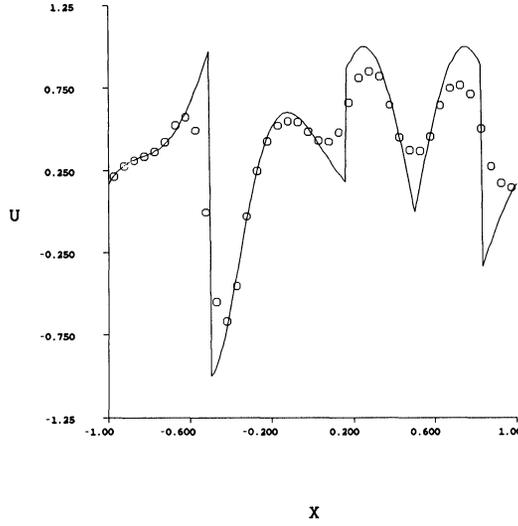


FIG. 5(b). Same as Fig. 4(b) except with given data at $x = -1$, and outflow at $x = 1$.

TABLE 3(a)

Numerical solutions of $u_t + uu_x = 0$, $u(x, 0) = \sin \pi x$, at $t = 0.15$ and $\tau/h = 0.5$ - UNO2 and TVD2.

N	L_∞ -ERROR		L_1 -ERROR	
	UNO2	TVD2	UNO2	TVD2
20	1.890×10^{-2}	2.238×10^{-2}	1.090×10^{-2}	1.854×10^{-2}
40	5.712×10^{-3}	1.054×10^{-2}	3.034×10^{-3}	5.051×10^{-3}
80	1.552×10^{-3}	4.422×10^{-3}	7.771×10^{-4}	1.340×10^{-3}
160	3.985×10^{-4}	1.837×10^{-3}	1.965×10^{-4}	3.621×10^{-4}

TABLE 3(b)

Same as Table 3(a) for FNO2 and FVD2.

N	L_∞ -ERROR		L_1 -ERROR	
	FNO2	FVD2	FNO2	FVD2
20	6.938×10^{-3}	2.091×10^{-2}	3.726×10^{-3}	1.322×10^{-2}
40	1.050×10^{-3}	1.054×10^{-2}	8.869×10^{-4}	3.835×10^{-3}
80	5.106×10^{-4}	4.424×10^{-3}	2.163×10^{-4}	1.072×10^{-3}
160	1.251×10^{-4}	1.837×10^{-3}	5.270×10^{-5}	2.946×10^{-4}

TABLE 3(c)

Same as Table 3(a) for ANO2.

N	L_∞ -ERROR	L_1 -ERROR
20	2.249×10^{-2}	1.221×10^{-2}
40	6.623×10^{-3}	3.243×10^{-3}
80	1.781×10^{-3}	8.259×10^{-4}
160	4.597×10^{-4}	2.079×10^{-4}

where $x_0(x_{j+1/2}, \tau/2)$ is (6.5c), i.e.,

$$(7.9c) \quad \hat{f}_{j+1/2} = \begin{cases} f\left(\nu_j^n + \frac{1}{2} \cdot \frac{1 - \lambda(\hat{a}_{j+1/2} + \hat{a}_{j-1/2})/2}{1 + \lambda(\hat{a}_{j+1/2} - \hat{a}_{j-1/2})/2} \cdot S_j^n\right) & \text{if } \hat{a}_{j+1/2} \geq 0, \\ f\left(\nu_{j+1}^n - \frac{1}{2} \cdot \frac{1 + \lambda(\hat{a}_{j+3/2} + \hat{a}_{j+1/2})/2}{1 + \lambda(\hat{a}_{j+3/2} - \hat{a}_{j+1/2})/2} \cdot S_{j+1}^n\right) & \text{if } \hat{a}_{j+1/2} \leq 0. \end{cases}$$

As we have remarked in § 2, ν_j^{n+1} in (7.9a) is not a cell-average of $\nu_n(x, \tau)$, but only an approximation to it. Therefore it is not necessary to take $\hat{a}_{j+1/2}$ in (7.8c) to be (6.4a). We choose $\hat{a}_{j+1/2}$ so that (7.9c) is continuous at $\hat{a}_{j+1/2} = 0$:

$$(7.9d) \quad \hat{a}_{j+1/2} = [f(\nu_{j+1}^n - \frac{1}{2}S_{j+1}^n) - f(\nu_j^n + \frac{1}{2}S_j^n)] / [(\nu_{j+1}^n - \frac{1}{2}S_{j+1}^n) - (\nu_j^n + \frac{1}{2}S_j^n)].$$

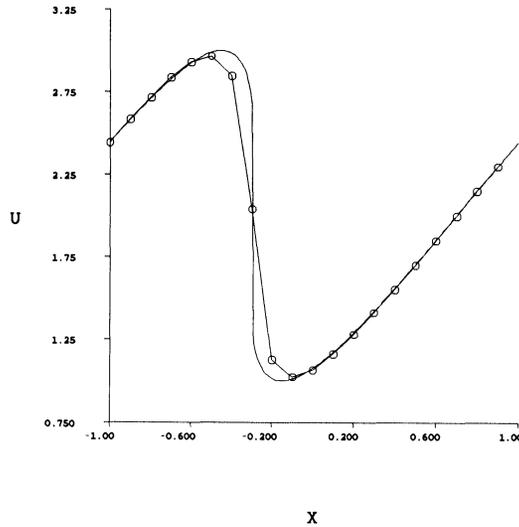


FIG. 6(a). Solution to (7.8) with $\alpha = 2, \beta = 1, CFL = .5, N = 20$, at $t = 0.35$ using UNO2.

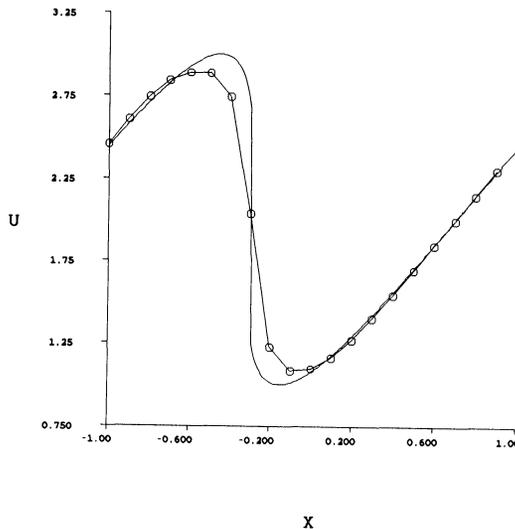


FIG. 6(b). Same as Fig. 6(a) using TVD2.

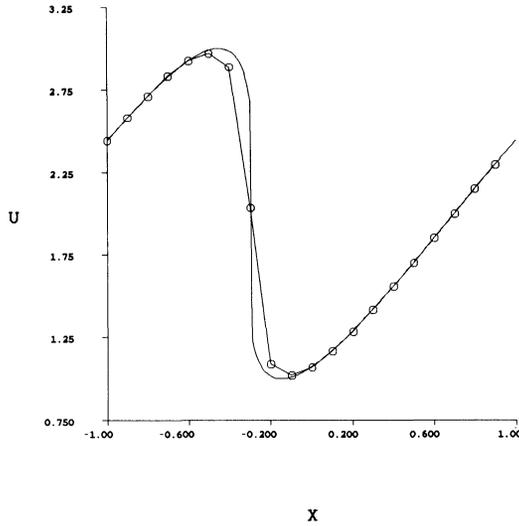


FIG. 6(c). Same as Fig. 6(a) using FNO2.

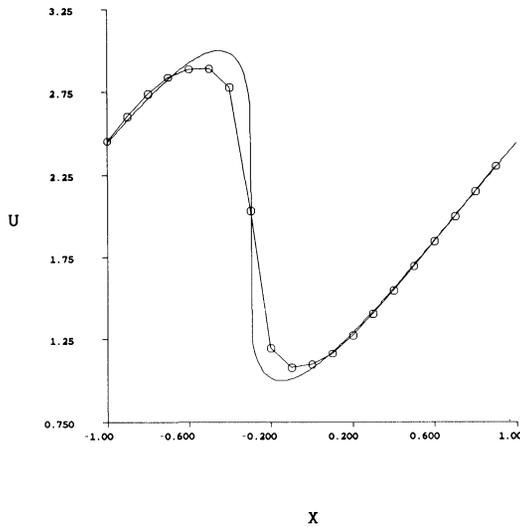


FIG. 6(d). Same as Fig. 6(a) using FVD2.

We denote the schemes (7.9) with S_j^n defined by (7.1c) and either (7.2) or (7.3) by FVD2 and FNO2, respectively. We note that (7.9) is identical to (7.1) in the constant coefficient case, and consequently FVD2 and FNO2 are nonoscillatory in the constant coefficient case. Figures 6(c) and 6(d) show that FNO2 and FVD2 are also nonoscillatory in the case (7.8). Furthermore, Table 3(b) shows that FNO2 is much more accurate than UNO2 (FVD2 is about the same as TVD2).

In all previous examples we have presented pointwise calculations; namely, we have initialized the numerical solution by taking v_j^0 to be the value of the initial data at x_j , and we have considered v_j^n to be an approximation to $u(x_j, t_n)$. (Surely this is an acceptable practice for second-order accurate schemes.) In Table 3(c) we repeat the calculation for UNO2 in Table 3(a), but now in a sense of cell-averages and denote it by ANO2. Now we initialize UNO2 for (7.8) with $\alpha = \beta = 0$ by cell-averages of the

initial data, i.e.,

$$(7.10a) \quad \nu_j^0 = -\frac{1}{\pi h} [\cos(\pi x_{j+1/2}) - \cos(\pi x_{j-1/2})] = \frac{\sin(\pi h/2)}{(\pi h/2)} \cdot \sin(\pi x_j),$$

and regard ν_j^n to represent cell-averages of $u(x, t_n)$. To obtain a pointwise approximation to $u(x, t_n)$ we first compute point values $V_{j+1/2}^n$ of its indefinite integral $u^n(x_{j+1/2}) = \int_{x_0}^{x_{j+1/2}} u(y, t_n) dy$ by

$$(7.10b) \quad V_{j+1/2}^n = h \sum_{i=i_0}^j \nu_i^n.$$

Next we obtain a global piecewise-linear approximation $\nu(x, t_n)$ to $u(x, t_n)$ by

$$(7.10c) \quad \nu(x, t_n) = \frac{d}{dx} Q(x; V^n)$$

where Q is the piecewise-parabolic interpolant of § 3. Finally we get

$$(7.10d) \quad \nu(x_j, t_n) = \frac{d}{dx} Q(x_j; V^n) = \frac{1}{h} (V_{j+1/2}^n - V_{j-1/2}^n) = \nu_j^n.$$

Thus the only difference between ANO2 in Table 3(c), and UNO2 in Table 3(a) is the initialization (7.10a), which itself differs only slightly from the mesh values of the initial data (since $\sin(\pi h/2)/(\pi h/2) = 1 - \frac{1}{6}(\pi h/2)^2 + O(h^4)$).

We remark that cell-averages do play a significant role when the initial data is discontinuous (since they provide information about the location of the discontinuity) and in higher-order Godunov-type schemes; this will be described elsewhere.

We turn now to present calculations with a formal extension of UNO2 and TVD2 for systems of conservation laws. We consider a Riemann problem for the Euler equations of gas dynamics

$$(7.11a) \quad u_t + f(u)_x = 0, \quad u(x, 0) = \begin{cases} u_L, & x < 0, \\ u_R, & x > 0, \end{cases}$$

$$(7.11b) \quad u = (\rho, m, E)^T, \quad f(u) = (m, m^2/\rho + P, m(E + P)/\rho)^T,$$

$$(7.11c) \quad P = (\gamma - 1)(E - \frac{1}{2}m^2/\rho).$$

Here ρ, m, E and P are the density, momentum, total energy and pressure, respectively; we take $\gamma = 1.4$.

In the following we apply the extension technique of [3] to UNO2 and TVD2. The idea is to extend UNO2 and TVD2 to systems in such a way that will be identical to (7.1) in the scalar case, and will decouple into (5.3) for each of the characteristic variables in the constant coefficient system case. To accomplish that we use Roe's averaging for (7.11) (see [13])

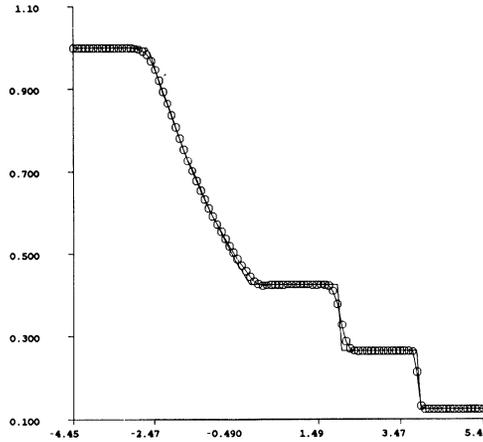
$$(7.12a) \quad \nu_{j+1/2} = V(\nu_j^n, \nu_{j+1}^n)$$

for which

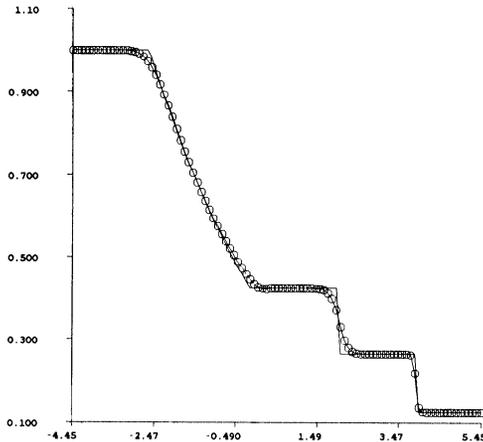
$$(7.12b) \quad f(\nu_{j+1}^n) - f(\nu_j^n) = A(\nu_{j+1/2})(\nu_{j+1}^n - \nu_j^n), \quad A(u) = \partial f / \partial u,$$

and define local characteristic variables with respect to the right-eigenvector system $\{R_{j+1/2}^k\}_{k=1}^3$ of $A(\nu_{j+1/2})$. We extend (6.11) to systems as follows:

$$(7.13a) \quad \nu_j^{n+1} = \nu_j^n - \lambda(\hat{f}_{j+1/2} - \hat{f}_{j-1/2}),$$



(a)



(b)

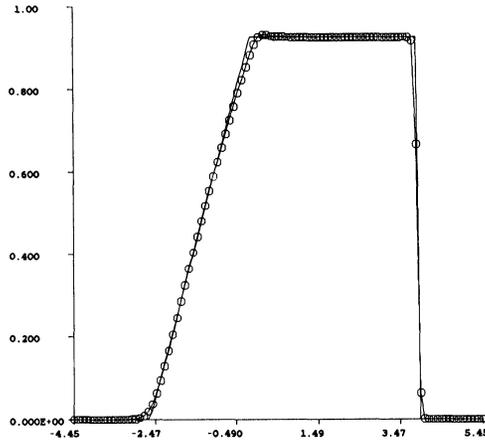
FIG. 7. Numerical solution of density in a Riemann problem for Euler's equations. (a) UNO2. (b) TVD2.

$$(7.13b) \quad \hat{f}_{j+1/2} = \frac{1}{2} \left[f(\nu_j^n) + f(\nu_{j+1}^n) - \sum_{k=1}^3 c_{j+1/2}^k R_{j+1/2}^k \right],$$

$$(7.13c) \quad c_{j+1/2}^k = |\bar{a}_{j+1/2}^k| d_{j+1/2}^k w - \max(0, \bar{a}_{j+1/2}^k) (1 - \lambda \bar{a}_{j-1/2}^k) \hat{S}_j^k + \min(0, \bar{a}_{j+1/2}^k) (1 + \lambda \bar{a}_{j+3/2}^k) \hat{S}_{j+1}^k.$$

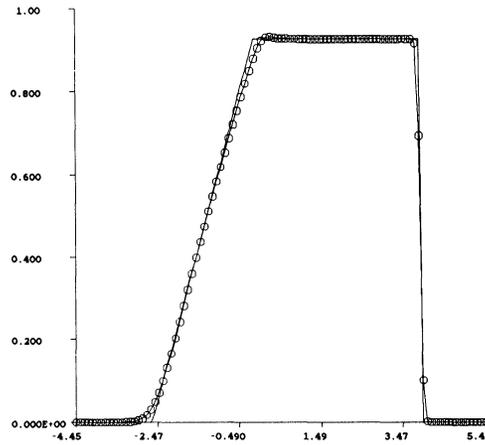
Here $\bar{a}_{j+1/2}^k$ is the k th eigenvalue of $A(\nu_{j+1/2})$ corresponding to $R_{j+1/2}^k$, and $d_{j+1/2}^k w$ denotes the component of $d_{j+1/2} \nu = \nu_{j+1}^n - \nu_j^n$ in the k th characteristic field, i.e.,

$$(7.13d) \quad d_{j+1/2} \nu = \sum_{k=1}^3 (d_{j+1/2}^k w) R_{j+1/2}^k.$$



X

(a)



X

(b)

FIG. 8. Numerical solution of velocity in a Riemann problem for Euler's equations. (a) UNO2. (b) TVD2.

Likewise \hat{S}_j^k denotes the component of the vector of slopes in the k th characteristic field, and is defined as follows:

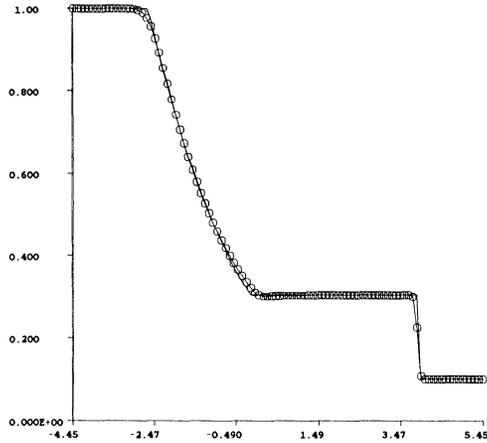
$$(7.13e) \quad \hat{S}_j^k = m(S_{-,j}^k, S_{+,j}^k) / [1 + \lambda(\bar{a}_{j+1/2}^k - \bar{a}_{j-1/2}^k)];$$

$m(x, y)$ is the min mod function (3.3). $S_{\pm,j}^k$ are different for VD2 and NO2:

$$(7.14) \quad \text{TVD2: } S_{\pm,j}^k = d_{j\pm 1/2}^k w,$$

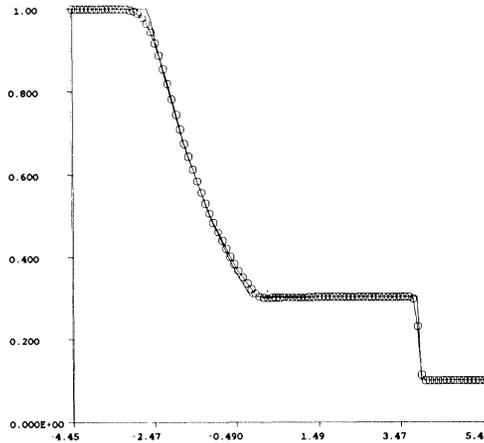
$$(7.15) \quad \text{UNO2: } S_{\pm,j}^k = d_{j+1/2}^k w \mp \frac{1}{2} D_{j\pm 1/2}^k w;$$

$$D_{i\pm 1/2}^k w = m(d_{i+3/2}^k w - d_{i+1/2}^k w, d_{i+1/2}^k w - d_{i-1/2}^k w).$$



X

(a)



X

(b)

FIG. 9. Numerical solution of pressure in a Riemann problem for Euler's equations. (a) UNO2. (b) TVD2.

In Figs. 7, 8 and 9 we show numerical solutions of UNO2 and TVD2, respectively, for the Riemann problem (7.1b) with

$$U_L = (1, 0, 2.5)^T, \quad U_R = (0.125, 0, 0.25).$$

These figures demonstrate that the formal extension to systems is nonoscillatory in this case. Since the solution to the Riemann problem is just constant states separated by waves we do not get to see here the extra resolution power of UNO2, except that its numerical solution is somewhat "crisper" than that of TVD2. In this calculation we have not employed any artificial compression in the linearly degenerate field and therefore the contact discontinuity smears like $n^{1/3}$, as expected. The interested reader is referred to [4], [5] and [10] for a detailed description of such compression techniques, as well as for details of entropy enforcement mechanisms.

REFERENCES

- [1] S. R. CHAKRAVARTHY AND S. OSHER, *A new class of high accuracy TVD schemes for hyperbolic conservation laws*, AIAA paper 85-0363, Reno, NV, 1985.
- [2] P. COLELLA AND P. R. WOODWARD, *The piecewise-parabolic method (PPM) for gas-dynamical simulations*, J. Comput. Phys., 54 (1984), pp. 174-201.
- [3] A. HARTEN, *On a class of high resolution total-variation-stable finite-difference schemes*, this Journal, 21 (1984), pp. 1-23.
- [4] ———, *High resolution schemes for hyperbolic conservation laws*, J. Comput. Phys., 49 (1983), pp. 357-393.
- [5] ———, *On second-order accurate Godunov-type schemes*, to appear.
- [6] A. HARTEN AND J. M. HYMAN, *A self-adjusting grid for the computation of weak solutions of hyperbolic conservation laws*, J. Comput. Phys., 50 (1983), pp. 235-269.
- [7] A. HARTEN AND P. D. LAX, *A random choice finite difference scheme for hyperbolic conservation laws*, this Journal, 18 (1981), pp. 289-315.
- [8] A. JAMESON AND P. D. LAX, *Conditions for the construction of multi-point total variation diminishing difference schemes*, Princeton Univ., MAE Report 1650, 1984.
- [9] S. OSHER, *Convergence of generalized MUSCL schemes*, NASA Langley Contractor Report 172306 1984; this Journal, 22 (1985), pp. 947-961.
- [10] S. OSHER AND S. R. CHAKRAVARTHY, *High-resolution schemes and the entropy condition*, this Journal, 21 (1984), pp. 955-984.
- [11] ———, *Very high order accurate TVD schemes*, ICASE Report #84-44, 1984.
- [12] S. OSHER AND E. TADMOR, *On the convergence of difference approximations to conservation laws*, Math. Comp., submitted.
- [13] P. L. ROE, *Approximate Riemann solvers, parameter vectors, and difference schemes*, J. Comput. Phys., 43 (1981), pp. 357-372.
- [14] P. K. SWEBY, *High resolution schemes using flux limiters for hyperbolic conservation laws*, this Journal, 21 (1984), pp. 995-1011.
- [15] B. VAN LEER, *Towards the ultimate conservative finite difference scheme, II. Monotonicity and conservation combined in a second-order scheme*, J. Comput. Phys., 14 (1974), pp. 361-376.
- [16] ———, *Towards the ultimate conservative finite difference scheme, IV. A new approach to numerical convection*, J. Comput. Phys., 23 (1977), pp. 276-298.