MATHEMATICS OF COMPUTATION
VOLUME 49, NUMBER 179
JULY 1987, PAGES 123–134

TVB Boundary Treatment for Numerical Solutions of Conservation Laws

By Chi-Wang Shu

Abstract. In the computation of hyperbolic conservation laws $u_t + f(u)_x = 0$, TVD (total-variation-diminishing) and TVB (total-variation-bounded) schemes have been very successful for initial value problems. But most of the existing boundary treatments are only proved to be linearly stable, hence the combined initial-boundary scheme may not be TVB. In this paper we describe a procedure of boundary treatment which uses the original high-order scheme up to the boundary, plus extrapolation and upwind treatment at the boundary. The resulting scheme is proved to be TVB for the scalar nonlinear case and for linear systems.

1. Introduction. In this paper we consider the numerical solutions to the hyperbolic conservation law

$$u_t + f(u)_x = 0, \tag{1.1a}$$

$$u(x,0) = u_0(x). \tag{1.1b}$$

Here $u = (u_1, \ldots, u_s)^T$, and the Jacobian matrix $A(u) = \partial f/\partial u$ has s real eigenvalues

$$\lambda_1(u) \leqslant \lambda_2(u) \leqslant \cdots \leqslant \lambda_s(u)$$

and a complete set of eigenvectors.

On a computational grid $x_j = j\Delta x$, $t_n = n\Delta t$, we use u_j^n to denote the computed approximation to the exact solution $u(x_j, t_n)$ of (1.1).

For pure initial value problems, i.e., problems with $u_0(x)$ in (1.1b) to be either periodic or to have a compact support, the recently introduced TVD (total-variation-diminishing) and TVB (total-variation-bounded) schemes have been very successful. See, e.g., [1], [2], [3], [4], [7], and the references listed there. The total variation of a discrete scalar solution is defined by

$$\mathrm{TV}(u) = \sum_j |u_{j+1} - u_j|, \tag{1.2}$$

and if

$$\mathrm{TV}(u^{n+1}) \leqslant \mathrm{TV}(u^n) \tag{1.3}$$

we say the scheme is TVD; while if

$$\mathrm{TV}(u^n) \leqslant B \tag{1.4}$$

Received July 7, 1986; revised September 22, 1986.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65N10, 35L65.
Key words and phrases. Conservation law, TVD, TVB, boundary condition.

0025-5718/87 $1.00 + $.25 per page

for some fixed $B > 0$ depending only on u^0 and $\mathrm{TV}(u^0)$ and all possible n and Δt such that $n\Delta t \leqslant T$, we say that the scheme is TVB in $0 \leqslant t \leqslant T$. Clearly TVD implies TVB.

One major advantage of TVB schemes is that there is a convergent (in L_1^{local}) subsequence as $\Delta x \to 0$ to a weak solution of (1.1). If an additional entropy condition, which implies uniqueness of weak solution to (1.1), is satisfied, then the scheme is convergent. See, e.g., [2].

For initial-boundary value problems we hope the boundary treatment can still retain the TVB property of the scheme. The usual method is to use a lower-order scheme near the boundary. This not only reduces the order of the scheme near the boundary, but also makes any theoretical results about TVB of the initial-boundary scheme very hard to prove. In this paper we present an approach to the treatment of boundaries which uses the same high-order scheme up to the boundary, plus extrapolation and an upwind treatment at the boundary. The resulting scheme is proved to be TVB for the scalar nonlinear case (Section 2) and for linear systems (Section 3).

Our boundary treatment is based on the globally high-order TVB schemes discussed in [4], [6], [7], and [8]. These schemes have natural upwind-downwind decompositions which help us to implement and prove the TVB boundary treatments.

We include several numerical results in the appendix to demonstrate the usefulness of the TVB boundary treatment in Sections 2 and 3.

2. Scalar Case. In this paper we use the following three equivalent forms of an rth order (in space *and* time) TVB scheme (see [4], [6], [7], and [8] for details):

$$u_j^{n+1} = \sum_{k=0}^{m} \left[\alpha_k u_j^{n-k} + |\beta_k| \lambda \left(C_{j+1/2}^{(n-k)} \Delta_+ u_j^{n-k} - D_{j-1/2}^{(n-k)} \Delta_- u_j^{n-k} \right)\right] + \mathrm{cor}_j^n, \tag{2.1a1}$$

where

$$C_{j+1/2} \geqslant 0, \quad D_{j+1/2} \geqslant 0, \quad 1 - \lambda\left(C_{j+1/2} + D_{j+1/2}\right) \geqslant 0, \tag{2.1a2}$$

$$\left|\mathrm{cor}_j^n\right| \leqslant B\Delta x^2. \tag{2.1a3}$$

Or,

$$u_j^{n+1} = \sum_{k=0}^{m} \left[\alpha_k u_j^{n-k} + |\beta_k| \lambda \left(\tilde{C}_{j+1/2}^{(n-k)} \Delta_+ u_j^{n-k} - \tilde{D}_{j-1/2}^{(n-k)} \Delta_- u_j^{n-k} \right)\right], \tag{2.1b1}$$

where

$$\tilde{C}_{j+1/2}^{(n-k)} \Delta_+ u_j^{n-k} = C_{j+1/2}^{(n-k)} \Delta_+ u_j^{n-k} + \mathrm{cor}\, 1_j^{n-k}, \tag{2.1b2}$$

$$-\tilde{D}_{j-1/2}^{(n-k)} \Delta_- u_j^{n-k} = -D_{j-1/2}^{(n-k)} \Delta_- u_j^{n-k} + \mathrm{cor}\, 2_j^{n-k}, \tag{2.1b3}$$

with

$$\left|\mathrm{cor}\, 1_j^{n-k}\right| \leqslant B\Delta x^2, \quad \left|\mathrm{cor}\, 2_j^{n-k}\right| \leqslant B\Delta x^2, \quad \mathrm{cor}\, 1_j^{n-k} + \mathrm{cor}\, 2_j^{n-k} = \mathrm{cor}_j^{n-k}. \tag{2.1b4}$$

Or

$$u_j^{n+1} = \sum_{k=0}^{m} \left[\alpha_k u_j^{n-k} + |\beta_k| \left(\left(Df^-\right)_j^{n-k} - \left(Df^+\right)_j^{n-k} \right)\right] \tag{2.1c1}$$

with

$$Df_j^- \equiv \lambda \tilde{C}_{j+1/2}\Delta_+ u_j, \qquad Df_j^+ \equiv \lambda \tilde{D}_{j-1/2}\Delta_- u_j. \tag{2.1c2}$$

In (2.1), α_k and β_k are such that

$$\alpha_k \geqslant 0, \qquad k = 0, 1, \ldots, m, \tag{2.2}$$

$$\sum_{k=0}^{m} \alpha_k = 1, \tag{2.3}$$

$$\beta_0 - \sum_{k=1}^{m} (k\alpha_k - \beta_k) = 1, \tag{2.4}$$

$$(-1)^l \sum_{k=1}^{m} k^{l-1}(k\alpha_k - l\beta_k) = 1, \qquad l = 2, 3, \ldots, r. \tag{2.5}$$

The CFL condition is

$$\lambda \leqslant \lambda_0 \cdot \min_k \left(\frac{\alpha_k}{|\beta_k|} \right), \tag{2.6}$$

where λ_0 satisfies (2.1a2).

We now consider the equation (1.1) defined on

$$0 \leqslant x < +\infty, \qquad t \geqslant 0, \tag{2.7}$$

and still assume that $u_0(x)$ in (1.1b) is zero in $x > L$.

Notice that (2.1) has a more than three-point stencil in x only because of the C's and the D's. Near the boundary $x = 0$ we may not have the necessary $u_{-1}, u_{-2}, \ldots$ etc., to define C and D. In this paper we always use *extrapolation with order r* to get the necessary u_{-j}. For example, if $r = 2$, we may use $u_{-1} = 3u_0 - 3u_1 + u_2$, etc. It is well known that extrapolation may cause instability, but due to the upwind-biased property of our scheme (2.1), we shall prove that the resulting initial-boundary scheme is still TVB. With the help of extrapolation, the boundary problem simplifies to that of determining u_0 only. From the well-known properties of the hyperbolic equation (1.1), we know that we should prescribe $u(0, t) = g(t)$ on the boundary $x = 0$ if $f'(u(0, t)) > 0$ (corresponding to case (a) of Theorem 2.1 below), and should prescribe nothing at $x = 0$ if $f'(u(0, t)) < 0$ (corresponding to case (b) of the theorem).

THEOREM 2.1. *In the region defined by* (2.7):

(a) *The scheme* (2.1) *for* u_j^n, $j \geqslant 1$, *with the above-mentioned extrapolations for computing the C's and D's, and initial condition* $u_j^0 = u(x_j, 0)$ *and boundary condition*

$$u_0^n = g(t^n), \tag{2.8}$$

where $g(t)$ *is a function of bounded variation in* $0 \leqslant t \leqslant T$, *is TVB in* $0 \leqslant t \leqslant T$.

(b) *Define*

$$\overline{Df}_0^+ = \min\left(|Df_0^+|, |Df_0^-| + K\Delta t\right) \cdot \operatorname{sign}\left(Df_0^+\right) \tag{2.9}$$

in (2.1c), *where* $K > 0$ *is a constant. Then the scheme* (2.1), *as in part* (a) *where* u_0^{n+1} *is computed by* (2.1c) *with* Df_0^+ *replaced by* $\overline{Df}_0^+$, *is TVB in* $0 \leqslant t \leqslant T$.

Proof. (a) We use form (2.1a):

$$\Delta_+ u_j^{n+1} = \sum_{k=0}^{m} \left(\alpha_k \left(1 - \frac{|\beta_k|}{\alpha_k} \lambda \left(C_{j+1/2}^{(n-k)} + D_{j+1/2}^{(n-k)} \right) \right) \left(\Delta_+ u_j^{n-k} \right) \right.$$
$$\left. + |\beta_k| \lambda C_{j+3/2}^{(n-k)} \left(\Delta_+ u_{j+1}^{n-k} \right) + |\beta_k| \lambda D_{j-1/2}^{(n-k)} \left(\Delta_+ u_{j-1}^{n-k} \right) \right) + \mathrm{cor}_{j+1}^n - \mathrm{cor}_j^n$$
$$\text{for } j = 1, 2, \ldots$$

and

$$\Delta_+ u_0^{n+1} = u_1^{n+1} - g(t^{n+1})$$
$$= \sum_{k=0}^{m} \left[\alpha_k u_1^{n-k} + |\beta_k| \lambda \left(C_{3/2}^{(n-k)} \Delta_+ u_1^{n-k} - D_{1/2}^{(n-k)} \Delta_- u_1^{n-k} \right) \right] + \mathrm{cor}_1^n - g(t^{n+1})$$
$$= \sum_{k=0}^{m} \left[\alpha_k \left(u_1^{n-k} - u_0^{n-k} \right) + |\beta_k| \lambda \left(C_{3/2}^{(n-k)} \Delta_+ u_1^{n-k} - D_{1/2}^{(n-k)} \Delta_- u_1^{n-k} \right) \right]$$
$$+ \mathrm{cor}_1^n - \sum_{k=0}^{m} \alpha_k \left(g(t^{n+1}) - g(t^{n-k}) \right).$$

Hence, by (2.1a2) and the CFL condition (2.6), we have

$$\mathrm{TV}(u^{n+1}) = \sum_{j=0}^{\infty} \left| \Delta_+ u_j^{n+1} \right|$$
$$\leqslant \sum_{j=1}^{\infty} \sum_{k=0}^{m} \left\{ \alpha_k \left[1 - \frac{|\beta_k|}{\alpha_k} \lambda \left(C_{j+1/2}^{(n-k)} + D_{j+1/2}^{(n-k)} \right) \right] \left| \Delta_+ u_j^{n-k} \right| \right.$$
$$+ |\beta_k| \lambda C_{j+3/2}^{(n-k)} \left| \Delta_+ u_{j+1}^{n-k} \right| + |\beta_k| \lambda D_{j-1/2}^{(n-k)} \left| \Delta_+ u_{j-1}^{n-k} \right|$$
$$\left. + \left| \mathrm{cor}_{j+1}^n \right| + \left| \mathrm{cor}_j^n \right| \right\}$$
$$+ \sum_{k=0}^{m} \left[\alpha_k \left(1 - \frac{|\beta_k|}{\alpha_k} \lambda D_{1/2}^{(n-k)} \right) \left| \Delta_+ u_0^{n-k} \right| + |\beta_k| C_{3/2}^{(n-k)} \left| \Delta_+ u_1^{n-k} \right| \right]$$
$$+ \sum_{k=0}^{m} \alpha_k \left| g(t^{n+1}) - g(t^{n-k}) \right|$$
$$\leqslant \sum_{j=0}^{\infty} \sum_{k=0}^{m} \alpha_k \left| \Delta_+ u_j^{n-k} \right| + 2BN\Delta x^2 + \sum_{k=0}^{m} \alpha_k \left| g(t^{n-k+1}) - g(t^{n-k}) \right|$$
$$\leqslant \sum_{k=0}^{m} \alpha_k \mathrm{TV}(u^{n-k}) + \bar{B}\Delta t + (m+1) \sum_{k=0}^{m} \left| g(t^{n-k+1}) - g(t^{n-k}) \right|,$$

where $N = L/\Delta x$, $\bar{B} = 2BL/\lambda$.

Hence clearly

$$\mathrm{TV}(u^n) \leqslant \max_{0 \leqslant k \leqslant m} \mathrm{TV}(u^k) + \bar{B}T + (m+1)^2 \mathrm{TV}(g)$$

for all $n\Delta t \leqslant T$.

(b) Similarly to part (a), we have

$$\mathrm{TV}(u^{n+1}) \leqslant \sum_{j=0}^{\infty} \sum_{k=0}^{m} \alpha_k \left|\Delta_+ u_j^{n-k}\right| + \sum_{k=0}^{m} \alpha_k \left(\overline{Df}_0^{+} - Df_0^{-}\right) + 2BN\Delta x^2$$

$$\leqslant \sum_{k=0}^{m} \alpha_k \mathrm{TV}(u^{n-k}) + (\bar{B} + K)\Delta t$$

(see (2.9)).

Hence clearly

$$\mathrm{TV}(u^n) \leqslant \max_{0 \leqslant k \leqslant m} \mathrm{TV}(u^k) + (\bar{B} + K)T$$

for all $n\Delta t \leqslant T$. □

Remark 2.1. From Definition (2.9) we see that

$$\overline{Df}_0^{+} = Df_0^{+} \tag{2.10}$$

if and only if

$$|Df_0^{+}| \leqslant |Df_0^{-}| + K\Delta t. \tag{2.11}$$

Since scheme (2.1) is upwind-biased, we should expect the "downwind" part $|Df_0^+|$ to be less than the "upwind" part $|Df_0^-|$ if $f'(u(0,t)) < 0$. With the help of $K\Delta t$, we may be very safe to expect (2.11), hence (2.10), to be almost always valid. Hence, since we used extrapolation of sufficient accuracy to approximate u_{-j}, the accuracy remains rth order up to the boundary. This is superior to the usual way of using a lower-order scheme with narrower stencil near the boundary.

3. Linear Systems. The techniques discussed in Section 2 can be generalized, via Godunov's, Osher's or Roe's field-by-field decompositions, to nonlinear systems (1.1). See, e.g., [4], [5]. Even for pure initial value problems, at present there is no theory about total variation boundedness for general nonlinear systems. For a linear system

$$u_t + Au_x = 0, \tag{3.1}$$

where, for simplicity, A is assumed a constant matrix (we may generalize the theory to the case $A = A(x)$), we have a similar theory as in Theorem 2.1.

Assume A has *s nonzero* real eigenvalues

$$\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_{s_1} < 0 < \lambda_{s_1+1} \leqslant \cdots \leqslant \lambda_s \tag{3.2}$$

with a complete set of right eigenvectors $r_1, r_2, \ldots, r_s$. Then we can write

$$u(x,t) = \sum_{p=1}^{s} \delta^{(p)}(x,t) r_p, \tag{3.3a}$$

$$Au(x,t) = \sum_{p=1}^{s} \lambda_p \delta^{(p)}(x,t) r_p, \tag{3.3b}$$

where the scalar functions $\delta^{(p)}$ satisfy the scalar conservation laws

$$\delta_t^{(p)} + \lambda_p \delta_x^{(p)} = 0. \tag{3.4}$$

Hence a generalization of the scalar TVB scheme (2.1) is just using (2.1) on (3.4) for each p:

$$u_j^{n+1} = \sum_{p=1}^{s} \left\{ \sum_{k=0}^{m} \left[\alpha_k \left(\delta^{(p)}\right)_j^{n-k} + |\beta_k| \left(\left(\left(D\delta^{(p)}\right)^- \right)_j^{n-k} - \left(\left(D\delta^{(p)}\right)^+ \right)_j^{n-k} \right) \right] \right\} r_p. \tag{3.5}$$

For pure initial value problems, the scheme (3.5) decouples to s scalar schemes, hence TVB of (3.5), with TV defined by

$$\mathrm{TV}(u) = \sum_j \sum_{p=1}^{s} \left| \delta_{j+1}^{(p)} - \delta_j^{(p)} \right|, \tag{3.6}$$

is an immediate consequence of TVB of each of the decoupled scalar schemes.

Now we consider the initial-boundary value problem defined on (2.7).

By (3.2), on the differential equation level, it is well known that a boundary condition

$$\begin{pmatrix} \delta^{(s_1+1)}(0,t) \\ \vdots \\ \delta^{(s)}(0,t) \end{pmatrix} = S(t) \begin{pmatrix} \delta^{(1)}(0,t) \\ \vdots \\ \delta^{(s_1)}(0,t) \end{pmatrix} + g(t) \tag{3.7}$$

is well posed.

We obtain the discrete form of (3.7) via the following procedure:

(i) Use extrapolation of order r to find the necessary $\delta_{-j}^{(p)}$ in order to compute C's and D's near $j = 0$;

(ii) Let

$$\overline{D\delta_0^{(p)+}} = \min\left(\left| D\delta_0^{(p)+} \right|, (1-\varepsilon) \left| D\delta_0^{(p)-} \right| + K_1 \Delta t \right) \cdot \operatorname{sign}\left(D\delta_0^{(p)+} \right), \tag{3.8}$$

where $K_1 > 0$, $0 < \varepsilon \ll 1$ are constants and $p = 1, 2, \ldots, s_1$.

Notice that (i) and (ii) are similar to the scalar case treatment. By Remark 2.1 we should expect that most of the time $\overline{D\delta_0^{(p)+}} = D\delta_0^{(p)+}$.

(iii) Define

$$\begin{pmatrix} \delta^{(s_1+1)} \\ \vdots \\ \delta^{(s)} \end{pmatrix}_0^n = S(t^n) \begin{pmatrix} \delta^{(1)} \\ \vdots \\ \delta^{(s_1)} \end{pmatrix}_0^n + g(t^n), \tag{3.9}$$

where $S(t) = (S_{ij}(t))$ is a $(s - s_1) \times s_1$ matrix function with each of its elements being Lipschitz continuous in $0 \leqslant t \leqslant T$:

$$\left| S_{ij}(t_1) - S_{ij}(t_2) \right| \leqslant L |t_1 - t_2| \quad \text{for } 0 \leqslant t_1, t_2 \leqslant T \tag{3.10}$$

and $g(t)$ is a $(s - s_1) \times 1$ vector function of bounded variation in $0 \leqslant t \leqslant T$.

(3.9) is just the discretization of (3.7).

Notice that the scheme (3.5) is coupled by the boundary condition (3.9).

THEOREM 3.1. *The scheme* (3.5) *with boundary treatment* (i)–(iii) *is TVB in* $0 \leqslant t \leqslant T$.

Proof. Let

$$Q = \max\left(1, \frac{2}{\varepsilon} \max_{0 \leqslant k \leqslant m} \frac{|\beta_k|}{\alpha_k} \|S\|_1\right) = \max\left(1, \frac{2}{\varepsilon} \max_{0 \leqslant k \leqslant m} \frac{|\beta_k|}{\alpha_k} \max_{\substack{0 \leqslant t \leqslant T \\ 1 \leqslant p \leqslant s_1}} \sum_{i=1}^{s-s_1} |S_{ip}(t)|\right)$$

and define

$$\widetilde{\mathrm{TV}}(u) = \sum_j \left(Q \sum_{p=1}^{s_1} + \sum_{p=s_1+1}^{s} \right) \left| \delta_{j+1}^{(p)} - \delta_j^{(p)} \right|. \tag{3.11}$$

Clearly,

$$\mathrm{TV}(u) \leqslant \widetilde{\mathrm{TV}}(u) \leqslant Q \cdot \mathrm{TV}(u). \tag{3.12}$$

Since

$$(\delta^{s_1+p})_0^n = \sum_{i=1}^{s_1} S_{pi}^n (\delta^{(i)})_0^n + g_p(t^n),$$

we have

$$\begin{aligned}
&\Delta_+ (\delta^{(s_1+p)})_0^{n+1} \\
&\quad = (\delta^{(s_1+p)})_1^{n+1} - (\delta^{(s_1+p)})_0^{n+1} \\
&\quad = \sum_{k=0}^{m} \Big[\alpha_k (\delta^{(s_1+p)})_1^{n-k} + |\beta_k| \lambda \Big((C^{(s_1+p)})_{3/2}^{n-k} \Delta_+ (\delta^{(s_1+p)})_1^{n-k} \\
&\qquad\qquad - (D^{(s_1+p)})_{1/2}^{n-k} \Delta_- (\delta^{(s_1+p)})_1^{n-k} \Big) \Big] \\
&\qquad - \left[\sum_{i=1}^{s_1} S_{pi}^{n+1} (\delta^{(i)})_0^{n+1} + g_p(t^{n+1}) \right] + \mathrm{cor}_1^n \\
&\quad = \sum_{k=0}^{m} \left[\alpha_k \left(1 - \frac{|\beta_k|}{\alpha_k} \lambda (D^{(s_1+p)})_{1/2}^{n-k} \right) \Delta_+ (\delta^{(s_1+p)})_0^{n-k} \right. \\
&\qquad\qquad \left. + |\beta_k| \lambda (C^{(s_1+p)})_{3/2}^{n-k} \Delta_+ (\delta^{(s_1+p)})_1^{n-k} \right] + \mathrm{cor}_1^n \\
&\qquad - \left[\sum_{i=1}^{s_1} \sum_{k=0}^{m} \alpha_k \left(S_{pi}^{n+1} (\delta^{(i)})_0^{n+1} - S_{pi}^{n-k} (\delta^{(i)})_0^{n-k} \right) \right. \\
&\qquad\qquad \left. + \sum_{k=0}^{m} \alpha_k \left(g_p(t^{n+1}) - g_p(t^{n-k}) \right) \right].
\end{aligned}$$

Notice that

$$\begin{aligned}
&S_{pi}^{n+1} (\delta^{(i)})_0^{n+1} - S_{pi}^{n-k} (\delta^{(i)})_0^{n-k} \\
&\quad = S_{pi}^{n+1} \left((\delta^{(i)})_0^{n+1} - (\delta^{(i)})_0^{n-k} \right) + (\delta^{(i)})_0^{n-k} \left(S_{pi}^{n+1} - S_{pi}^{n-k} \right),
\end{aligned}$$

and, since $u_0(x)$ is zero for large x and the numerical solution has a finite speed of propagation, we have $(\delta^{(i)})_J^n = 0$ for large enough J, so

$$\begin{aligned}
\sum_{i=1}^{s_1} \left| (\delta^{(i)})_0^n \right| &= \sum_{i=1}^{s_1} \left| (\delta^{(i)})_J^n - (\delta^{(i)})_0^n \right| \leqslant \sum_{i=1}^{s_1} \sum_{j \geqslant 0} \left| (\delta^{(i)})_{j+1}^n - (\delta^{(i)})_j^n \right| \\
&\leqslant \mathrm{TV}(u^n) \leqslant \widetilde{\mathrm{TV}}(u^n).
\end{aligned}$$

(This is the maximum principle for TVB schemes. See Remark 3.1 below.)

So, following the lines of the proof of Theorem 2.1, we have

$$\begin{aligned}
\widetilde{\mathrm{TV}}(u^{n+1}) &= \sum_{j\geqslant 0}\left(Q\sum_{p=1}^{s_1}+\sum_{p=s_1+1}^{s}\right)\left|\Delta_+(\delta^{(p)})_j^{n+1}\right| \\
&\leqslant Q\sum_{p=1}^{s_1}\left[\sum_{j\geqslant 0}\left(\sum_{k=0}^{m}\alpha_k\left|\Delta_+(\delta^{(p)})_j^{n-k}\right|\right)+K_1\Delta t-\lambda\varepsilon(C^{(p)})_{1/2}^{n-k}\left|\Delta_+(\delta^{(p)})_0^{n-k}\right|\right] \\
&\quad+\sum_{p=1}^{s-s_1}\left\{\sum_{j\geqslant 0}\left(\sum_{k=0}^{m}\alpha_k\left|\Delta_+(\delta^{(s_1+p)})_j^{n-k}\right|\right)\right. \\
&\qquad+\sum_{i=1}^{s_1}\left[\left|S_{pi}^{n+1}\right|\sum_{k=0}^{m}|\beta_k|\lambda\left((C^{(i)})_{1/2}^{n-k}\left|\Delta_+(\delta^{(i)})_0^{n-k}\right|\right.\right. \\
&\qquad\qquad\left.+(D^{(i)})_{-1/2}^{n-k}\left|\Delta_-(\delta^{(i)})_0^{n-k}\right|\right) \\
&\qquad\quad+\sum_{k=0}^{m}\alpha_k\left|(\delta^{(i)})_0^{n-k}\right|\left|S_{pi}^{n+1}-S_{pi}^{n-k}\right|\Bigg] \\
&\qquad\left.+\left(\sum_{k=0}^{m}\alpha_k\left|g_p(t^{n+1})-g_p(t^{n-k})\right|\right)\right\}+2NB\Delta x^2 \\
&\leqslant \sum_{k=0}^{m}\alpha_k\widetilde{\mathrm{TV}}(u^{n-k}) \\
&\quad-\sum_{p=1}^{s_1}\sum_{k=0}^{m}\alpha_k\left[\left(Q\varepsilon-2\frac{|\beta_k|}{\alpha_k}\sum_{i=1}^{s-s_1}\left|S_{ip}^{n+1}\right|\right)\lambda(C^{(p)})_{1/2}^{n-k}\left|\Delta_+(\delta^{(p)})_0^{n-k}\right|\right] \\
&\quad+(m+1)L\Delta t\sum_{k=0}^{m}\alpha_k\widetilde{\mathrm{TV}}(u^{n-k}) \\
&\quad+(m+1)\sum_{p=1}^{s-s_1}\sum_{k=0}^{m}\left|g_p(t^{n-k+1})-g_p(t^{n-k})\right|+H\Delta t \\
&\leqslant \sum_{k=0}^{m}\alpha_k(1+\tilde{L}\Delta t)\widetilde{\mathrm{TV}}(u^{n-k}) \\
&\quad+(m+1)\sum_{p=1}^{s-s_1}\sum_{k=0}^{m}\left|g_p(t^{n-k+1})-g_p(t^{n-k})\right|+H\Delta t.
\end{aligned}$$

Hence, clearly,

$$\begin{aligned}
\widetilde{\mathrm{TV}}(u^n) &\leqslant (1+\tilde{L}\Delta t)^n\max_{0\leqslant k\leqslant m}\widetilde{\mathrm{TV}}(u^k) \\
&\quad+\sum_{i=k}^{n-1}(1+\tilde{L}\Delta t)^{n-1-i}\left[(m+1)\sum_{p=1}^{s-s_1}\sum_{k=0}^{m}\left|g_p(t^{i-k+1})-g_p(t^{i-k})\right|+H\Delta t\right] \\
&\leqslant e^{\tilde{L}T}\max_{0\leqslant k\leqslant m}\widetilde{\mathrm{TV}}(u^k)+e^{\tilde{L}T}\left[(m+1)^2\mathrm{TV}(g)+HT\right].
\end{aligned}$$

This, together with (3.12), proves that

$$\mathrm{TV}(u^n)\leqslant B\cdot\max_{0\leqslant k\leqslant m}\mathrm{TV}(u^k)+\bar{B}$$

for some B, $\bar{B}>0$. □

Remark 3.1. With the help of the conservation form, we do not need to require that $u_0(x)$ have compact support to prove the maximum principle for a TVB scheme in the half-plane (2.7). Taking a scalar conservation TVB scheme

$$u_j^{n+1} = u_j^n - \left(h_{j+1/2}^n - h_{j-1/2}^n\right)$$

as an example (the same idea certainly works for the general TVB scheme (3.5)), we have

$$\max_{j\geqslant 0}\left(u_j^n\right) - \min_{j\geqslant 0}\left(u_j^n\right) \leqslant \mathrm{TV}(u^n) \leqslant B\cdot \mathrm{TV}(u^0). \tag{3.13}$$

Let

$$s_N = \frac{1}{N}\sum_{j=0}^{N-1} u_j^n;$$

then

$$\begin{aligned} s_N &= \frac{1}{N}\sum_{j=0}^{N-1}\left[u_j^{n-1} - \left(h_{j+1/2}^{n-1} - h_{j-1/2}^{n-1}\right)\right] \\ &= \frac{1}{N}\left[\sum_{j=0}^{N-1}\left(u_j^{n-1}\right) + h_{-1/2}^{n-1} - h_{N-1/2}^{n-1}\right] \\ &= \cdots = \frac{1}{N}\left[\sum_{j=0}^{N-1} u_j^0 + \sum_{k=0}^{n-1}\left(h_{-1/2}^k - h_{N-1/2}^k\right)\right]. \end{aligned}$$

Since h is Lipschitz continuous in its arguments and we may assume (inductively) that u_j^k is bounded for $k < n$, the sum

$$\sum_{k=0}^{n-1}\left(h_{-1/2}^k - h_{N-1/2}^k\right)$$

is bounded independent of N. Hence

$$\overline{\lim_{N\to\infty}}\,|s_N| = \overline{\lim_{N\to\infty}}\left|\frac{1}{N}\sum_{j=0}^{N-1} u_j^0\right| \leqslant \|u^0\|_\infty.$$

Clearly,

$$\min_j\left(u_j^n\right) \leqslant s_N \leqslant \max_j\left(u_j^n\right).$$

So

$$\max_j\left(u_j^n\right) \leqslant B\cdot \mathrm{TV}(u^0) + \min_j\left(u_j^n\right) \leqslant B\cdot \mathrm{TV}(u^0) + s_N \leqslant B\cdot \mathrm{TV}(u^0) + \|u^0\|_\infty.$$

Similarly,

$$\begin{aligned} \min_j\left(u_j^n\right) &\geqslant \max_j\left(u_j^n\right) - B\cdot \mathrm{TV}(u^0) \geqslant s_N - B\cdot \mathrm{TV}(u^0) \\ &\geqslant -\left[B\cdot \mathrm{TV}(u^0) + \|u^0\|_\infty\right]. \end{aligned}$$

Hence

$$\max_j\left|u_j^n\right| \leqslant B\cdot \mathrm{TV}(u^0) + \|u^0\|_\infty.$$

Acknowledgment. The author would like to thank Professor Stanley Osher for his valuable help and suggestions.

Appendix: *Numerical Results.* The numbers in this appendix are often written in exponential forms, e.g., 4.2 (−3) means 4.2×10^{-3}.

All the tables are collected at the end.

Example 1. We use the boundary treatment (2.8)–(2.9) and 3-3 (third order in space and time) TVD and TVB schemes discussed in [4], [6], [7], and [8] to solve the following scalar initial-boundary value problem:

$$\text{(A.1)} \qquad \begin{aligned} &u_t + u_x = 0, \qquad -1 \leqslant x \leqslant 1, \\ &u(-1,t) = \sin \pi t, \qquad u(x,0) = \sin \pi x. \end{aligned}$$

The boundary condition given is well posed.

The 3-3 TVD and TVB schemes we use are (2.1c) with

$$\text{(A.2)} \qquad m = 3; \quad \alpha_i = \frac{16}{27}, 0, 0, \frac{11}{27}, \quad \beta_i = \frac{16}{9}, 0, 0, \frac{44}{27};$$

$$\text{(A.3a)} \qquad Df_j^- = -\lambda\left[df_{j-1/2}^+ + \frac{1}{3}\Delta_-\left(df_{j+1/2}^+\right)^{(-)} + \frac{1}{6}\Delta_-\left(df_{j-1/2}^+\right)^{(+)}\right],$$

$$\text{(A.3b)} \qquad Df_j^+ = -\lambda\left[df_{j+1/2}^- - \frac{1}{3}\Delta_+\left(df_{j-1/2}^-\right)^{(-)} + \frac{1}{6}\Delta_+\left(df_{j-1/2}^-\right)^{(+)}\right],$$

where

$$\text{(A.4)} \qquad df_{j+1/2}^+ = f(u_{j+1}) - h_{j+1/2}; \qquad df_{j+1/2}^- = h_{j+1/2} - f(u_j),$$

with

$$\text{(A.5)} \qquad \begin{aligned} h_{j+1/2} &= \frac{1}{2}\left[f(u_{j+1}) + f(u_j) - \alpha\Delta_+ u_j\right], \\ \alpha &= \max_{\min_j u_j \leqslant u \leqslant \max_j u_j} |f'(u)|, \end{aligned}$$

being the first-order monotone Lax-Friedrichs flux (we may also use any other smooth monotone flux here), and the limited quantities defined by

$$\text{(A.6)} \qquad y_{j+1/2}^{(+)} = m\left(y_{j+1/2}, by_{j+3/2}\right); \qquad y_{j+1/2}^{(-)} = m\left(y_{j+1/2}, by_{j-1/2}\right).$$

Here $b = 4$, and the function "m" is defined by

$$\text{(A.7a)} \qquad m(\alpha,\beta) = \operatorname{minmod}(\alpha,\beta) \equiv (\operatorname{sign}\alpha)\max\left(0, \min\left(|\alpha|, \beta \operatorname{sign}\alpha\right)\right)$$

in the TVD case, and

$$\text{(A.7b)} \qquad m(\alpha,\beta) = \operatorname{minmod}\left(\alpha, \beta + M\Delta x^2 \operatorname{sign}\alpha\right)$$

in the TVB case. (We choose $M = 50$.)

We use (2.8) at the left boundary $x = -1$ and (2.9) at the right boundary $x = 1$. The necessary extrapolations are done to third order. We use $K = 1.0$ in (2.9).

The numerical results are listed in Table 1.

By comparing Table 1 with the results in [7] (same scheme for pure initial value problem), we see that the boundary treatments (2.8) and (2.9) work quite well: We get almost the same accuracy as in the pure initial value calculations.

Example 2. The same 3-3-TVB scheme as in Example 1 with the TVB boundary treatment (3.8)–(3.9) is applied to the problem

$$\text{(A.8)}\qquad \begin{aligned} &\begin{pmatrix} u \\ v \end{pmatrix}_t = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}_x, \qquad 0 \leqslant x \leqslant 1, \\ &u(x,0) = v(x,0) = \sin 2\pi x, \\ &u(0,t) = -v(0,t), \qquad v(1,t) = -u(1,t). \end{aligned}$$

We apply (3.8)–(3.9) to both boundaries. ε and K_1 in (3.8) are taken to be $1.(-2)$ and 1, respectively. The exact solutions of (A.8) are

$$u(x,t) = \sin 2\pi(x - t); \qquad v(x,t) = \sin 2\pi(x + t).$$

We list the numerical errors at $t = 2$ in Table 2.

Table 2 shows that the boundary treatment (3.8)–(3.9) works very well for smooth problems.

Example 3. The same scheme and the same boundary treatment are applied to the same equation as in Example 2 with a discontinuous initial-boundary condition:

$$\text{(A.9)}\qquad u(x,0) = \begin{cases} 1, & \text{if } \frac{1}{3} \leqslant x \leqslant \frac{2}{3}, \\ 0, & \text{otherwise,} \end{cases} \qquad v(x,0) \equiv 0 \quad \text{for } x \geqslant 0;$$

$$u(0,t) = v(0,t).$$

We use the 3-3-TVD scheme (in order to see how the boundary treatment itself affects the total variation) and the 3-3-TVB scheme with $M = 200$ in (A.7b). For $\Delta x = 1/20$ the solutions at $t = 0.5$ and $t = 1$ are printed (the exact solutions have values 0 and 1 at two sides of the star*):

(i) 3-3-TVD:

$t = 0.5$, u: .98, .97, .87, .57*, .25, 4.3(−2), 0. 0, ...

$t = 0.5$, v: .98, .98, .85, .59*, .29, 7.1(−2), 8.3(−3), 8.8(−4), 0. 0, ...

$t = 1.0$, u: −6.4(−4), 1.1(−3), 1.6(−3), 3.4(−3), 1.4(−2), 9.6(−2), .32, *.60, .84, .95, .96, .94, .81, .57*, .30, 9.7(−2), 7.2(−3), 0. 0, ...

(ii) 3-3-TVB with $M = 200$:

$t = 0.5$, u: 1.06, 1.08, 0.85, 0.53*, 0.23, 5.2(−2), −1.8(−2), −2.3(−2), −1.0(−2), −1.1(−3), 1.4(−3), 9.0(−4), ...

$t = 0.5$, v: 1.09, 1.01, 0.84, 0.60*, 0.33, 8.3(−2), −4.9(−2), −5.8(−2), −1.4(−2), 1.2(−2), 7.8(−3), ...

$t = 1.0$, u: 4.2(−3), −7.9(−3), −5.1(−2), −7.1(−2), −2.1(−2), .12, .33, *.59, .84, 1.03, 1.11, 1.02, .81, .54*, .28, 9.7(−2), −9.0(−4), −3.1(−2), −2.6(−2), −1.2(−2)

We can see that 3-3-TVD has essentially no overshoots or undershoots. This implies that the boundary treatment itself does not increase the total variation in this case. For an M as big as 200 and a Δx which is not too small, we still get reasonable results for the 3-3-TVB scheme.

TABLE 1 (*Example* 1)

r: numerical order

	L_∞-error				L_1-error			
Δx	TVD	r	TVB	r	TVD	r	TVB	r
1/10	6.5(−2)		9.4(−3)		1.7(−2)		3.5(−3)	
1/20	2.2(−2)	1.59	1.3(−3)	2.80	3.5(−3)	2.30	5.1(−4)	2.77
1/40	7.1(−3)	1.61	2.0(−4)	2.75	6.9(−4)	2.34	8.2(−5)	2.65

TABLE 2 (*Example* 2)

L_∞: L_∞-error; L_1: L_1-error; r: numerical order

	u				v			
Δx	L_∞	r	L_1	r	L_∞	r	L_1	r
1/10	6.5(−2)		3.6(−2)		6.5(−2)		3.8(−2)	
1/20	1.1(−2)	2.53	7.1(−3)	2.35	1.1(−2)	2.53	7.0(−3)	2.42
1/40	1.7(−3)	2.72	1.1(−3)	2.70	1.7(−3)	2.72	1.1(−3)	2.69
1/80	2.4(−4)	2.83	1.5(−4)	2.84	2.4(−4)	2.83	1.5(−4)	2.84

Institute for Mathematics and Its Applications
University of Minnesota
Minneapolis, Minnesota 55455

1. A. HARTEN, "High resolution schemes for hyperbolic conservation laws," *J. Comput. Phys.*, v. 49, 1983, pp. 357–393.

2. A. HARTEN, "On a class of high resolution total-variation-stable finite difference schemes," *SIAM J. Numer. Anal.*, v. 21, 1984, pp. 1–23.

3. S. OSHER & S. CHAKRAVARTHY, "High resolution schemes and the entropy condition," *SIAM J. Numer. Anal.*, v. 21, 1984, pp. 955–984.

4. S. OSHER & S. CHAKRAVARTHY, *Very High Order Accurate TVD Schemes*, ICASE Report #84-44, 1984; IMA Volumes in Mathematics and its Applications, vol. 2, Springer-Verlag, 1986, pp. 229–274.

5. P. ROE, "Approximate Riemann solvers, parameter vectors, and difference schemes," *J. Comput. Phys.*, v. 43, 1981, pp. 357–372.

6. C. SHU, "TVD time discretization II—time dependent problems." (Preprint.)

7. C. SHU, "TVB uniformly high-order schemes for conservation laws," *Math. Comp.*, v. 49, 1987, pp. 105–121.

8. C. SHU, *Numerical Solutions of Conservation Laws*, Ph. D. dissertation, Department of Mathematics, University of California, Los Angeles, 1986.