JOURNAL OF COMPUTATIONAL PHYSICS 92, 422-455 (1991)

A Treatment of Discontinuities in Shock-Capturing Finite Difference Methods

DE-KANG MAO*

Department of Mathematics, University of California, Los Angeles, California 90024 and Department of Mathematics, Shanghai University of Science and Technology, Shanghai, China

Received February 27, 1989; revised December 7, 1989

In this paper a treatment to sharpen discontinuities for shock capturing methods is introduced. The treatment is a modification of the underlying scheme that makes the computation on each side of the discontinuity use information only from that side. The modification is done by adding specific artifical terms to the underlying scheme. The correctness of the discontinuity's location is guaranteed by some limitation of the artificial terms. Shock tracking ideas are involved in the treatment; however, no lower dimensional grid is needed to fit the discontinuity. A high resolution technique is set up to find out the location of the discontinuity within the cell. Several numerical examples including spontaneous shocks, linear discontinuity calculations, and the blast waves problem are presented.

1. INTRODUCTION

The initial value problem for hyperbolic conservation laws is as follows:

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

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

where $u = (u_1, u_2, ..., u_m)^T$ is a state vector and $f(u)$, the flux, is a vector-valued function of m components. The system is hyperbolic in the sense that the $m \times m$ Jacobian matrix

$$A(u) = \partial f(u)/\partial u$$

has m real eigenvalues

$$a_1(u) \leqslant a_2(u) \leqslant \cdots \leqslant a_m(u)$$

and a complete set of m linearly independent right-eigenvectors. A weak solution to (1.1) is a bounded measurable function $u(x, t)$ that satisfies

$$\int_0^\infty \int_{-\infty}^\infty (u\phi_t + f(u)\phi_x)\, dx\, dt + \int_{-\infty}^\infty u_0(x)\, \phi(x, 0)\, dx = 0 \tag{1.2}$$

for all $\phi \in C_0^2((-\infty, \infty) \times [0, \infty))$.

* Research partially supported by ONR Grant N00014-86-k-0691.

0021-9991/91 $3.00

The main cause of the numerical difficulty in shock capturing methods is the occurrence of discontinuities. The fluid state variables may jump across shocks, and they may have discontinuous derivatives across characteristics. Most currently used shock capturing methods ignore discontinuities by applying almost the same numerical scheme everywhere in the flow. As a consequence, the consistency of the difference scheme with the partial differential equation is based on the assumption that the exact solution to the problem is smooth; on the other hand, at the discontinuities, the exact solution lacks the smoothness necessary for the consistency. That is why oscillations and smearing of discontinuities often occur in the computations.

Many efficient finite difference approximations to (1.1) have been developed. Particularly, the ENO schemes lately developed by Harten, Osher, Engquist, and Chakravarthy (see [1–4]) have been very successful in dealing with shocks. These schemes use a local adaptive stencil to obtain information automatically from regions of smoothness when the solution develops discontinuities. Obviously, the idea of picking up information from the smooth parts is an attempt to prevent the computation from crossing a discontinuity; in other words, it tries to make the numerical flux use data only from one side of the discontinuity. As a result, approximations using these methods obtain uniformly high order accuracy right up to discontinuities, while keeping a sharp, essentially nonoscillatory shock transition.

Two improvements of ENO schmes that should be mentioned are:

In [6, 7] Shu and Osher constructed pointwise ENO schemes by applying the idea of an adaptive stencil to the numerical flux and using a TVD Runge–Kutta type time discretization. This greatly eases the implementation and simplifies the programming.

In [8] Yang designed an artifical compession method for ENO schemes by a modification of the slope in the reconstruction procedure. This technique efficiently improves the performance of the ENO schemes at contract discontinuities.

Recently, Harten introduced a concept of "subcell resolution" to the ENO schemes. The main ingredient is the observation that the information in the cell average of a discontinuous function contains the location of the discontinuity within the cell. Using this observation one can modify the ENO reconstruction to recover accurately any discontinuous function from its cell everage. The modification of the ENO reconstruction, which extends the reconstruction function in each left and right adjacent cell to the recovered location of the discontinuity, makes the numerical flux only depend upon the data from one side of the discontinuity. As a consequence, it effectively prevents the computation from crossing the discontinuity. The application of this technique to the linearly degenerate characteristic field greatly sharpens contact discontinuities.

Six years ago the author began to work on a treatment of discontinuities for shock capturing methods, which applies to arbitrary schemes (see [11]). Although the treatment is similar to Harten's "subcell resolution," it has a different origin. The essence of the treatment is that in the vicnity of the critical intervals that are suspected of harboring discontinuities, the underlying difference scheme is modified

so that the computation on each side of the discontinuity only involves the numerical solution and its extrapolated data from that side. The modification is implemented by adding specific artificial terms to the underlying scheme. Limitation of the artificial terms guarantees the exactness of the discontinuity's location. Since the computation is not done across the discontinuities, the spurious oscillation and the smearing of discontinuities are essentially eliminated. We believe that the basic idea presented here also applies to the multidimensional computations.

The paper is organized in the following manner: Section 2 describes the treatment in detail. Section 3 describes the incorporation of the treatment into difference schemes. Section 4 contains a numerical analysis of the treatment and obtains some theoretical results. Section 5 generalizes the treatment to the Euler equations of gas dynamics. Section 6 presents several numerical examples to show the performance of the treatment.

2. Treatment of Discontinuities in the Scalar Case

We begin with the scalar, convex problem; i.e., both u and f in (1.1) are scalar, and $f'' \geqslant 0$. The conservative difference scheme is

$$u_j^{n+1} = u_j^n - \lambda(\hat{f}_{j+1/2}^n - \hat{f}_{j-1/2}^n), \tag{2.1}$$

where $u_j^n = u_h(x_j, t_n)$ denotes a numerical approximation to the exact solution $u(x_j, t_n)$, and $\hat{f}_{j+1/2}^n = \hat{f}(u_{j-k+1}^n, \ldots, u_{j+k}^n)$ is the numerical flux dependent on $2k$ variables. The flux is consistent with (1.1a) in the sense that

$$\hat{f}(u, u, \ldots, u) = f(u), \tag{2.2}$$

λ is the mesh ratio; i.e., $\lambda = \tau/h$, where τ and h are the time and space increments, respectively.

There are two versions of the treatment, an x version and an $x-t$ version, and we first describe the $x-t$ version. Suppose that $[x_{j_1}, x_{j_1+1}]$ is a critical interval on level n suspected of harboring a shock (or a contact discontinuity), which we refer as a *generated interval*; in addition, the numerical solution on each side of it is "smooth" (as shown in Fig. 2.1). The treatment begins by extrapolating the numerical solution from each side of the interval to the other side, and obtaining a set of extrapolated data $u_{j_1-k}^{n,+}, u_{j_1-k+1}^{n,+}, \ldots, u_{j_1}^{n,+}, u_{j_1+1}^{n,-}, u_{j_1+2}^{n,-}, \ldots, u_{j_1+k+1}^{n,-}$ (see Fig. 2.1). Then the unknown u_j^{n+1} on the next level is computed in such a way that on each side of the shock it is evaluated only with u_j^n and the extrapolated data from that side. For example, if point (x_j, t_{n+1}) is considered on the left of the shock, u_j^{n+1} is computed as

$$u_j^{n+1} = u_j^n - \lambda(f_{j+1/2}^{n,-} - \hat{f}_{j-1/2}^{n,-}) \tag{2.3}$$

rather than (2.1), where

$$\hat{f}_{j+1/2}^{n,-} = \hat{f}(u_{j-k+1}^n, \ldots, u_{j_1}^n, u_{j_1+1}^{n,-}, \ldots, u_{j+k}^{n,-}). \tag{2.4}$$

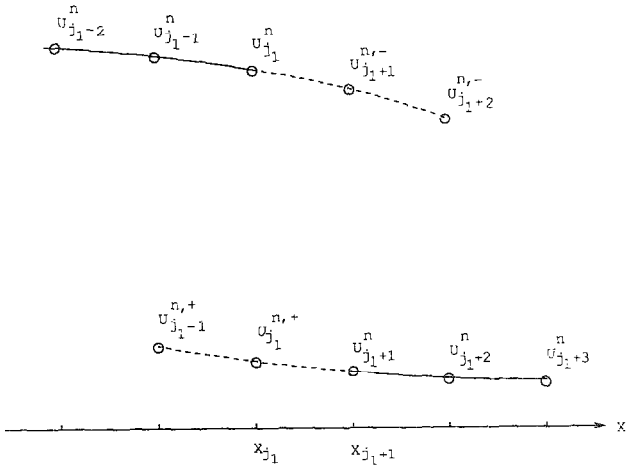


FIGURE 2.1

Also if the point (x_j, t_{n+1}) is on the right of the shock, u_j^{n+1} is computed as

$$u_j^{n+1} = u_j^n - \lambda(\hat{f}^{n,+}_{j+1/2} - \hat{f}^{n,+}_{j-1/2}) \tag{2.5}$$

rather than (2.1), where

$$\hat{f}^{n,+}_{j+1/2} = \hat{f}(u^{n,+}_{j-k+1}, \ldots, u^{n,+}_{j_1}, u^n_{j_1+1}, \ldots, u^n_{j+k}). \tag{2.6}$$

In doing so, the computation is completely prevented from crossing the shock; as a result, the oscillation and smearing are essentially eliminated.

If the shock has a nonzero speed, the generated interval must move to the left or to the right as time increases. Therefore, there are three possibilities for the generated interval on level $n+1$: it is the left adjacent cell $[x_{j_1-1}, x_{j_1}]$, it is the right adjacent cell $[x_{j_1+1}, x_{j_1+2}]$, or it is the same cell $[x_{j_1}, x_{j_1+1}]$, and the computation in each case is different. In the first case, the point (x_{j_1}, t_{n+1}) is on the right of the shock, and u_1^{n+1} is computed as

$$u^{n+1}_{j_1} = u^{n,+}_{j_1} - \lambda(\hat{f}^{n,+}_{j_1+1/2} - \hat{f}^{n,+}_{j_1-1/2}). \tag{2.7}$$

In the second case, the point (x_{j_1+1}, t_{n+1}) is on the left of the shock, and $u^{n+1}_{j_1+1}$ is computed as

$$u^{n+1}_{j_1+1} = u^{n,-}_{j_1+1} - \lambda(\hat{f}^{n,-}_{j_1+3/2} - \hat{f}^{n,-}_{j_1+1/2}). \tag{2.8}$$

In the third case, the point (x_{j_1}, t_{n+1}) is on the left, whereas the point (x_{j_1+1}, t_{n+1}) is on the right; thus, $u^{n+1}_{j_1}$ and $u^{n+1}_{j_1+1}$ are computed by (2.3) and (2.5), respectively (as shown in Fig. 2.2). A critical problem occurs when the generated interval moves to an adjacent cell—how to get the correct shock speed? In [1], this problem is

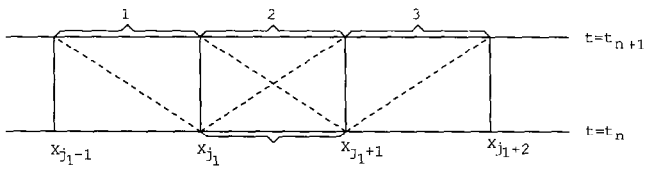


FIGURE 2.2

dealt with by using the location and the speed of the discontinuity; however, in this paper it is dealt with in a simpler way.

The modified scheme is not conservative across the cell $[x_{j_1}, x_{j_1+1}]$; nevertheless, it can be written into a conservation-like form by introducing two auxiliary variables; i.e., it can be written as

$$u_j^{n+1} = u_j^n - \lambda(\tilde{f}_{j+1/2}^n - \tilde{f}_{j-1/2}^n) + p_{j+1/2}^n - p_{j-1/2}^n + q_j^{n+1} - q_j^n, \tag{2.9}$$

where

$$\tilde{f}_{j+1/2}^n = \begin{cases} \hat{f}_{j+1/2}^{n,-}, & j \leqslant j_1 \\ \hat{f}_{j+1/2}^{n,+}, & j \geqslant j_1 + 1. \end{cases} \tag{2.10}$$

The two auxiliary variables, one along the spatial direction while the other along the temporal direction, are referred to as *artificial terms*, and they are different in the three cases. In the first case,

$$\begin{aligned} &p_{j+1/2}^n = 0 \qquad \text{for all} \quad j \neq j_1 \\ &p_{j_1-1/2}^n = -q_{j_1}^n + (u_{j_1}^n - u_{j_1}^{n,+}) + \lambda(\hat{f}_{j_1-1/2}^{n,-} - \hat{f}_{j_1-1/2}^{n,+}) \\ &q_j^{n+1} = 0 \qquad \text{for all} \quad j \neq j_1 - 1 \\ &q_{j_1-1}^{n+1} = -p_{j_1-1/2}^n. \end{aligned} \tag{2.11}$$

In the second case,

$$\begin{aligned} &p_{j+1/2}^n = 0 \qquad \text{for all} \quad j \neq j_1 \\ &p_{j_1+1/2}^n = q_{j_1}^n + \lambda(\hat{f}_{j_1+1/2}^{n,+} - \hat{f}_{j_1+1/2}^{n,-}) \\ &q_j^{n+1} = 0 \qquad \text{for all} \quad j \neq j_1 + 1 \\ &q_{j_1+1}^{n+1} = q_{j_1}^n + (u_{j_1+1}^{n,-} - u_{j_1+1}^n) + \lambda(\hat{f}_{j_1+3/2}^{n,+} - \hat{f}_{j_1-3/2}^{n,-}). \end{aligned} \tag{2.12}$$

In the third case,

$$\begin{aligned} &p_{j+1/2}^n = 0 \qquad \text{for all} \quad j \\ &q_j^{n+1} = 0 \qquad \text{for all} \quad j \neq j_1 \\ &q_{j_1}^{n+1} = q_{j_1}^n + \lambda(\hat{f}_{j_1+1/2}^{n,+} - \hat{f}_{j_1+1/2}^{n,-}). \end{aligned} \tag{2.13}$$

As opposed to u_j^n, the quantity $u_j^n - q_j^n$ is conserved in the modified scheme (2.9).

The artificial terms exist only in the close vicinity of the generated interval; in fact, they equal zero when $|j-j_1|>1$. Now let us study them in a weak sense. Suppose $v(x, t)$ is a piecewise smooth solution to IVP (1.1) that contains a unique shock, within the interval, $[x_{j_n}, x_{j_n+1}]$ $(n=1, 2, ...,)$, the generated interval on each time level. Denote $v(x_j, t_n)$ by v_j^n, the value of the solution at grid point (x_j, t_n). Suppose that $\phi(x, t)$ is a twice differentiable test function with its support in the upper half plane. Denote $\phi(x_j, t_n)$ by ϕ_j^n also. Consider the conservative difference scheme with numerical flux $\tilde{f}$ defined in (2.10). According to its definition, $\tilde{f}_{j+1/2}^n$ always contains smooth data; therefore,

$$\frac{v_j^{n+1}-v_j^n}{\tau}+\frac{\tilde{f}_{j+1/2}^n-\tilde{f}_{j-1/2}^n}{h}=O(h^r) \qquad \text{for all} \quad j\neq j_n, j_n+1 \tag{2.14}$$

if both the scheme (2.1) and the extrapolation are of the rth order. Thus, when the shock does not cross the support of ϕ,

$$\sum_{n=1}^{\infty}\sum_{j=-\infty}^{\infty}\left(\frac{\phi_j^n-\phi_j^{n-1}}{\tau}v_j^n+\frac{\phi_{j+1}^n-\phi_j^n}{h}\tilde{f}_{j+1/2}^n\right)\tau h$$
$$=-\sum_{n=0}^{\infty}\sum_{j=-\infty}^{\infty}\left(\frac{v_j^{n+1}-v_j^n}{\tau}+\frac{\tilde{f}_{j+1/2}^n-\tilde{f}_{j-1/2}^n}{h}\right)\phi_j^n\tau h=O(h^r). \tag{2.15}$$

At the grid points x_{j_n} and x_{j_n+1}, owing to the discontinuity,

$$\frac{v_j^{n+1}-v_j^n}{\tau}+\frac{\tilde{f}_{j+1/2}^n-\tilde{f}_{j-1/2}^n}{h}=O\left(\frac{1}{h}\right), \qquad j=j_n, j_n+1; \tag{2.16}$$

accordingly, the right member of (2.15) is $O(1)$ rather than $O(h^r)$ is the shock cuts the support of ϕ with a certain length. However, according to the definition of p and q,

$$\frac{v_j^{n+1}-v_j^n}{\tau}+\frac{\tilde{f}_{j+1/2}^n-\tilde{f}_{j-1/2}^n}{h}$$
$$=\frac{q_j^{n+1}-q_j^n}{\tau}+\frac{1}{\lambda}\frac{p_{j+1/2}^n-p_{j-1/2}^n}{h}+O(h^r), \qquad j=j_n, j_n+1; \tag{2.17}$$

consequently,

$$\sum_{n=1}^{\infty}\sum_{j=-\infty}^{\infty}\left(\frac{\phi_j^n-\phi_j^{n-1}}{\tau}v_j^n+\frac{\phi_{j+1}^n+\phi_j^n}{h}\tilde{f}_{j+1/2}^n\right)\tau h$$
$$=\sum_{n=1}^{\infty}\sum_{|j-j_n|\leqslant 1}\left(\frac{\phi_j^n-\phi_j^{n-1}}{\tau}q_j^n+\frac{1}{\lambda}\frac{\phi_{j+1}^n-\phi_j^n}{h}p_{j+1/2}^n\right)\tau h+O(h^r). \tag{2.18}$$

The left sum in (2.15) or (2.18) is the discretization of the integral in (1.2) with flux $\tilde{f}$, which always gets smooth data from each side of the shock. The right

member in (2.17) is the truncation error at the discontinuity in a weak sense, and the artificial terms are the principal part of it. This means that artificial terms should always take a relatively small value during the computation. Since p is related to q, we only take care of q. This gives the following criterion for the generated interval's movement. Calculate $q_{j_{n+1}}^{n+1}$ for each case, where j_{n+1} is the left endpoint of the generated interval on level $n+1$, compare them, and then choose the case that corresponds to the smallest $|q_{j_{n+1}}^{n+1}|$. The numerical examples presented in Section 6 show that this cirterion is effective.

The $x-t$ version of the treatment requires artifical terms along the spatial direction as well as the temporal direction, and the generated interval that contains the discontinuity includes only one cell. In contrast, the x version of the treatment involves only artificial terms along the spatial direction and the generated interval includes two cells. For a one-cell interval, it is impossible to let the computation on each side of it only use information from that side without involving an auxiliary term along the temporal direction. That is why the generated interval in the x version should have two cells.

Suppose $[x_{j_1-1}, x_{j_1+1}]$ is a generated interval. The treatment begins by extrapolating the numerical solution from each side of the interval to the other side and obtaining a set of extrapolated data $u_{j_1k-1}^{n,+}, \ldots, u_{j_1}^{n,+}, u_{j_1}^{n,-}, \ldots, u_{j_1+k+1}^{n,-}$. Likewise, the evaluation of the unknown u_j^{n+1} uses information only from one side of the discontinuity. If point (x_j, t_{n+1}) is on the left of the shock, u_j^{n+1} is computed by (2.3) or (2.8) with flux

$$\hat{f}_{j+1/2}^{n,-} = \hat{f}(u_{j-k+1}^n, \ldots, u_{j_1-1}^n, u_{j_1}^{n,-}, \ldots, u_{j+k}^{n,-}). \tag{2.19}$$

If point (x_j, t_{n+1}) is on the right of the shock, u_j^{n+1} is computed by (2.5) or (2.7) with flux (2.6).

The modified scheme can be written as

$$u_j^{n+1} = u_j^n - \lambda(\tilde{f}_{j+1/2}^n - \tilde{f}_{j-1/2}^n) + p_{j+1/2}^n - p_{j+1/2}^n, \tag{2.20}$$

where $\tilde{f}$ is defined in (2.10). Similarly, there are three possible locations for the generated intervals on level $n+1$, and the relevant artificial terms in each case are as follows: In the first case, in which the generated interval moves one cell to the left,

$$\begin{aligned} &p_{j+1/2}^n = 0 \qquad \text{for all} \quad j \neq j_1 - 1 \\ &p_{j_1-1/2}^n = u_{j_1} - u_{j_1}^{n,+} - \lambda(\hat{f}_{j_1-1/2}^{n,+} - \hat{f}_{j_1-1/2}^{n,-}). \end{aligned} \tag{2.21}$$

In the second case, in which the generated interval moves one cell to the right,

$$\begin{aligned} &p_{j+1/2}^n = 0 \qquad \text{for all} \quad j \neq j_1 \\ &p_{j_1+1/2}^n = u_{j_1}^{n,-} - u_{j_1}^n - \lambda(\hat{f}_{j_1+1/2}^{n,-} - \hat{f}_{j_1+1/2}^{n,+}). \end{aligned} \tag{2.22}$$

In the third case, in which the generated interval remains in the same cell.

$$p^n_{j+1/2}=0 \qquad \text{for all } j. \tag{2.23}$$

However, the value of the numerical solution at the middle point of the generated interval is computed using the information from both sides of the discontinuity. In the first case,

$$u^{n+1}_{j_1-1}=u^n_{j_1}+(u^n_{j_1-1}-u^{n,+}_{j_1})-\lambda(\hat f^{n,+}_{j_1-1/2}-\hat f^{n,-}_{j_1-3/2}); \tag{2.24}$$

in the second case,

$$u^{n+1}_{j_1+1}=u^n_{j_1}+(u^n_{j_1+1}-u^{n,-}_{j_1})-\lambda(\hat f^{n,+}_{j_1+3/2}-\hat f^{n,-}_{j_1+1/2}); \tag{2.26}$$

and in the third case,

$$u^{n+1}_{j_1}=u^n_{j_1}-\lambda(\hat f^{n,+}_{j_1+1/2}-\hat f^{n,-}_{j_1-1/2}). \tag{2.26}$$

In the x version, the value at the middle point plays the same role as q does in $x-t$ version. Being a transition point of the shock, it should be close to the arithmetic mean of the shock's left and right states; in other words,

$$|u^n_{j_1}-0.5(u^n_{j_1+1}+u^n_{j_1-1})| \tag{2.27}$$

should always take a relatively small value. This gives a criterion for the generated interval's movement.

Remark 2.1. The subcell resolution in [1] involves only the artificial terms along the spatial direction; therefore, it is an x version. If the basic scheme is the ENO scheme, the x version of this treatment is similar to the subcell resolution.

Remark 2.2. If the underlying scheme (2.1) and the extrapolation are of order r, the modified schemes of both versions still have an ordinary truncation error of the same order in the vicinity of the generated interval. This means that the treatment will cause little trouble if it captures a generated interval that does harbor a discontinuity.

3. Incorporation of the Treatment into Difference Schemes

(1) *Generation of Generated Intervals*

The treatment in this paper can be used as a front tracking technique. The front tracking methods equipped with this treatment do not need a lower dimensional adaptive grid for resolving the discontinuity as the ordinary front tracking does (see [13, 14]), and the relevant algorithm is much simpler than the ordinary one. To incroporate this treatment into a shock capturing difference scheme, a mechanism to detect discontinuities is necessary for finding out where and when a shock starts.

That is to say, a criterion for the generation of intervals must be included in the algortihm. This is very critical in dealing with a spontaneous shock.

The criterion needs a measure of nonsmoothness that is relatively big in candidate cells. A simple choice is

$$|u_{j+1}^n - u_j^n|. \tag{3.1}$$

The main drawback of this choice is that we could neglect weak shocks (or contact discontinuities) and apply the treatment to intervals that have relatively large slopes but still should be regarded as "smooth" according to a global observation of the numerical solution. A better measure of nonsmoothness involves high order difference quotients of numerical solution.

The criterion suggested in [1] is based on the ENO reconstruction and involves high order difference quotients. An important point in the criterion is that the measure of non-smoothness in a candidate interval attains a local maximum. If $|u_{j+1}^n - u_j^n|$ is the measure of non-smoothness, it requires that $|u_{j+1}^n - u_j^n|$ is the candidate cell be greater than both $|u_j^n - u_{j-1}^n|$ and $|u_{j+2}^n - u_{j+1}^n|$. This criterion mostly avoids neglecting weak shocks and contact discontinuities.

The criterion in this paper is a combination of the two above. The measure of non-smoothness $|u_{j+1}^n - u_j^n|$ in a candidate cell is greater than a constant α and attains a local maximum as well.

The generated interval in the $x-t$ version includes one cell; in contrast, the generated interval in the x version includes two cells. For this reason, the candidate cells in the x version should be extended to intervals of more than one cell. For adjacent cells, we merge them into one interval; for isolated cells, we extend them in the following way: Check the difference quotient in each left and right adjacent cell, then extend the candidate cell to the cell with greater difference quotient. In doing so we expect to eliminate spurious oscillations.

In the case that $f'' > 0$, the additional conditions $u_{j_1+1}^n - u_{j_1}^n < 0$ in the $x-t$ version and $u_{j_1+1}^n - u_{j_1-1}^n < 0$ in x version are required so that the treatment is only applied to shocks, not rarefaction waves.

After a generated interval is indentified, it is treated in a front tracking way. That is the interval that contains the discontinuity coming from the preceding level should be accepted as a generated interval. For example in the $x-t$ version, when $[x_{j_1}, x_{j_1+1}]$ is a generated interval on level n, $[x_{j_1-1}, x_{j_1}]$ in the first case, $[x_{j_1+1}, x_{j_1+2}]$ in the second case, and $[x_{j_1}, x_{j_1+1}]$ in the third case are the generated intervals on level $n+1$.

The criterion for identifying a discontinuity described in this paper is still preliminary and some more efficient criterions are under investigation. Another measure of nonsmoothness suggested is

$$\frac{|u_{j+1} - u_j|}{|u_j - u_{j-1}|}, \tag{3.2}$$

which is used by many flux-limiting schemes (see [9]). By comparison with (3.1),

which is a dimensional measure and so must be adjusted for each individual problem, (3.2) is a nondimensional measure and, therefore, independent of the individual problem. Also, a criterion based on the idea of a local adaptive stencil in ENO schemes (see [1–4]) is under investigation. For each cell, we choose a stencil that has the smallest divided differences of the numerical solution up to a certain order. In a candidate cell, its stencil and the stencil of its adjacent cell do not overlap.

No criterion can distinguish discontinuities strictly from the smooth parts of the numerical solution especially when spontaneous shocks are involved. This means that no one can avoid the "accidental effect" of the treatment to the smooth parts. However, Remark 2.2 indicates that this accidental effect is tolerable.

(2) *Performance of the Treatment*

In Section 2 we showed that the treatment can be implemented by adding the artificial terms defined in (2.11)–(2.13) or (2.21)–(2.23) to the difference scheme with flux $\hat{f}$. Similarly, in this section we show that the treatment also can be implemented by adding specific artificial terms to the underlying scheme (2.1). This makes the algorithm easy to program.

First, we discuss the $x-t$ version, for which the modified cheme can be written as

$$u_j^{n+1} = u_j^n - \lambda(\hat{f}_{j+1/2}^n - \hat{f}_{j-1/2}^n) + p_{j+1/2}^n - p_{j-1/2}^n + q_j^{n+1} - q_j^n. \tag{3.3}$$

The artificial terms in each case are as follows: In the first case, in which the generated interval moves to the left,

$$p_{j+1/2}^n = \lambda(\hat{f}_{j+1/2}^n - \hat{f}_{j+1/2}^{n,-}), \qquad j_1 - k + 1 \leqslant j \leqslant j_1 - 2, \tag{3.4}$$

$$p_{j_1-1/2}^n = -q_{j_1}^n + (u_{j_1}^n - u_{j_1}^{n,+}) + \lambda(\hat{f}_{j_1-1/2}^n - \hat{f}_{j_1-1/2}^{n,+}), \tag{3.5}$$

$$p_{j+1/2}^n = \lambda(\hat{f}_{j+1/2}^n - \hat{f}_{j+1/2}^{n,+}), \qquad j_1 \leqslant j \leqslant j_1 + k - 1, \tag{3.6}$$

and

$$\begin{aligned} q_j^{n+1} &= 0 \qquad \forall j \neq j_1 - 1 \\ q_{j_1-1}^{n+1} &= q_{j_1}^n + (u_{j_1}^n) + \lambda(\hat{f}_{j_1-1/2}^{n,+} - \hat{f}_{j_1-1/2}^{n,-}). \end{aligned} \tag{3.7}$$

In the second case, in which the generated interval moves to the right,

$$p_{j+1/2}^n = \lambda(\hat{f}_{j+1/2}^n - \hat{f}_{j+1/2}^{n,-}), \qquad j_1 - k + 1 \leqslant j \leqslant j_1 - 1, \tag{3.8}$$

$$p_{j_1+1/2}^n = q_{j_1}^n + \lambda(\hat{f}_{j_1+1/2}^n - \hat{f}_{j_1+1/2}^{n,-}), \tag{3.9}$$

$$p_{j+1/2}^n = \lambda(\hat{f}_{j+1/2}^n - \hat{f}_{j+1/2}^{n,+}), \qquad j_1 + 1 \leqslant j \leqslant j_1 + k - 1, \tag{3.10}$$

and

$$\begin{aligned} q_j^{n+1} &= 0, \qquad \forall j \neq j_1 + 1 \\ q_{j_1+1}^{n+1} &= q_{j_1}^n + (u_{j_1+1}^{n,-} - u_{j_1-1}^n) + \lambda(\hat{f}_{j_1+3/2}^{n,+} - \hat{f}_{j_1+3/2}^{n,-}). \end{aligned} \tag{3.11}$$

In the third case, in which the generated interval remains as the same cell,

$$p^n_{j+1/2} = \lambda(\hat{f}^n_{j+1/2} - \hat{f}^{n,-}_{j+1/2}), \qquad j_1 - k + 1 \leqslant j \leqslant j_1 - 1, \tag{3.12}$$

$$p^n_{j+1/2} = \lambda(\hat{f}^n_{j+1/2} - \hat{f}^{n,+}_{j+1/2}), \qquad j_1 \leqslant j \leqslant j_1 + k - 1, \tag{3.13}$$

and

$$\begin{aligned} q^{n+1}_j &= 0, \qquad \forall j \neq j_1 \\ q^{n+1}_{j_1} &= q^n_{j_1} + \lambda(\hat{f}^{n,+}_{j_1+1/2} - \hat{f}^{n,-}_{j_1+1/2}). \end{aligned} \tag{3.14}$$

The artificial terms along the temporal direction in (3.7), (3.11), and (3.14) are just the same as those in the preceding section; therefore, they determine the movement of the discontinuity.

The modified scheme for the x version can be written as

$$u^{n+1}_j = u^n_j - \lambda(\hat{f}^n_{j+1/2} - \hat{f}^n_{j-1/2}) + p^n_{j+1/2} - p^n_{j-1/2}. \tag{3.15}$$

The artificial terms in each case are as follows: In the first case, in which the generated interval moves to the left,

$$p^n_{j+1/2} = \lambda(\hat{f}^n_{j+1/2} - \hat{f}^{n,-}_{j+1/2}), \qquad j_1 - k \leqslant j \leqslant j_1 - 2, \tag{3.16}$$

$$p^n_{j_1-1/2} = u^n_{j_1} - u^{n,+}_{j_1} + \lambda(\hat{f}^n_{j_1-1/2} - \hat{f}^{n,+}_{j_1-1/2}), \tag{3.17}$$

and

$$p^n_{j+1/2} = \lambda(\hat{f}^n_{j+1/2} - \hat{f}^{n,+}_{j+1/2}), \qquad j_1 \leqslant j \leqslant j_1 + k - 1. \tag{3.18}$$

In the second case, in which the generated interval moves to the right,

$$p^n_{j+1/2} = \lambda(\hat{f}^n_{j+1/2} - \hat{f}^{n,-}_{j+1/2}), \qquad j_1 - k \leqslant j \leqslant j_1 - 1, \tag{3.19}$$

$$p^n_{j_1+1/2} = u^{n,-}_{j_1} - u^n_{j_1} + \lambda(\hat{f}^n_{j_1+1/2} - \hat{f}^{n,-}_{j_1+1/2}), \tag{3.20}$$

and

$$p^n_{j+1/2} = \lambda(\hat{f}^n_{j+1/2} - \hat{f}^{n,+}_{j+1/2}), \qquad j_1 + 1 \leqslant j \leqslant j_1 + k - 1. \tag{3.21}$$

In the third case, in which the generated interval remains the same interval,

$$p^n_{j+1/2} = \lambda(\hat{f}^n_{j+1/2} - \hat{f}^{n,+}_{j+1/2}), \qquad j_1 - k \leqslant j \leqslant j_1 - 1, \tag{3.22}$$

and

$$p^n_{j+1/2} = \lambda(\hat{f}^n_{j+1/2} - \hat{f}^{n,+}_{j+1/2}), \qquad j_1 \leqslant j \leqslant j_1 + k - 1. \tag{3.23}$$

The value of the numerical solution at the middle point of the generated interval, $u^n_{j_1}$, determines the movement of the generated interval.

(3) *High Resolution Technique*

The movement of a discontinuity fully depends on the artificial term along the temporal direction in the $x-t$ version, or on the value of the numerical solution at the middle point of the generated interval in the x version. This means that these data contain the location of the discontinuity; thus, they can locate the discontinuity within the interval. We refer to this technique as a "high resolution technique." In the $x-t$ version the coordinate of the discontinuity's location, s^n, is computed by

$$s^n = x_{j_1} + h(\tfrac{1}{2}(u^n_{j_1} - u^n_{j_1+1}) - q^n_{j_1}))/(u^n_{j_1} - u^n_{j_1+1}), \tag{3.24}$$

while in the x version it is computed by

$$s^n = x_{j_1} - 0.5h + h(u^n_{j_1} - u^n_{j_1+1})/(u^n_{j_1-1} - u^n_{j_1+1}). \tag{3.25}$$

It is easy to verify that when the problem solved is a Riemann problem containing a moving shock between two constant states, both the numerical solution and the shock location obtained by the high resolution technique are exact. In fact, the early idea of this treatment was motivated from this example.

When the initial value of the problem contains jumps that develop shocks or contact discontinuities, (3.24) and (3.25) also give the initial data for the artificial term q in the $x-t$ version and the middle point value of the numerical solution in the x version. At this moment s^n is known and $q^n_{j_1}$ or $u^n_{j_1}$ are calculated. But when the shock develops spontaneously in the computation, the artificial term q of the first detected generated interval in the $x-t$ version is zero, while the middle point value of the first detected generated interval in the x version is the original value.

(4) *Interaction of Generated Intervals*

Until now we have only studied the case that involves a single discontinuity. However, the IVP (1.1) may involve several shocks and contact discontinuities, which may collide with each other. For this reason, the interactions of generated intervals must be studied.

Only the interaction in the $x-t$ version is studied in this paper. The interaction in the x version can be treated similarly. Without loss of generality only the interaction of two generated intervals is considered. Suppose that $[x_{j_1}, x_{j_1+1}]$ and $[x_{j_2}, x_{j_2+1}]$ are two generated intervals on level n (as shown in Fig. 3.1). If the extrapolation in the treatment is of the rth order, the evaluation of the extrapolated data needs $r+1$ data from one side of the discontinuity. When the two generated intervals are so close that the number of grid points between them is less than $r+1$, the rth order extrapolation is impossible. In this case, the order of the extrapolation is reduced; in other words, $u^{n,+}_{j_1-k,m}, \ldots, u^{n,+}_{j_1,m}$ and $u^{n,-}_{j_2+1,m}, \ldots, u^{n,-}_{j_2+k+1,m}$ are evaluated by an extrapolation of the order lower than r, where $u^{n,+}_{j,m}$ and $u^{n,-}_{j,m}$ are the extrapolated data obtained from the numerical solution between the two generated intervals (see Fig. 3.1).

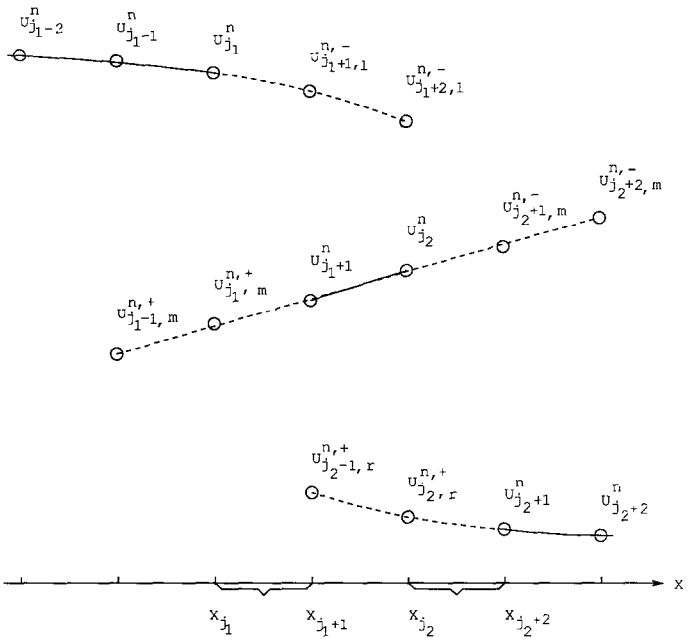


FIGURE 3.1

When the number of grid points between the two generated intervals are less than $2k$, the stencils of some of them will cross both the two generated intervals. At this moment, the numerical flux

$$\hat{f}(u^{n,+}_{j-k+1,m}, ..., u^{n,+}_{j_1,m}, u^n_{j_1+1}, ..., u^n_{j_2}, u^{n,-}_{j_2+1,m}, ..., u^{n,-}_{j+k,m}) \tag{3.26}$$

will be used so that the computation consistently picks up information from one side of the discontinuities.

The collision of generated intervals happens when the generated intervals meet, by which we mean that $j_2 = j_1 + 1$. Two kinds of situations should be studied.

(1) The left (right) generated interval moves to the right (left), while the right (left) one remains in the original cell (as shown in Figs. 3.2a, b). As a result, there are two overlapped generated intervals on level $n+1$. Treat each generated interval individually, and denote the corresponding artificial terms along the temporal direction by $q^{n+1}_{j_1,l}$ and $q^{n+1}_{j_1,r}$ ($q^{n+1}_{j_1-1,l}$ and $q^{n+1}_{j_1-1,r}$), respectively. Then take $q^{n+1}_{j_1}$ ($q^{n+1}_{j_1-1}$) = $q^{n+1}_{j_1,l} + q^{n+1}_{j_1,r}$ ($q^{n+1}_{j_1-1,l} + q^{n+1}_{j_1-1,r}$), and accept the two overlapped generated intervals on level $n+1$ as a single one.

(2) The left generated interval moves to the right while the right generated interval moves to the right (see Fig. 3.2c). There are two different way to deal with this case. The first way is to accept $[x_{j_1-1}, x_{j_1+1}]$ on level n as a single generated

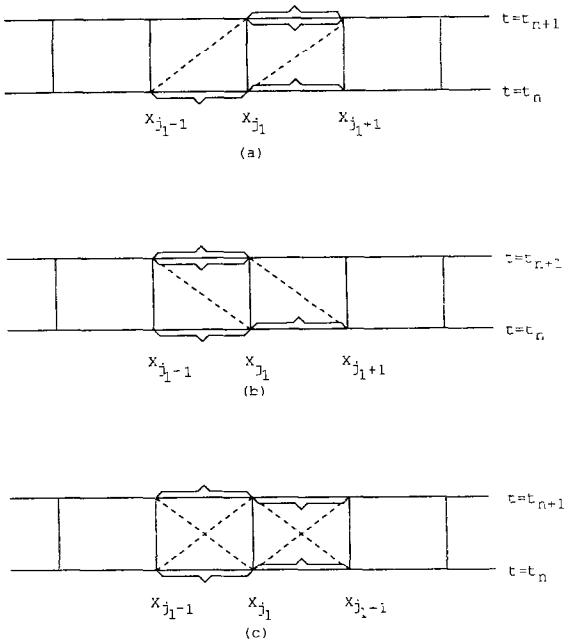


FIGURE 3.2

interval; however, it includes two cells. For this reason, the treatment of the generated interval with two cells will be described. The treatment is a simple extension of the treatment with one cell. First, the numerical solution is extrapolated from each side of the interval to the other side. Then the numerical solution on the next level is computed by (2.3)–(2.8). Nevertheless, unlike the case of a generated interval containing one cell, there are four possible cases for the movement; i.e., the generated interval on level $n+1$ could be one of $[x_{j_1-2}, x_{j_1-1}]$, $[x_{j_1-1}, x_{j_1}]$, $[x_{j_1}, x_{j_1+1}]$, and $[x_{j_1+1}, x_{j_1+2}]$. The artificial terms in each case are calculated by solving the following system:

$$\begin{aligned} u_j^{n+1} &= u_j^n - \lambda(\bar{f}_{j+1/2}^n - \bar{f}_{j-1/2}^n) + p_{j+1/2}^n - p_{j-1/2}^n - q_j^n, && j<k \\ u_k^{n+1} &= u_k^n - \lambda(\bar{f}_{k+1/2}^n - \bar{f}_{k-1/2}^n) + p_{k+1/2}^n - p_{k-1/2}^n + q_k^{n+1} - q_k^n, && j=k \\ u_j^{n+1} &= u_j^n - \lambda(\bar{f}_{j+1/2}^n - \bar{f}_{j-1/2}^n) + p_{j+1/2}^n - p_{j-1/2}^n - q_j^n, && j>k, \end{aligned} \tag{3.27}$$

where $\bar{f}_{j+1/2}^n$ is defined by (2.10), $p_{j_1-3/2}^n = p_{j_1+3/2} = 0$, and k is j_1-2, j_1-1, j_1, or j_1+1. Compare $|q_k^{n+1}|$'s, and then choose the case that has the smallest $|q_k^{n+1}|$. The implementation by adding specific terms to the underlying scheme also can be obtained in a same way as before.

The second way, which is simpler than the first, is to hold the left generated interval and move the right, or conversely. Then the case is treated in the first way.

4. Analysis

The numerical analysis in this section focuses on the influence of the treatment on the underlying scheme. Only the $x-t$ version is discussed. The x version can be discussed similarly.

Assume u_j^n is a numerical solution to (1.1) with space increment h, obtained by the modified scheme. Denote by $v_h(x, t)$ the piecewise constant extension of u_j^n, which is defined as

$$v_h(x, t) = u_j^n, \qquad (j-\tfrac{1}{2})h \leqslant x < (j+\tfrac{1}{2})h,\ n\tau \leqslant t < (n+1)\tau. \tag{4.1}$$

In addition, a discrete discontinuity is a string of generated intervals located on each time level in which the latter results from the former.

First, we prove a theorem, which is parallel to the results in [10, 15].

Theorem 4.1. *Assume that $\lambda \geqslant \alpha > 0$, where $\lambda = \tau/h$ is the mesh ratio and α is fixed. Assume $v_h(x, t)$ and its total variation are uniformly bounded with respect to h. Assume that the artificial terms along the temporal direction are uniformly bounded with respect to h. Assume that the number of discrete discontinuities involved in the solution is uniformly finite with respect to h. Then there exists a weak solution solution to* (1.1), *which is the limit of a convergent sequence selected from $\{v_h\}$.*

Proof. The convergent sequence is obtained by Helly's theorem in the same way as in [15]. Without confusion it is still denoted by $\{v_h\}$ and the corresponding numerical solution is denoted by u_j^n. Denote by $u(x, t)$ the limit function of the sequence, which is bounded and has bounded total variation.

Multiply (2.9) by $\phi_j^n \tau$ and sum it with respect to j and n. Apply summation by parts to it. We obtain

$$\sum_{n=1}^{\infty}\sum_{j=-\infty}^{\infty}\left(\frac{\phi_j^n-\phi_j^{n-1}}{\tau}u_j^n+\frac{\phi_{j+1}^n-\phi_j^n}{h}\bar{f}_{j+1/2}^n\right)\tau h+\sum_{j=-\infty}^{\infty}\phi_j^0 u_j^0 h$$

$$=\sum_{n=1}^{\infty}\sum_{j=-\infty}^{\infty}\left(\frac{\phi_j^\phi-\phi_j^{n-1}}{\tau}q_j^n+\frac{1}{\lambda}\frac{\phi_{j+1}^n-\phi_j^n}{h}p_{j+1/2}^n\right)\tau h+\sum_{n=-\infty}^{\infty}\phi_j^0 q_j^0 h. \tag{4.2}$$

According to its definition, $\bar{f}$ is also consistent with $f(u)$ is the sense (2.2). Since the sequence converges to $u(x, t)$, the left member of (4.2) converges to the integral in (1.2) as h tends to zero. According to the discussion in the last section, the artificial terms p are related to the artificial terms q and the numerical solution; thus, p is also uniformly bounded since q and the numerical solution are uniformly bounded. Because of the last assumption, the area of the region in which the artificial terms are nonzero is of $O(h)$. Therefore, the right member of (4.2) converges to zero as h tends to zero.

The theorem assumes that the involved discrete discontinuities are uniformly finite. This implies that the exact solution is piecewise smooth. A theoretic result without this assumption is under investigation.

The assumption of the uniform boundedness of the numerical solution and its total variation and the assumption of the uniform boundedness of artificial term q are critical in the above theorem. The following part of this section studies the influence of the treatment on the variation of numerical solutions and artificial term q.

In the study of the influence on the variation, we start with a TVD scheme that can be written as

$$u_j^{n+1} = u_j^n + C_{j+1/2}^{n,+} \Delta_{j+1/2}^n u - C_{j-1/2}^{n,-} \Delta_{j-1/2}^n u, \tag{4.3}$$

where

$$\begin{gathered} C_{j+1/2}^{n,+} \geqslant 0, \qquad C_{j+1/2}^{n,-} \geqslant 0, \\ C_{j+1/2}^{n,+} + C_{j+1/2}^{n,-} \leqslant 1, \end{gathered} \tag{4.4}$$

and

$$\begin{aligned} C_{j+1/2}^{n,+} &= \lambda \frac{f(u_j^n) - \hat{f}_{j+1/2}^n}{\Delta_{j+1/2}^n u}, \\ C_{j+1/2}^{n,-} &= \lambda \frac{f(u_{j+1}^n) - \hat{f}_{j+1/2}^n}{\Delta_{j+1/2}^n u}. \end{aligned} \tag{4.5}$$

These schemes do not increase the total variation of a numerical solution (refer to [5, 16]). Furthermore, we assume

$$\lambda \left| (f(v_j) - \hat{f}_{j+1/2}) + (f(v_{j+1}) - \hat{f}_{j+1/2}) \right| \leqslant |v_{j+1} - v_j|. \tag{4.6}$$

It is easy to see that (4.6) implies

$$\hat{f}(v_{j-k+1}, \ldots, v_{j-1}, w, w, v_{j+2}, \ldots, v_{j+k}) = f(w). \tag{4.7}$$

Such essetially 3-point schemes include, beside the standard 3-point schemes, several recently constructed second-order accurate converging schemes (refer to [16]).

The rth-order treatment is the treatment that uses the rth-order extrapolation.

THEOREM 4.2. *If the underlying scheme is a TVD scheme, the* 0th *order treatment keeps the TVD property in the vicinity of a generated interval when there is no interaction of discontinuities.*

The term "vicinity" indicates the region the treatment affects. When the under-

lying scheme is of $(2k+1)$-point and $[x_{j_1}, x_{j_1+1}]$ is the generated interval, it is the interval $[x_{j_1-k}, x_{j_1+k+1}]$. Thus, it is required to show that

$$\begin{aligned}\sum_{j=j_1-k}^{j_1+k} |\Delta^{n+1}_{j+1/2}u| \leqslant\ & C^{n,+}_{j_1+k+1/2}\,|\Delta^n_{j_1+k+1/2}u| \\ & + (1 - C^{n,-}_{j_1+k-1/2})\,|\Delta^n_{j_1+k-1/2}u| \\ & + \sum_{j=j_1-k+1}^{j_1+k-1} |\Delta^n_{j+1/2}u| + (1 - C^{n,+}_{j_1-k+1/2})\,|\Delta^n_{j_1-k+1/2}u| \\ & + C^{n,-}_{j_1-k-1/2}\,|\Delta^n_{j_1-k-1/2}u|. \end{aligned} \tag{4.8}$$

Proof. Assume that $[x_{j_1}, x_{j_1+1}]$ is a generated interval on level n. When $j \leqslant j_1 - 1$, or $j \geqslant j_1 + 2$, u_j^{n+1} is evaluated by (2.3) or (2.5). Since the underlying scheme has a TVD form of (4.3),

$$u_j^{n+1} = u_j^n + C^{n,+}_{j+1/2,l}\,\Delta^n_{j+1/2}u - C^{n,-}_{j-1/2,l}\,\Delta^n_{j-1/2}u, \qquad j \leqslant j_1 - 1, \tag{4.9}$$

and

$$u_j^{n+1} = u_j^n + C^{n,+}_{j+1/2,r}\,\Delta^n_{j+1/2}u - C^{n,-}_{j-1/2,r}\,\Delta^n_{j-1/2}u, \qquad j \geqslant j_1 + 2, \tag{4.10}$$

where

$$\begin{aligned} C^{n,+}_{j+1/2,l} &= \lambda\,\frac{f(u_j^n) - \hat{f}^{n,-}_{j+1/2}}{\Delta^n_{j+1/2}u} \\ C^{n,-}_{j+1/2,l} &= \lambda\,\frac{f(u_{j+1}^n) - \hat{f}^{n,-}_{j+1/2}}{\Delta^n_{j+1/2}u}, \end{aligned} \tag{4.11}$$

and

$$\begin{aligned} C^{n,+}_{j+1/2,r} &= \lambda\,\frac{f(u_j^n) - \hat{f}^{n,+}_{j+1/2}}{\Delta^n_{j+1/2}u} \\ C^{n,-}_{j+1/2,r} &= \lambda\,\frac{f(u_{j+1}^n) - \hat{f}^{n,+}_{j+1/2}}{\Delta^n_{j+1/2}u}. \end{aligned} \tag{4.12}$$

Equations (4.11) and (4.12) satisfy (4.4). This means that the modified scheme still has a TVD form at these points but with different coefficients.

When the generated interval on level $n+1$ remains in $[x_{j_1}, x_{j_1+1}]$, $u_{j_1}^{n+1}$ and $u_{j_1+1}^{n+1}$ are still given by (2.3) and (2.5), respectively. Since the extrapolation is of zeroth order,

$$\begin{aligned} u_{j_1}^{n+1} &= u_{j_1}^n - C^{n,-}_{j-1/2,l}\,\Delta_{j_1-1/2}u \\ u_{j_1+1}^{n+1} &= u_{j_1+1}^n + C^{n,+}_{j+3/2,r}\,\Delta^n_{j_1+3/2}u; \end{aligned} \tag{4.13}$$

thus,

$$\begin{aligned}
\Delta^{n+1}_{j_1-1/2}u &= (1-C^{n,+}_{j_1-1/2,l}-C^{n,-}_{j_1-1/2,l})\,\Delta^n_{j_1-1/2}u + C^{n,-}_{j_1-3/2,l}\,\Delta^n_{j_1-3/2}u\\
\Delta^{n+1}_{j_1+1/2}u &= \Delta^n_{j_1+1/2}u + C^{n,-}_{j_1+3/2,r}\,\Delta^n_{j_1+3/2}u + C^{n,-}_{j_1-1/2,l}\,\Delta^n_{j_1-1/2}u\\
\Delta^{n+1}_{j_1+3/2}u &= C^{n,+}_{j_1+5/2,r}\,\Delta^n_{j_1+5/2}u + (1-C^{n,+}_{j_1+3/2,r}-C^{n,-}_{j_1+3/2,r})\,\Delta^n_{j_1+3/2}u.
\end{aligned} \tag{4.14}$$

When the generated interval on level $n+1$ moves to the right,

$$\begin{aligned}
u^{n+1}_{j_1} &= u^n_{j_1} - C^{n,-}_{j_1-1/2,l}\,\Delta^n_{j_1-1/2}u\\
u^{n+1}_{j_1+1} &= u^n_{j_1};
\end{aligned} \tag{4.15}$$

thus,

$$\begin{aligned}
\Delta^n_{j_1-1/2}u &= (1-C^{n,+}_{j_1-1/2,l}-C^{n,-}_{j_1-1/2,l})\,\Delta^n_{j_1-1/2}u + C^{n,-}_{j_1-3/2,l}\,\Delta^n_{j_1-3/2}u\\
\Delta^{n+1}_{j_1+1/2}u &= C^{n,-}_{j_1-1/2,l}\,\Delta^n_{j_1-1/2}u\\
\Delta^{n+1}_{j_1+3/2} &= C^{n,+}_{j_1+5/2,r}\,\Delta^n_{j_1+5/2}u + (1-C^{n,+}_{j_1+3/2,r})\,\Delta^n_{j_1+3/2}u + \Delta^n_{j_1+1/2}u.
\end{aligned} \tag{4.16}$$

Obviously, both (4.14) and (4.16) with (4.9) and (4.10) lead to (4.8). The same conclusion is also true for the case when the gerated interval moves to the left. Thus, the proof is complete.

When the order of the extrapolation is higher than 0, (4.13) and (4.15) are replaced by

$$\begin{aligned}
u^{n+1}_{j_1} &= u^n_{j_1} + C^{n,+}_{j_1+1/2,l}\,\Delta^n_{j_1+1/2,l}u - C^{n,-}_{j_1-1/2,l}\,\Delta^n_{j_1-1/2}u\\
u^{n+1}_{j_1+1} &= u^n_{j_1+1} + C^{n,+}_{j_1+3/2,r}\,\Delta^n_{j_1+3/2}u - C^{n,-}_{j_1+1/2,r}\,\Delta^n_{j_1+1/2,r}u
\end{aligned} \tag{4.17}$$

and

$$\begin{aligned}
u^{n+1}_{j_1} &= u^n_{j_1} + C^{n,+}_{j_1+1/2,l}\,\Delta^n_{j_1+1/2,l}u - C^{n,-}_{j_1-1/2,l}\,\Delta^n_{j_1-1/2}u\\
u^{n+1}_{j_1+1} &= u^n_{j_1} + (u^{n,-}_{j_1+1} - u^n_{j_1}) + C^{n,+}_{j_1+3/2,l}\,\Delta^n_{j_1+3/2,l}u - C^{n,-}_{j_1+1/2,l}\,\Delta^n_{j_1+1/2,l}u,
\end{aligned} \tag{4.18}$$

where

$$\begin{aligned}
\Delta^n_{j_1+3/2,l}u &= u^{n,-}_{j_1+2} - u^{n,-}_{j_1+1}\\
\Delta^n_{j_1+1/2,l}u &= u^{n,-}_{j_1+1} - u^n_{j_1}\\
\Delta^n_{j_1+1/2,r}u &= u^n_{j_1+1} - u^{n,+}_{j_1}.
\end{aligned} \tag{4.19}$$

The appearance of the terms in (4.19) kills the TVD property in the vicinity of the generated interval. Nevertheless, these terms are only of $O(h)$ if the numerical solution is "smooth" on each side of the discontinuity. Hence, the error caused by them is also of $O(h)$ on each time level. Therefore, the total variation of the numerical solution will be uniformly bounded if the solution is piecewise smooth.

Finally, when the underlying scheme is not a TVD scheme, the numerical solution obtained by the modified scheme still has uniformly bounded total variation if it is piecewise smooth without interactions of discontinuities. In fact, for each j,

$$|u_{j+1}^{n+1} - u_j^{n+1}| \leqslant |u_{j+1}^n - u_j^n| + \lambda\,|\hat{f}_{j+3/2}^n - 2\hat{f}_{j+1/2}^n + \hat{f}_{j-1/2}^n|. \tag{4.20}$$

The second term on the right in (4.20) is of $O(h^2)$ if the numerical solution is smooth. Therefore,

$$\sum_{j=-\infty}^{\infty} |u_{j+1}^{n+1} - u_j^{n+1}| \leqslant \sum_{j=-\infty}^{\infty} |u_{j+1}^n - u_j^n| + O(h) \tag{4.21}$$

if the support of the solution is bounded. This implies that the solution has uniformly bounded total variation.

Suppose that the solution has one discrete discontinuity, contained in the generated interval $[x_{j_1}, x_{j_1+1}]$ on level n. When the generated interval on level $n+1$ remains in the interval,

$$|u_{j+1}^{n+1} - u_j^{n+1}| \leqslant |u_{j+1}^n - u_j^n| + O(h^2) \tag{4.22}$$

still holds for $j \neq j_1$ because the computation uses only the extrapolated data across the interval. For $j = j_1$,

$$\begin{aligned} |u_{j_1+1}^{n+1} - u_{j_1}^{n+1}| = |u_{j_1+1}^n - u_{j_1}^n| &+ \lambda\,|\hat{f}_{j_1+3/2}^{n,+} - \hat{f}_{j_1+1/2}^{n,+}| \\ &+ \lambda\,|\hat{f}_{j_1+1/2}^{n,-} - \hat{f}_{j_1-1/2}^{n,-}|. \end{aligned} \tag{4.23}$$

The last two terms on the right in (4.23) are of $O(h)$ due to the smoothness of the solution on each side. As a consequence, (4.21) still holds in this case. The same conclusion is true when the generated interval on level $n+1$ moves to the left or to the right. Thus, the total variation of the numerical solution is bounded. The result is easily extended to the case of piecewise smooth solution without interactions.

The case that involves interactions of discontinuities remains to be studied. It is quite possible that the treatment of interactions does not affect much the boundedness of the total variation either.

Now we study the artificial term q, starting with the zeroth-order treatment. According to (2.11)–(2.13), q on level $n+1$ in each case at this moment is

$$q_{j_1-1}^{n+1} = q_{j_1}^n + (u_{j_1+1}^n - u_{j_1}^n) + \lambda(f(u_{j_1+1}^n) - f(u_{j_1}^n)) + \lambda(f(u_{j_1}^n) - \hat{f}_{j_1-1/2}^{n,-}), \tag{4.24}$$

$$q_{j_1+1}^{n+1} = q_{j_1}^n + (u_{j_1}^n - u_{j_1+1}^n) + \lambda(f(u_{j_1+1}^n) - f(u_{j_1}^n)) + \lambda(\hat{f}_{j_1+3/2}^{n,+} - f(u_{j_1+1}^n)), \tag{4.25}$$

or

$$q_{j_1}^{n+1} = q_{j_1}^n + \lambda(f(u_{j_1+1}^n) - f(u_{j_1}^n)), \tag{4.26}$$

where $[x_{j_1}, x_{j_1+1}]$ is the generated interval on level n. The last terms on the right

in (4.24) and (4.25) are of $O(h)$ if the numerical solution is smooth on each side of the interval. When the CFL-condition is satisfied for λ; i.e., $\lambda |f'(u)| < 1$, and h is small enough, one of $q_{j_1-1}^{n+1}$, $q_{j_1+1}^{n+1}$, and $q_{j_1}^{n+1}$ is bigger than $q_{j_1}^n$, and one of them is smaller than $q_{j_1}^n$. The movement of the generated interval is determined by choosing the smallest one among $q_{j_1-1}^{n+1}$, $q_{j_1-1}^{n+1}$, and $q_{j_1}^{n+1}$; therefore, q is uniformly bounded with respect to h.

If the order of the treatment is higher than 0, the artificial term q in each case is given as in (2.11)–(2.13). However, the difference between (4.24)–(4.26) and (2.11)–(2.13) are of $O(h)$; thus, the above discussion still holds.

5. Application of the Treatment to Euler Equations of Gas Dynamics

The Euler equations of gas dynamics for a polytropic gas are

$$u_t + f(u)_x = 0, \tag{5.1a}$$

$$u = (\rho, m, E)^T, \tag{5.1b}$$

$$f(u) = qu + (0, p, qp)^T, \tag{5.1c}$$

$$p = (\gamma - 1)(E - \tfrac{1}{2} pq^2), \tag{5.1d}$$

where ρ, q, p, and E are the density, velocity, pressure, and total energy, respectively, $m = \rho q$ is the momentum and γ is the ratio of specific heats.

The eigenvalues of the Jacobian matrix $A(u) = \partial f / \partial u$ are

$$a_1(u) = q - u, \qquad a_2(u) = q, \qquad a_3(u) = q + u, \tag{5.2}$$

where $c = (\gamma p/\rho)^{1/2}$ is the sound speed. In this section the application of the treatment to the above system is described.

In [1] the "subcell resolution" is applied to the linear degenerate characteristic field; i.e., it is only applied to the second locally defined characteristic variable in a (linear) field-by-field way. However, the treatment in this paper is applied to discontinuities of all different kinds in a nonlinear way. The treatment is only used to track the discontinuities; put more precisely, all the discontinuities start from the initial level, and no new discrete discontinuity is generated during the computation. Correspondingly, all the numerical examples of the Euler system reported in the next section are of piecewise smooth solutions with finite discontinuities. No spontaneous shock is involved.

Only the $x - t$ version is concerned. We begin with a simple case that contains one discontinuity, on each side of which the numerical solution is smooth. There are three different kinds of discontinuities in the Euler system: the left shock, the right shock, and the contact discontinuity. Therefore, the generated intervals are classified into three different kinds also: the left shock generated interval (LSGI),

the right shock generated interval (RSGI), and the contact discontinuity generated interval (CDGI), each of which is identified by solving the Riemann problem that has $u^n_{j_1}$ and $u^n_{j_1+1}$ as its left and right states, where $[x_{j_1}, x_{j_1+1}]$ is the corresponding generated interval. We denote the Riemann problem by $RP(u^n_{j_1}, u^n_{j_1+1})$. The solution to the $RP(u^n_{j_1}, u^n_{j_1+1})$ in an LSGI or an RSGI has a relatively strong left or right shock, while the solution in a CDGI has a relatively strong contact discontinuity.

Assume $[x_{j_1}, x_{j_1+1}]$ is the generated interval on level n. As before, the first step is to extrapolate the numerical solution from each side of the generated interval to the other side. The extrapolation data are $u^{n,+}_{j_1-k}$, $u^{n,+}_{j_1-k+1}$, ..., $u^{n,+}_{j_1}$, $u^{n,-}_{j_1+1}$, $u^{n,-}_{j_1+2}$, ..., $u^{n,-}_{j_1+k+1}$. However, if the numerical solution is evaluated by (2.3)–(2.8), the treatment would affect the fields of other characteristics. The revision is as follows: If $[x_{j_1}, x_{j_1+1}]$ is an LSGI, we solve the Riemann problem $RP(u^n_{j_1-i}, u^{n,+}_{j_1-i})$ $(i=0, ..., k)$, and obtain a set of left middle states $u^{n,l}_{j_1-i,*}$ $(i=0, ..., k)$ (as shown in Fig. 5.1). We replace the $u^{n,+}_{j_1-i}$ in (2.3)–(2.8) by these left middle states. If $[x_{j_1}, x_{j_1+1}]$ is a RSGI, we solve the Riemann problem $RP(u^{n,-}_{j_1+i}, u^n_{j_1+i})$ $(i=1, ..., k+1)$, and replace the $u^{n,-}_{j_1+i}$ in (2.3)–(2.8) by the corresponding right middle states $u^{n,r}_{j_1+i,*}$.

If the original system (5.1a) is linear with constant coefficients, such a handling is equivalent to applying the treatment in a field-by-field way; i.e., it is to do a field-by-field decomposition at first, and then to apply the treatment to each characteristic variable. But in the nonlinear case (such as the Euler system) they are different, especially when strong discontinuities are involved. The present one is more precise, for it involves a "nonlinear" field-by-field decomposition by solving an exact Riemann problem.

The treatment of a CDGI should be as follows: Solve the Riemann problem $RP(u^n_{j_1-i}, u^{n,+}_{j_1-i})$ and get the left middle states $u^{n,l}_{j_1-i,*}$ $(i=0, ..., k)$, and solve the Riemann problems $RP(u^{n,-}_{j_1+i}, u^n_{j_1+i})$ and get the right middle states $u^{n,r}_{j_1+i,*}$ $(i=1, ..., k+1)$. Replace $u^{n,+}_{j_1-i}$ by $u^{n,l}_{j_1-i,*}$, and $u^{n,-}_{j_1+i}$ by $u^{n,r}_{j_1+i,*}$ in (2.3)–(2.8). However, the practical used treatment is simpler than this one. The density is discontinuous across the contact discontinuity, but the velocity and the pressure are

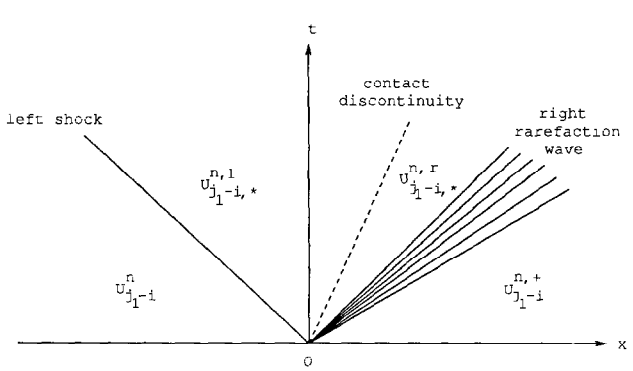


FIGURE 5.1

continuous. Therefore, we only extrapolate ρ and take the original data for the velocity q and the pressure p across the CDGI. The examples presented in Section 6 show that the treatment works well.

The movement of a generated interval is determined by keeping $|q_{j_1}^n|$ small, where $q_{j_1}^n$ is the artificial term along the temporal direction. Because $q_{j_1}^n$ is a vector with three components now, $|q_{j_1}^n|$ must be some norm of $q_{j_1}^n$. A natural candidate is

$$|q_{j_1}^n| = \sum_{l=1}^{3} |q_{j_1}^{n,(l)}|, \tag{5.3}$$

where $q_{j_1}^{n,(l)}$ is a component of $q_{j_1}^n$. The three components of the system (5.1) may have quite different scales; for this reason, the presently used $|q_{j_1}^n|$ is

$$|q_{j_1}^n| = \sum_{l=1}^{3} |q_{j_1}^{n,(l)}/(u_{j_1}^{n,(l)} - u_{j_1+1}^{n,(l)})|, \tag{5.4}$$

where $u_j^{n,(l)}$ is the component of u_j^n.

The solution to the Riemann problem $RP(u_{j_1}^n, u_{j_1+1}^n)$ may have more than one strong discontinuity. This situation occurs when two generated intervals of different kinds collide with each other. The corresponding generated interval is referred as a node. A typical case is that the $RP(u_{j_1}^n, u_{j_1+1}^n)$ contains a left and a right shock with a middle state u_*. The main idea in treating nodes is the consideration that there are more than one different generated intervals in $[x_{j_1}, x_{j_1+1}]$, which overlap each other, and the treatment should be precisely applied to each one. Let's use the above case to describe the algorithm.

Suppose $[x_{j_1}, x_{j_1}]$ is the generated interval. We introduce two auxiliary initial value problems, of which the initial value of the first one has the left part of u^n as its left part and the middle state u_* as its right part with an LSGI $[x_{j_1}^n, x_{j_1+1}^n]$, and the initial value of the second one has the right part of u^n as its right part and the middle state u_* as its left part with an RSGI $[x_{j_1}, x_{j_1+1}]$. The numerical results of these two problems determine the movements of the two generated intervals, which are located in the same cells for the first several time-steps until they separate, and then the middle state appears between them. Before the two generated intervals separate, the numerical solution is computed just as only one generated interval exists.

If the original system is a linear system with constant coefficients, such a handling is also equivalent to applying the treatment in a field-by-field way. The treatment is easily extended to the more complicated nodes for which the $RP(u_{j_1}^n, u_{j_1+1}^n)$'s involve three different discontinuities.

6. Numerical Experiments

In this section some numerical results are presented to show the performance of our method.

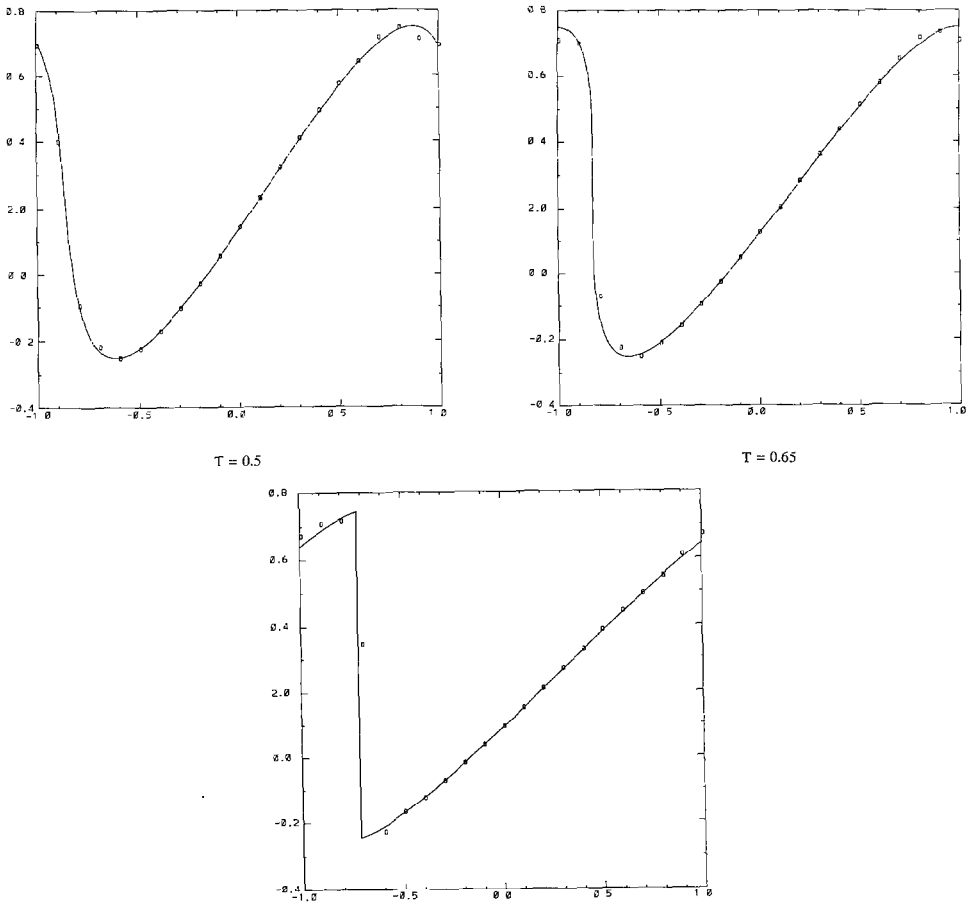


FIG. 6.1a. $T=0.5$; $T=0.65$; $T=1.1$.

EXAMPLE 1. The following initial value problem

$$u_t+\left(\frac{u^2}{2}\right)_x=0, \qquad -1\leqslant x\leqslant 1$$
$$u(x,0)=\tfrac{1}{4}+\tfrac{1}{2}\sin \pi x \tag{6.1}$$

is solved. The exact solution is smooth up to $t=2/\pi$, then it develops a moving shock, which interacts with the rarefaction waves. The exact solution is obtained by Newton iteration. For details, see [4].

The Lax–Wendroff scheme is chosen as the underlying scheme, and λ is taken to be 0.5. The parameter α, which is used in the generation of generated intervals, is taken to be 0.1. The zeroth-order treatment is tested.

In Figs. 6.1a and b the numerical solution obtained with x and $x-t$ version for

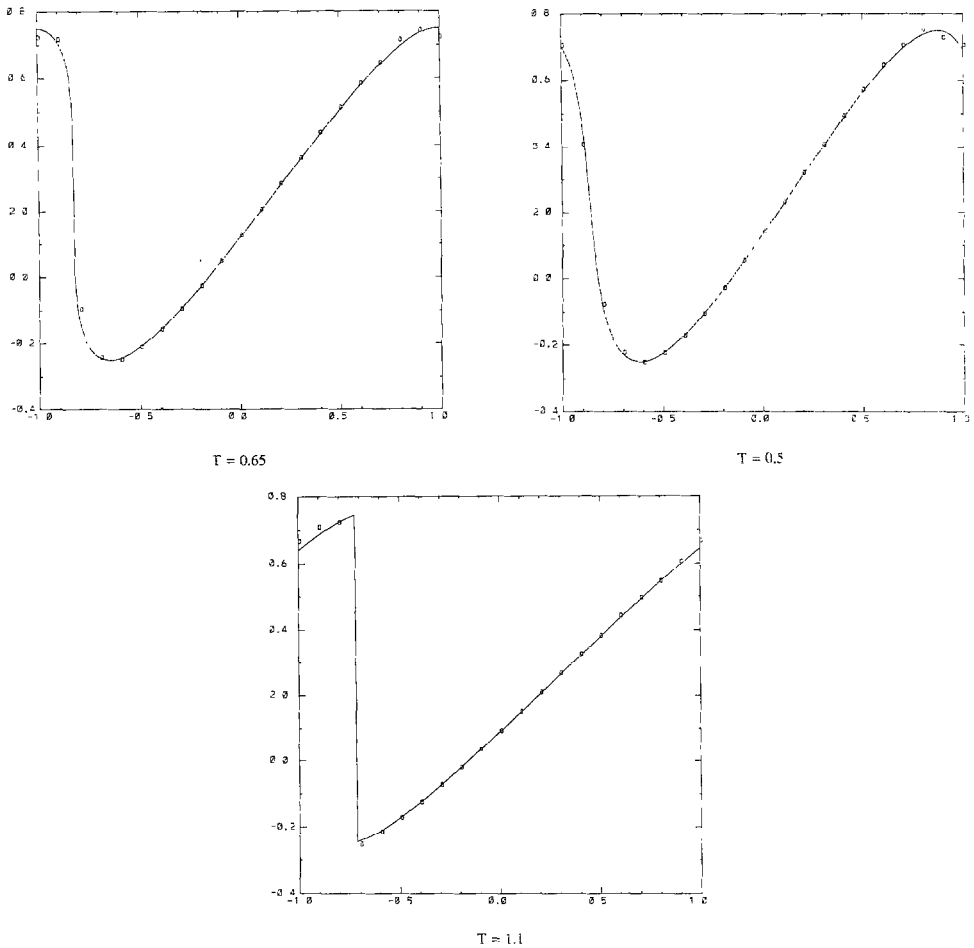


FIG. 6.1b. $T=0.5$; $T=0.65$; $T=1.1$.

$\Delta x = \frac{1}{10}$ are displayed, where t is taken to be 0.5, 0.65 ($\approx 2/\pi$, and 1.1, respectively. In Figs. 6.2a and b the results for $\Delta x = \frac{1}{40}$ are displayed.

In this example, the numerical solutions approximate the exact solution very well in both smooth and nonsmooth parts. The discrete shock (or shocks) develops before $t = 2/\pi$, and sometimes there are interactions of generated intervals; however, the treatment of interactions causes little damage to the smooth parts.

The location of the discontinuity on each time level is computed by the high resolution technique described in Section 3. Tables I and II compare the locations of the discrete and exact shocks. Table I displays the results of the x version for $\Delta x = \frac{1}{10}$ from the 14th to the 22nd time-step, while Table II displays the results of the $x-t$ version for $\Delta x = \frac{1}{40}$ from the 52nd to 88th time-step. The locations of the discrete shocks in both examples are accurate.

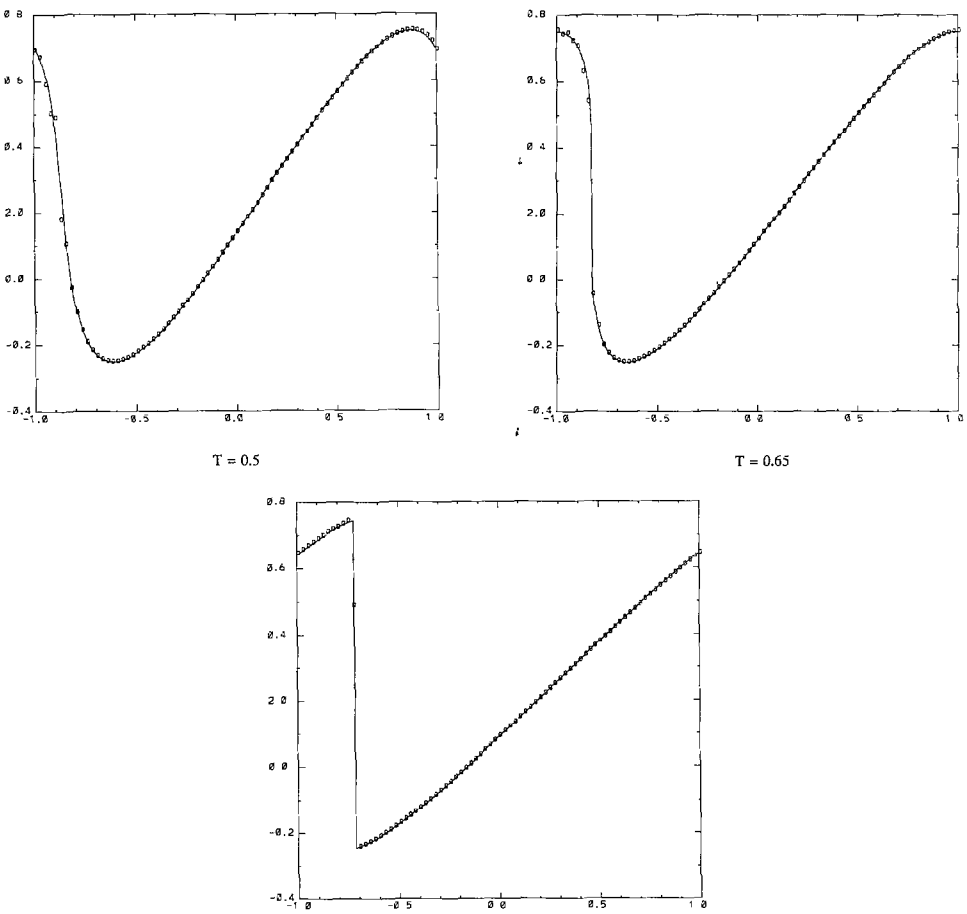


FIG. 6.2a. $T=0.5$; $T=0.65$; $T=1.1$.

The $x-t$ version has no shock transition point; in contrast, the x version has one shock transition point. Therefore, the $x-t$ version has higher resolution for a shock than the x version. In addition, the $x-t$ version is easier to program than the x version, especially in dealing with interactions of generated intervals.

Treatments higher than the zeroth-order one were also tested, andthe results are the same as for the zeroth-order treatment; occasionally, the results of the zeroth-order are a little better than those of high order. According to Remark 2.2, the zeroth-order treatment kills the consistency for it causes an L^{∞} truncation error of $O(1)$. However, owing to the converging characteristics, a shock has a self-sharpening mechanism, which presents the $O(1)$ error to be transported to the smooth region. That is why the zeroth-order treatment causes little damage to the smooth parts.

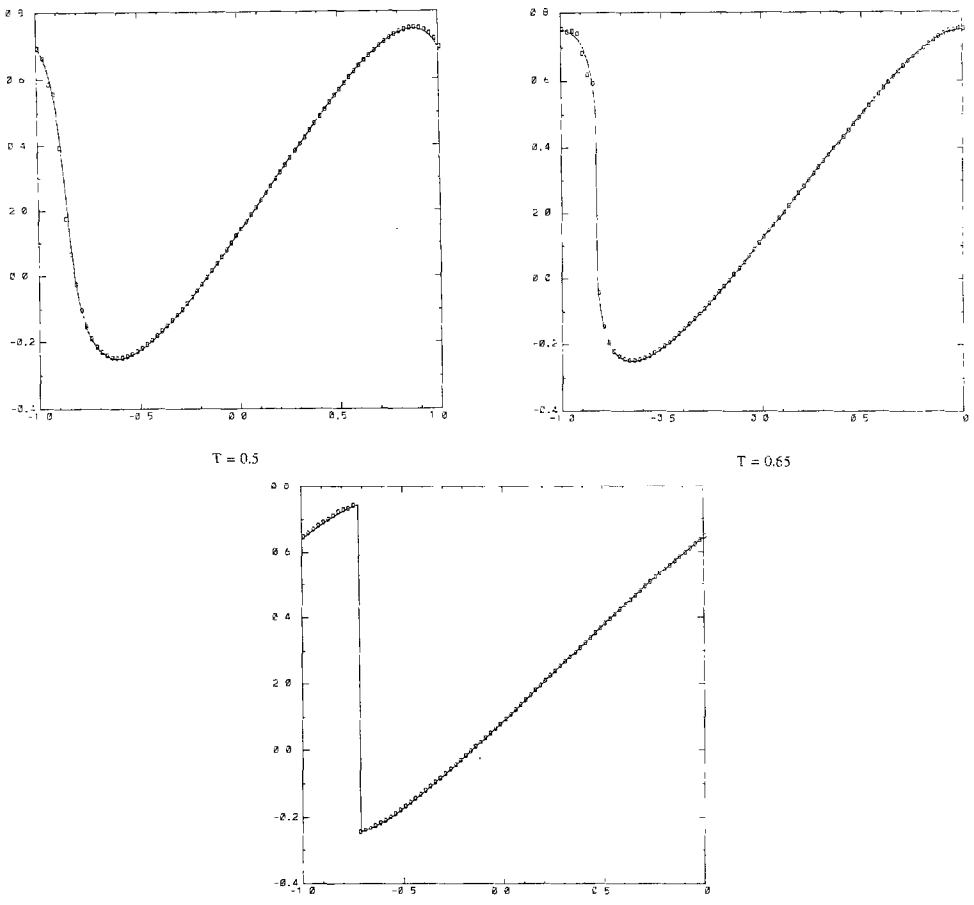


FIG. 6.2b. $T = 0.5$; $T = 0.65$; $T = 1.1$.

TABLE I

Locations of Discrete and Exact Shocks

n	Discrete	Exact
14	−0.8241019	−0.8250000
15	−0.8116784	−0.8125000
16	−0.7993561	−0.8000000
17	−0.7870850	−0.7875000
18	−0.7748046	−0.7750000
19	−0.7624528	−0.7625000
20	−0.7499743	−0.7500000
21	−0.7369952	−0.7375000
22	−0.7254016	−0.7250000

TABLE II

Locations of Discrete and Exact Shocks

n	Discrete	Exact	n	Discrete	Exact
52	−0.8384687	−0.8375000	71	−0.7787599	−0.781250
53	−0.8350775	−0.8343570	72	−0.7756485	−0.7750000
57	−0.8203193	−0.8218750	76	−0.7628636	−0.7625000
58	−0.8174989	−0.8187500	77	−0.7596936	−0.7593750
59	−0.8176763	−0.8156250	78	−0.7565193	−0.7562500
60	−0.8118547	−0.8125000	79	−0.7533406	−0.7531250
61	−0.8090336	−0.8093750	80	−0.7501586	−0.7500000
62	−0.8062074	−0.8062500	81	−0.7470106	−0.7468750
63	−0.8033639	−0.8031250	82	−0.7438717	−0.7437500
64	−0.8004860	−0.8000000	83	−0.7407333	−0.7406250
65	−0.7967674	−0.7968750	84	−0.7375951	−0.7375000
66	−0.7938213	−0.7937500	85	−0.7344573	−0.7343750
67	−0.7908556	−0.7906250	86	−0.7313206	−0.7312500
68	−0.7878712	−0.7875000	87	−0.7281857	−0.7281250
69	−0.7848648	−0.7843750	88	−0.7250538	−0.7250000
70	−0.7818301	−0.7812500			

EXAMPLE 2. The following linear IVP is solved to test the $x-t$ version treatment:

$$u_t + u_x = 0 \tag{6.2a}$$

$$u_0(x+0.5) = \begin{cases} -x\sin(\frac{3}{2}\pi x^2), & -1 < x < -\frac{1}{3} \\ |\sin(2\pi x)| & |x| < \frac{1}{3} \\ 2x - 1 - \sin(3\pi x)/6 & \frac{1}{3} < x < 1 \end{cases} \tag{6.2b}$$

$$u_0(x+2) = u_0(x). \tag{6.2c}$$

The solution to the IVP contains three contact discontinuities and one weak discontinuity (i.e., discontinuity of derivative). The numerical results without treatment are almost unacceptable.

The Lax–Wendroff scheme is again chosen as the underlying scheme, and $\lambda = 0.8$. The treatment is used to track the discontinuities. The treatment is applied to the contact disontinuities as well as the weak discontinuity. The motivation for applying the treatment to the weak discontinuity is based on the consideration that preventing computation from crossing the discontinuity is also applicable to it. Because of the lack of converging characteristics, the numerical results are sensitive to the order of the treatment. The higher the order, the better the resolution of the

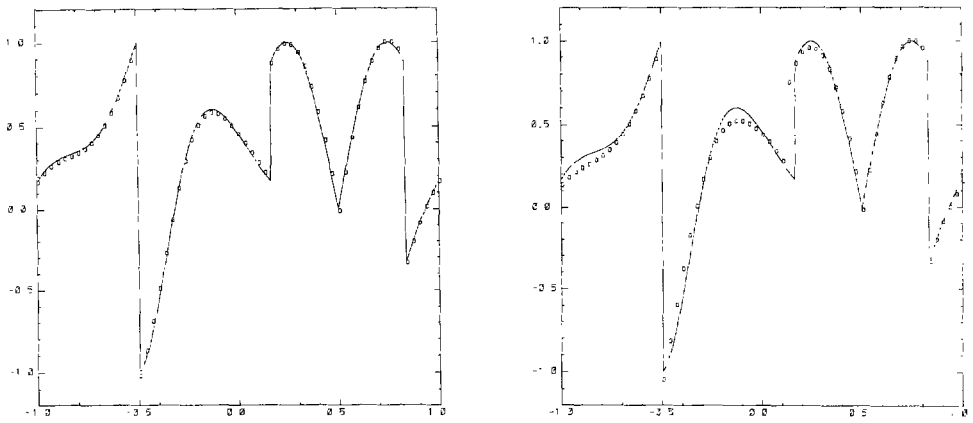


FIGURE 6.3

discontinuities. Figures 6.3 and 6.4 present the results obtained with the second-order treatment for $\Delta x = \frac{1}{30}$ and $\Delta x = \frac{1}{60}$, respectively. The figures show that the results are quite good, especially the numerical solutions for $\Delta x = \frac{1}{60}$ at both $t = 2$ and $t = 8$ are very close to the exact solution.

The numerical solution for $\Delta x = \frac{1}{60}$ at $t = 32$ and 64 are computed in order to test the long-term performance of the treatment. The results are also presented in Fig. 6.4. They show that the long-term performance of the treatment is also good. Particularly at $t = 64$, except for the second one, the discontinuities are well resolved, even though the smooth part of the numerical solution has been seriously damped.

There is an obvious deviation of the second discontinuity and "sinking" on its right. Seemingly, the "sinking" causes the deviation of the discontinuity. The problem is not understood. A possible cause of this "sinking" midht be the "wrong up-wind computation" at the right endpoint of the generated interval. An appropriate evaluation of u_j^n should mainly obtain information from the left side because of the positive speed of wave propagation. But at this endpoint, the use of the extrapolation data makes the evaluation essentially obtain information from the right side.

In the following examples of the Euler system, $\gamma = 1.4$, the second-order $x-t$ version is used to track the discontinuities, and the Lax–Wendroff scheme is chosen as the underlying scheme.

EXAMPLE 3. The Riemann problems for the Euler equations of gas dynamics (5.1) with the following two initial conditions are solved:

$$\begin{aligned} (\rho_l, q_l, p_l) &= (1., 0., 1.) \\ (\rho_r, q_r, p_r) &= (0.125, 0., 0.1) \end{aligned} \tag{6.3}$$

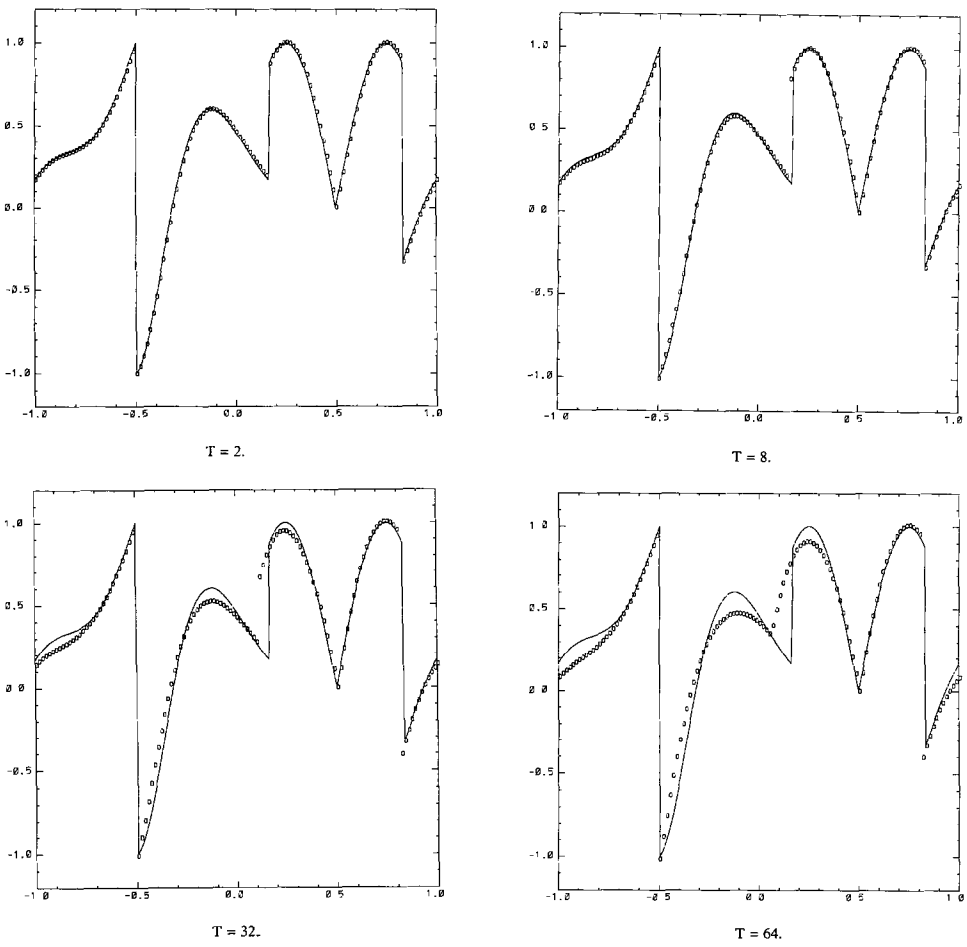


FIGURE 6.4

and

$$(\rho_l, q_l, p_l) = (0.445, 0.698, 3.528)$$
$$(\rho_r, q_r, p_r) = (0.5, 0., 0.571). \tag{6.4}$$

They are known as Sod's problem and Lax's problem, respectively. Both problems have a left rarefaction wave, a right shock, and a contact discontinuity.

The node treatment works in the first several time-steps to separate each different wave. Inspired by Example 2, in which the treatment is also applied to the weak discontinuity, we apply the treatment to the edges of the rarefaction wave as well. The naive application is unacceptable; if we did, we could get a rarefaction shock.

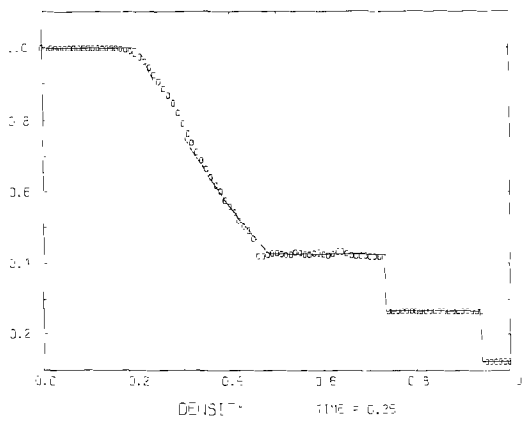
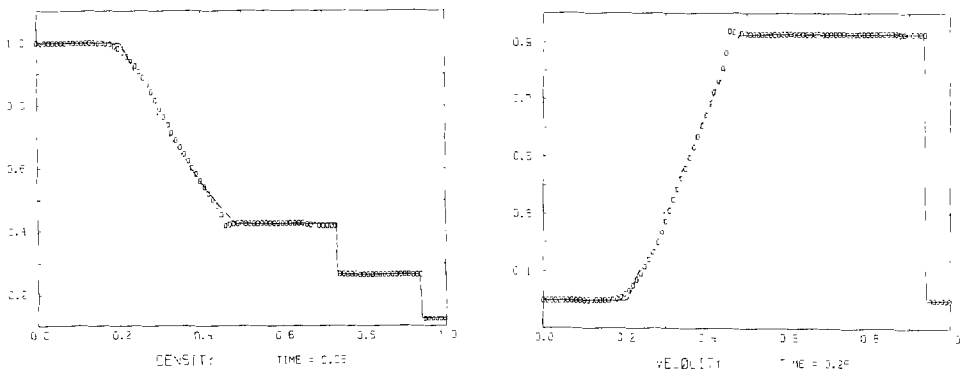


FIGURE 6.5a



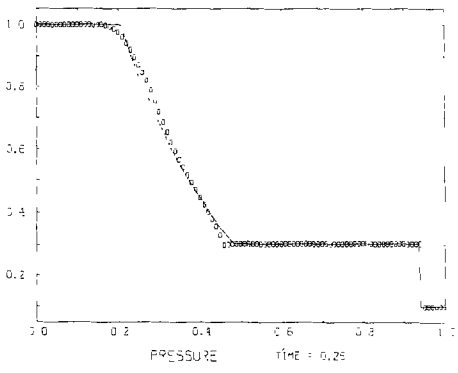


FIGURE 6.5b

Instead, suppose that the back edge of the left rarefaction is treated. Since the underlying scheme is a 3-point scheme, the extrapolated data used on each side of the generated interval are only $u_{j_1}^{n,+}$ and $u_{j_1+1}^{n,-}$. In order to expand the rarefaction wave, we replace the $u_{j_1+1}^{n,-}$ by the original datum $u_{j_1+1}^{n}$. The back edge of the right rarefaction wave can be treated in the same way. The locations of the back edges captured by the treatment are not very accurate; moreover, the analogue treatment to the front edge of the rarefaction wave fails to capture the weak discontinuities (it disappears after several time-steps). A more proper treatment of the rarefaction wave is under investigation.

For Sod's problem, $\Delta x = 0.01$ and $CFL = 0.4$. The numerical result for the density is displayed in Fig. 6.5a. Some weak oscillation appears in the smooth region. When $CFL = 0.8$, the oscillation is stronger. It is obviously caused by the underlying scheme, which is sensitive to oscillation. To reduce it, a second-order viscosity term introduced in [4] is added to the smooth part. The corresponding numerical result, which has no oscillation, is displayed in Fig. 6.5b.

For the Lax problem, $\Delta x = 0.01$ and $CFL = 0.8$. The numerical result is displayed in Fig. 6.6. No oscillation occurs in this example.

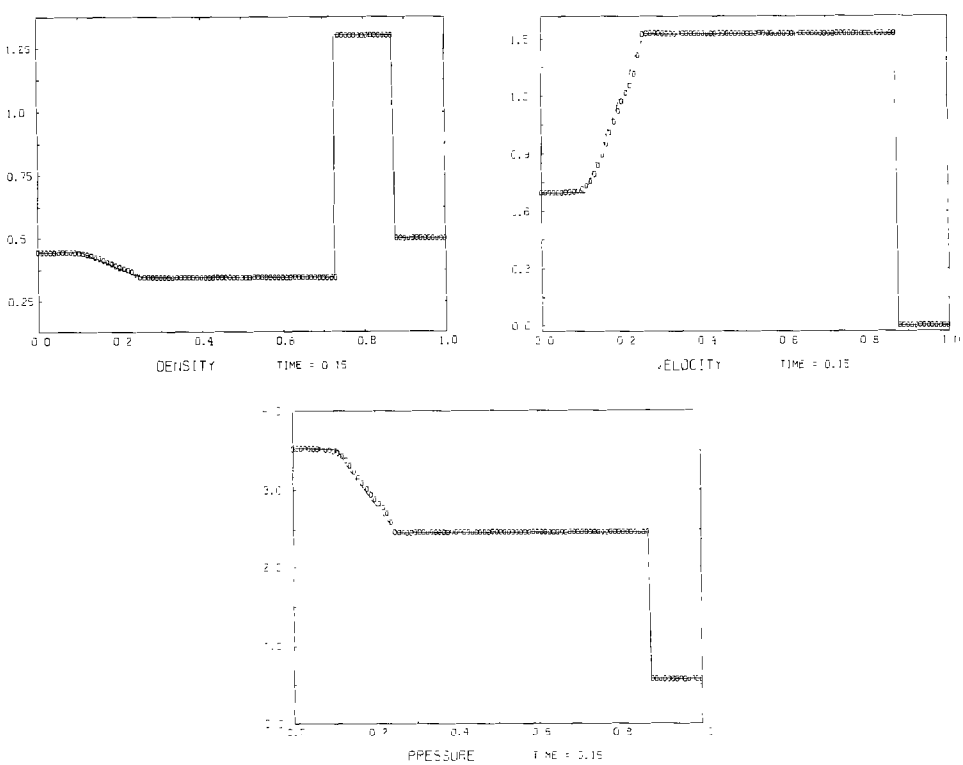


FIGURE 6.6

EXAMPLE 4 (The blast waves problem). The Euler system with the initial conditions

$$u_0 = \begin{cases} u_l, & 0. \leqslant x < 0.1 \\ u_m, & 0.1 \leqslant x < 0.9 \\ u_r, & 0.9 \leqslant x \leqslant 1. \end{cases} \tag{6.5}$$

is solved, where

$$\begin{aligned} &\rho_l = \rho_m = \rho_r = 1., \qquad q_l = q_m = q_r = 0 \\ &p_l = 10^3, \qquad p_m = 10^{-2}, \qquad p_r = 10^2, \end{aligned} \tag{6.6}$$

and the two boundaries are assumed to be solid walls. See [12] for details of the solution and comparison of the performance of various schemes.

$\Delta x = 0.005$, and the treatment is also applied to the back edges of the rarefaction waves near the waves' centers to expand them. $CFL = 0.7$ rather than 0.8; the reason is that the treatment of the interactions of discontinuities involves middle states, which do not appear in the numerical solution and may damage the CFL condition if CFL is too big. The numerical solutions at $t = 0.026$ and 0.038 are presented in Figs. 6.7a and b. The solid lines in these figures are the numerical

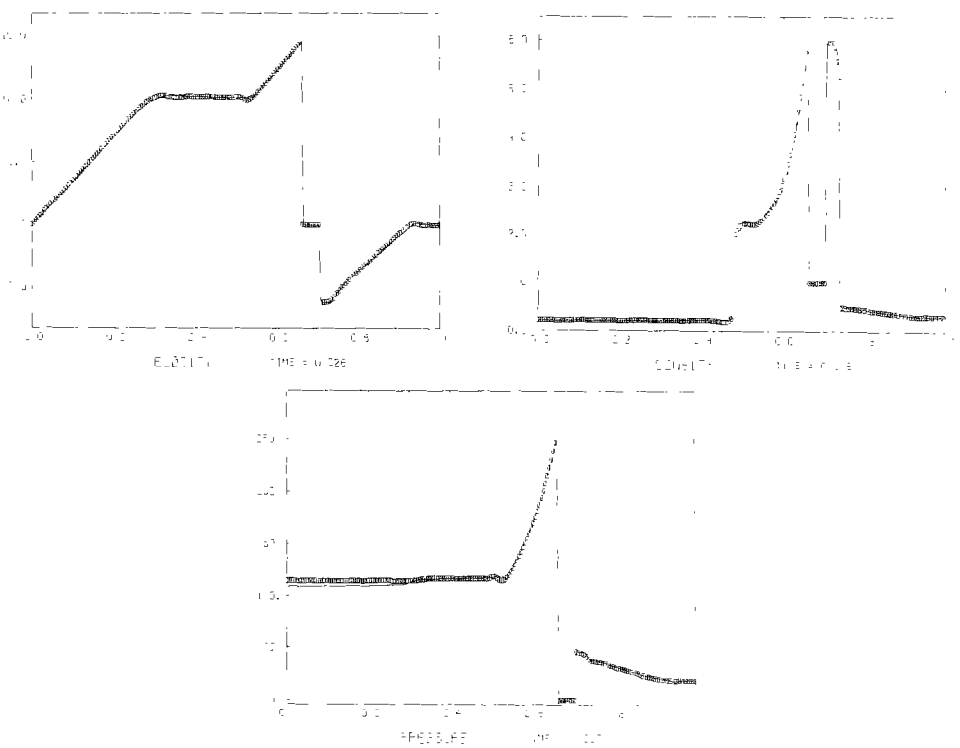


FIGURE 6.7a

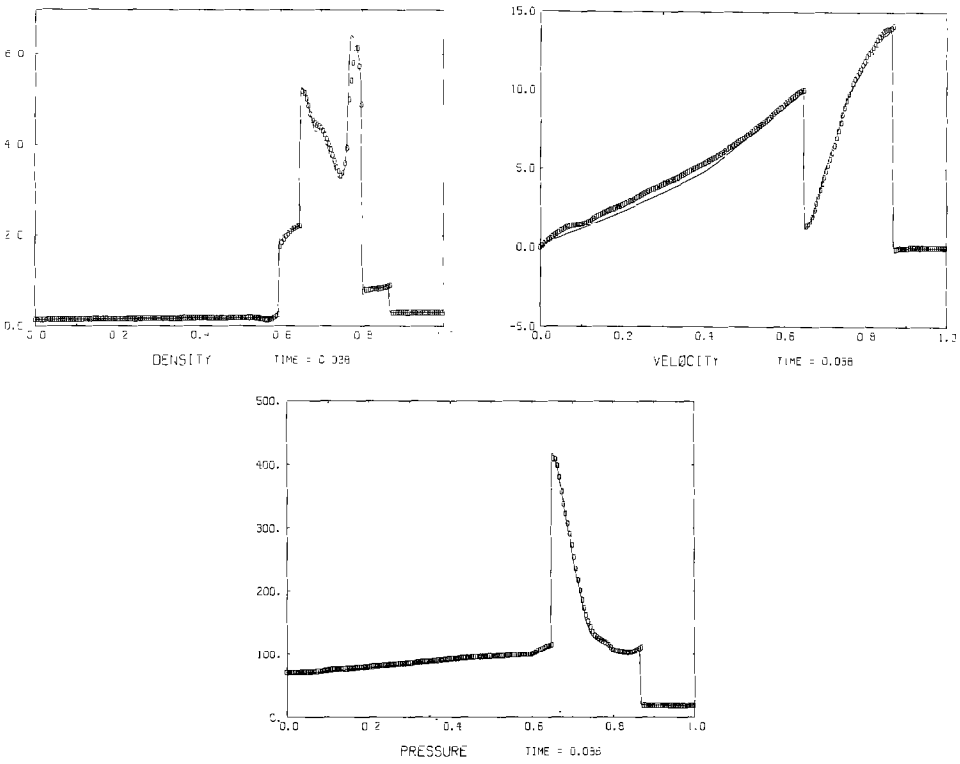


FIGURE 6.7b

solutions of a fourth-order ENO scheme with 800 space points, which is considered as a "converged solution."

At around $t = 0.032$ there isa rarefaction wave with two front edges (its right edge movesto the right), which results from the collision of a shock and a contact discontinuity. The failure of the treatment to capture the weak discontinuities causes a rarefaction shock. In order to expand the rarefaction wave, a first order viscosity

of the treatment is reduced to the first order for 20 time-steps there. This creates some error in the smooth part, especially in the velocity. The "sinking" phenomenon is still seen in the right sides of some contact discontinuities, which might be caused by the so-called "wrong up-wind computation."

ACKNOWLEDGMENTS

The author thanks Stanley Osher, Ami Harten, Chi-Wang Shu, Huanan Yang, Emad Fatemi, Rosa Donat, and an unknown referee for helpful discussion and/or for suggesting numerical examples, as well as correcting English errors in the manuscript.

REFERENCES

1. H. HARTEN, *J. Comput. Phys.* **83**, 148 (1989).
2. A. HARTEN, "Preliminary Results on the Extension of ENO Schemes to Two-Dimensional Problems," Proceedings of the International Conference on Hyperbolic Problems, Saint-Etienne, January 1986.
3. A. HARTEN AND S. OSHER, *SIAM J. Numer. Anal.* **24**, 179 (1987).
4. A. HARTEN, B. ENGQUIST, S. OSHER, AND CHAKRAVARTHY, *J. Comput. Phys.* **71**, 231 (1987).
5. A. HARTEN, *J. Comput. Phys.* **49**, 357 (1983).
6. C-W. SHU AND S. OSHER, *J. Comput. Phys.* **77**, 439 (1988).
7. C-W. SHU AND S. OSHER, *J. Comput. Phys.* **83**, 32 (1989).
8. H. YANG, *J. Comput. Phys.* **89**, 125 (1990).
9. P. K. SWBY, *SIAM J. Numer. Anal.* **21**, 995 (1984).
10. P. LAX AND B. WENDROFF, *Commun. Pure Appl. Math.* **13**, 217 (1960).
11. D.-K. MAO, *J. Comput. Math.* No. 3, 356 (1985). [Chinese]
12. P. WOODWARD AND P. COLLELA, *J. Comput. Phys.* **54**, 115 (1984).
13. I.-L. CHEN, J. GLIMM, O. MCBRYAN, B. PLOHR, AND S. YANIV, *J. Comput. Phys.* **62**, 83 (1986).
14. R. RICHTMYER AND K. MORTON, *Difference Methods for Initial Value Problems* (Interscience, New York, 1976), p. 308.
15. A. OLEINIK, *Uspehki Mat. Nauk.* **12**, 169 (1957).
16. E. TADMOR, *Math. Comput.* **43**, 369 (1984).