JOURNAL OF COMPUTATIONAL PHYSICS 103, 359–369 (1992)

A Treatment of Discontinuities for 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, People's Republic of China

Received September 19, 1990; revised July 19, 1991

In this paper we develop a new version of the treatment of discontinuities developed earlier by this author. It is a variant of the $x-t$ version in (*J. Comput. Phys.* **83**, 148 (1989)). The new version differs from the $x-t$ version in the sense that it tracks discontinuity positions instead of the local conservation errors. One of the main considerations of this paper is to show the equivalence between the two versions, by revealing the link between the discontinuity positions and the local conservation errors. In addition, a "stacking treatment" is developed to deal with discontinuities that locate in the same grid cell. Numerical experiments for applying the new version to a two-step scheme are presented, and the results are qualitatively the same as for applying the $x-t$ version to one-step schemes.

1. INTRODUCTION

The partial differential equations considered in this paper are the hyperbolic system of conservation laws

$$\begin{aligned} u_t + f(u)_x &= 0 \\ u(x, 0) &= u_0(x), \end{aligned} \tag{1.1}$$

where $u = (u_1, u_2, ..., u_m)^T$, and the Jacobian matrix of f has m real eigenvalues and a complete set of m linearly independent right-eigenvectors. A weak solution to (1.1) is a bounded measurable function $u(x, t)$ satisfying

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

for all smooth test functions.

This paper continues to discuss the treatment of discontinuities developed in [7, 8] for finite difference methods for (1.1). As is described in [7], the treatment in some respects is similar to Harten's subcell resolution (see [6]). However, it is a shock tracking technique, whose main idea is this: in the field of a discontinuity the computation only uses information from the same side of the discontinuity. In the scalar case, this is done by doing the following: at the grid points on the other side of the discontinuity, we replace the original data by the extrapolated data from the previous side for the numerical flux. In the system case, Riemann problems related to the original and extrapolated data are solved to produce the data that replace the original data. By doing this, the whole computation still proceeds on the regular grid and there are no states on the discontinuities and their corresponding computation. The treatment applies to any difference schemes, and the corresponding algorithm is much simpler than the traditional shock-tracking methods.

The method that will be developed in this paper is a variant of the $x-t$ version in [7]. As described in [7], the $x-t$ version of the treatment is not conservative, since in some cells different flux terms are used. The differences of the numerical flux are accumulated and recorded and compensate the numerical solution in a certain way. In [7], we call the accumulated flux differences artificial terms along the t direction; however, in this paper we call them local conservation errors, since the present terminology seems to be more proper. The grid cells that contain discontinuities are called critical cells. The movement of critical cells in the $x-t$ version is based on the magnitude of the conservation errors; in other words, the $x-t$ version tracks the conservation errors.

The movement of critical cells in the version developed in this paper is based on the positions of the discontinuities; in other words, this version tracks discontinuity positions rather than conservation errors, just as was done in traditional front-tracking methods (see [1, 5, 9, 12]). One of the main considerations of this paper is to show the equivalence between this version and the $x-t$ version when the solution is piecewise smooth, by revealing the link between the discontinuity positions and the local conservation errors, that is, Proposition 3.2 and Proposition 4.1.

The author would like to point out that the method developed in [2] has the same basic idea as our treatment. For an isolated discontinuity in the one-dimensional case, it

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

0021-9991/92 $5.00

also extrapolates the solution from each side of the computation to evade the discontinuity. However, its extensions to the system case are different from ours. Moreover, we have developed many treatments for interactions of discontinuities; this makes our treatment far more effective than that in [2].

We also point out that the idea of recording local conservation errors and compensating the numerical solution by them also occurred in [4]. Actually, the terminology "local conservation error" was first introduced in that paper. However, the method there is essentially a shock-capturing method and there is no discussion concerning the link between the conservation errors and the discontinuity positions.

The paper is organized in the following way: Section 2 describes the treatment of isolated discontinuities. Section 3 shows the equivalence between the present version and the $x-t$ version, i.e., the conservation feature of the treatment. Section 4 describes the treatment of discontinuity interactions and also shows its conservation feature. A so-called "stacking treatment" is developed in this section. Section 5 extends the treatment to Euler equations of gas dynamics. Section 6 presents the numerical results obtained by applying the treatment to a two-step scheme.

2. TREATMENT OF ISOLATED DISCONTINUITIES

Since the version that will be developed in this paper is similar to the $x-t$ version in [7], we only describe it briefly and emphasize the differences between the two versions. As usual, the discussion starts with the simplest case, i.e., the one-dimensional scalar problem; the corresponding equation is

$$u_t + f(u)_x = 0, \tag{2.1}$$

where both u and f are scalar. The underlying difference scheme is a general conservative scheme,

$$u_j^{n+1} = u_j^n - \lambda(\hat{f}_{j+1/2}^n - \hat{f}_{j-1/2}^n), \tag{2.2}$$

where u_j^n denotes the numerical solution at the grid point (x_j, t^n),

$$\hat{f}_{j+1/2}^n = \hat{f}(u_{j-k+1}^n, ..., u_{j+k}^n) \tag{2.3}$$

is the numerical flux depending on $2k$ variables; $\lambda = \tau/h$ is the mesh ratio, where τ and h are the time and space increments, respectively. The numerical flux is consistent with the flux in (2.1) in the sense that

$$\hat{f}(u, u, ..., u) = f(u). \tag{2.4}$$

The cells that contain discontinuities are called *critical cells*, on which the treatment is applied. (We change the terminology for these cells, considering that the present one is more proper than the old one used in [7], which is *generated interval*).

In general, the treatment is performed in the following four steps for an isolated discontinuity:

(1) Extrapolate the numerical solution from each side of the critical cell to the other side, and obtain a set of extrapolated data:

$$u_{j_1-k}^{n,+}, ..., u_{j_1}^{n,+}, u_{j_1+1}^{n,-}, ..., u_{j_1+k+1}^{n,-}, \tag{2.5}$$

where $[x_{j_1}, x_{j_1+1}]$ is the critical cell, and the data with "$-$" are from the left to the right and the data with "$+$" are from the right to the left (as shown in Fig. 2.1).

(2) Compute u^{n+1}, the numerical solution on the following level: When x_j is on the left side of the critical cell, i.e., $j \leqslant j_1$, we compute u_j^{n+1} by the flux with "$-$"; or,

$$u_j^{n+1} = u_j^n - \lambda(\hat{f}_{j+1/2}^{n,-} - \hat{f}_{j-1/2}^{n,-}), \tag{2.6}$$

where

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

When x_j is on the right side of the critical cell, i.e., $j \geqslant j_1 + 1$, we compute u_j^{n+1} by the flux with "$+$," defined as

$$\hat{f}_{j+1/2}^{n,+} = \hat{f}(u_{j-k+1}^{n,+}, ..., u_{j_1}^{n,+}, u_{j_1+1}^n, ..., u_{j+k}^n). \tag{2.8}$$

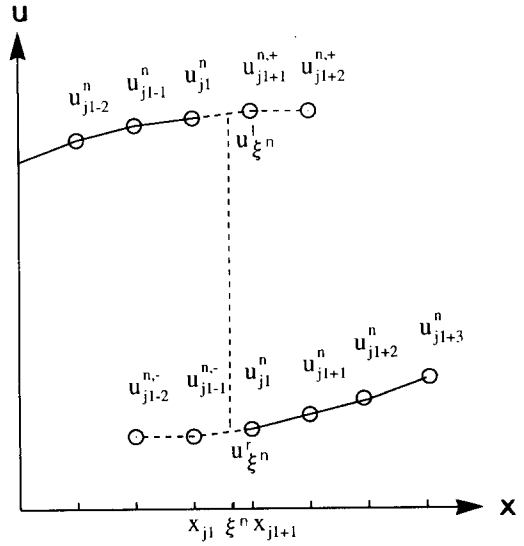

FIG. 2.1. The numerical solution on the level n has a jump in the cell $[x_{j_1}, x_{j_1+1}]$; circles denote the original data of the numerical solution and "*"s denote the extrapolated data from the two sides of the jump.

(3) Compute ξ^{n+1}, the discontinuity position on the following level. First, we extrapolate u^n from each side to ξ^n and obtain two extrapolated data $u^l_{\xi^n}, u^r_{\xi^n}$. Then, we compute the speed of the discontinuity by the Hugoniot condition

$$s=\frac{f(u^r_{\xi^n})-f(u^l_{\xi^n})}{u^r_{\xi^n}-u^l_{\xi^n}}. \tag{2.9}$$

Finally, we compute ξ^{n+1} by

$$\xi^{n+1}=\xi^n+s\tau. \tag{2.10}$$

(4) Set up the critical cell on the following level. If ξ^{n+1} remains in $[x_{j_1}, x_{j_1+1}]$, the same cell will be the critical cell on the new level; otherwise, ξ^{n+1} should move into one of the adjacent cells, in which case this adjacent cell to becomes the critical dell, and the numerical solution at the grid point crossed over by the discontinuity should be updated either by the datum computed from the side the point now belongs to or by the extrapolated datum also from this side.

We make the following remarks:

(a) The propagation of the discontinuity in this version is different from that in the $x-t$ version; how the discontinuity is moved depends on its position.

(b) In step (3), the discontinuity position is computed only with first-order accuracy. Higher order accuracy can be obtained either by using information on the new level to compute s in (2.9) or by a Runge–Kutta procedure. The algorithm we used to compute the numerical examples in Section 6 is based on a second-order Runge–Kutta procedure, which will be described later.

(c) The speed of the discontinuity computed in (2.9) satisfies

$$|s|\leqslant \max|f'(u)|.$$

Thus, the discontinuity may only move to the left or right adjacent cell, if the mesh ratio λ satisfies the CFL-condition; it will not move far away from the original cell.

3. CONSERVATION FEATURE OF THE TREATMENT OF ISOLATED DISCONTINUITIES

In order to discuss the conservation feature, we write the overall algorithm in a conservation-like form,

$$u_j^{n+1}=u_j^n-\lambda(\tilde f^n_{j+1/2}-\tilde f^n_{j-1/2}) + p^n_{j+1/2}-p^n_{j-1/2}+q_j^{n+1}-q_j^n, \tag{3.1}$$

where

$$\tilde f^n_{j+1/2}=\begin{cases}\hat f^{n,-}_{j+1/2}, & j\leqslant j_1,\\ \hat f^{n,+}_{j+1/2}, & j\geqslant j_1+1.\end{cases} \tag{3.2}$$

The q^n's are the local conservation errors, and they balance the different numerical flux terms used in some cells. As described in [7], p^n and q^n are nonzero only in the vicinity of critical cells. For an isolated discontinuity, if the numerical solution at the grid point crossed over by the discontinuity is updated by the datum computed from the same side of the critical cell, then the nonzero p's and q's are as follows: If ξ^{n+1} remains in the original critical cell,

$$q^{n+1}_{j_1}=q^n_{j_1}+\lambda(\hat f^{n,+}_{j_1+1/2}-\hat f^{n,-}_{j_1+1/2}); \tag{3.3}$$

if ξ^{n+1} moves to the left adjacent cell,

$$\begin{aligned} p^n_{j_1-1/2}&=-q^n_{j_1}+(u^n_{j_1}-u^{n,+}_{j_1})+\lambda(\hat f^{n,-}_{j_1-1/2}-\hat f^{n,+}_{j_1-1/2}),\\ q^{n+1}_{j_1-1}&=-p^n_{j_1-1/2};\end{aligned} \tag{3.4}$$

and, if ξ^{n+1} moves to the right adjacent cell,

$$\begin{aligned} p^n_{j_1+1/2}&=q^n_{j_1}+\lambda(\hat f^{n,+}_{j_1+1/2}-\hat f^{n,-}_{j_1+1/2}),\\ q^{n+1}_{j_1+1}&=q^n_{j_1}+(u^{n,-}_{j_1+1}-u^n_{j_1+1})+\lambda(\hat f^{n,+}_{j_1+3/2}-\hat f^{n,-}_{j_1+3/2}),\end{aligned} \tag{3.5}$$

where $[x_{j_1}, x_{j_1+1}]$ is the critical cell. If the numerical solution is updated by the extrapolated data from the same side, (3.3)–(3.5) are accurate up to $O(h^r)$, where r is the accuracy order of the underlying scheme and the treatment. We recommend readers see [7] for the derivation of (3.3)–(3.5).

PROPOSITION 3.1. *Assume that the solution to* (1.1) *is piecewise smooth. Then for an isolated discontinuity, if both the underlying scheme* (2.2) *and the treatment are of the first order,*

$$q^n_{j_1}=\frac{(\xi^n-x_{j_1+1/2})(u^n_{j_1+1}-u^n_{j_1})}{h}+O(1); \tag{3.6}$$

if both the underlying scheme and the treatment are of the second order, then

$$q^n_{j_1}=\frac{\begin{pmatrix}(u^n_{j_1}-u^{n,+}_{j_1})(x_{j_1+1}-\xi^n)^2\\ +(u^n_{j_1+1}-u^{n,-}_{j_1+1})(\xi^n-x_{j_1})^2\end{pmatrix}}{2h^2}+O(h), \tag{3.7}$$

where $[x_{j_1}, x_{j_1+1}]$ *is the critical cell on the level* n, *and* $x_{j_1+1/2}=\frac{1}{2}(x_{j_1}+x_{j_1+1})$.

Here we call the treatment of the rth order if both the extrapolation and the algorithm to compute the discontinuity positions are of rth order.

The proposition indicates that if the numerical solution is uniformly bounded, p and q are also uniformly bounded. Since the principal feature of the $x-t$ version is that q stays uniformly bounded (see [7]), this shows the equivalence between this version and the $x-t$ version when the solution to (1.1) is piecewise smooth. Also, we have from (3.1)

$$\sum_j u_j^n h = \sum_j u_j^0 h + O(h), \tag{3.8}$$

which implies that the numerical solution is almost conserved. By the same argument in [7] we can obtain the convergence of the method under certain conditions. This is why we call the proposition the conservation feature of the treatment.

A geometric interpretation for (3.6) is as follows: The common interpretation of the numerical solution of a conservative scheme is that it is a piecewise constant function with each constant u_j^n covering an interval $[x_{j-1/2}, x_{j+1/2}]$. If we consider that at the two endpoints of the critical cell the constants $u_{j_1}^n$ and $u_{j_1+1}^n$ cover the intervals $[x_{j_1-1/2}, \xi^n]$ and $[\xi^n, x_{j_1+3/2}]q$, respectively, by taking the discontinuity into account, (3.6) indicates that $q_{j_1}^n h$ is equal to the difference resulting from the different interpretations of the solution up to $O(h)$, as shown in Fig. 3.1. Equation (3.7) also can be interpreted in this way if we see the numerical solution as a continuous piecewise linear function and take the discontinuity into account.

When both the underlying scheme and the treatment are of second order, (3.7) differs from (3.6) only by $O(h)$, and this can be shown by writing (3.7) as

$$\begin{aligned} q_{j_1}^n = {} & \frac{1}{h}(\xi^n - x_{j_1+1/2})(u_{j_1+1}^n - u_{j_1}^n) \\ & + \frac{1}{2h^2}\{(u_{j_1+1}^n - u_{j_1}^{n,+})(x_{j_1+1} - \xi^n)^2 \\ & + (u_{j_1}^n - u_{j_1+1}^{n,-})(\xi^n - x_{j_1}^n)^2\} + O(h). \end{aligned} \tag{3.9}$$

Proof of (3.6). Denote

$$S^n = (\xi^n - x_{j_1+1/2})(u_{j_1+1}^n - u_{j_1}^n) - hq_{j_1}^n, \tag{3.10}$$

and we will prove

$$S^{n+1} - S^n = O(h^2), \tag{3.11}$$

from which the conclusion of the proposition follows easily. The proof should be completed for all three cases: ξ^{n+1} either remains in the original critical cell or moves into one of its adjacent cells. We will only present the proof for the first case, since the proof for the other two cases is similar.

By (3.3), we have

$$\begin{aligned} S^{n+1} - S^n = {} & d(u_{j_1+1}^{n+1/2} - u_{j_1}^{n+1/2}) \\ & - \tau\left\{\hat{f}_{j_1+1/2}^{n,+} + (\xi^{n+1/2} - x_{j_1+1/2}) \right. \\ & \times \frac{\hat{f}_{j_1+3/2}^{n,+} - \hat{f}_{j_1+1/2}^{n,+}}{h} \\ & - \hat{f}_{j_1+1/2}^{n,-} - (\xi^{n+1/2} - x_{j_1+1/2}) \\ & \left. \times \frac{\hat{f}_{j_1+1/2}^{n,-} - \hat{f}_{j_1-1/2}^{n,-}}{h}\right\}, \end{aligned} \tag{3.12}$$

where $d = \xi^{n+1} - \xi^n$, $\xi^{n+1/2} = \frac{1}{2}(\xi^{n+1} + \xi^n)$, and $u_j^{n+1/2} = \frac{1}{2}(u_j^{n+1} + u_j^n)$. The right side of (3.12) is an approximation to the left side of the Hugoniot condition

$$dx - dt[f]/[u] = 0 \tag{3.13}$$

at the point $(\xi^{n+1/2}, t^{n+1/2})$. Since both the underlying scheme and the treatment are of the first order, the accuracy of both the numerical solution u^n and the discontinuity position ξ^n are of the first order. Thus, (3.11) follows. This completes the proof.

Proof of (3.7). Essentially, (3.7) can be proved in the same way as (3.6); however, it is more complicated, since we

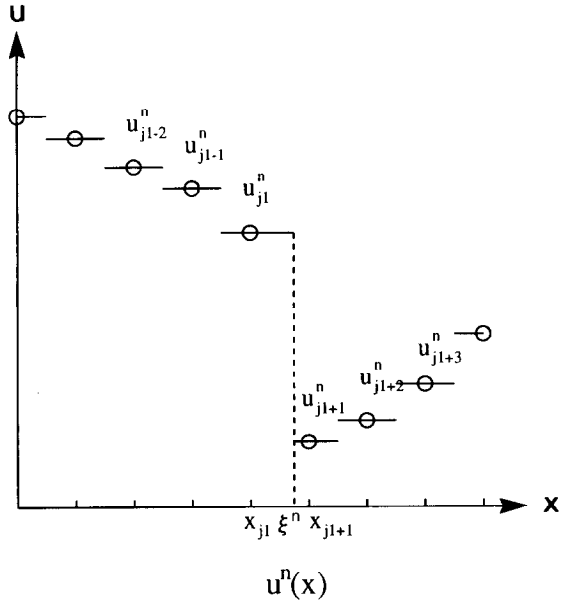

FIG. 3.1. Piecewise constant function $U(x)$ defined by (3.10) and (3.11).

need to show the second-order accuracy. (3.7) can be written as

$$(u^{n,+}_{j_1+1/2}-u^{n,-}_{j_1+1/2})(\xi^n-x_{j_1+1/2})-q^n_{j_1}h$$
$$-\frac{1}{4h}(u^{n,-}_{j_1+1}-u^n_{j_1})\{(x_{j_1+1}-\xi^n)^2-(\xi^n-x_{j_1})^2\}$$
$$+\frac{1}{4h}(u^n_{j_1+1}-u^{n,+}_{j_1})\{(x_{j_1+1}-\xi^n)^2+(\xi^n-x_{j_1})^2\}$$
$$=O(h^2), \tag{3.14}$$

where $u^{n,-}_{j_1+1/2}=\frac{1}{2}(u^n_{j_1}+u^{n,-}_{j_1+1})$ and $u^{n,+}_{j_1+1/2}=\frac{1}{2}(u^{n,+}_{j_1}+u^n_{j_1+1})$. Denote by Q^n the last two terms on the left in (3.14); we have

$$Q^n=-\frac{1}{2h}(u^{n,-}_{j_1+1}-u^n_{j_1}-u^n_{j_1+1}+u^{n,+}_{j_1})$$
$$\left((x_{j_1+1/2}-\xi^n)^2+\frac{h^2}{4}\right). \tag{3.15}$$

Denote by S^n the left side of (3.14). At this time we will show that (3.11) is accurate up to $O(h^3)$ (rather than $O(h^2)$). Also we only present the proof for the case where ξ^{n+1} remains in the original critical cell. In this case we have

$$S^{n+1}-S^n=d(u^{n+1/2,+}_{j_1+1/2}-u^{n+1/2,-}_{j_1+1/2})$$
$$+(\xi^{n+1/2}-x_{j_1+1/2})$$
$$\times(u^{n+1,+}_{j_1+1/2}-u^{n+1,-}_{j_1+1/2}-u^{n,+}_{j_1+1/2}+u^{n,-}_{j_1+1/2})$$
$$-h(q^{n+1}_{j_1}-q^n_{j_1})+(Q^{n+1}-Q^n), \tag{3.16}$$

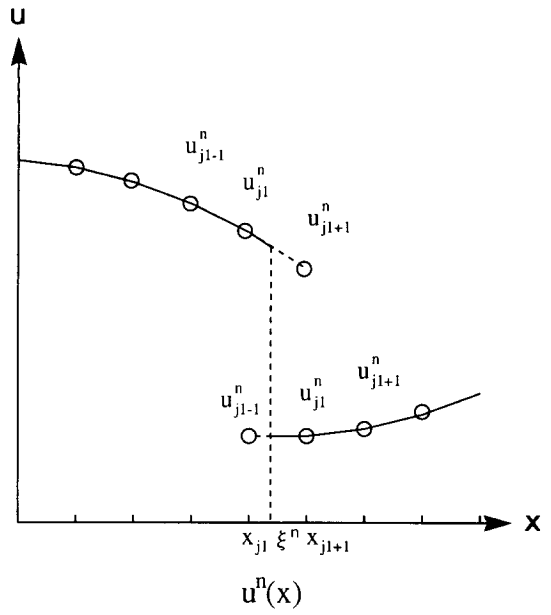

FIG. 3.2. Piecewise linear function $U(x)$ defined by (3.12) and (3.13).

where $u^{n+1/2,+}_{j_1+1/2}=\frac{1}{2}(u^{n+1,+}_{j_1+1/2}+u^{n,+}_{j_1+1/2})$, $u^{n+1/2,-}_{j_1+1/2}=\frac{1}{2}(u^{n+1,-}_{j_1+1/2}+u^{n,-}_{j_1+1/2})$, and d and $\xi^{n+1/2}$ are defined as before. Since the underlying scheme and the treatment are of the second order,

$$u^{n+1,-}_{j_1+1}=u^{n,-}_{j_1+1}-\lambda(\hat f^{n,-}_{j_1+3/2}-\hat f^{n,-}_{j_1+1/2})+O(h^3) \tag{3.17}$$

and

$$u^{n+1,+}_{j_1}=u^{n,+}_{j_1}-\lambda(\hat f^{n,+}_{j_1+1/2}-\hat f^{n,+}_{j_1-1/2})+O(h^3). \tag{3.18}$$

Therefore,

$$u^{n+1,-}_{j_1+1/2}-u^{n,-}_{j_1+1/2}=-\frac{\lambda}{2}(\hat f^{n,-}_{j_1+3/2}-\hat f^{n,-}_{j_1-1/2})+O(h^3) \tag{3.19}$$

and

$$u^{n+1,+}_{j_1+1/2}-u^{n,+}_{j_1+1/2}=-\frac{\lambda}{2}(\hat f^{n,+}_{j_1+3/2}-\hat f^{n,+}_{j_1-1/2})+O(h^3). \tag{3.20}$$

Moreover, we have

$$(x_{j_1+1/2}-\xi^{n+1})^2=\frac{d^2}{4}+d(\xi^{n+1/2}-x_{j_1+1/2})$$
$$+(\xi^{n+1/2}-x_{j_1+1/2})^2 \tag{3.21}$$

and

$$(x_{j_1+1/2}-\xi^n)^2=\frac{d^2}{4}-d(\xi^{n+1/2}-x_{j_1+1/2})$$
$$+(\xi^{n+1/2}-x_{j_1+1/2})^2. \tag{3.22}$$

Therefore,

$$Q^{n+1}-Q^n=d\left(\frac{u^{n+1/2}_{j_1+1}-u^{n+1/2,+}_{j_1}}{h}-\frac{u^{n+1/2,-}_{j_1+1}-u^{n+1/2}_{j_1}}{h}\right)$$
$$\times(\xi^{n+1/2}-x_{j_1+1/2})+R^{n+1/2}, \tag{3.23}$$

where

$$R^{n+1/2}=\{(u^{n+1,-}_{j_1+1}-u^{n+1}_{j_1})-(u^{n,-}_{j_1+1}-u^n_{j_1})$$
$$-(u^{n+1}_{j_1+1}-u^{n+1,+}_{j_1+1})+(u^n_{j_1+1}-u^{n,+}_{j_1})\}$$
$$\times\frac{1}{2h}\left\{\frac{d^2}{4}+\frac{h^2}{4}+(\xi^{n+1/2}-x_{j_1+1/2})^2\right\}. \tag{3.24}$$

Obviously, $R^{n+1/2} = O(h^3)$. Substitute (3.17)–(3.24) into (3.16); we have

$$\begin{aligned} S^{n+1} - S^n = d \Bigg\{ & u^{n+1/2,+}_{j_1+1/2} + \frac{u^{n+1/2}_{j_1+1} - u^{n+1/2,+}_{j_1}}{h} \\ & \times (\xi^{n+1/2} - x_{j_1+1/2}) \\ & - u^{n+1/2,-}_{j_1+1/2} - \frac{u^{n+1/2,-}_{j_1+1} - u^{n+1/2}_{j_1}}{h} \\ & \times (\xi^{n+1/2} - x_{j_1+1/2}) \Bigg\} \\ & - \tau \Bigg\{ \hat{f}^{n,+}_{j_1+1/2} + \frac{\hat{f}^{n,+}_{j_1+3/2} - \hat{f}^{n,+}_{j-1/2}}{2h} \\ & \times (\xi^{n+1/2} - x_{j_1+1/2}) \\ & - \hat{f}^{n,-}_{j_1+1/2} - \frac{\hat{f}^{n,-}_{j_1+3/2} - \hat{f}^{n,-}_{j_1-1/2}}{2h} \\ & \times (\xi^{n+1/2} - x_{j_1+1/2}) \Bigg\} + O(h^3). \end{aligned} \tag{3.25}$$

The right side of (3.25) is again an approximation to the left side of the Hugoniot condition (3.13) at the point $(\xi^{n+1/2}, t^{n+1/2})$. Since both the underlying scheme and the treatment are of second order, the accuracy for both the numerical solution and the discontinuity position is second order. With $\hat{f}^n_{j+1/2} = f|_{x=x_{j+1/2},\, t=t^{n+1/2}} + O(h^2)$, the conclusion follows.

4. TREATMENT OF INTERACTIONS OF DISCONTINUITIES

When two critical cells are close to each other, they are treated in the same way as in [7]; hence, our considerations focus only on the case when they have moved to the same cell and when they merge. A so-called "stacking treatment" will be developed (Fig. 4.1).

In the $x-t$ version we merge the two critical cells that have moved to the same cell and add the two local conservation errors together to give the local conservation error for the newly-formed discontinuity. However, the two discontinuity positions usually have not crossed over each other at the moment when we merge the critical cells; that is, the merging has taken place a little too early.

In this paper we will stack the critical cells in the same cell, but with a middle state between them when their discontinuity positions have not crossed over each other. The middle state can always be obtained in a natural way, for example, in the cases (a) and (b) in Fig. 4.2. The non-updated numerical solution at the grid point (x_{j_2}, t^{n+1}) before the step 4 can be chosen as the middle state for the

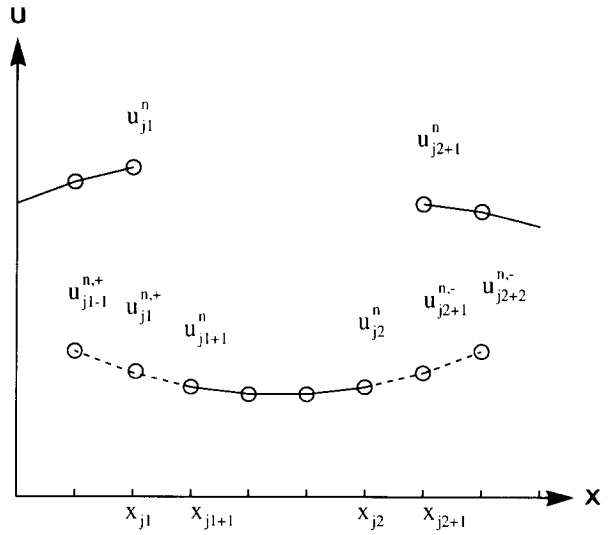

FIG. 4.1. Two critical cells on level n separate the numerical solution into three smooth parts; circles denote the original data of the numerical solution and "*"s denote the extrapolated from the three parts.

two stacked critical cells on level $n+1$. The algorithm we used to compute the numerical examples in Section 6 does not let case (c) happen by the handling described in [7], which holds one of the critical cells. The middle state stands as the right state for the left critical cell and the left state stands for the right critical cell.

This stacking treatment can stack more than two critical cells in the same cell with several middle states. In doing so, we actually store more information than usual in the cell. Such a treatment is efficient in dealing with the small region structure if we do not care too much about the structure details but only the discontinuity positions.

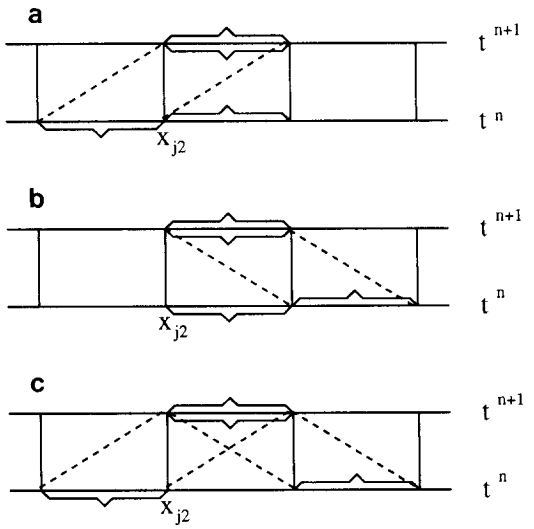

FIG. 4.2. (a) Critical cell denoted by upward brace moves one cell to the left and critical cell denoted by downward brace remains in the same cell. (b) Critical cell denoted by upward brace remains in the same cell and critical cell denoted by downward brace moves one cell to the right. (c) Two critical cells move across each other.

The two stacked critical cells coexist until the two discontinuity positions cross over each other; then we merge them to form a new critical cell. In the present version the discontinuity position for the new-formed one is computed as

$$\xi^n = \frac{\xi_1^n(u_* - u_{j_1}^n) + \xi_2^n(u_{j_1+1}^n - u_*)}{u_{j_1+1}^n - u_{j_1}^n}, \tag{4.1}$$

where $[x_{j_1}, x_{j_1+1}]$ is the critical cell, ξ_1 and ξ_2 are the two discontinuity positions before the merging, and u_* is the middle state in between.

PROPOSITION 4.1. *If the solution to* (1.1) *is piecewise smooth,* (4.1) *has second-order accuracy.*

Proof. First we prove that (4.1) is exact when the solution is piecewise constant. In this case, the discontinuities starting from ξ_1^{n-1} and ξ_2^{n-1} on the level $n-1$ are both straight lines with speeds s_1 and s_2, respectively. They meet at a point (ξ^*, t^*) and then merge to form a new discontinuity, with speed s_0, which reaches the level n at ξ^n (as shown in Fig. 4.3). Obviously,

$$\xi^n = \zeta^* + (t^n - t^*)\, s_0, \tag{4.2a}$$

$$\xi_1^n = \xi^* + (t^n - t^*)\, s_1, \tag{4.2b}$$

$$\xi_2^n = \xi^* + (t^n - t^*)\, s_2, \tag{4.2c}$$

where

$$s_0 = \frac{f(u_r) - f(u_l)}{u_r - u_l}, \tag{4.3a}$$

$$s_1 = \frac{f(u_*) - f(u_l)}{u_* - u_l}, \tag{4.3b}$$

$$s_2 = \frac{f(u_r) - f(u_*)}{u_r - u_*}; \tag{4.3c}$$

u_l, u_r, and u_* are the left, right, and middle states of the numerical solution, respectively. It is easy to obtain

$$s_0 = \frac{(u_* - u_l)\, s_1 + (u_r - u_*)\, s_2}{u_r - u_l} \tag{4.4}$$

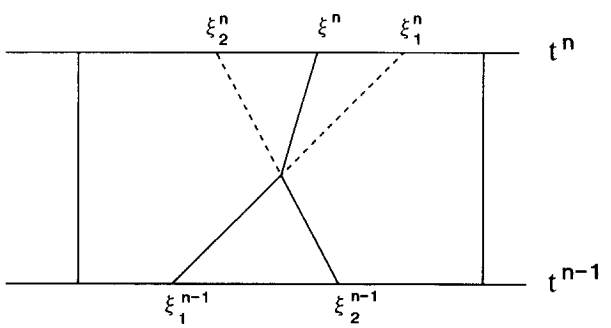

FIG. 4.3. Two discontinuities start from ξ_1^{n-1} and ξ_2^{n-1}, meet at point (ξ^*, t^*), and form a new discontinuity.

from (4.3); thus, by (4.2),

$$\xi^n = \frac{\xi_1^n(u_* - u_l) + \xi_2^n(u_r - u_*)}{u_r - u_l}, \tag{4.5}$$

which indicates that (4.1) is exact.

If the numerical solution is piecewise smooth, (4.3) is correct up to $O(h)$; therefore, (4.2) is correct up to $O(h^2)$. Thus the conclusion follows.

We add the two local conservation errors together and take the sum to be the local conservation error for the new critical cell.

PROPOSITION 4.2. *If the solution to* (1.1) *is piecewise smooth, then* (3.6) *and* (3.7) *are still true for the new critical cell under the conditions of Proposition* 3.1.

Proposition 4.2 is the extension of Proposition 3.1 to the interaction case. It shows the conservation feature of this merging treatment and the equivalence between it and that of the $x-t$ version.

Proof. When both the underlying scheme and the treatment are of the first order, we have from Proposition 3.1

$$q_{j_1,1}^n = \frac{(\xi_1^n - x_{j_1+1/2})(u_* - u_{j_1}^n)}{h} + O(1) \tag{4.6}$$

and

$$q_{j_1,2}^n = \frac{(\xi_2^n - x_{j_1+1/2})(u_{j_1+1}^n - u_*)}{h} + O(1), \tag{4.7}$$

where, $q_{j_1,1}^n$ and $q_{j_1,2}^n$ are the local conservation errors of the two stacked critical cells. Equation (3.6) is obtained for the newly-formed critical cell by adding (4.6) and (4.7) together and using (4.1).

When both the underlying scheme and the treatment are of the second order, by (3.9), (3.7) differs from (3.6) only by $O(h)$; hence (3.7) follows easily.

5. EXTENSION OF THE TREATMENT TO EULER EQUATIONS OF GAS DYNAMICS

The Euler equations of gas dynamics for a polytropic gas are

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

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

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

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

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

$$a_1(u) = q - u, \quad a_2(u) = q, \quad a_3(u) = q + u, \qquad (5.2)$$

where $c = (\gamma p/\rho)^{1/2}$ is the sound speed.

As in [7], there are three types of critical cells, the left shock critical cells (LSCC), right shock critical cells (RSCC), and contact discontinuity critical cells (CDCC), corresponding to the left shocks, right shocks, and contact discontinuities, respectively. For a single discontinuity, the treatment is still carried out in the four steps outlined above, i.e., (1) compute the data that will replace the data on the other side of the discontinuity for the numerical flux; (2) compute the numerical solution on each side of the discontinuity; (3) compute the new position of the discontinuity; (4) set up the critical cell on the new level and update the numerical solution at the grid point crossed over by the discontinuity (if there is such a grid point).

As was done in [7], the data in step (1) are computed by solving Riemann problems related to the original and the extrapolated data and picking the proper states in the Riemann separations. In doing so, the treatment acts only on the corresponding field of the discontinuity.

The speed in step (3) is also computed by solving a Riemann problem related to the extrapolated data at the discontinuity position. The treatment can also be analyzed in terms of p^n and q^n as was done in Section 3 to show the conservative feature; the conclusion of Proposition 3.1 is still true, only p^n and q^n are vectors and all the expressions in the proof should be understood in vector forms.

Most of the treatment of discontinuity interactions in the scalar case can be formally transformed to the system case. The only complication is that there are three waves resulting from the collision of two discontinuities. Assume that the two critical cells with discontinuity positions ξ_1^n and ξ_2^n collide in the cell $[x_{j_1}, x_{j_1+1}]$ at level n, we treat the collision as follows:

(1) Solve the Riemann problem $RP(u_{j_1}^n, u_{j_1+1}^n)$ and obtain two middle states u_*^l and u_*^r. The wave that connects $u_{j_1}^n$ and u_*^l or u_*^r and $u_{j_1+1}^n$ can be either a shock or rarefaction wave; nevertheless, the wave that connects u_*^l and u_*^r is always a contact discontinuity.

(2) Denote by Q^n the quantity $(u_* - u_{j_1}^n)\,\xi_1^n + (u_{j_1+1}^n - u_*)\,\xi_2^n$. Represent Q^n as a linear combination of the vectors $u_*^l - u_{j_1}^n$, $u_*^r - u_*^l$, and $u_{j_1+1}^n - u_*^r$; that is,

$$Q^n = \bar{\xi}_1^n(u_*^l - u_{j_1}^n) + \bar{\xi}_2^n(u_*^r - u_*^l) + \bar{\xi}_3^n(u_{j_1+1}^n - u_*^r) \qquad (5.3)$$

The representation is unique if the three vectors are linearly independent.

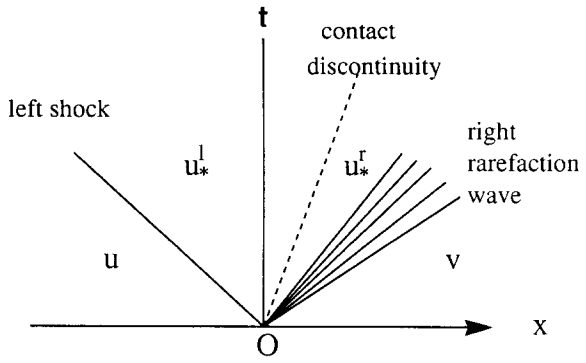

FIG. 5.1. Diagram for the Riemann problem with u and v as its left and right states.

(3) Set a CDCC in the cell $[x_{j_1}, x_{j_1+1}]$ with $\bar{\xi}_2^n$ as its discontinuity position. When the wave between $u_{j_1}^n$ and u_*^l is a shock, set an LSCC in the same cell with $\bar{\xi}_1^n$ as its discontinuity position, which with the CDCC set before are stacked in $[x_{j_1}, x_{j_1+1}]$ with u_*^l as their hidden middle state. When the connecting wave is a rarefaction wave, no critical cell is set; meanwhile $u_{j_1}^n$ should be updated by $u_{j_1}^n - \bar{\xi}_1^n(u_*^l - u_{j_1}^n)$. The wave between u_*^r and $u_{j_1+1}^n$ is treated similarly.

If a resulting k-wave is a discontinuity, the conclusion of Proposition 4.1 with respect to this wave is still true, only ξ^n should be replaced by $\bar{\xi}_k^n$, the discontinuity position of the k-wave, and (4.3a) should be understood as the k-component of the linear representation.

Also, when a resulting wave is a discontinuity, the conclusion of Proposition 4.2 is still true with respect to it. When a resulting wave is a rarefaction, the conservation error allocated to it should be transferred back to the numerical solution for the sake of conservation, since there is no critical cell set for it. This is why we update the numerical solution at the endpoint of the critical cell. A geometric interpretation of this is as follows; Since the collision happens at a certain moment between the two time levels, the updated $u_{j_1}^n$ represents the position of the rarefaction and how much the rarefaction has been spanned since the collision.

6. NUMERICAL EXAMPLES

We use the version of the treatment developed in this paper to recompute Example 2, and Example 4 in [7], and find that the results are qualitatively same.

The underlying scheme used for the numerical experiments in this section is a second-order TVD scheme

described in [10] with a Runge–Kutta-type time discretization. The second-order Runge–Kutta method is two-step and can be written as

$$u_j^{n+1/2} = u_j^n - \lambda(\hat{f}_{j+1/2}^n - \hat{f}_{j-1/2}^n), \tag{6.1a}$$

$$\begin{aligned} \bar{u}_j^{n+1} &= u_j^{n+1/2} - \lambda(\hat{f}_{j+1/2}^{n+1/2} - \hat{f}_{j-1/2}^{n+1/2}), \\ u_j^{n+1} &= \tfrac{1}{2}(u_j^n + \bar{u}_j^{n+1}) \end{aligned} \tag{6.1b}$$

with a numerical flux $\hat{f}$ satisfying

$$\frac{1}{h}(\hat{f}_{j+1/2}^n - \hat{f}_{j-1/2}^n) = f_x|_{x=x_j} + O(h^2), \tag{6.2}$$

where (6.1a) is the predictor step and (6.1b) is the corrector step. Since the underlying scheme is two-step, there is some complication in incorporating the treatment into it. The following remarks address how we deal with the complication:

(1) The treatment should be performed in each of the predictor and corrector steps; however, only the predictor step may move critical cells to their adjacent cells according to their discontinuity positions. The corrector step does not move critical cells, even though the final discontinuity positions might be a little bit out of the critical cells;

(2) for a critical cell, the corrector step gives a discontinuity position $\bar{\xi}^{n+1}$. The final discontinuity position on the new level is

$$\xi^{n+1} = \tfrac{1}{2}(\xi^n + \bar{\xi}^{n+1}); \tag{6.3}$$

(3) when a critical cell moves to its adjacent cell in the predictor step, u_j^n at the grid point crossed over by the discontinuity in the second formula in (6.1b) is replaced by the corresponding extrapolated value in the scalar case or the corresponding middle state in the system case. For example, when a critical cell $[x_{j_1}, x_{j_1+1}]$ moves to its left adjacent cell $[x_{j_1-1}, x_{j_1}]$, $u_{j_1}^n$ in the second formula in (6.1b) is replaced by $u_{j_1}^{n,+}$ in the scalar case or $u_{j_1,*}^{n,l}$ in the system case.

EXAMPLE 1. This example is the recomputing of Example 2 in [7], where

$$u_t + u_x = 0, \tag{6.4a}$$

$$u_0(x+0.5) = \begin{cases} -x\sin\left(\frac{3}{2}\pi x^2\right), & -1 < x < -\frac{1}{3} \\ |\sin(2\pi x)|, & |x| < \frac{1}{3}, \\ 2x - 1 - \frac{\sin(3\pi x)}{6}, & \frac{1}{3} < x < 1, \end{cases} \tag{6.4b}$$

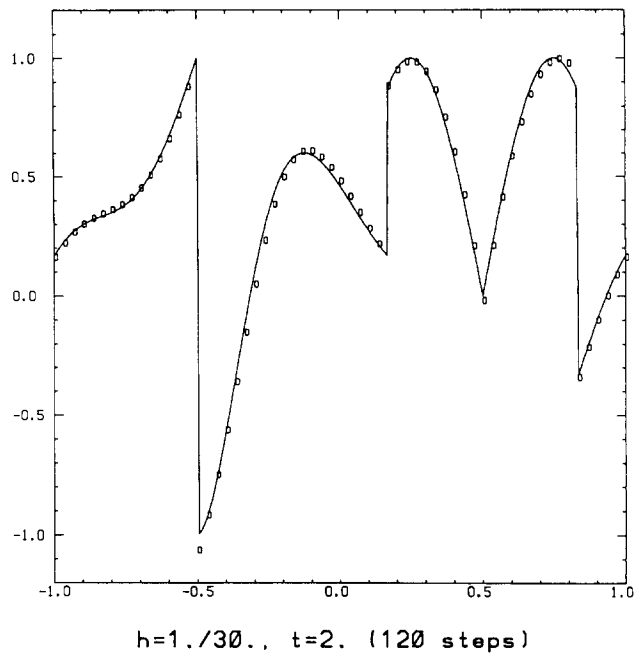

FIG. 6.1. Linear scalar equation with highly discontinuous initial data, 60 grid points.

and

$$u_0(x+2) = u_0(x). \tag{6.4c}$$

The mesh ratio λ is 0.5, and a TVB modification developed by Shu in [11] is made to the underlying scheme to improve

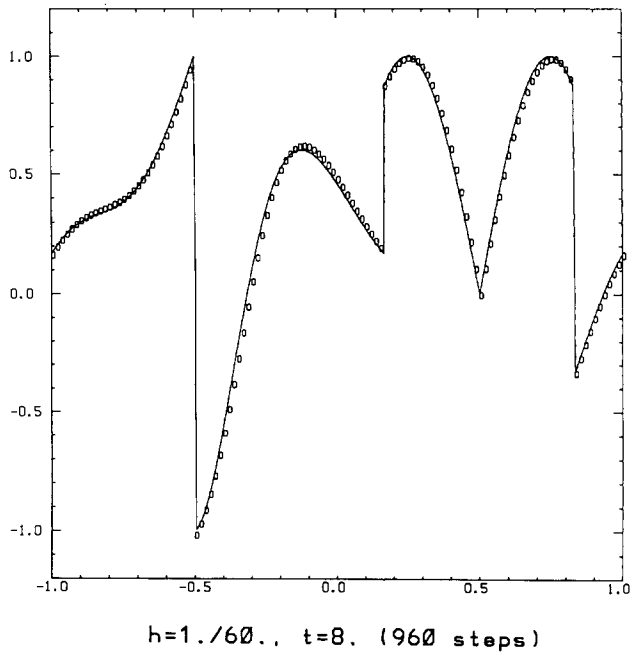

FIG. 6.2. Linear scalar equation with highly discontinuous initial data, 120 grid points.

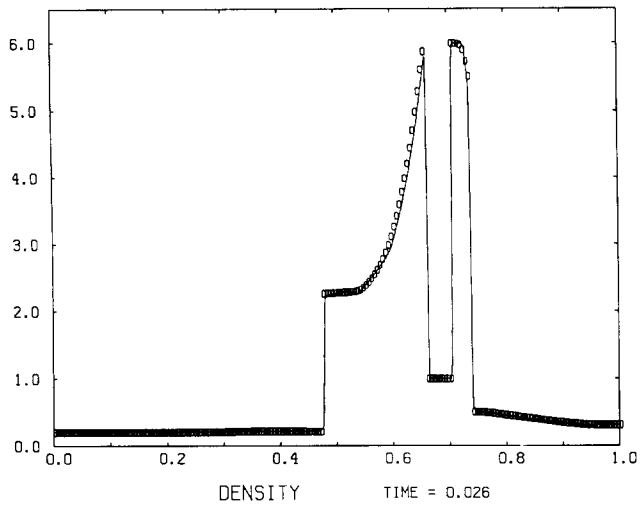

FIG. 6.3. Euler equations, blast wave problem, $t = 0.026$.

the computation near extremum points. Figure 6.1 presents the numerical result with $h = \frac{1}{30}$ (60 grid points) at $t = 2$. Figure 6.2 presents the numerical result with $h = \frac{1}{60}$ (120 grid points) at $t = 8$. We compare the results with those without treatment, which are not presented here. Figures 6.1 and 6.2 show quite an improvement in discontinuity transition and overall resolution.

EXAMPLE 2. This example is the recomputing of Example 4 in [7], which is for a Euler system with the initial values

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

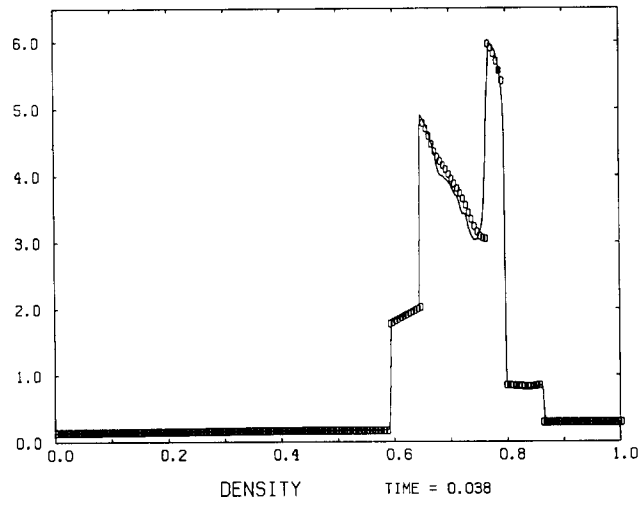

FIG. 6.4. Euler equations, blast wave problem, $t = 0.038$.

and solid wall boundary conditions, where

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

λ is still 0.5. The numerical results for density at $t = 0.026$ and $t = 0.038$ are presented in Fig. 6.3 and Fig. 6.4, respectively, with the solid lines representing the numerical solution obtained by an ENO scheme with 800 grid points, considered as the exact solution here. At around $t = 0.032$, a very strong rarefaction wave is generated from the interaction of a shock and a contact discontinuity. The underlying scheme based on Roe's approximate Riemann solver cannot span the rarefaction; therefore, the viscosity term

$$\Delta_+(\Delta_-\Delta_+)\, u_j^n$$

with a coefficient of 0.2 is added to the underlying scheme to obtain a physical solution.

7. CONCLUSIONS

We have developed a new version of the treatment of discontinuities, which is a variant of the $x - t$ version in [7]. The difference between the two versions is that they track different objects. The $x - t$ version tracks the local conservation errors, while the present one tracks the discontinuity positions. We have shown the equivalence between the two-version piecewise smooth solutions. We have also developed a stacking treatment for interaction of the discontinuities. Numerical experiments show that our treatment can be efficiently used in a two-step scheme.

ACKNOWLEDGMENTS

The author thanks Professor Ami Harten and Professor Stanley Osher for helpful discussions.

REFERENCES

1. I-L Chen, J. Glimm, U. McBryan, B. Plohr, and S. Yaniv, *J. Comput. Phys.* **62**, 83 (1986).
2. P. Charrier and B. Tessieras, *SIAM J. Numer. Anal.* **23**, 461 (1986).
3. A. Chorin and J. Marsden, *A Mathematical Introduction to Fluid Mechanics* (Springer-Verlag, New York/Berlin, 1979).
4. B. Sjogreen and B. Engquist, in *Procedings of the Third International*

Conference on Hyperbolic Problems, Uppsala, Sweden, 1990, edited by B. Engquist and B. Gustafsson (Studentlitteratur, Sweden, 1991), p. 848.

5. J. Glimm, C. Klingenberg, O. McBryan, B. Plohr, D. Sharp, and S. Yaniv, *Advs. Appl. Math.* **6**, 259 (1985).
6. A. Harten, *J. Comput. Phys.* **83**, 148 (1989).
7. D. Mao, *J. Comput. Phys.* **92**, 422 (1991).
8. D. Mao, *J. Comput. Math.* **3**, 256 (1985). [Chinese]
9. Moretti, *Thoughts and Afterthoughts about Shock Computations*, Report No. PIBAL-72-37, Polytechnic Institute of Brooklyn, 1972.
10. S. Osher and S. Chakravarthy, *Very High Order Accurate TVD Schemes*, ICASE Report 84.44, 1984, IMA Volume in Mathematics and Its Applications, Vol. 2 (Springer-Verlag, New York/Berlin, 1986), p. 229.
11. C-W. Shu, *Math. Comput.* **49**, 105 (1987).
12. B. K. Swartz and B. Wendroff, *Appl. Numer. Math.* **2**, 385 (1986).