90-19

A Treatment for Discontinuities for Finite Difference Methods

Mao De-kang [1]
Department of Mathematics
University of California
Los Angeles, CA 90024
and
Department of Mathematics
Shanghai University of Science and Technology

Abstract

This paper continue to discuss the treatment for discontinuities in [11] and [12], whose main idea is that on each side of a discontinuity the computation draws the information that jumps at the discontinuity only from the same side. A numerical method for ordinary differential equations is incorporated into the algorithm to compute the locations of the discontinuity. Analysis shows that this is equivalent to record the conservation errors and compensate them in the following time-steps; therefore, the treatment still has a conservation feature. Finally, some numerical examples for both scalar and system cases are displayed.

Subject Classification: 65M05, 35L65
Key Words: treatment of discontinuity, conservation feature, interaction of discontinuities

[1]Research was supported by ONR Grant No. N00014-86-k-0691.

1. Introduction

The partial differential equations concerned in this paper is 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 $u = (u_1, u_2, \cdots, u_m)^T$, and the Jacobian matrix of f has m real eigenvalues and a complete set of m linearly independent right-eigenvectors. A weak solution of (1.1) is a bounded measurable function $u(x,t)$ satisfying

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

for all smooth test functions. This paper continues to discuss the treatment of discontinuities introduced in [11] and [12] for finite difference methods for (1.1).

As we claimed in [11], the treatment is a shock tracking technique, whose main idea is that the computation on each side of a discontinuity draws the information that jumps at the discontinuity only from the same side. In the scalar case, this is done by using extrapolated data from the same side at the grid points on the other side. In the system case, Riemann problems related to the original and extrapolated data are solved to obtain the data that replace the original data at the grid points on the other side. By doing this, the whole computation still proceeds on the regular grid, and there is no need for a lower dimensional grid to resolve the discontinuity; therefore, the treatment applies to any difference schemes, and the algorithm is much simpler than the traditional shock tracking methods.

The idea that computation uses information only from one side of a discontinuity has been behind many recently developed numerical schemes and techniques. Except [12], an early paper by the author, I first saw this idea in [2] which uses the extrapolated data on each side of a discontinuity it tracks. [2] also explains that the idea makes sense since the extrapolated data are "virtual" due to the characteristic-converging feature for shocks and the characteristic-paralleling feature for contact discontinuities.

ENO schemes, constructed by Harten, Osher, Engquist and Chakravarthy (see [6, 7], [8], [9]), uses a local adaptive stencil to obtain information automatically from regions of the smoothness when the solution develops discontinuities. As a result, near an isolated strong discontinuity, the stencil

of an ENO flux will choose grid points only from one side of it. However, ENO schemes do not make the computation completely get information from one side of a discontinuity. For example, when a discontinuity locates right in a cell $[x_j, x_{j+1}]$, the stencil of an ENO flux corresponding to the cell at least contains its two endpoints, i.e. x_j and x_{j+1}; therefore, the flux uses information from at least one point on the other side.

Later, Harten developed a technique called "subcell resolution", which recovers the location of a discontinuity from the cell averages of the numerical solution and uses it in computation. In doing this, the subcell uses extrapolated data in constructing the numerical fluxes on the two sides to prevent their stencils to be across the discontinuity. Harten's subcell keeps conservation feature for the algorithm; as a cost, there is at least one transition point in the discontinuity. It is impossible for a conservative scheme to always keep a discontinuity in one cell, since a true discontinuity almost never coincides with a cell boundary and, therefore, at least one transition point is needed to represent the discontinuity's location within a cell. Another feature of Harten's subcell is that, being a shock capturing technique, it does not save locations of the discontinuities for the use in the following timesteps; instead, it recovers the locations once every time.

This paper continues to discuss the $x - t$ version of the treatment in [11]. This version always keeps a discontinuity in one cell, so that it is geometrically simple and easy to be extended to the high dimensional case, a paper about its extension to two dimensional case is in preparation. This version is not conservative; however, a conservation feature in a more general sense is discovered for it.

The paper is organized in the following manner: §2. describes the treatment for a single discontinuity in scalar case, which is a little different from but still almost equivalent to that described in [11]. §3. analyses the treatment by writing it into the form of (3.1), where q^n is the local conservation error recorded at every timestep. The idea that the algorithm records the local conservation errors and compensates them in the following timesteps also occurs in [4]. However, this paper reveals the relation between the conservation error and the location of a discontinuity, which is the theorem 3.1. As a result, the uniform boundedness of the conservation error follows; which is the conservation feature in a more general sense we mentioned before. §4. extends the treatment to the case of interactions of discontinuities and also shows its conservation feature. §5. extends the treatment to the system case.

§6. displays numerical examples for the linear scalar and Euler system cases.

2. Treatment for a Single Discontinuity

As usual, the discussion starts with the simplest case, i.e. the one dimensional scalar problem, for which the corresponding equation is

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

where both u and f are scalar. The underlying scheme used to compute the numerical solution is a general conservative difference 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 datum of the numerical solution at a grid point (x_j, t^n),

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

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 flux is consistent with the flux in (2.1) in the sense that

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

The following example is used to show the performance of the treatment for a single discontinuity.

Assume that on the level n the numerical solution just has a jump in a cell $[x_{j_1}, x_{j_1+1}]$, on each side of which the numerical solution is supposed to be smooth (as shown in Figure 2.1). Also assume that the position of the discontinuity within the cell is known as ξ^n. The things now we are going to compute are the numerical solution and the position of the discontinuity on the latter level, i.e., u^{n+1} and ξ^{n+1}. The cell that contains the discontinuity is called a *critical cell* , on which the treatment is going to be applied. (We change the name for the cell that contains a discontinuity, considering that the new name is more proper than the old one used in [12], which is *generated interval*).

As mentioned in §1., the treatment is a shock tracking technique based on the idea that on each side of the discontinuity the computation uses information only from the same side. In the present example it performs in the following four steps:

1) This step prepares the data needed in the computation. It extrapolates the numerical solution from each side of the critical cell to the other side, and gets a set of extrapolated data:

$$u^{n,+}_{j_1-k}, \cdots, u^{n,+}_{j_1}, u^{n,-}_{j_1+1}, \cdots, u^{n,-}_{j_1+k+1}, \tag{2.5}$$

where the data with "$-$" are from the left to the right and the data with "$+$" are from the right to the left (as shown in Figure 2.1).

2) This step computes u^{n+1}, the numerical solution on the new level. When x_j is on the left side of the critical cell, by which mean that $j \leq j_1$, it computes u^{n+1}_j by the fluxes with "$-$"; i.e.,

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

where

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

When x_j is on the right side of the critical cell, it computes u^{n+1}_j by fluxes with "$+$", which are defined as

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

3) This step computes ξ^{n+1}, the position of the discontinuity on the new level. First, it extrapolates u^n from each side to ξ^n and gets two extrapolated data $u^l_{\xi^n}$, $u^r_{\xi^n}$. Then, it computes the speed of the discontinuity by 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, it computes ξ^{n+1} as

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

4) This step determines the critical cell on the level $n+1$. If ξ^{n+1} is still in $[x_{j_1}, x_{j_1+1}]$, it takes the same cell as the critical cell on the new level. If ξ^{n+1}

moves into the left adjacent cell, it takes $[x_{j_1-1}, x_{j_1}]$ as the critical cell on the new level, meanwhile updates $u_{j_1}^{n+1}$ either by the datum computed only using information from the right side or, more simply, by the extrapolated datum of u^{n+1} at x_{j_1}. If ξ^{n+1} moves into the right adjacent cell, it takes $[x_{j_1+1}, x_{j_1+2}]$ as the new critical cell and updates $u_{j_1+1}^{n+1}$.

Thus ends the computation from the level n to the level $n+1$.

The treatment is a shock tracking technique; however, it does not define a lower dimensional adaptive grid for the discontinuity as the traditional tracking method does (cf [1], [5]), nor computes the numerical solution at grid points on that. Therefore, it does not need adaptive schemes at the points near the discontinuity and the computation still proceeds on the regular grid. The whole algorithm is quite simple and it is possible to use the treatment to capture a spontaneous shock. This paper focuses only on the tracking feature of the treatment; nevertheless, readers can find some discussion about its capturing feature in[11].

Several comments on the above four steps are as follows:

a) In [11], a conservation error is introduced by writing the overall algorithm into (3.1) (in [11] the conservation error is called artificial term), and the treatment is performed based on the observation of the conservation error. However, the treatment in this paper is performed based on the observation of the position of the discontinuity. Nevertheless, Theorem 3.1 in the following section shows the relation between the conservation error and the position of the discontinuity, which indicates that the two versions of the treatment are essentially equivalent to each other when the solution is piecewise smooth.

b) In order to keep the same order of accuracy as the underlying scheme's for the overall algorithm, the order of the extrapolation used in step 1) must be the same as or higher than the order of accuracy of the underlying scheme. For example, if the underlying scheme is of the second order, the extrapolation should be at least the second order (cf[11]).

c) In step 3), the position of the discontinuity is computed in a first oder accurate way; however, the algorithm can be improved to be of the higher order accuracy, either by using information on the level $n+1$ in computing s in (2.9) or by a Runge-Kutta procedure. The algorithm that computes the numerical examples in §6 is built upon a second order Runge-Kutta procedure, which will be described later.

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

$$|s| \leq \max |f'(u)|;$$

therefore, the position of the discontinuity will only moves to the left or right adjacent cell if the mesh ratio λ satisfies CFL-condition. It will not jump to a cell far from the original critical cell.

e) The treatment does not keep the conservation feature for the overall algorithm; in other words, the numerical solution u^n obtained by using the treatment does not satisfy

$$\sum_j u_j^n h = \sum_j u_j^0 h. \tag{2.11}$$

The reason is that in some cell different numerical fluxes are used in the computation. For example, when ξ^{n+1} is still in the critical cell on the level n, the flux corresponding to the cell $[x_{j_1}, x_{j_1+1}]$ in computing $u_{j_1}^{n+1}$ is with "$-$", while the flux corresponding to the same cell in computing $u_{j_1+1}^{n+1}$ is with "$+$", and they are not equal. The similar situations also happen in the other two cases.

3. Conservation Feature of the Treatment

The overall algorithm for the example studied in the last section can be written in a conservation-like form,

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

where

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

and p^n and q^n are introduced to balance the numerical fluxes in the cell where different fluxes are used. The p^n and q^n are nonzero only in the vicinity of the critical cell and are as follows in each case: When ξ^{n+1} moves into the left adjacent cell

$$\begin{aligned} p_{j_1-1/2}^n &= -q_{j_1}^n + (u_{j_1}^n - u_{j_1}^{n,+}) + \lambda(\hat{f}_{j_1-1/2}^{n,-} - \hat{f}_{j_1-1/2}^{n,+}) \\ q_{j_1-1}^{n+1} &= -p_{j_1-1/2}^n, \end{aligned} \tag{3.3}$$

when ξ^{n+1} remains in the original critical cell

$$q_{j_1}^{n+1} = q_{j_1}^n + \lambda(\hat{f}_{j_1+1/2}^{n,+} - \hat{f}_{j_1+1/2}^{n,-}) , \tag{3.4}$$

and when ξ^{n+1} moves into the right adjacent cell

$$\begin{aligned} p_{j_1+1/2}^n &= q_{j_1}^n + \lambda(\hat{f}_{j_1+1/2}^{n,+} - \hat{f}_{j_1+1/2}^{n,-}) \\ q_{j_1+1}^{n+1} &= q_{j_1}^n + (u_{j_1+1}^{n,-} - u_{j_1+1}^n) + \lambda(\hat{f}_{j_1+3/2}^{n,+} - \hat{f}_{j_1+3/2}^{n,-}). \end{aligned} \tag{3.5}$$

As mentioned in the last section, u_j^n is not conserved; however, (3.1) indicates that $u_j^n - q_j^n$ is conserved, i.e.

$$\sum_j u_j^n h - q_{j_1}^n h = \sum_j u_j^0 h - q_{j_0}^0 h, \tag{3.6}$$

where x_{j_0} is the left endpoint of the critical cell on the initial level. Therefore, q^n is the local conservation error caused by the treatment.

In this section two theorems are proved. The first theorem shows the relation between the position of the discontinuity and q^n, from which the uniform boundedness of q^n follows. The second theorem, which is similar to the Lax-Wendroff theorem in [10], says that if the numerical solution converges, it converges to a weak solution of (2.1). Although the discussion in this section is only for the scalar case with a single discontinuity, the latter sections will extend it to more general cases.

In this paper we call a treatment to be of the rth order if it uses the rth order extrapolation in step 1), 3), and 4).

Theorem 3.1 *If both the underlying scheme* (2.2) *and the treatment are of the first order, then*

$$q_{j_1}^n = \frac{(\xi^n - x_{j_1+1/2})(u_{j_1+1}^n - u_{j_1}^n)}{h} + O(1), \tag{3.7}$$

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

$$q_{j_1}^n = \frac{(u_{j_1}^n - u_{j_1}^{n,+})(x_{j_1+1} - \xi^n)^2 + (u_{j_1+1}^n - u_{j_1+1}^{n,-})(\xi^n - x_{j_1})^2}{2h^2} + O(h), \tag{3.8}$$

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

There are several comments on the theorem.

a) The $q_{j_1}^n$ in (3.7) or (3.8) is uniformly bounded if the numerical solution is uniformly bounded; therefore, (3.6) gives that

$$\sum_j u_j^n h = \sum_j u_j^0 h + O(h), \tag{3.9}$$

which indicates that the numerical solution computed by using the treatment is almost conserved.

b) A geometrical interpration for (3.7) and (3.8) is as follows: Suppose that $U^n(x)$ is a piecewise constant function defined as

$$U^n(x) = \begin{cases} u_j^n & x_j \le x < x_{j+1/2} \\ u_{j+1}^n & x_{j+1/2} \le x < x_{j+1} \end{cases}, \tag{3.10}$$

in a cell $[x_j, x_{j+1}]$ that is not the critical cell with $x_{j+1/2}$ as its midpoint, and

$$U^n(x) = \begin{cases} u_{j_1}^n & x_{j_1} \le x < \xi^n \\ u_{j_1+1}^n & \xi^n \le x < x_{j_1+1} \end{cases}. \tag{3.11}$$

in the critical cell (as shown in Figure 3.1). We integrate the function $U^n(x)$ over $(-\infty, \infty)$ and let the integration equal to $\sum_j u_j^n h - q_{j_1}^n h$ (if there is no critical cell, the integration is just $\sum_j u_j^n h$), then is (3.7). If $U^n(x)$ is piecewise linear function defined as

$$U^n(x) = [u_j^n(x_{j+1} - x) + u_{j+1}^n(x - x_j)]/h \tag{3.12}$$

in a non-critical cell, and

$$U^n(x) = \begin{cases} (u_{j_1}^n(x_{j_1+1} - x) + u_{j_1+1}^{n,-}(x - x_{j_1}))/h & x_{j_1} \le x < \xi^n \\ (u_{j_1}^{n,+}(x_{j_1+1} - x) + u_{j_1+1}^n(x - x_{j_1}))/h & \xi^n < x < x_{j_1+1} \end{cases} \tag{3.13}$$

in the critical cell (as shown in Figure 3.2), and let the corresponding integration equal to $\sum_j u_j^n h - q_{j_1}^n h$ again, then is (3.8).

c) If the numerical solution on each side of the critical cell is smooth enough, (3.7) and (3.8) only have a difference of $O(h)$ by ignoring the last error terms in them. This can be shown by rewritting (3.8) as

$$\begin{aligned} q_{j_1}^n = \; & \frac{1}{h}(\xi^n - x_{j_1+1/2})(u_{j_1+1}^n - u_{j_1}^n) \\ & + \frac{1}{2h^2}\{(u_{j_1+1}^n - u_{j_1}^{n,+})(x_{j_1+1} - \xi^n)^2 + (u_{j_1}^n - u_{j_1+1}^{n,-})(\xi^n - x_{j_1}^n)^2\} \\ & + O(h). \end{aligned} \tag{3.14}$$

Proof of (3.7). Denote

$$S^n = (\xi^n - x_{j_1+1/2})(u^n_{j_1+1} - u^n_{j_1}) - hq^n_{j_1}, \tag{3.15}$$

and we are going to show for each case

$$S^{n+1} - S^n = O(h^2), \tag{3.16}$$

from which conclusion of the theorem follows easily. First, for the case that ξ^{n+1} remains in the original critical cell,

$$\begin{aligned} S^{n+1} - S^n = \; & d(u^{n+1/2}_{j_1+1} - u^{n+1/2}_{j_1}) \\ & -\tau\{\hat{f}^{n,+}_{j_1+1/2} + (\xi^{n+1/2} - x_{j_1+1/2})\frac{\hat{f}^{n,+}_{j_1+3/2} - \hat{f}^{n,+}_{j_1+1/2}}{h} \\ & -\hat{f}^{n,-}_{j_1+1/2} - (\xi^{n+1/2} - x_{j_1+1/2})\frac{\hat{f}^{n,-}_{j_1+1/2} - \hat{f}^{n,-}_{j_1-1/2}}{h}\}, \end{aligned} \tag{3.17}$$

where $d = \xi^{n+1} - \xi^n$, $\xi^{n+1/2} = \frac{1}{2}(\xi^{n+1} + \xi^n)$, and $u^{n+1/2}_j = \frac{1}{2}(u^{n+1}_j + u^n_j)$. The right side of (3.17) is an approximation to the left side of the following Hugoniot condition

$$dx - dt[f]/[u] = 0 \tag{3.18}$$

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 for both the numerical solution u^n and the position of the discontinuity ξ^n are of the first order too; therefore, the (3.16) follows. For the case in which ξ^n moves to the right adjacent cell,

$$\begin{aligned} S^{n+1} - S^n = \; & d(u^{n+1/2}_{j_1+1/2} - u^{n+1/2}_{j_1-1/2}) \\ & -\tau\{\hat{f}^{n,+}_{j_1-1/2} + (\xi^{n+1/2} - x_{j_1-1/2})\frac{\hat{f}^{n,+}_{j_1+1/2} - \hat{f}^{n,+}_{j_1-1/2}}{h} \\ & \hat{f}^{n,-}_{j_1-1/2} - (\xi^{n+1/2} - x_{j_1-1/2})\frac{\hat{f}^{n,-}_{j_1-1/2} - \hat{f}^{n,-}_{j_1-3/2}}{h}\} + O(h^2), \end{aligned} \tag{3.19}$$

where $u^{n+1/2}_{j+3/2} = \frac{1}{2}(u^{n+1}_{j+1} + u^n_j)$. Again (3.16) follows. Finally, for the case in which ξ^n moves to the left adjacent cell, a similar analysis shows that (3.16) is also true. Thus ends the proof.

Proof of (3.8). (3.8) can be written as follows:

$$\begin{aligned}
&(u^{n,+}_{j_1+1/2} - u^{n,-}_{j_1+1/2})(\xi^n - x_{j_1+1/2}) - q^n_{j_1} h \\
&-\frac{1}{4h}(u^{n,-}_{j_1+1} - u^n_{j_1})\{(x_{j_1+1} - \xi^n)^2 - (\xi^n - x_{j_1})^2\} \\
&+\frac{1}{4h}(u^n_{j_1+1} - u^{n,+}_{j_1})\{(x_{j_1+1} - \xi^n)^2 + (\xi^n - x_{j_1})^2\} \\
&= O(h^2),
\end{aligned} \tag{3.20}$$

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 (3.20), and we have

$$Q^n = -\frac{1}{2h}(u^{n,-}_{j_1+1} - u^n_{j_1} - u^n_{j_1+1} + u^{n,+}_{j_1})((x_{j_1+1/2} - \xi^n)^2 + \frac{h^2}{4}), \tag{3.21}$$

Denote by S^n the left side of (3.20), and we are going to show that (3.16) with $O(h^3)$ being its right side is true. We still start with the case in which ξ^{n+1} remains in the original critical cell, and have

$$\begin{aligned}
S^{n+1} - S^n = \;& d(u^{n+1/2,+}_{j_1+1/2} - u^{n+1/2,-}_{j_1+1/2}) \\
&+(\xi^{n+1/2} - x_{j_1+1/2})(u^{n+1,+}_{j_1+1/2} - u^{n+1,-}_{j_1+1/2} - u^{n,+}_{j_1+1/2} + u^{n,-}_{j_1+1/2}) \\
&-h(q^{n+1}_{j_1} - q^n_{j_1}) + (Q^{n+1} - Q^n),
\end{aligned} \tag{3.22}$$

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

$$u^{n+1,-}_{j_1+1} = u^{n,-}_{j_1+1} - \lambda(\hat{f}^{n,-}_{j_1+3/2} - \hat{f}^{n,-}_{j_1+1/2}) + O(h^3), \tag{3.23}$$

and

$$u^{n+1,+}_{j_1} = u^{n,+}_{j_1} - \lambda(\hat{f}^{n,+}_{j_1+1/2} - \hat{f}^{n,+}_{j_1-1/2}) + O(h^3). \tag{3.24}$$

Therefore,

$$u^{n+1,-}_{j_1+1/2} - u^{n,-}_{j_1+1/2} = -\frac{\lambda}{2}(\hat{f}^{n,-}_{j_1+3/2} - \hat{f}^{n,-}_{j_1-1/2}) + O(h^3), \tag{3.25}$$

and

$$u_{j_1+1/2}^{n+1,+} - u_{j_1+1/2}^{n,+} = -\frac{\lambda}{2}(\hat{f}_{j_1+3/2}^{n,+} - \hat{f}_{j_1-1/2}^{n,+}) + O(h^3). \tag{3.26}$$

Moreover, we have

$$(x_{j_1+1/2} - \xi^{n+1})^2 = \frac{d^2}{4} + d(\xi^{n+1/2} - x_{j_1+1/2}) + (\xi^{n+1/2} - x_{j_1+1/2})^2, \tag{3.27}$$

and

$$(x_{j_1+1/2} - \xi^{n})^2 = \frac{d^2}{4} - d(\xi^{n+1/2} - x_{j_1+1/2}) + (\xi^{n+1/2} - x_{j_1+1/2})^2. \tag{3.28}$$

Therefore,

$$\begin{aligned} Q^{n+1} - Q^n = \; & d\left(\frac{u_{j_1+1}^{n+1/2} - u_{j_1}^{n+1/2,+}}{h} - \frac{u_{j_1+1}^{n+1/2,-} - u_{j_1}^{n+1/2}}{h}\right)(\xi^{n+1/2} - x_{j_1+1/2}) \\ & + R^{n+1/2}, \end{aligned} \tag{3.29}$$

where,

$$\begin{aligned} R^{n+1/2} = \; & \{(u_{j_1+1}^{n+1,-} - u_{j_1}^{n+1}) - (u_{j_1+1}^{n,-} - u_{j_1}^{n}) \\ & -(u_{j_1+1}^{n+1} - u_{j_1+1}^{n+1,+}) + (u_{j_1+1}^{n} - u_{j_1}^{n,+})\} \\ & \frac{1}{2h}\{\frac{d^2}{4} + \frac{h^2}{4} + (\xi^{n+1/2} - x_{j_1+1/2})^2\}. \end{aligned} \tag{3.30}$$

Obviously, $R^{n+1/2} = O(h^3)$. Substitute (3.23)–(3.30) into (3.22), we have

$$\begin{aligned} S^{n+1} - S^n \; = \; & d\{u_{j_1+1/2}^{n+1/2,+} + \frac{u_{j_1+1}^{n+1/2} - u_{j_1}^{n+1/2,+}}{h}(\xi^{n+1/2} - x_{j_1+1/2}) \\ & -u_{j_1+1/2}^{n+1/2,-} - \frac{u_{j_1+1}^{n+1/2,-} - u_{j_1}^{n+1/2}}{h}(\xi^{n+1/2} - x_{j_1+1/2})\} \\ & -\tau\{\hat{f}_{j_1+1/2}^{n,+} + \frac{\hat{f}_{j_1+3/2}^{n,+} - \hat{f}_{j-1/2}^{n,+}}{2h}(\xi^{n+1/2} - x_{j_1+1/2}) \\ & -\hat{f}_{j_1+1/2}^{n,-} - \frac{\hat{f}_{j_1+3/2}^{n,-} - \hat{f}_{j_1-1/2}^{n,-}}{2h}(\xi^{n+1/2} - x_{j_1+1/2})\} + O(h^3), \end{aligned} \tag{3.31}$$

the right side of which is again an approximation to the left side of the Hugoniot condition (3.18) at the point $(\xi^{n+1/2}, t^{n+1/2})$. Since both the underlying scheme and the treatment are of the second order, the accuracy for both the numerical solution and the position of the discontinuity is of the second order too, and also $\hat{f}^n_{j+1/2} = f|_{x=x_{j+1/2},t=t^{n+1/2}} + O(h^2)$. Thus, the conclusion follows.

For the case in which ξ^{n+1} moves to the left adjacent cell,

$$\begin{aligned} S^{n+1} - S^n = \; & d\{\bar{u}^{n+1/2,+}_{j_1} + (\xi^{n+1/2} - x_{j_1-1/2})\frac{(u^{n+1}_{j_1} - u^{n+1,+}_{j_1-1})}{2h} \\ & +(\xi^{n+1/2} - x_{j_1+1/2})\frac{(u^{n}_{j_1+1} - u^{n,+}_{j_1})}{2h} \\ & -\bar{u}^{n+1/2,-}_{j_1} - (\xi^{n+1/2} - x_{j_1-1/2})\frac{(u^{n+1,-}_{j_1} - u^{n+1}_{j_1-1})}{2h} \\ & -(\xi^{n+1/2} - x_{j_1+1/2})\frac{(u^{n,-}_{j_1+1} - u^{n}_{j_1})}{2h}\} \\ & -\tau\{\hat{f}^{n,+}_{j_1-1/2} + \frac{(\hat{f}^{n,+}_{j_1+1/2} - \hat{f}^{n,+}_{j_1-3/2})}{2h}(\xi^{n+1/2} - x_{j_1-1/2}) \\ & -\hat{f}^{n,-}_{j_1-1/2} - \frac{(\hat{f}^{n,-}_{j_1+1/2} - \hat{f}^{n,-}_{j_1-3/2})}{2h}(\xi^{n+1/2} - x_{j_1-1/2})\} + O(h^3), \end{aligned} \tag{3.32}$$

where $\bar{u}^{n+1/2,-}_{j_1} = \frac{1}{2}(u^{n+1,-}_{j_1-1/2} + u^{n,-}_{j_1+1/2})$ and $\bar{u}^{n+1/2,+}_{j_1} = \frac{1}{2}(u^{n+1,+}_{j_1-1/2} + u^{n,+}_{j_1+1/2})$. For the case in which ξ^n moves to the right adjacent cell, a similar equation holds. Therefore, the conclusion is true for these two cases too by the same analysis. Thus ends the proof.

The following theorem is similar to the Lax-Wendroff theorem in [10].

Theorem 3.2 *Assume that as h and τ tend to zero with $\lambda = \tau/h \geq c_0 > 0$, the numerical solution computed by the first or second order underlying scheme with the first or second order treatment, which involves a single discontinuity, converges boundedly to some function $u(x,t)$, then $u(x,t)$ is a weak solution of (2.1).*

Proof. Multiply (3.1) by any test function ϕ, sum with respect to j and

n, and by almost the same argument in [10], we get the indentity

$$\sum_n \sum_j \{\frac{\phi_j^n - \phi_j^{n-1}}{\tau} u_j^n + \frac{\phi_{j+1}^n - \phi_j^n}{h} \tilde{f}_{j+1/2}^n\} \tau h + \sum_j u_j^0 \phi_j^0$$

$$= \sum_n \sum_j \{\frac{\phi_j^n - \phi_j^{n-1}}{\tau} q_j^n + \frac{1}{\lambda} \frac{\phi_{j+1}^n - \phi_j^n}{h} p_{j+1/2}^n\} \tau h + q_{j_0}^0 \phi_{j_0}^0. \tag{3.33}$$

Since the numerical solution converges to $u(x,t)$, the left side of (3.33) tends to the the left side of (1.2) as h and τ tend to zero. Since p^n and q^n are nonzero only in the vicinity of the critical cell, and q^n is uniformly bounded, which implies that p^n is uniformly bounded too, the right side of (3.33) tends to zero. Thus the conclusion follows.

The critical point in the proof is that the artificial terms p^n and q^n are uniformly bounded and locally (i.e. only in the vicinity of the critical cell) nonzero, which indicates that the algorithm using the treatment deviates from a conservative scheme only by a difference that tends to zero when the mesh size tends to zero. The latter sections will show that this is also true for more general cases.

4. Treatment for Interactions of Discontinuities

Till now only the treatment for a single discontinuity has been studied. Nevertheless, the solution to (2.1) may have several discontinuities, which may interact with each other. This section describes the treatment for the interaction of discontinuities. Only the interaction of two discontinuities is considered, and the method applies to more general cases.

Assume $[x_{j_1}, x_{j_1+1}]$ and $[x_{j_2}, x_{j_2+1}]$ are two critical cells on the level n, which divide the whole x region into three parts, the left, right, and middle parts. The numerical solution in each part is supposed to be smooth (as shown in Figure 4.1). First, when the two critical cells get close to each other, the stencil of the numerical flux at some grid points in the middle part might cover both of the two critical cells. As a result, the following numerical flux

$$\hat{f}(u_{j-k+1}^{n,+}, \cdots, u_{j_1}^{n,+}, u_{j_1+1}^n, \cdots, u_{j_2}^n, u_{j_2+1}^{n,-}, \cdots, u_{j+k}^{n,-}) \tag{4.1}$$

might have to be used at these points since we insist that the computation in each part use information only from the same part, where the data with"+"

and "−" are the extrapolated data from the middle part to the left and right respectively. Also the order of the extrapolation used in step 1) and step 3) might have to be lowered since we might not have enough points in the middle part to keep the same order when the two critical cells get close to each other.

When the two critical cells moves to a same cell, a handling about when and how to merge them to form a new critical cell is needed. The treatment in [11] merges the two critical cells immediately after they get together; nevertheless, the positions of this two discontinuities usually still have a short distance between them at this moment. This means that the mergence takes place a little early. In this paper, a more precise way is described to handle the case.

First, we must avoid to having the two critical cells to move across each other, i.e. we can not let the case (c) in Figure 4.2 happen. When it happens in the third step of the treatment, we change the algorithm by holding one of the critical cells, even though its position of the discontinuity moves out of the corresponding cell.

When the two critical cells get together in a same cell with their positions of discontinuities satisfying $\xi_1^n < \xi_2^n$, they are not going to be merged, but are still reguarded as two individual critical cells with a middle state between them. In other words, they are considered to be overlapped by each other in the same cell. The middle state, which does not appear in the numerical solution, stands as the right state for the left critical cell and the left state for the right one; that is to say, the computation proceeds as the left critical cell has the middle state on its right and the right critical cell has the middle state on its left. The middle state can be defined in a proper way. For example, in the cases (a) and (b) in Figure 4.2, The un-updated datum of the numerical solution at the grid point (x_{j_2}, t^{n+1}) before the fourth step can be chosen as the middle state on the level $n+1$.

The two overlapped critical cells coexist until $\xi_1^n > \xi_2^n$, then they have to be merged to form a new critical cell. The position of the discontinuity for the new one is computed as follows:

$$\xi^n = \frac{\xi_1^n(u_* - u_{j_1}^n) + \xi_2^n(u_{j_1+1}^n - u_*)}{u_{j_1+1}^n - u_{j_1}^n}, \tag{4.2}$$

where u_* is the middle state in between. Denote by s_1^{n-1}, s_2^{n-1} and s^n the exact positions of the discontinuities for the left and right critical cells on the

level $n-1$ and the new critical cell on the level n respectively. The following theorem shows that the treatment has a second order accuracy.

Theorem 4.1 *Assume that ξ_1^{n-1} and ξ_2^{n-1} are exact on the level $n-1$, to be precise, $\xi_1^{n-1} = s_1^{n-1}$ and $\xi_2^{n-1} = s_2^{n-1}$, then*

$$\xi^n = s^n + O(h^2) \tag{4.3}$$

if both the underlying scheme and the treatment are at least of the first order.

Proof. First we prove that ξ^n computed by (4.2) is exact when the numerical solution in each of the three parts is constant. Since the solution is piecewise constant, both the discontinuities starting from the level $n-1$ are straight lines. They start at ξ_1^{n-1} and ξ_2^{n-1} on the level $n-1$, with speeds s_1 and s_2, meet at a point (ξ^*, t^*), and then merge to form a new one, with a speed s_0, which reachs the level n at s^n (as shown in Figure 4.3). Extend the left and right discontinuities starting from the level $n-1$ to the level n at ξ_1^n and ξ_2^n respectively, which are the positions of the discontinuities computed without considering interaction. Now we are going to show that ξ_1^n, ξ_2^n and s^n satisfy (4.2) with s^n standing for ξ^n.

Obviously,

$$s^n = \xi^* + (t^n - t^*)s_0, \tag{4.4a}$$

$$\xi_1^n = \xi^* + (t^n - t^*)s_1, \tag{4.4b}$$

$$\xi_2^n = \xi^* + (t^n - t^*)s_2, \tag{4.4c}$$

where

$$s_0 = \frac{f(u_r) - f(u_l)}{u_r - u_l}, \tag{4.5a}$$

$$s_1 = \frac{f(u_*) - f(u_l)}{u_* - u_l}, \tag{4.5b}$$

$$s_2 = \frac{f(u_r) - f(u_*)}{u_r - u_*}, \tag{4.5c}$$

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

$$s_0 = \frac{(u_* - u_l)s_1 + (u_r - u_*)s_2}{u_r - u_l} \tag{4.6}$$

from (4.5); thus, by(4.4)

$$s^n = \frac{\xi_1^n(u_* - u_l) + \xi_2^n(u_r - u_*)}{u_r - u_l}, \tag{4.7}$$

which means that (4.2) is exact now.

Now we prove that ξ^n has the second order accuracy if the numerical solution is piecewise smooth. As we argued before, the numerical solution in each of the three parts has at least a first order accuracy since both the underlying scheme and the treatment are at least of the first order. As a result, each identity in (4.5) has an error at least of $O(h)$; therefore, each identity in (4.4) has an error at least of $O(h^2)$. Thus the theorem follows.

The treatment also can be analysed in terms of p^n and q^n to show its conservative feature. Assume that the two critical cells merge on the level n and the local conservation errors for them before they merge are $q^n_{j_1,1}$ and $q^n_{j_1,2}$ respectively.

Theorem 4.2 *If both the underlying scheme and the treatment are of the first order, (3.7) holds for the new-formed critical cell on the level n with $q^n_{j_1,1} + q^n_{j_1,2}$ standing for $q^n_{j_1}$. If both the underlying scheme and the treatment are of the second order, (3.8) holds for the new-formed critical cell with $q^n_{j_1,1} + q^n_{j_1,2}$ standing for $q^n_{j_1}$.*

Proof. According to theorem 3.1,

$$q^n_{j_1,1} = \frac{(\xi_1^n - x_{j_1+1/2})(u_* - u^n_{j_1})}{h} + O(1) \tag{4.8}$$

and

$$q^n_{j_1,2} = \frac{(\xi_2^n - x_{j_1+1/2})(u^n_{j_1+1} - u_*)}{h} + O(1) \tag{4.9}$$

if the underlying scheme and the treatment are both of the first order, where u_* is the middle state defined as before. The first conclusion of the theorem is obtained by adding (4.8) and (4.9) and substituting (4.2). Also according to theorem 3.1 and its comment c),

$$q^n_{j_1,1} = \frac{(\xi_1^n - x_{j_1+1/2})(u_* - u^n_{j_1})}{h} + O(h) \tag{4.10}$$

and

$$q^n_{j_1,2} = \frac{(\xi_1^n - x_{j_1+1/2})(u^n_{j_1+1} - u_*)}{h} + O(h) \tag{4.11}$$

if the underlying scheme and the treatment are both of the second order. The second conclusion is obtained by the same argument and comment c).

Theorem 4.2 means that the conclusion of theorem 3.1 still holds for the local-conservation error of the new-formed critical cell $q^n_{j_1} = q^n_{j_1,1} + q^n_{j_1,2}$. Thus p^n and q^n are still uniformly bounded and locally nonzero. As results, (3.9) and the following corollary hold.

Corollary 4.1 *The conclusion of theorem 3.2 is still true if the solution is piecewise smooth with a finite number of the interactions of discontinuities*

5. Treatment for Euler Equations of Gas Dynamics

The Euler equations of gas dynamics for a polytropic gas are

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

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

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

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

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

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

where $c = (\gamma p/\rho)^{\frac{1}{2}}$ is the sound speed. This section describes the extension of the treatment to the Euler system case.

The underlying scheme is still assumed to be a general conservative scheme defined by (2.2) with the numerical flux (2.3), only both u^n and f^n are vectors, and the discussion also starts with the case involving a single discontinuity. It is well known that there are three different kinds of discontinuities in the Euler system, the left shock, right shock and contact discontinuity; therefore, there should be correspondingly three different kinds of critical cells, which are called as the left shock critical cell (LSCC), right shock critical cell (RSCC) and contact discontinuity critical cell (CDCC).

A Riemann problem with u and v as its left and right states, which we denote by $RP(u, v)$, has three waves and two (left and right) middle states (as shown in Figure 5.1). An LSCC is a critical cell that the left shock in the $RP(u^l_{\xi^n}, u^r_{\xi^n})$ is much stronger than the other two waves, where $u^l_{\xi^n}$ and $u^r_{\xi^n}$ are defined as in §2. The RSCC and CDCC can be defined in similar ways.

The treatment for a single critical cell still proceeds in the four steps described in §2; i.e., 1) compute the extrapolated data on the two sides of the critical cell; 2) compute the numerical solution on the two sides of the critical cell; 3) compute the speed of the discontinuity; and 4) determine the critical cell on the new level and update the data of the numerical solution on some grid points if it is needed. However, the system case has some particular feature the scalar case does not have, which obstructs us from naively extending the treatment. Let's take the left shock as an example. It is well known that the left shock is associated to the first characteristic field. In a flow, it catches the first characteristic lines and let off the other two characteristic lines. Hence, the information carried by the other two characteristic lines on the left side of the shock will be propagated across the shock to the right side. However, the extrapolated data on the right side of the critical cell does not contain this information; therefore, if we still use the extrapolated data in step 2), 3) and 4) on the right side, we will miss the information going from the left side to the right side. For this problem, the traditional shock tracking treats the discontinuity as a moving boundary by solving Riemann problems and imposing boundary conditions on it; however, the treatment in this paper treats it in a much simpler way.

For an LSCC, we replace the original extrapolated data by some other data as follows: In step 2), we solve Riemann problems $RP(u^n_j, u^{n,+}_j)$ for all $j_1 - k \le j \le j_1$ to get a set of left middle states, which we denote by $\{u^{n,l}_{j,*}\}$, and use them to replace the original extrapolated data in computing the numerical solution. In step 3), we solve a Riemann problem $RP(u^l_{\xi^n}, u^r_{\xi^n})$ to get a left middle state $u^{r,l}_{\xi^n,*}$ and use it to replace the $u^r_{\xi^n}$ in computing the speed of the shock. In step 4), if the datum of the numerical solution at the point x_{j_1} needs to be updated, we solve a Riemann problem $RP(u^{n+1}_{j_1}, u^{n+1,+}_{j_1})$ to get a left middle state $u^{n+1,l}_{j_1,*}$, and updated the $u^{n+1}_{j_1}$ by it. In doing so, the critical cell only keeps the information that will be catched by the shock and let off the information that will be propagated to the right. The extrapolated data from the left to the right in an RSCC and the extrapolated data on two

sides in a CDCC are replaced by the corresponding middle states in the same way as it is done for an LSCC. The following theorem provides a theoretical basis for such a handling with replaced data.

Theorem 5.1 *Assume that u(x,t) is a solution to (5.1) with a single discontinuity, on which the Hugoniot condition*

$$s(u_r(x_0,t) - u_l(x_0,t)) = f(u_r(x_0,t)) - f(u_l(x_0,t)) \tag{5.3}$$

is satisfied, where s is the speed of the discontinuity, x_0 is the position of the discontinuity at the time t, and $u_l(x,t)$ and $u_r(x,t)$ are the solution on the left and right sides of the discontinuity. Also assume that $u_l(x,t)$ and $u_r(x,t)$ are continuous with continuous $(n+1)$th derivates on the two sides. We extend each of $u_l(x,t)$ and $u_r(x,t)$ to their opposite sides by an nth order extrapolation and still denote them by $u_l(x,t)$ and $u_r(x,t)$ respectively. Then the following relation

$$\begin{aligned} &s(u_r(x_0+dx,t) - u_l(x_0+dx,t)) \\ &= f(u_r(x_0+dx,t)) - f(u_l(x_0+dx,t)) + O(dx^{n+1}) \end{aligned} \tag{5.4}$$

holds for any small dx.

Proof. Differentiate (5.1a) with respect to x, we get

$$u_{xt} + f(u)_{xx} = 0. \tag{5.5}$$

Since the solution on each side of the discontinuity is smooth, a jump condition concerning its x derivative at the discontinuity can be obtained in the same way as Hugoniot condition is obtained (refer to [3]), which is

$$s(\frac{\partial u_r}{\partial x}|_{x=x_0} - \frac{\partial u_l}{\partial x}|_{x=x_0}) = \frac{\partial f(u_r)}{\partial x}|_{x=x_0} - \frac{\partial f(u_l)}{\partial x}|_{x=x_0}. \tag{5.6}$$

By the same argument, we can obtain for any $0 \le i \le n$

$$s(\frac{\partial^i u_r}{\partial x^i}|_{x=x_0} - \frac{\partial^i u_l}{\partial x^i}|_{x=x_0}) = \frac{\partial^i f(u_r)}{\partial x^i}|_{x=x_0} - \frac{\partial^i f(u_l)}{\partial x^i}|_{x=x_0} \tag{5.7}$$

iteratively. Multiply (5.7) by dx^i, sum them over all the i for $0 \le i \le n$, by taking into account the Taylor expension of $u_l(x,t)$ and $u_r(x,t)$, we get (5.4). Thus ends the proof.

Theorem 5.1 indicates that the middle states obtained by solving Riemann problems on each side of the discontinuity are identical with the original extrapolated data on the same side to a high order truncation error if the solution on each side is smooth enough, from which the consistency of the numerical flux involving middle states with the flux function $f(u)$ follows. However, when the solution contains weak discontinuities (such as edges of rarefaction waves) associated with other characteristic fields, using the original extrapolated data and using the corresponding middle states will have different senses as we discussed before, the former may miss the information crossing the discontinuity while the later will not.

The treatment also can be analysed in terms of p^n and q^n as we did in §3 to show the conservative feature, and the conclusion of Theorem 3.1 is still true, only p^n and q^n are vectors and all the identities in the corresponding proof are in vector forms.

Now we discuss the treatment for interactions of discontinuities and still focus on the case involving two critical cells. When the two critical cells are close to each other, they are handled in the same way as we described in §4 for the scalar case before they are merged. However, the handling of their mergence is a little complicated than the scalar case since there are three different kinds of discontinuities. Assume that the two critical cells overlap in a cell $[x_{j_1}, x_{j_1+1}]$ with the two positions of the discontinuities satisfying $\xi_1^n > \xi_2^n$ and a hidden middle state u_* between them. The treatment for their mergence is described as follows:

1) Solve a Riemann problem $RP(u_{j_1}^n, u_{j_1+1}^n)$ and obtain two middle states u_*^l and u_*^r. The waves that connect $u_{j_1}^n$ and u_*^l or u_*^r and $u_{j_1+1}^n$ can be either shocks or rarefaction waves; 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 $u_{j_1+1}^n - u_*^r$, such as

$$Q^n = \bar{\xi}_1^n(u_*^l - u_{j_1}^n) + \bar{\xi}_2^n(u_*^r - u_*^l) + \bar{\xi}_3^n(u_{j_1+1}^n - u_*^r) \tag{5.8}$$

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

3) Define a CDCC on the cell $[x_{j_1}, x_{j_1+1}]$ with $\bar{\xi}_2^n$ as its position of the contact discontinuity. When the wave connecting $u_{j_1}^n$ and u_*^l is a shock, define an LSCC with $\bar{\xi}_1^n$ as its position of the shock, which and the CDCC defined before overlap 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 defined, meanwhile $u_{j_1}^n$ should be updated by $u_{j_1}^n - \bar{\xi}_1^n(u_*^l - u_{j_1}^n)$. The case for the wave connecting u_*^r and $u_{j_1+1}^n$ is handled similarly.

We update $u_{j_1}^n$ when the connecting wave is a rarefaction wave, in order to keep the conservation feature for the treatment, which will be shown in the later discussion. A geometrical interpretation of this handling is that the updated $u_{j_1}^n$ represents where the rarefaction wave is centered and how much it has been spanned. Denote by $\bar{s}_k^n$ the position of the k-wave on the level n if it is a discontinuity. A theorem concerning the consistency of the treatment , which corresponds to Theorem 4.1, is follows:

Theorem 5.2 *Assume that ξ_1^{n-1} and ξ_2^{n-1} are exact on the level $n-1$ and the k-wave in the Riemann problem is a discontinuity, then its position of the discontinuity computed in the treatment satisfies*

$$\bar{\xi}_k^n = \bar{s}_k^n + O(h^2) \tag{5.9}$$

if both the underlying scheme and the treatment are at least of the first order.

Proof. By the same argument in the proof of Theorem 4.1, we only need to prove that (5.9) is exact for piecewise constant solutions. We denote by $(x)_k$ the kth term in the linear representation of a vector x by vectors $u_*^l - u_l$, $u_*^r - u_*^l$, and $u_r - u_*^r$, and adopt the notations in the proof of Theorem 4.1 and the Figure 4.3, except that the ξ^n should be replaced by $\bar{\xi}_k^n$. Still we have

$$s^n = \xi^* + (t^n - t^*)s_0, \tag{5.10a}$$

$$\xi_1^n = \xi^* + (t^n - t^*)s_1, \tag{5.10b}$$

$$\xi_2^n = \xi^* + (t^n - t^*)s_2, \tag{5.10c}$$

where

$$s_0(u_r - u_l)_k = (f(u_r) - f(u_l))_k, \tag{5.11a}$$

$$s_1(u_* - u_l) = f(u_*) - f(u_l), \tag{5.11b}$$

$$s_2(u_r - u_l) = f(u_r) - f(u_*), \tag{5.11c}$$

Represent $s_1(u_* - u_l) + s_2(u_r - u_*)$ as a linear combination of the vectors $u_*^l - u_l$, $u_*^r - u_*^l$, and $u_r - u_*^r$. It is easy to obtain that s_0 is the coefficient of $(u_r - u_l)_k$ in the representation; therefore, the conclusion follows from (5.10).

A theorem concerning the conservation feature of the treatment, which corresponding to Theorem 4.2 is follows:

Theorem 5.3 *If both the underlying scheme and the treatment are of the first order, then if the k-wave is a discontinuity, (3.7) holds for the corresponding new-formed critical cell on the level n with* $(q^n_{j_1,1} + q^n_{j_2,2})_k$ *standing for* $q^n_{j_1}$ *and* $\bar{\xi}^n_k$ *for* ξ^n. *otherwise, the change of the numerical solution in treating the rarefaction wave is* $(q^n_{j_1,1} + q^n_{j_1,2})_k + O(1)$. *If both the underlying scheme and the treatment are of the second order, then if the k-wave is a discontinuity,(3.8) holds for the corresponding new-formed critical cell with* $(q^n_{j_1,1} + q^n_{j_1,2})_k$ *standing for* $q^n_{j_1}$ *and* $\bar{\xi}^n_k$ *for* ξ^n, *otherwise, the change of the numerical solution in treating the rarefaction wave is* $(q^n_{j_1,1} + q^n_{j_1,2})_k + O(h)$

Proof. According to the discussion before, (4.8) and (4.9) for the first order case and (4.10) and (4.11) for the second order case are still true. Therefore, by noticing the way $\bar{\xi}^n_k$ is computed, the conclusion for the discontinuity case follows almost along the same proof for Theorem 4.2. The conclusion for the rarefaction wave follows easily from the way the numerical solution is updated and (4.8) and (4.9) or (4.10) and (4.11).

Due to the above theorem, (3.9) is still true by the same argument in §4, only it is in a vector form now. Finally, we have the following theorem as a summary of this section.

Theorem 5.4 *The conclusion of Theorem 3.2 is still true for the Euler system case when the solution is piecewise smooth with a finite number of interactions of discontinuities.*

6. Numerical examples

The underlying scheme used for the numerical experiments in this section is a second order TVD scheme described in [13] with a Runge-Kutta type time discretization. Since the Runge-Kutta procedure can be described as a two step procedure, the predicator-corrector procedure, the underlying scheme also can be written as two step scheme, the predictor step

$$u^{n+1/2}_j = u^n_j - \lambda(\hat{f}^n_{j+1/2} - \hat{f}^n_{j-1/2}), \tag{6.1a}$$

and the corrector step

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

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

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

The treatment is incorporated into the underlying scheme in the following way:

1) The treatment should perform in each of the predictor and corrector steps, however, only the predictor step moves critical cells to their adjacent cells according to their positions of discontinuities. The corrector step does allow critical cells to move, even though their positions of the discontinuities are little bit out of the critical cells.

2) For a critical cell, the predictor step gives a position of the discontinuity $\xi^{n+1/2}$, and the corrector step gives a position of the discontinuity $\bar{\xi}^{n+1}$. The position of the discontinuity on the new level is finally given by

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

3) When a critical cell moves to its adjacent cell in the predictor step, at the point at which the numerical solution is updated, u^n_j in the second formula in (6.1b) should be replaced accordingly by the extrapolated datum 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 $[x_{j_1-1}, x_{j_1}]$ in the predictor step, $u^n_{j_1}$ in the second formula in (6.1b) should be replaced by $u^{n,+}_{j_1}$ in the scalar case or $u^{n,l}_{j_1,*}$ in the system case.

EXAMPLE 1. This is an example for the linear scalar case.

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

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

and

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

The exact solution to this problem contains three discontinuities and a week discontinuity (a discontinuity of the first derivative). The second order treatment is applied to the contact discontinuities as well as to the weak discontinuity by a considering that taking information only from the same side will also sharpen the corner of the weak discontinuity. The mesh ratio λ is 0.5, and a TVB modification introduced by Shu in [14] is made on the underlying scheme to improve the computation near extremum points. Figure 6.1-a. and 6.1-b. present the numerical results with $h = \frac{1}{30}$ (60 grid points) at $t = 2$ and $t = 8$ respectively. Figure 6.2-a. and Figure 6.2-b. present the numerical results with $h = \frac{1}{60}$ (120 grid points) at $t = 2$ and $t = 8$ respectively. Comparing with the results without treatment, which we didn't present here, Figure 6.1 and 6.2 show a quite improvement in discontinuity transition and overall resolution.

EXAMPLE 2. This is an example for the Euler system with initial conditions:

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

and solid wall boundary conditions, where

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

λ is still 0.5 and h is 0.005 (two hundred grid points). The second order treatment is used. The numerical results at $t = 0.026$ and $t = 0.038$ are presented in Figure 6.3 and Figure 6.4 respectively, where the solid lines are the numerical solution obtained by an ENO scheme with 800 grid points and are reguarded 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 can not span the rarefaction; therefore, the following 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 introduced a treatment for discontinuities, which is a new shock tracking technique, however, without a lower dimensional grid to resolve the discontinuities. The main idea of the treatment is that the computation on each side draws information that jumps at the discontinuity only from the same side. A numerical method for ordinary differential equation is incorporated into the algorithm to compute the locations of the discontinuity. The analysis shows that this is equivalent to record the conservation error and compensate it in the following timesteps; therefore, the treatment still has a conservation feature. Numerical examples shows that the treatment works quite well in both one dimensional scalar and system cases. A paper about the extension of the treatment to the two dimensional case is in preparation. A further development of the treatment is how to use it to capture a spontaneous shock; in fact, some of this work has been done in [11].

Acknowelegememts

We thank professor Ami Harten 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(1979).

[4] B. Engquist and B. Sjogreen, *Numerical approximation of hyperbolic conservation laws with stiff terms.* UCLA CAM Report 89-07.

[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] A. Harten, *Preliminary results on the extension of ENO schemes to two-dimensional problems.* In Proceedings of the International Conference of Hyperbolic Problems, Siant-Etienne, January, 1986.

[8] A. Harten and S. Osher, SIAM J. Numer. Anal. **24** 279(1987)

[9] A. Harten, B. Engquist, S. Osher and S. Chakravarthy, J. Comput. Phys. **71** 231(1987).

[10] P. Lax and B. Wendroff, Commun. Pure Appl. Math. **13** 217(1960).

[11] D. Mao, *A treatment of discontinuities in shock-capturing finite difference methods.* to appear in J. Comput. Phys.

[12] D. Mao, J. Comput. Math. **No. 3**, 256(1985) (in Chinese).

[13] 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, 229(1986).

[14] Chi-Wang Shu, Math. Comp. **49**, 105(1987).

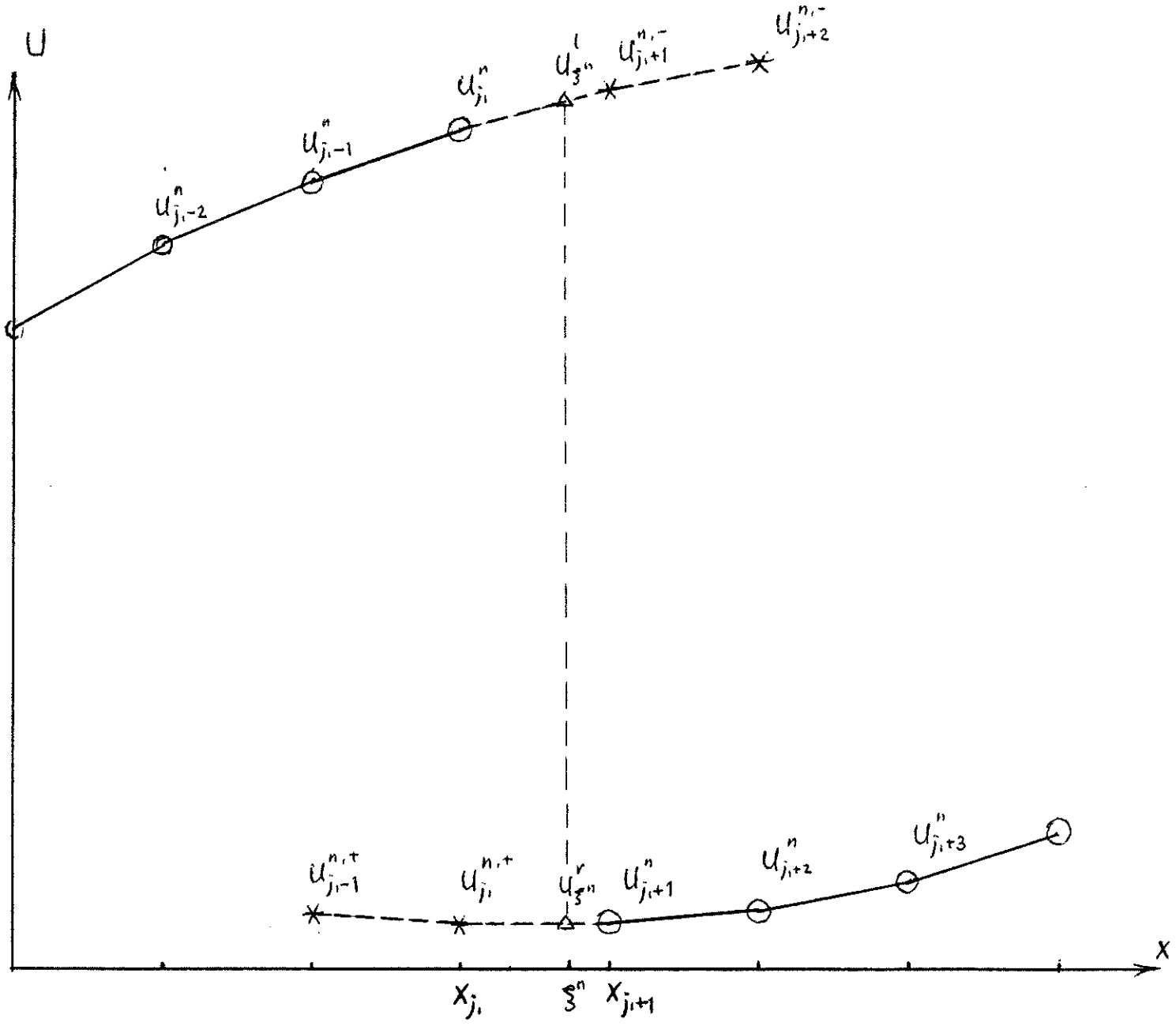

U
$u^{n,-}_{j_1+2}$
$u^{l}_{\xi^n}$
$u^{n,-}_{j_1+1}$
$u^{n}_{j_1}$
$u^{n}_{j_1-1}$
$u^{n}_{j_1-2}$
$u^{n,+}_{j_1-1}$
$u^{n,+}_{j_1}$
$u^{r}_{\xi^n}$
$u^{n}_{j_1+1}$
$u^{n}_{j_1+2}$
$u^{n}_{j_1+3}$
x
x_{j_1}
ξ^n
x_{j_1+1}

Figure 2.1

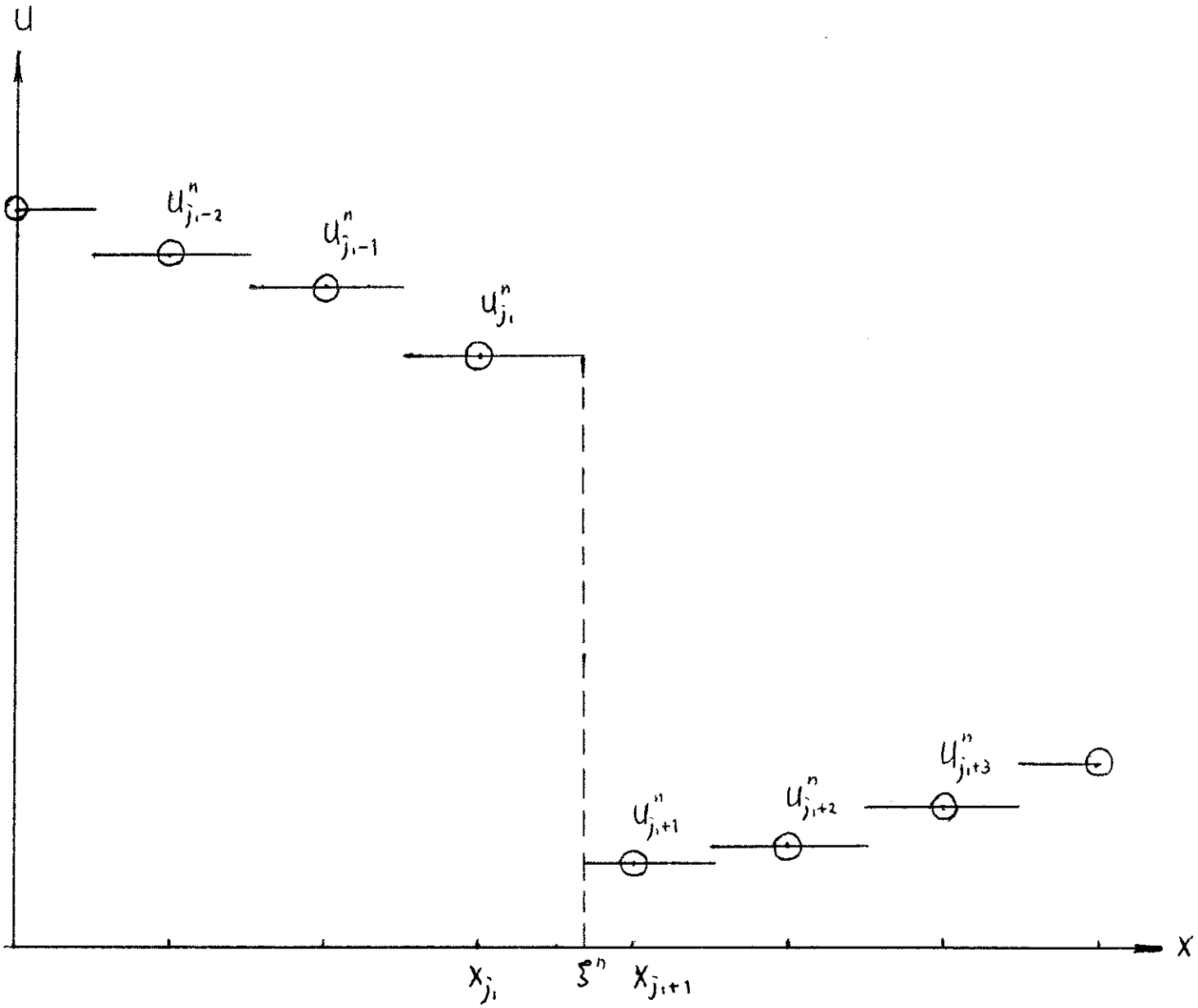
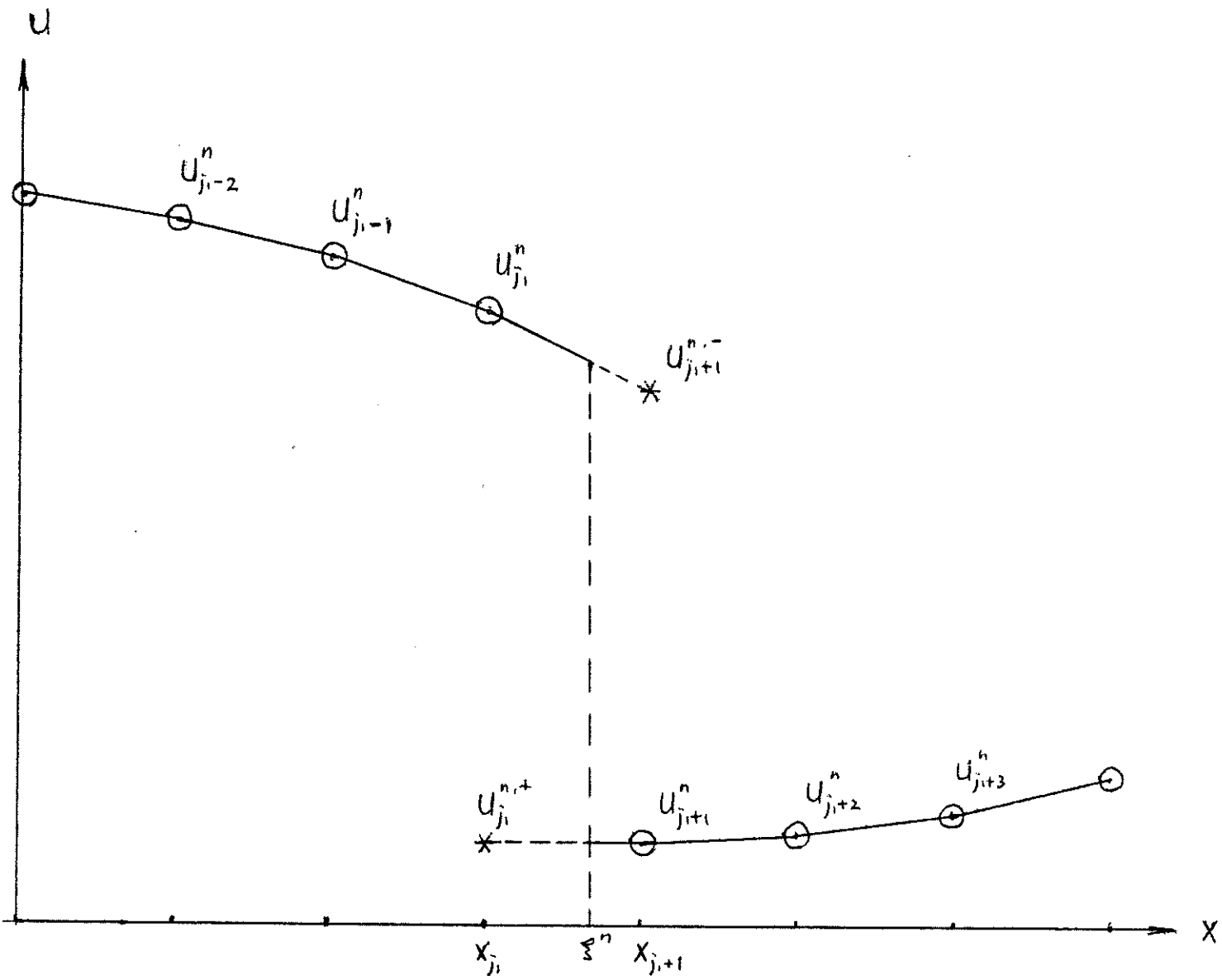


Figure 3.1



$U^n(x)$

Figure 3.2

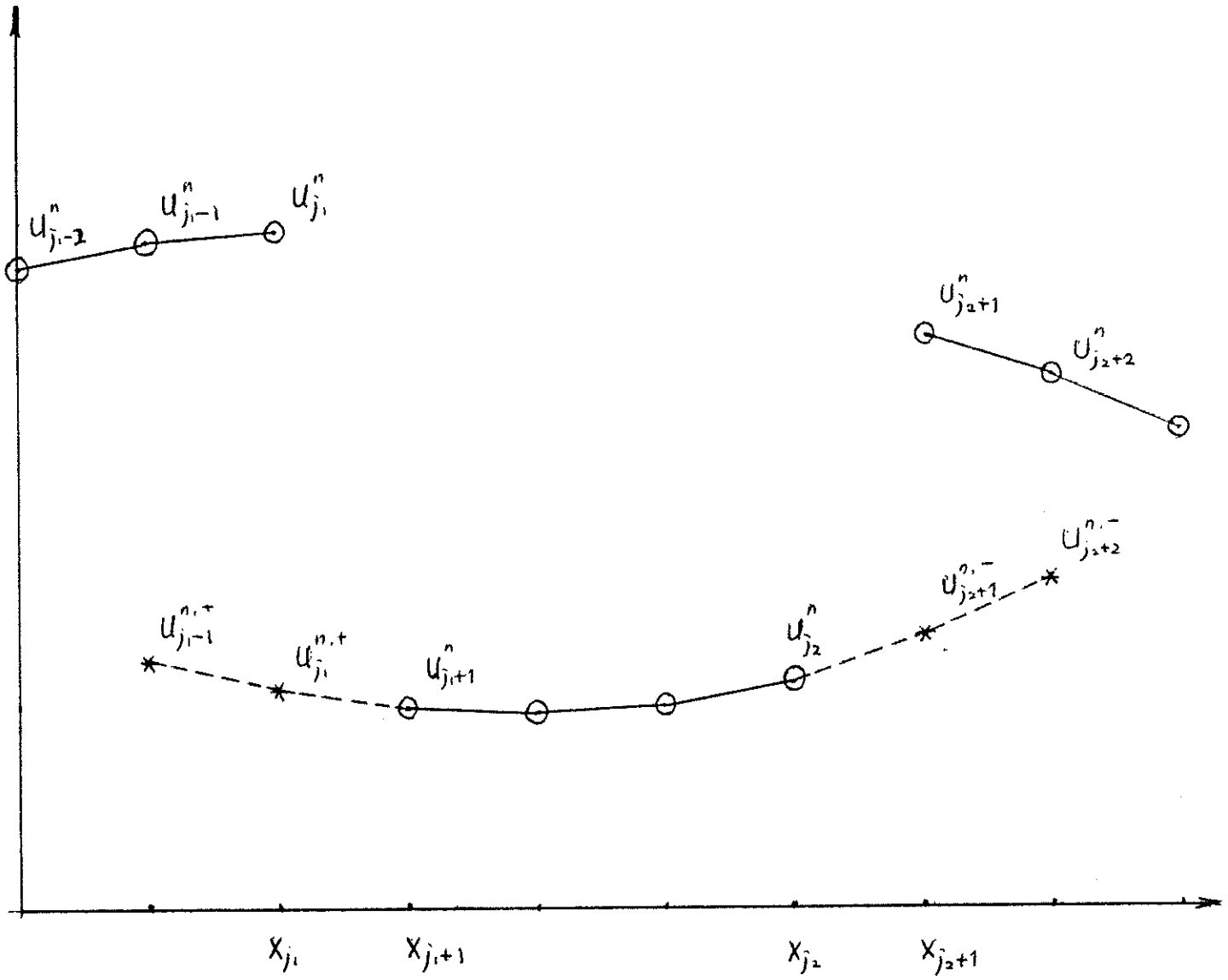
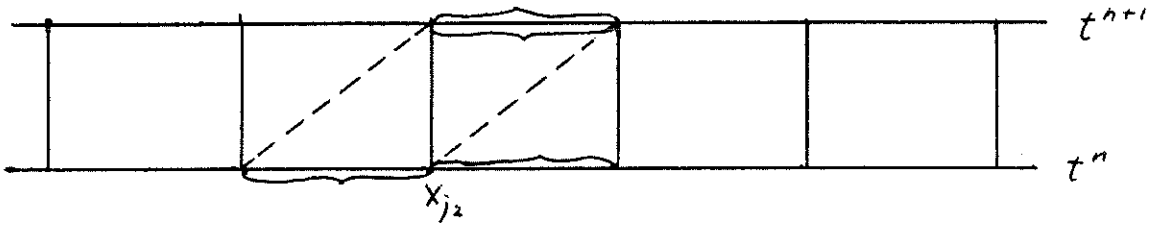
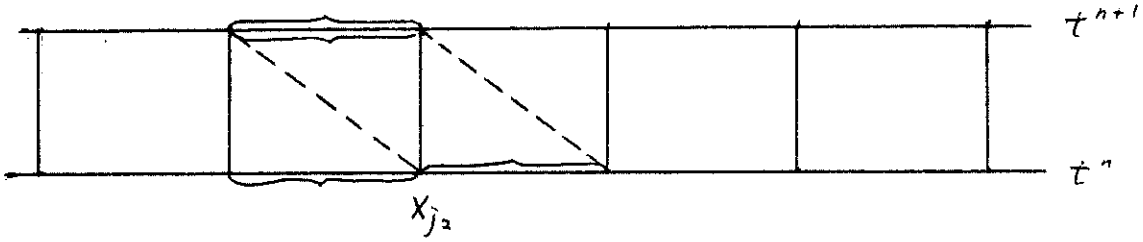

$U^n_{j_1-2}$
$U^n_{j_1-1}$
$U^n_{j_1}$
$U^n_{j_2+1}$
$U^n_{j_2+2}$
$U^{n,-}_{j_2+2}$
$U^{n,-}_{j_2+1}$
$U^{n,+}_{j_1-1}$
$U^{n,+}_{j_1}$
$U^n_{j_1+1}$
$U^n_{j_2}$
x_{j_1}
x_{j_1+1}
x_{j_2}
x_{j_2+1}

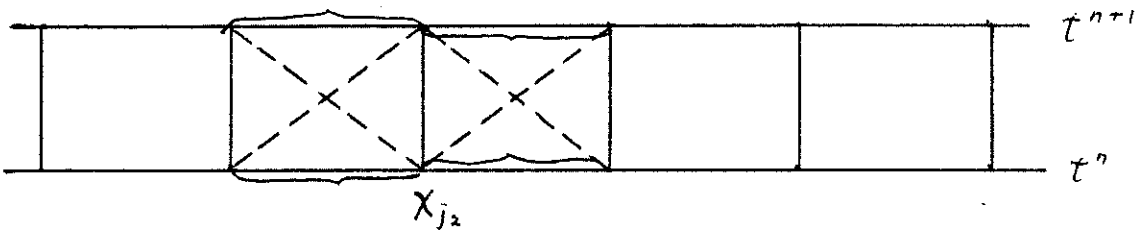
Figure 4.1



(a)



(b)



(c)

Figure 4.2

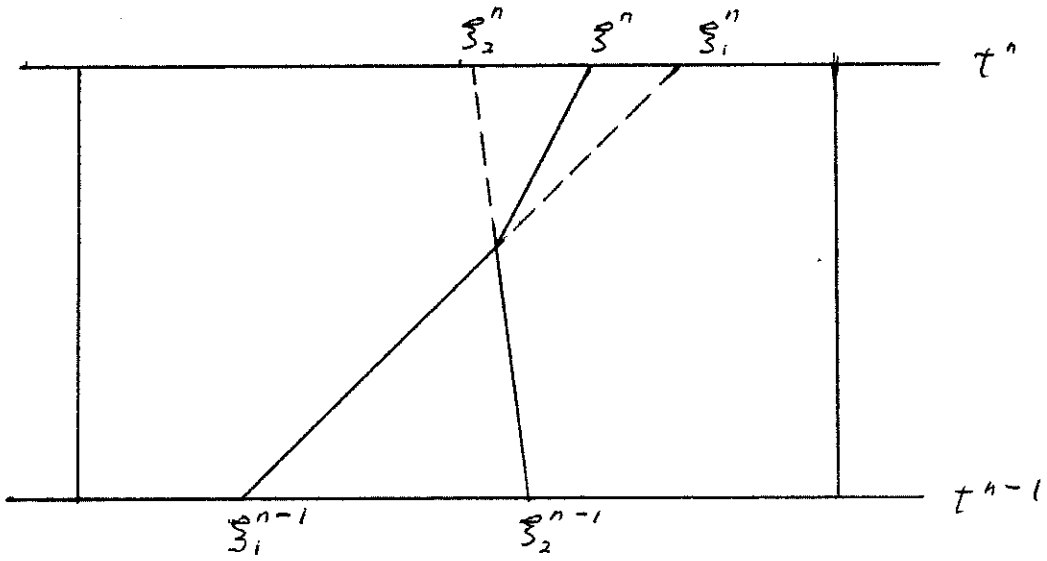


Figure 4.3

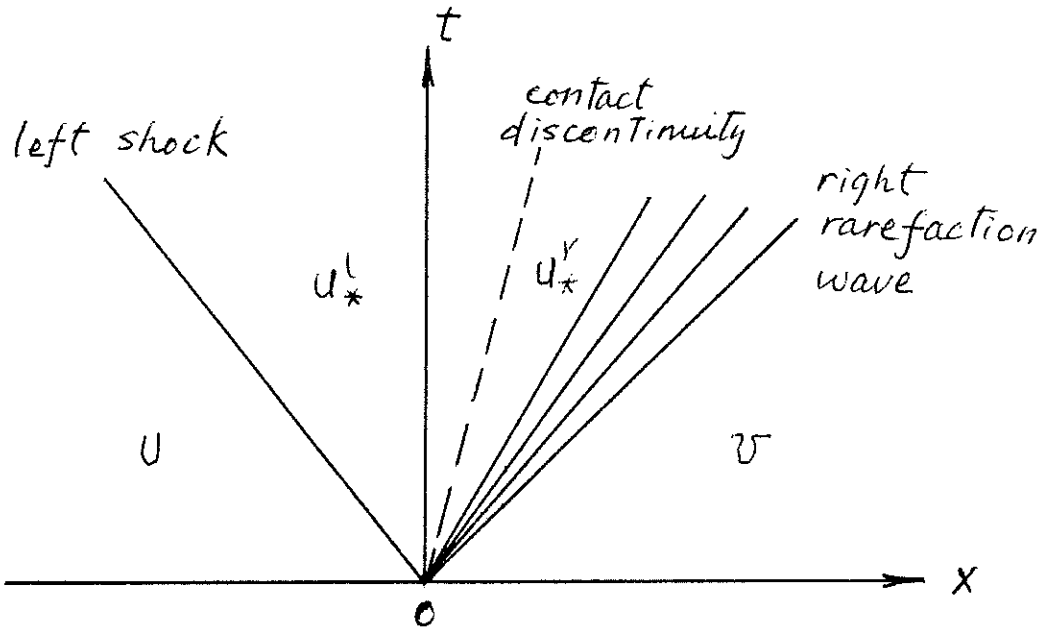

t
left shock
contact
discontinuity
right
rarefaction
wave
u_*^l
u_*^r
u
v
x
0

Figure 5.1

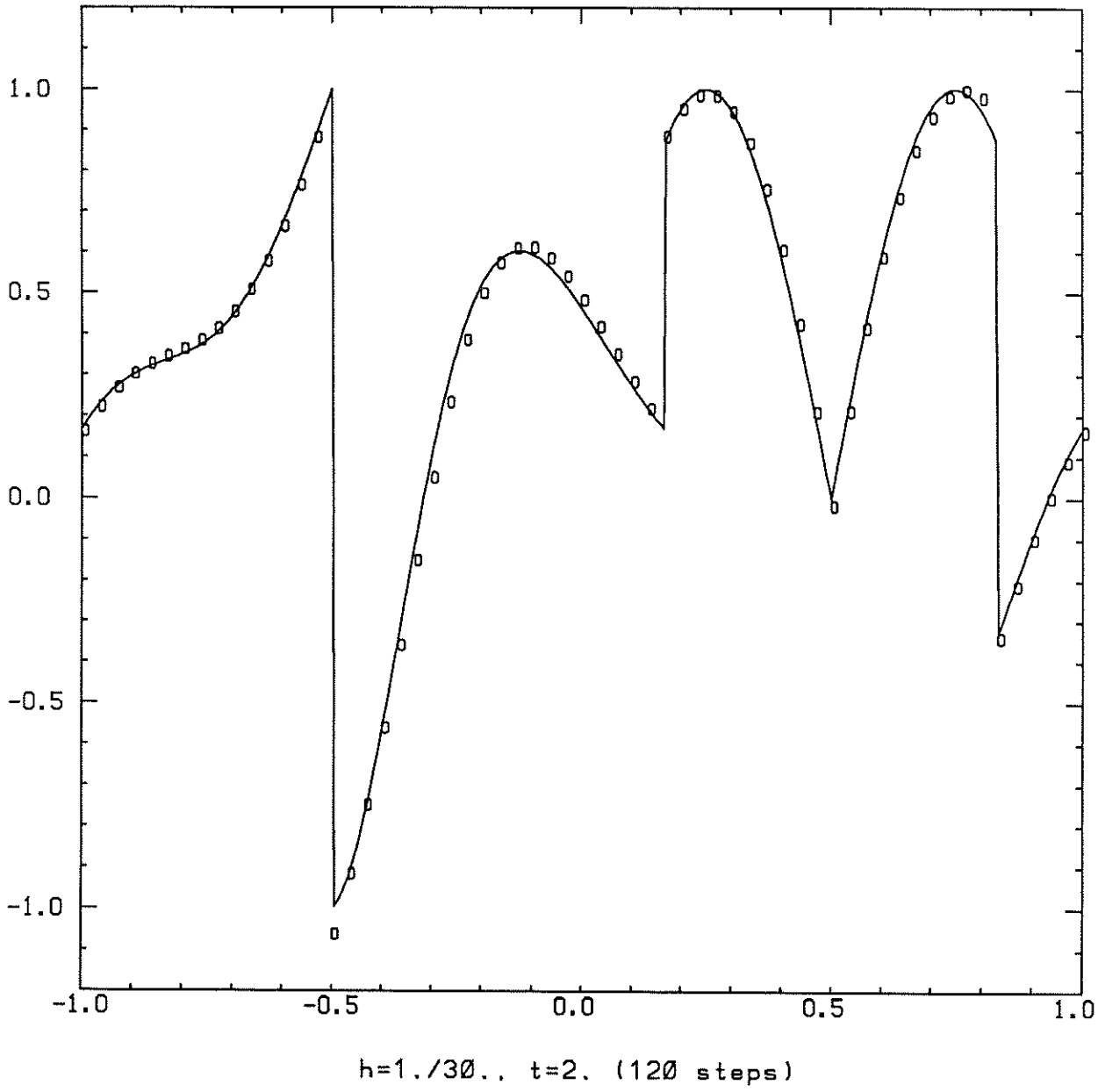


Figure 6.1 - a.

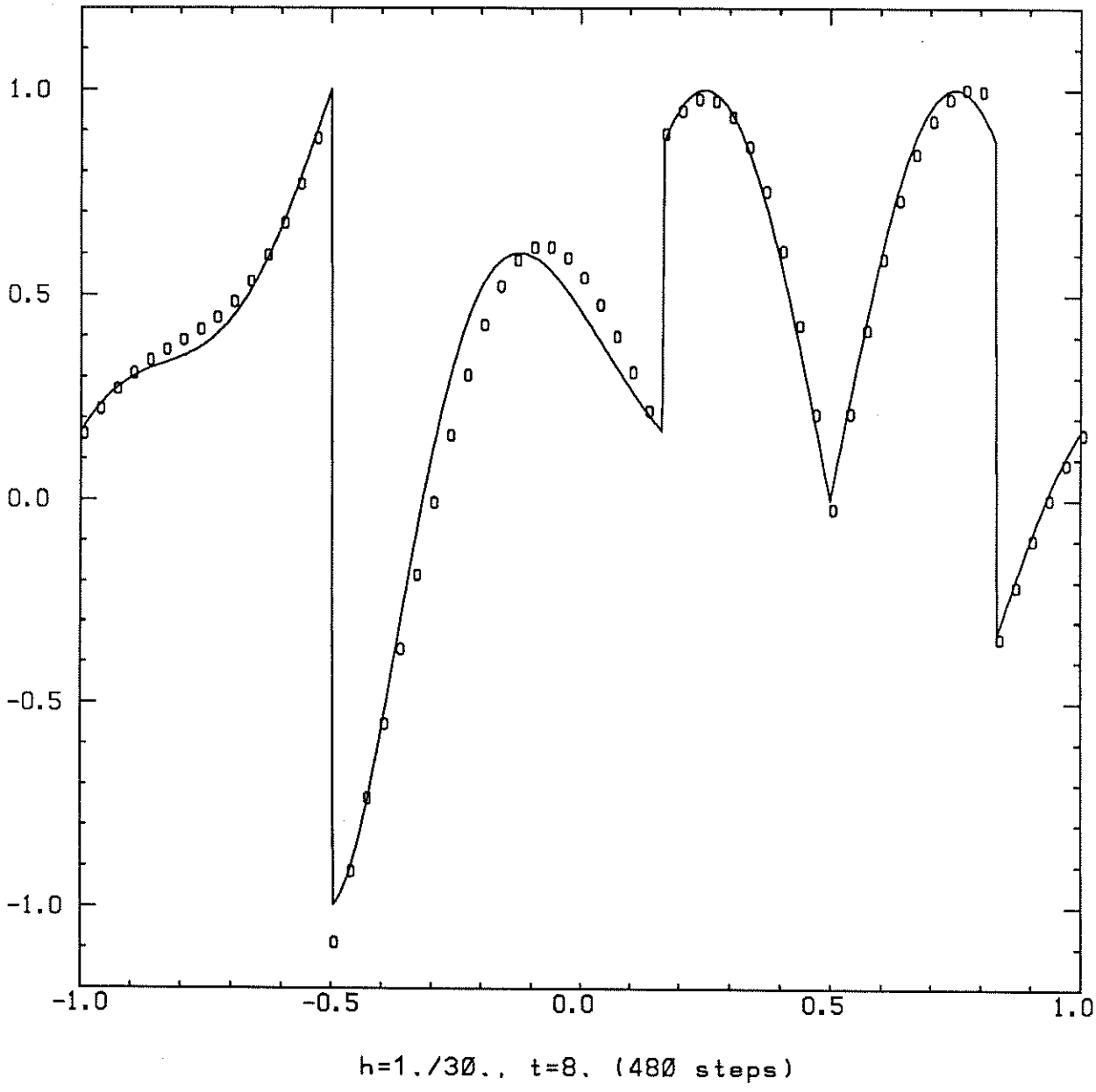


Figure 6.1 - b.

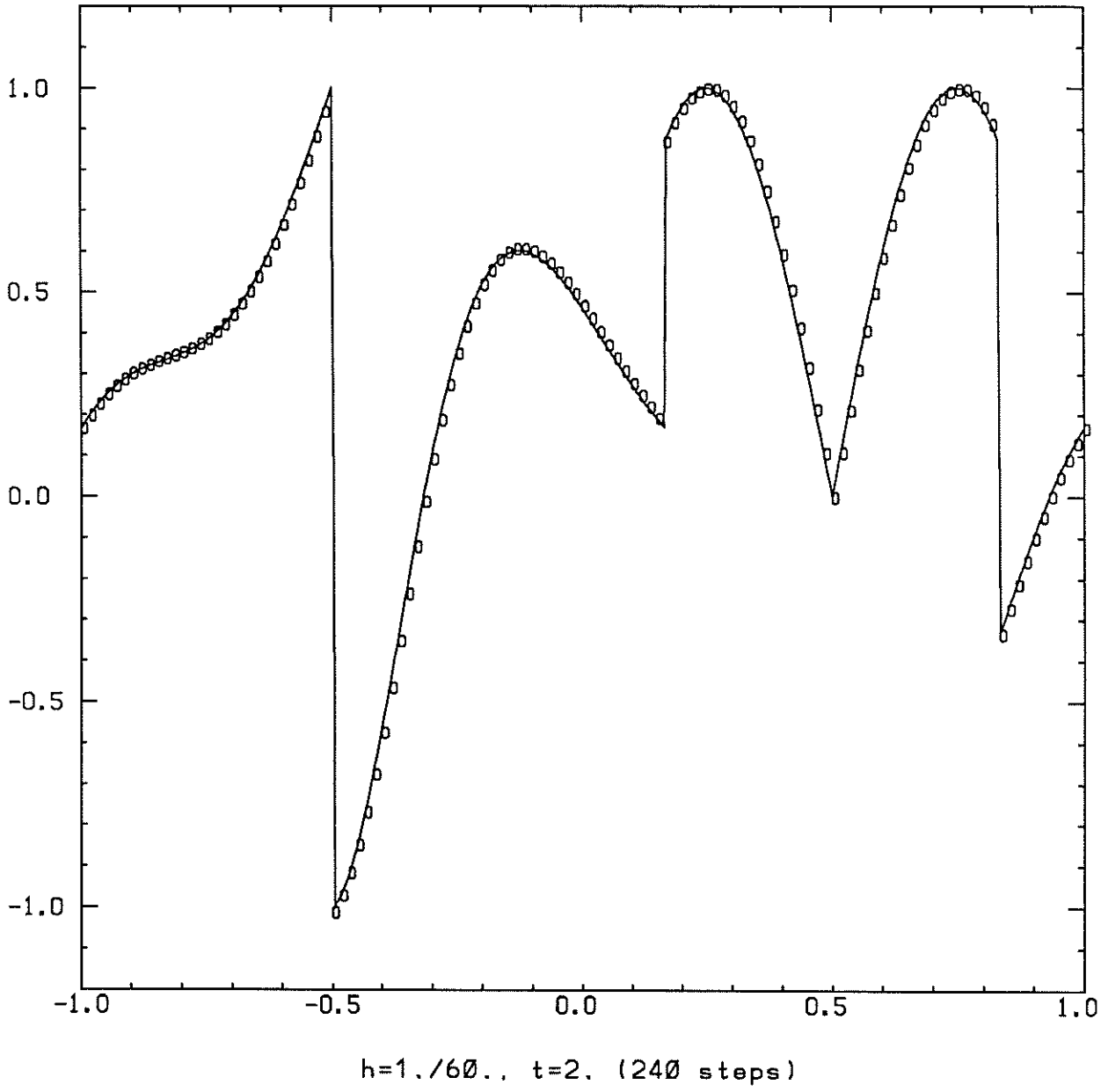


Figure 6.2 - a.

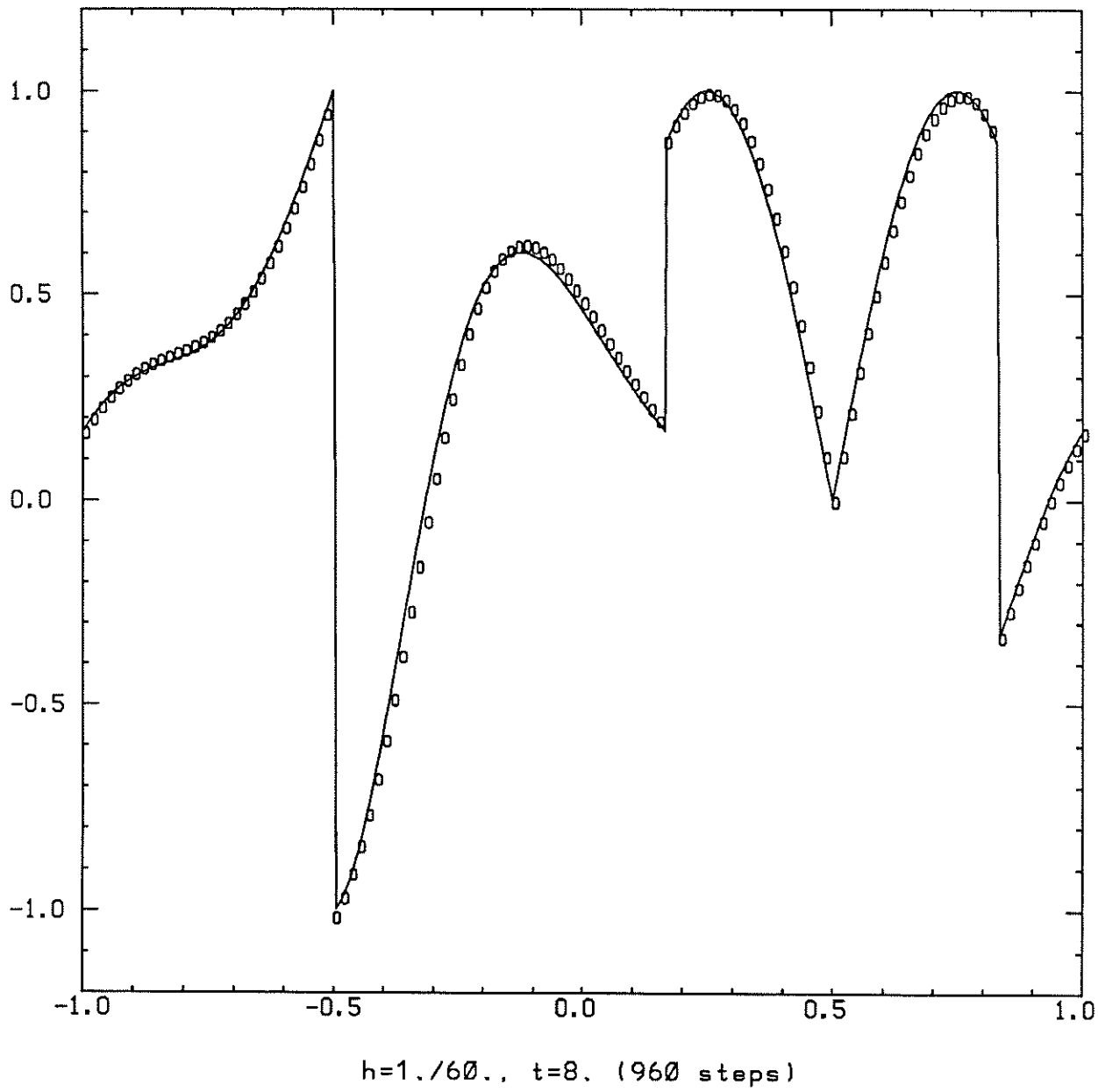


Figure 6.2 - b.

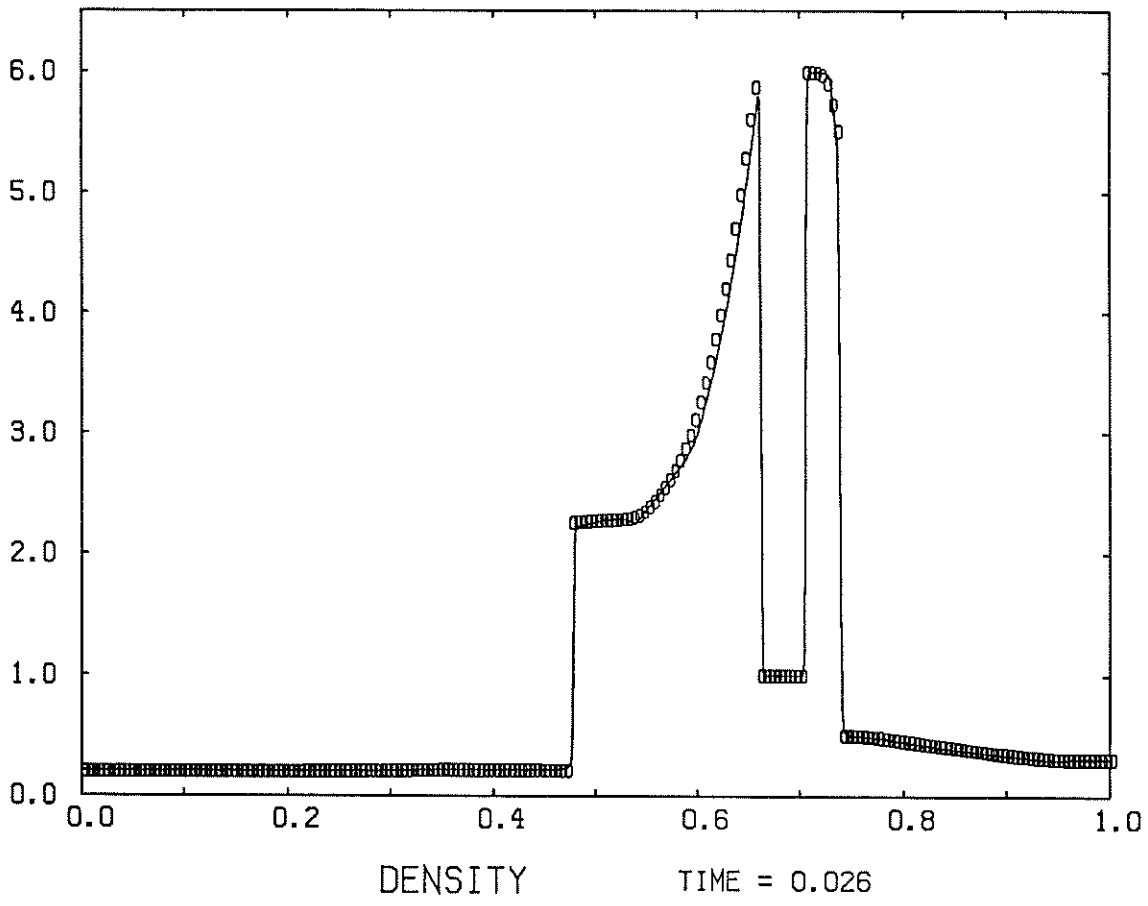


Figure 6.3 - a.

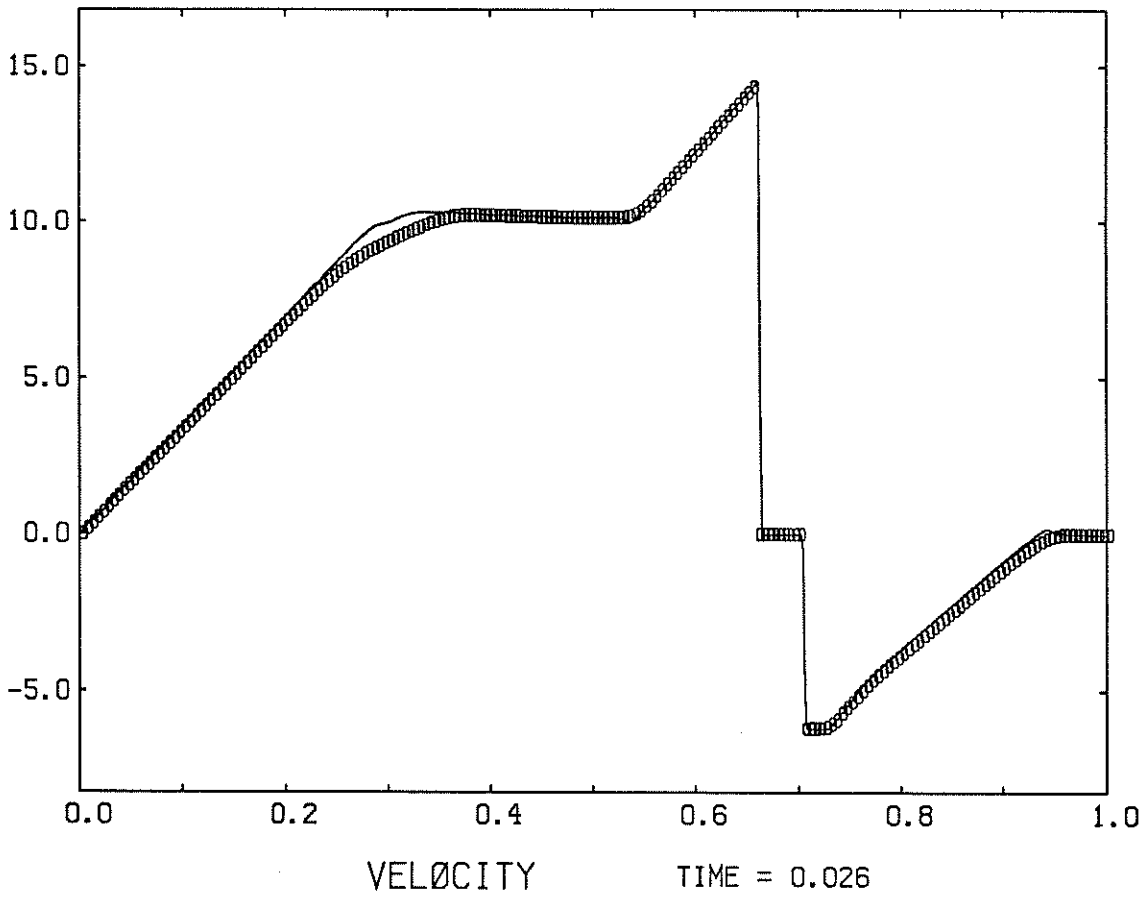


Figure 6.3 - b.

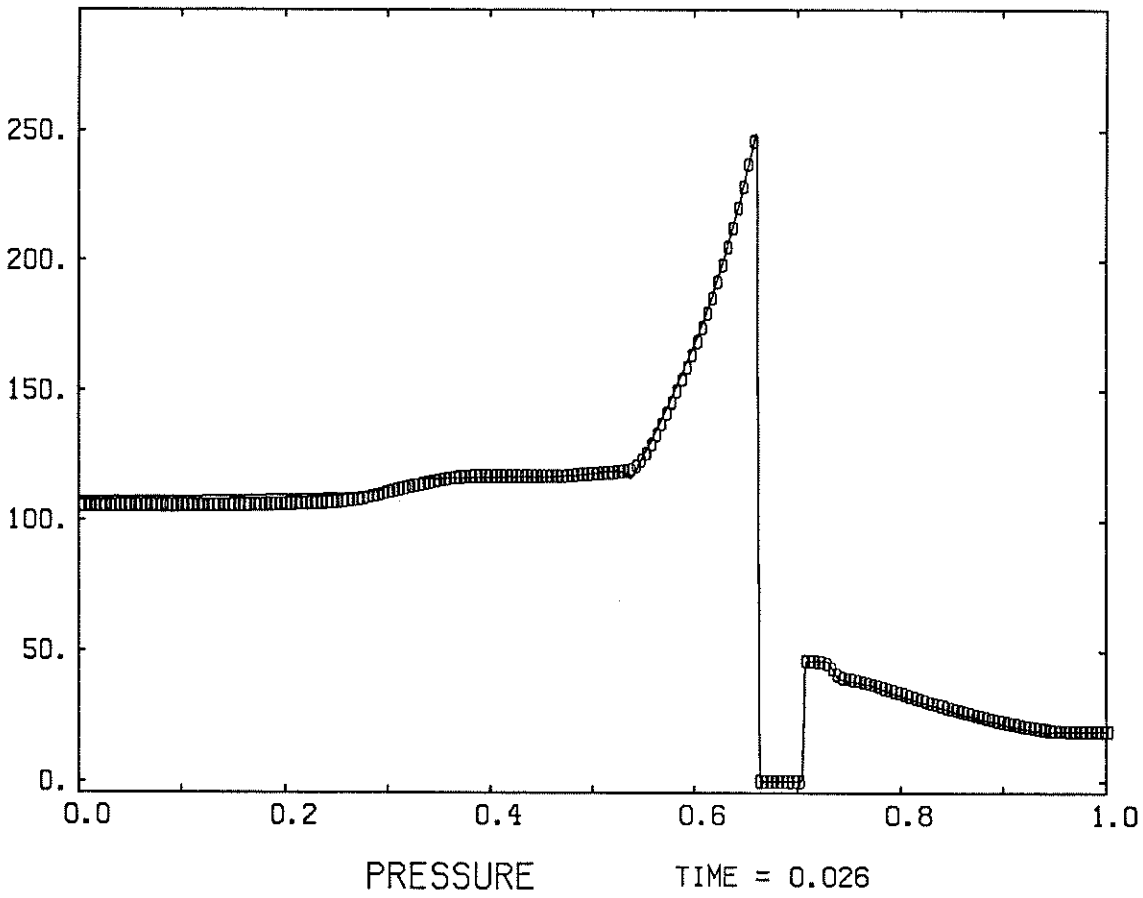


Figure 6.3 - c.

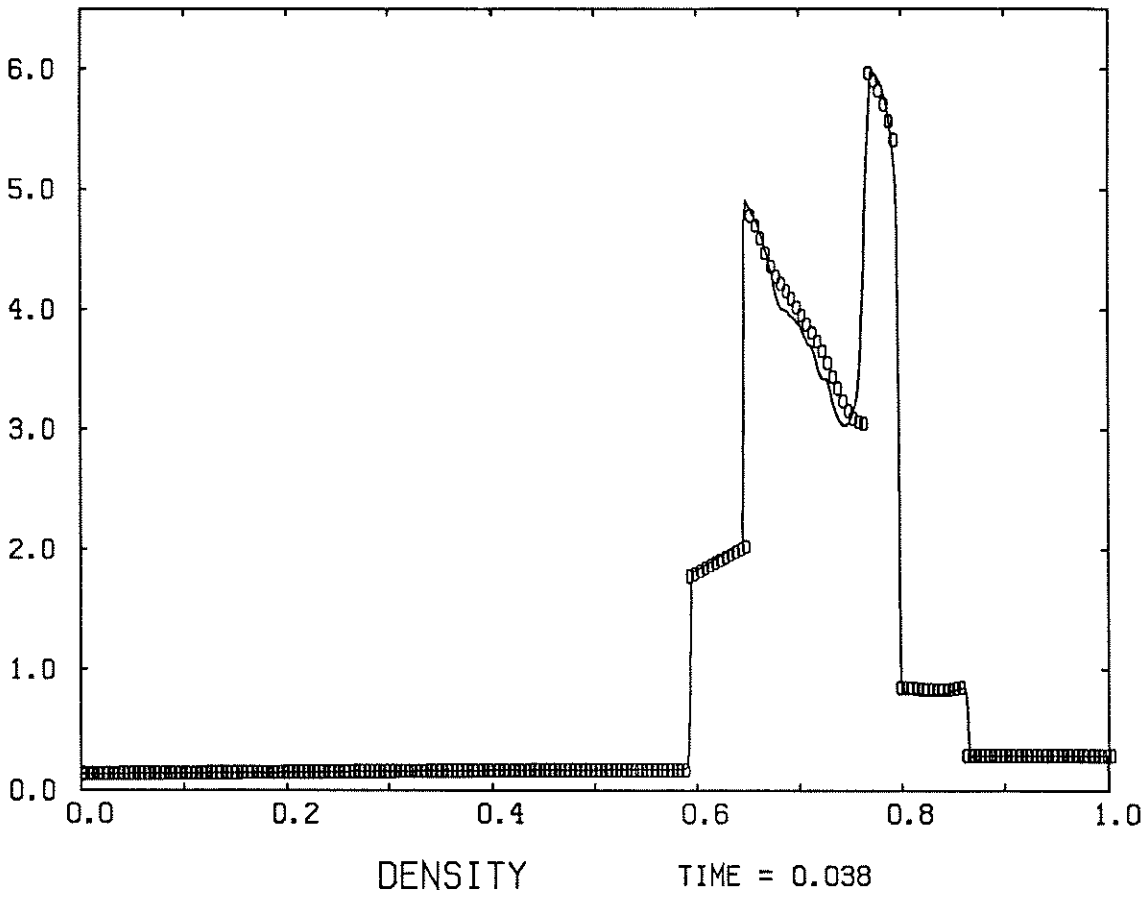


Figure 6.4 - a.

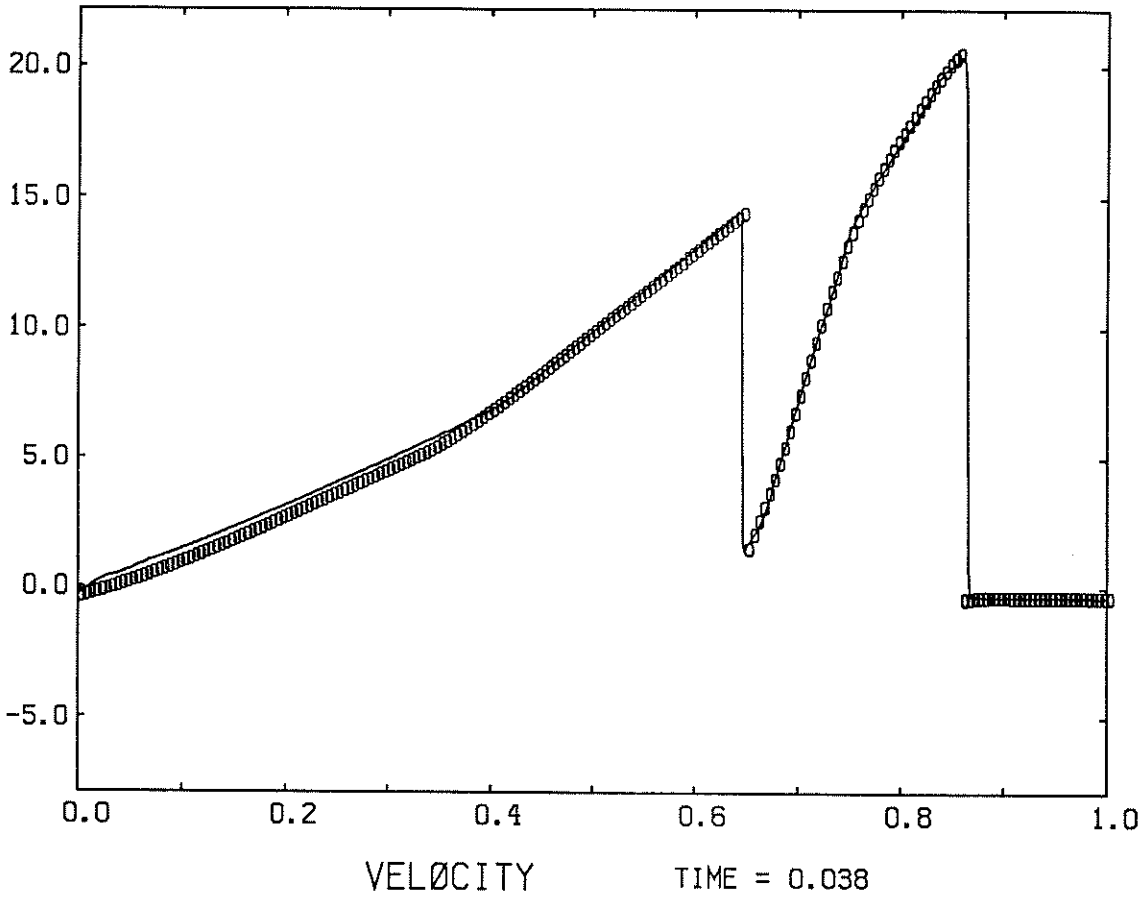


Figure 6.4 - b.

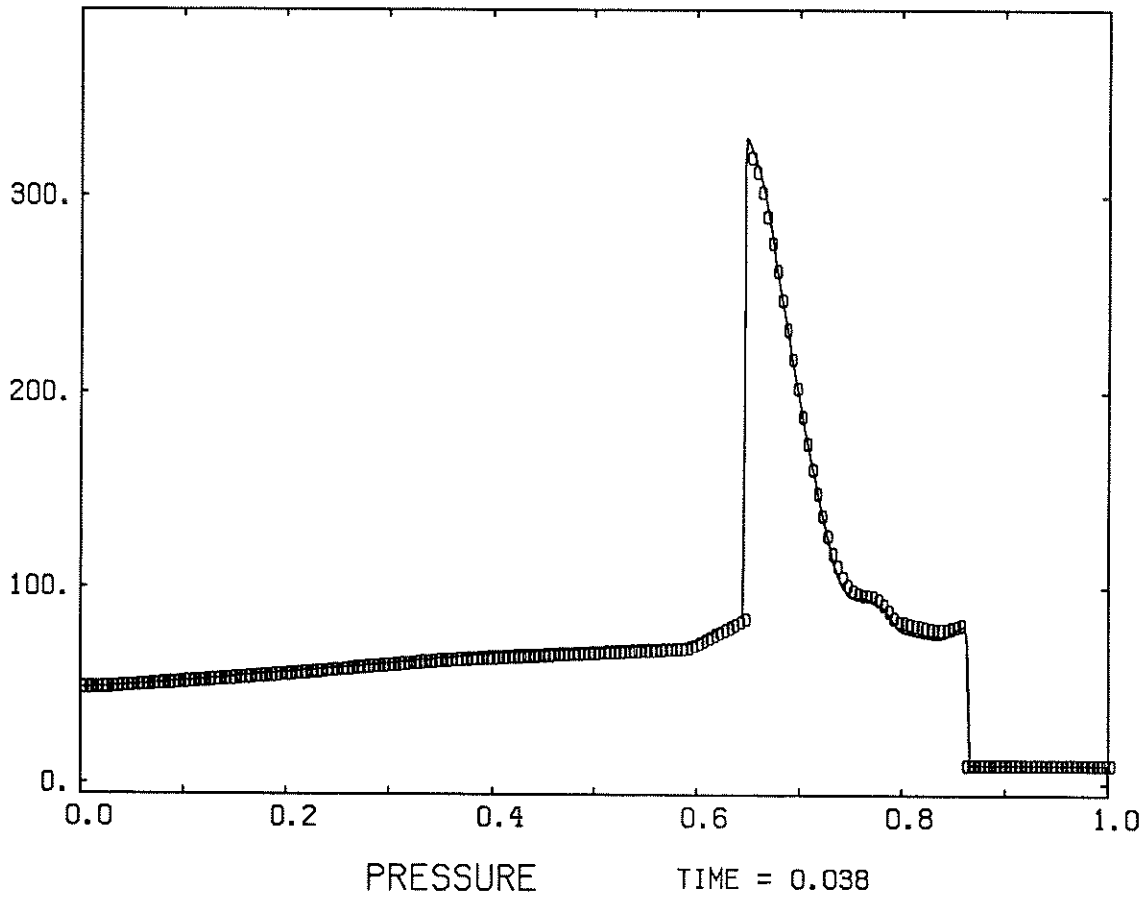


Figure 6.4 - c.