A NUMERICAL METHOD FOR FIRST ORDER NONLINEAR SCALAR HYPERBOLIC CONSERVATION LAWS IN ONE DIMENSION

Helge Holden[1], Lars Holden[2,*] and Raphael Høegh-Krohn[3]

1 Matematisk institutt, Universitetet i Trondheim-NTH, 7034 Trondheim-NTH

2 Norwegian Computing Center, Blindern, Oslo 3, Norway

3 Matematisk Institutt, Universitetet i Oslo, Blindern, Oslo 3, Norway

Abstract

A numerical method for first order nonlinear scalar hyperbolic conservation laws in one dimension is presented, using an idea by Dafermos. In this paper it is proved that it may be used as a numerical method for a general flux function and a general initial value. It is possible to give explicit error estimates for the numerical method. The error in the method is far smaller than in any other method. The numerical method is illustrated in an example.

* Supported by the Royal Norwegian Council for Technical and Industrial Research.

Supported by the VISTA program for fundamental research and the Norwegian Science Research Council NAVF.

1 INTRODUCTION

In this paper we suggest a numerical method for the Cauchy problem

$$(1.1) \qquad u_t + f(u)_x = 0$$

where f is locally Lipschitz continuous, with initial condition

$$(1.2) \qquad u(x,0) = u_0(x)$$

assumed to be bounded and of locally bounded variation in R. Using the convergence of the method it is easy to prove existence, uniqueness and other well-known properties of the solution of the Cauchy problem.

It is well-known that there exist a unique weak solution of (1.1) and (1.2), see e.g. Oleinik [4] and [5], Vol'pert [9] and Kruzkov [2]. The usual approach to the equation is by finite differences. This was first done by Lax [3].

The method presented in this paper was first mentioned by Dafermos [1] as a new method to study the initial value problem. Using the method he proved some properties of the solution. He stated that it may be used as a numerical method for f convex or f concave. Unaware of Corollary 2.3 in this paper it is reasonable to believe that the numerical algorithm contains an infinite number of steps when f is not convex or concave. Lucier [6] proved that the method may be used as a numerical scheme for a special f and initial value and that the method has optimal convergence. LeVeque [4] used the method and for each timestep projected the solution back on a grid. Most other numerical methods for (1.1) and (1.2) are not optimal, see Lucier [5].

Dafermos' obervation is that with f piecewise linear and the initial value piecewise constant, then the solution is piecewise constant. There are no rarefaction waves, only shocks. The solution u(.,t) is piecewise constant for all t and the solution is found by solving Riemann problems and following the shocks.

The method is to approximate f by a piecewise linear function and the initial value by a piecewise constant function. Solving the perturbed problem, we get an approximation to the solution of the original problem.

This may be used as a numerical scheme if there is a finite number of constant states. Of course this is only possible if the initial value has a finite number of constant states. In this paper we prove that there is only a finite number of constant states even in infinite time. Therefore it is possible with a finite algorithm to approximate the solution in infinite time. Usually a numerical method is only finite in finite time. Thus we may describe the method as superfast.

In chapter 5 the numerical method is illustrated in a simple example.

2 THE NUMERICAL METHOD

We will first describe the well-known solution to the Riemann problem. In the following we assume f to be piecewise linear and the initial value to be piecewise constant since this is the only case needed in the paper. This is the simplest problem in the form (1.1) and (1.2).

$$u_0 = \begin{cases} u_- & \text{when} \quad x < 0 \\ u_+ & \text{when} \quad x > 0. \end{cases}$$

First we need some notation. Assume $u_- < u_+$. We define f_c to be the convex envelope of f relative to the interval (u_-,u_+). Since f is continuous and piecewise linear, f_c also is continuous and piecewise linear. We define u_i, i=1, 2,...,N by

$$u_1 = u_-, \quad u_N = u_+ \quad \text{and} \quad u_i < u_{i+1}$$

and

$$f_c \text{ is linear in the interval } (u_i,u_{i+1}).$$

Then the exact solution of the Riemann problem is

$$u(x,t) = u_i \text{ for } t\, f_c'(u_{i-1}+) = t\, f_c'(u_i-) <= x < \; t\, f_c'(u_i+) = t\, f_c'(u_{i+1}+)$$

where

$$f_c'(u_0+) = f_c'(u_1-) = -\infty$$

and

$$f_c'(u_N+) = f_c'(u_{N+1}-) = \infty.$$

See Figure 2.1. If $u_- > u_+$, the solution is found by substituting x with -x and f with -f.

Observe that we do not encounter rarefaction waves with our assumptions on f. As we have seen it is easy to solve the Riemann problem exactly and represent it numerically with our restriction on f. This is the cornerstone of the numerical method.

Given arbitrary u and f, we may approximate u with a piecewise constant function and f with a continuous and piecewise linear function. Then the solution of the Riemann problem on each discontinuity in u is used. It is then possible to solve (1.1) and (1.2) until two discontinuity lines collide. u(.,t) is still piecewise constant and it is possible to start over again. See Figure 2.2 for a typical solution u(x,t).

To prove that this is a well-defined procedure it is necessary to prove that if u_0 has a finite number of jumps, then u(x,t) is constant on a finite number of domains for t < T. In fact we are going to prove that u(x,t) is constant on a finite number of domains for all t.

First we need a proposition.

Proposition 2.1

Assume that u_0 is a step function with a finite number of jumps and that f is continuous and piecewise linear with a finite number of breakpoints. Then for fixed t $u(x,t)$ is a step function and

$$u(x,t) \in \{u_0(x);\ x \in R\} \cup \{u;\ f' \text{ is discontinuous at } u\}.$$

Proof

New values only arise in the Riemann problems. In the Riemann problems with the assumption on f, the only values that may arise are the values where f' is discontinuous.●

We now need some definitions. Using Proposition 2.1 we let

$$w_1 < w_2 < \ldots < w_M$$

be the values u can take.

A curve of discontinuity for u is called a shock front. A point where two or more shock fronts collide is called a shock collision. A shock front where u has the value w_i on one side and w_j on the other side is said to contain $|i-j|$ shock lines. The total number of shock lines in $u(.,t)$ is a measure of the total variation of $u(.,t)$.

Theorem 2.2

Let N be the number of intervals where f is linear, let $L(t)$ be the number of shock lines in $u(.,t)$ and $F(t)$ be the number of shock fronts in $u(.,t)$. Then the function

$$G(t) = L(t)*N + F(t)$$

is strictly decreasing for every shock collision for $t > 0$.

Proof

Assume $u_1, u_2, \ldots, u_K$ are the values of $u(x,t)$ which meet in a shock collision. Then $K-1$ shock fronts meet in the collision. We distinguish two cases. In the first case at least one of $u_2,\ldots,u_{K-1}$ is not between u_1 and u_K, and the second case all $u_2,\ldots,u_{K-1}$ are between u_1 and u_K.

In the first case we obtain after the shock collision at most N shock fronts and the value to the left of the left shock front is u_1 and the value to the right of the right shock front is u_K. If there are more than one shock front after the collision, then the values between the shock fronts form a monotone sequence. Thus the values which were not between u_1 and u_K have disappeared. Then the number of shock lines has decreased and the number of shock fronts has at most increased with $N-2$. Thus $G(t)$ has decreased.

In the second case we can disregard the values that make the sequence $\{u_i\}$ non monotone. For a monotone sequence $\{u_i\}$ it is easy to prove that after the shock collision all the $K-1$ shock fronts are united into one shock front. Then the number of shock fronts decreases and the number of shock lines does not increase, thus $G(t)$ decreases.●

Corollary 2.3

If u_0 has a finite number of jumps and f is linear on a finite number of intervals, then u(x,t) is constant on a finite number of domains.

Assume $|f'|<C$. Then the maximum speed of a shock front in a Riemann problem is C and therefore the numerical solution in (x_0,t) is independent of $u_0(x)$ for $x < x_0 - C\,t$ and for $x > x_0 + C\,t$. Therefore it is possible to use this method even when u_0 has a infinite number of constant states. We only need to care about the number of states in an interval which depend on the interval where we want the solution and C.

In ordinary numerical methods it is necessary to bound the timestep compared to the spacestep in order to keep the numerical method stable. This is related to the demand that neighbouring Riemann problem should not interact in the same timestep. In the method presented in this paper, this is not any problem. In fact the numerical method only considers interactions between neighbouring Riemann problems. Therefore we may say that this method has longer timestep than other methods. The last timestep is infinite. Thus this method is faster than other numerical methods. However an optimal implementation in regard to computer time, uses more storage than most other methods.

In the next chapter we will prove that the error in this method is far smaller than in any other numerical method for this problem. The error estimate is also very precise and simple.

3 ERROR ESTIMATES

In Lucier [6] the following theorem is proved

Theorem 3.1

If f and g are Lipschitz continuous functions, u_0 and $v_0 \in BV(R)$ and u and v are solutions of

$$\begin{aligned} u_t + f(u)_x &= 0 && \text{for } x\in R \text{ and } t>0 \\ u(x,0) &= u_0(x) && \text{for } x\in R \end{aligned}$$

and

$$\begin{aligned} v_t + g(v)_x &= 0 && \text{for } x\in R \text{ and } t>0 \\ v(x,0) &= v_0(x) && \text{for } x\in R, \end{aligned}$$

then for any t>0

$$\|u(.,t)-v(.,t)\|_{L_1(R)} \le \|u_0(.)-v_0(.)\|_{L_1(R)} + t\,\|f-g\|_{Lip}\min(|u_0|_{BV(R)},|v_0|_{BV(R)}),$$

where we have defined

$$\|g\|_{Lip} = \sup_{x\neq y}\left|\frac{g(x)-g(y)}{x-y}\right| .$$

Since the numerical method is the exact solution of a perturbed problem, this theorem may be used as an error estimate. Let g be a piecewise linear approximation to f and v_0 a piecewise constant approximation to u_0. Then Theorem 3.1 gives an error estimate. It is well-known that if g is a piecewise linear interpolant of f with breakpoints in ih, for iεZ, then

$$||f-g||_{Lip} \leq \frac{h}{2} ||f''||_{L_\infty(R)},$$

for sufficiently smooth f. Therefore Theorem 3.1 proves optimal asymptotic convergence.

Using the convergence of the numerical method it is possible to improve this estimate. First we need two definitions.

<u>Definition 3.1</u>

Assume that the step function u(.,t) satisfies

$$u(x,t) = u_i \quad \text{for } x\varepsilon(a_i,a_{i+1}) \quad i=1,2,\ldots,M$$

for a fixed t. Then we define $u_c(.,t)$ by

$$u_c(x,t) = \begin{cases} u_i & \text{for } x\varepsilon(a_i,a_{i+1}-\varepsilon) \\ u_i + \dfrac{(x-a_{i+1}+\varepsilon)}{\varepsilon}(u_{i+1}-u_i) & \text{for } x\varepsilon(a_{i+1}-\varepsilon,a_{i+1}) \end{cases}$$

where

$$\varepsilon = \frac{1}{3} \min_i (a_{i+1} - a_i).$$

$u_c(x,t)$ is a continuous piecewise linear function. See Figure 3.1.

<u>Definition 3.2</u>

$T.V._x(f(u(x)))$ is the total variation of f when xεR, i.e.

$$T.V._x(f(u(x))) = \sup_{\{x_i\}} \sum_{i=1}^{N} |f(u(x_{i+1}))-f(u(x_i))|$$

where $\{x_i\}$ is a finite set of real numbers. Here it is essential that u is continuous.

We may then state the theorem

<u>Theorem 3.2</u>

If f is an absolutely continuous function with |f'| bounded by C, $u_0\varepsilon BV(R)$ and u is the solution of

$$u_t + f(u)_x = 0 \qquad \text{for } x\varepsilon R \text{ and } t>0$$
$$u(x,0) = u_0(x) \qquad \text{for } x\varepsilon R$$

and g is a continuous piecewise linear function with |g'| bounded by C, v_0 is piecewise constant with a finite number of constant states and v is the solution of

$$v_t + g(v)_x = 0 \qquad \text{for } x\varepsilon R \text{ and } t>0$$
$$v(x,0) = v_0(x) \qquad \text{for } x\varepsilon R,$$

then for any t>0

$$\int_{a+Ct}^{b-Ct} |u(x,t) - v(x,t)|dx \leq \int_a^b |u_0(x) - v_0(x)|dx + t \; T.V._{x\varepsilon[a,b]}(f(v_{0,c}(x)) - g(v_{0,c}(x)))$$

where $a < b$.

This theorem differs from Theorem 3.1 in two ways.
First the Theorem 3.2 emphasizes the local behaviour of the solution.
Second the increase in the L_1 norm is sharper in Theorem 3.2 since

$$T.V._{x\varepsilon[a,b]}f(x) \leq ||f||_{Lip}(b-a).$$

4 SOME PROPERTIES OF THE SOLUTION OF INITIAL VALUE PROBLEM

Dafermos[1] presented his idea as a new method to study the initial value problem. Many of the well-known properties to the initial value problem can easily be proved by this method. We will list some of these properties. Dafermos notices some of them, but his proofs are possible to simplify using corollay 2.3 in this paper.

Exsistence and uniqueness for the initial value problem are well-known. These properties are easy to prove using convergence of the numerical method, since the number of constant states is finite.

Dafermos proved that if the initial value is bounded then the solution of the initial value problem is bounded by the same bounds as the inital value. This is easily seen from the numerical method.

The local behaviour of the solution follow directly from the local behaviour of the numerical method.

The stability Theorems 3.1 and 3.2 are easily proved using the convergence of the numerical method.

Dafermos proved in a special case that the total variation of the solution is not increasing in time. This follows in the general case from Theorem 2.2. The function L(t) is equivalent to the total variation. The total variation is decreasing in every shock collision when the middle value(s) is not between the the values on each side. The total variation is constant in shock collisions when the middle value(s) is between the values on each side. In this latter case the number of shock fronts decrease.

Since the number of shock collision is finite in the numerical method, there is for each pair of f and initial value a time T for which there is no shock collisions for t>T. This means that all shocks to the left of one shock A has lower speed than the shock A. This gives some indication on the behaviour of the solution when t increase to infinity.

5 A NUMERICAL EXAMPLE

In this chapter the numerical method is shown in an example. The flux function is a piecewise linear interpolant to

$$f(u) = (u - a)^3 + b.$$

The flux function is shown in figure 5.1. For simplicity we take the initial function to be piecewise constant, see figure 5.2. Then this problem with the perturbed flux function is solved exactly. Figure 5.3, and enlarged for small t in figure 5.4, shows the piecewise constant solution in the (x,t) plane. Figure 5.4 shows that for t small the solutions from the individual Riemann problems do not iteract. For t large there is however only one shock in this example. Figures 5.5 - 5.10 show the solution for different t values. We see that the solution is piecewise constant also where the exact solution of original problem is a rarefaction solution. With the natural approximation to the flux function the exact solution crosses the numerical solution between the endpoints in each interval where the numerical solution is constant. By improving the approximation to the flux function, the solution will of course converge to the exact solution.

Acknowledgement

The authors would like to thank Tore Gimse for carrying out the computation presented in chapter 5.

References

[1] Dafermos, C. M., Polygonal Approximation of Solutions of the Initial Value Problem for a Conservation Law, J. Math. Anal. Appl., 38 (1972), pp. 33-41.

[2] Kruzkov, S. N., First order quasilinear equations in several independent variables, Math. USSR Sbornik, 10 (1970), pp. 217-243.

[3] Lax P. D., Weak solutions of hyperbolic equations and their numerical computation. Comm. Pure Appl. Math. 7 (1954) pp. 159-193

[4] LeVeque, R. J., A Large Time Step shock-capturing techniques for Scalar Conservation Laws, SIAM J. Numer. Anal., 19 (1982) pp. 1051-1073.

[5] Lucier, L. J., Error Buonds for the Methods of Glimm, Godunov and LeVeque, SIAM J. Numer. Anal., 22 (1985), pp. 1074-1081.

[6] Lucier, L. J., A Moving Mesh Numerical Method for Hyperbolic Conservation Laws, Math. Comp., 46 (1986), pp. 59-69.

[7] Oleinik O. A., Discountinuous solutions of non-linear differential equations, Usp. Mat. Nauk. (N.S.), 12 (1957) 3-73, English transl. Amer. Math. Soc. Transl. Ser. 2, 26 (1963) pp. 95-172

[8] Oleinik O. A., Uniqueness and a stability of the generalized solution of the Chauchy problem for a quasilinear equation, Usp. Mat. Nauk. 14 (1959) pp. 165-170, English transl. Amer. Math. Soc. Transl. Ser. 2, 33 (1964) pp. 285-290

[9] Volpert, A. I., The Spaces BV and quasilinear equations, Math. USSR Sbornik, 2 (1967), pp. 225-267.

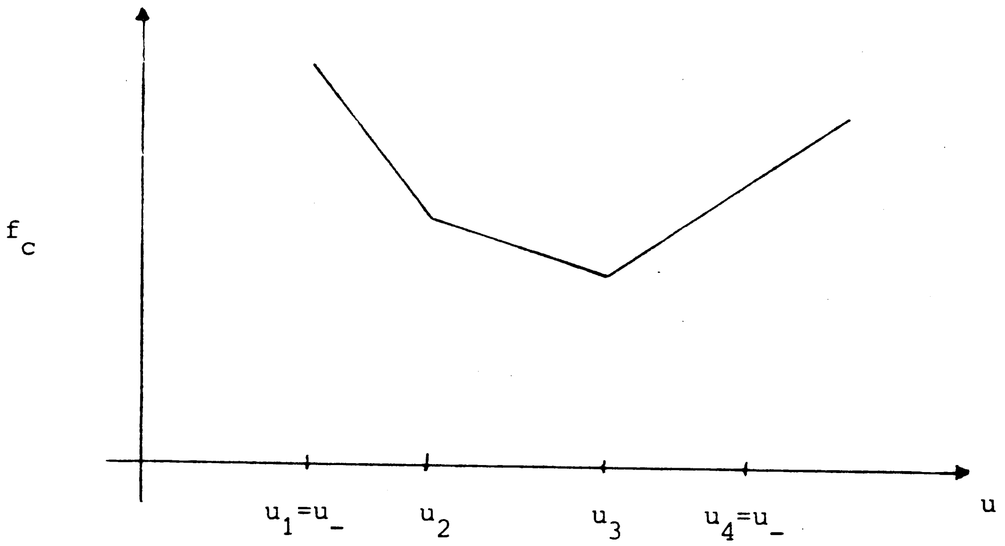

Figure 2.1a

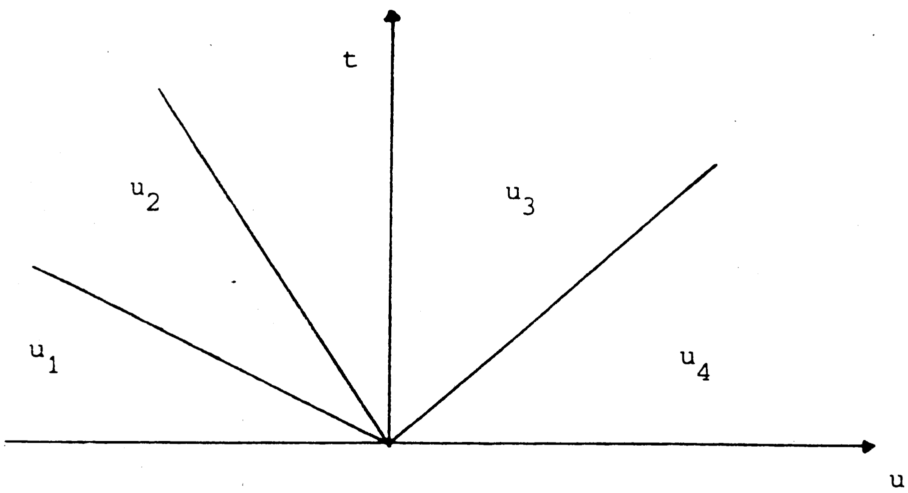

Figure 2.1b

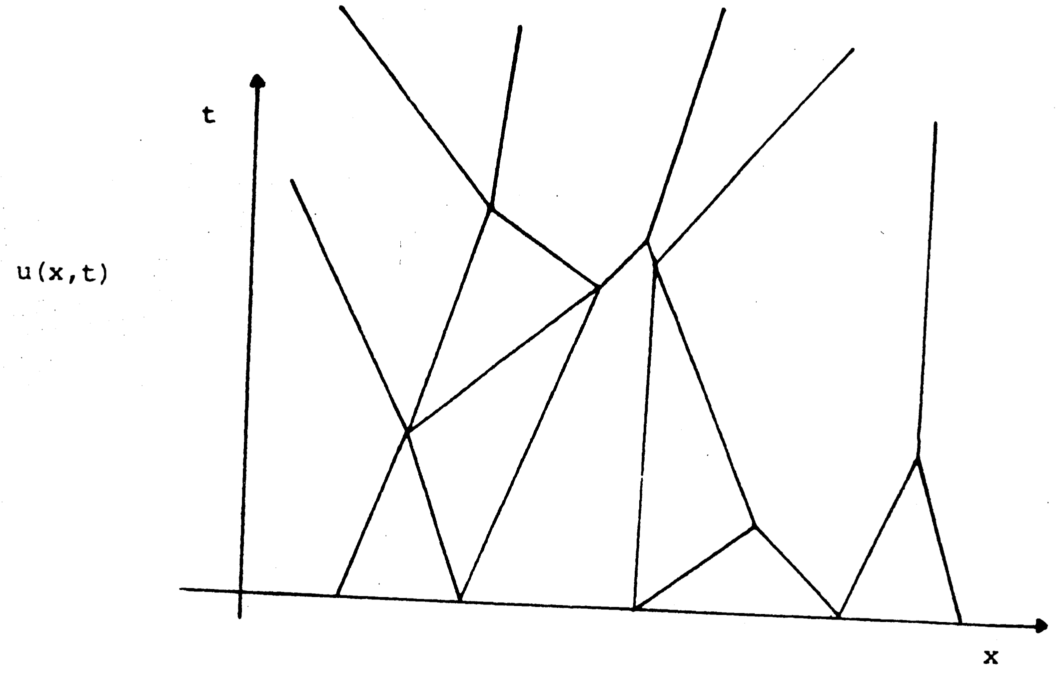

Figure 2.2

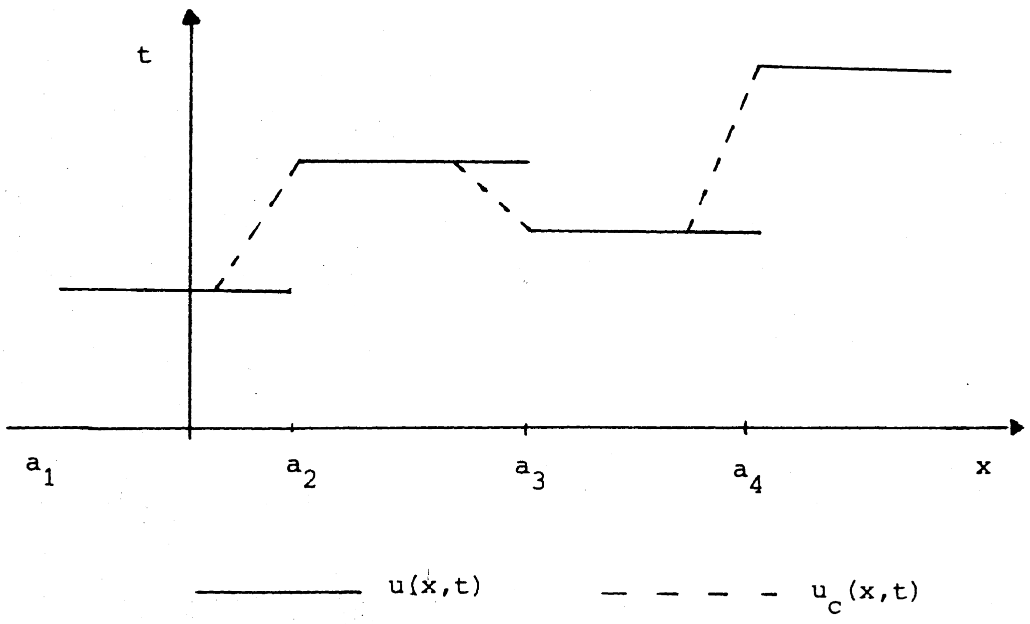

Figure 3.1

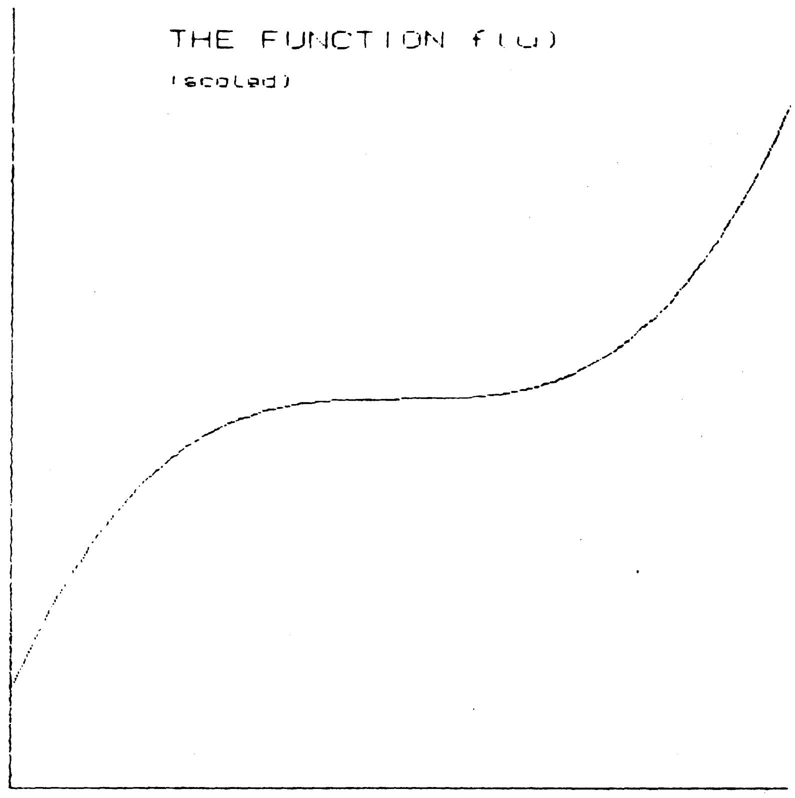

Figure 5.1

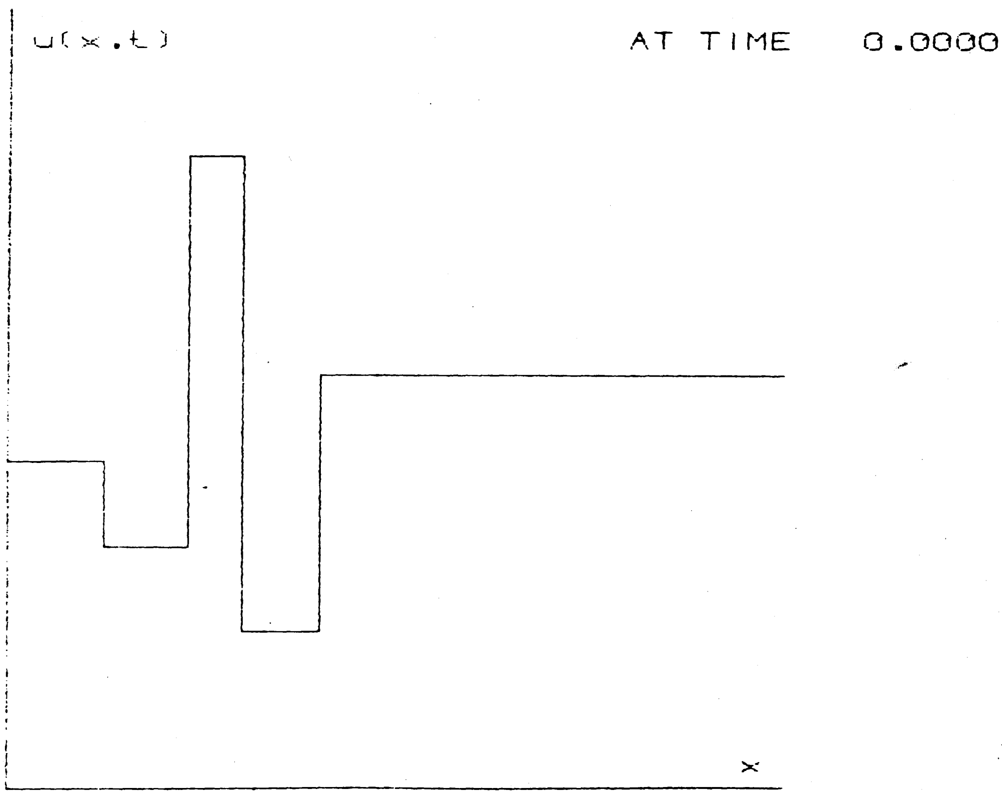

Figure 5.2

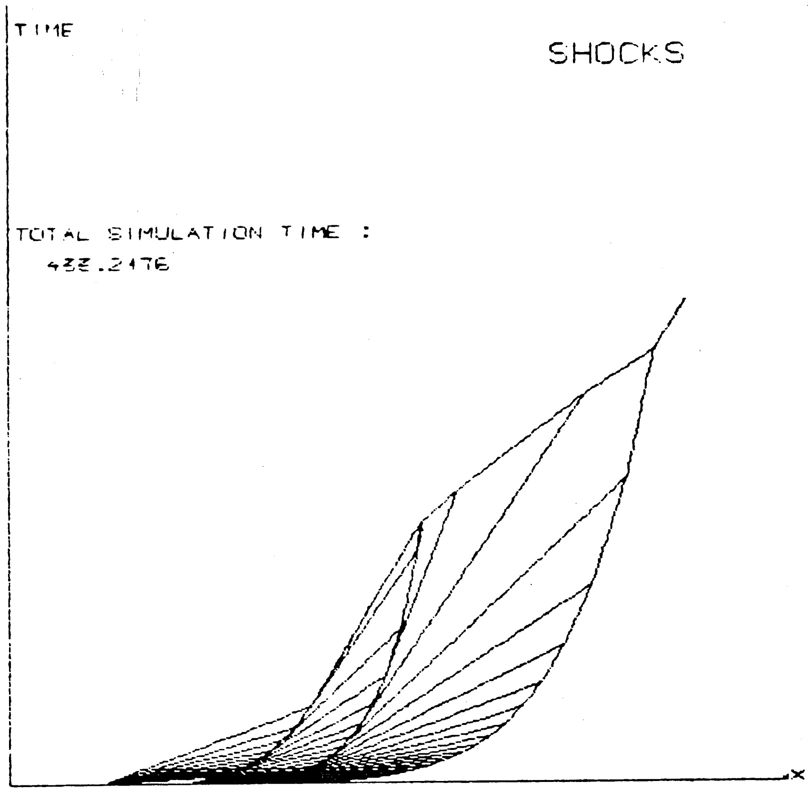

Figure 5.3

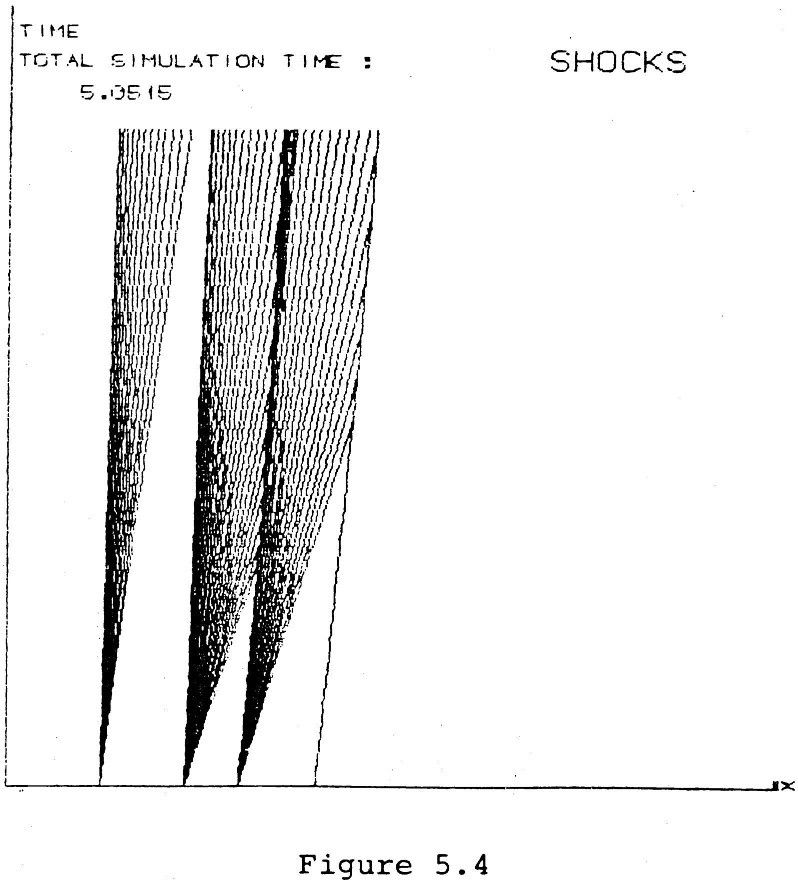

Figure 5.4

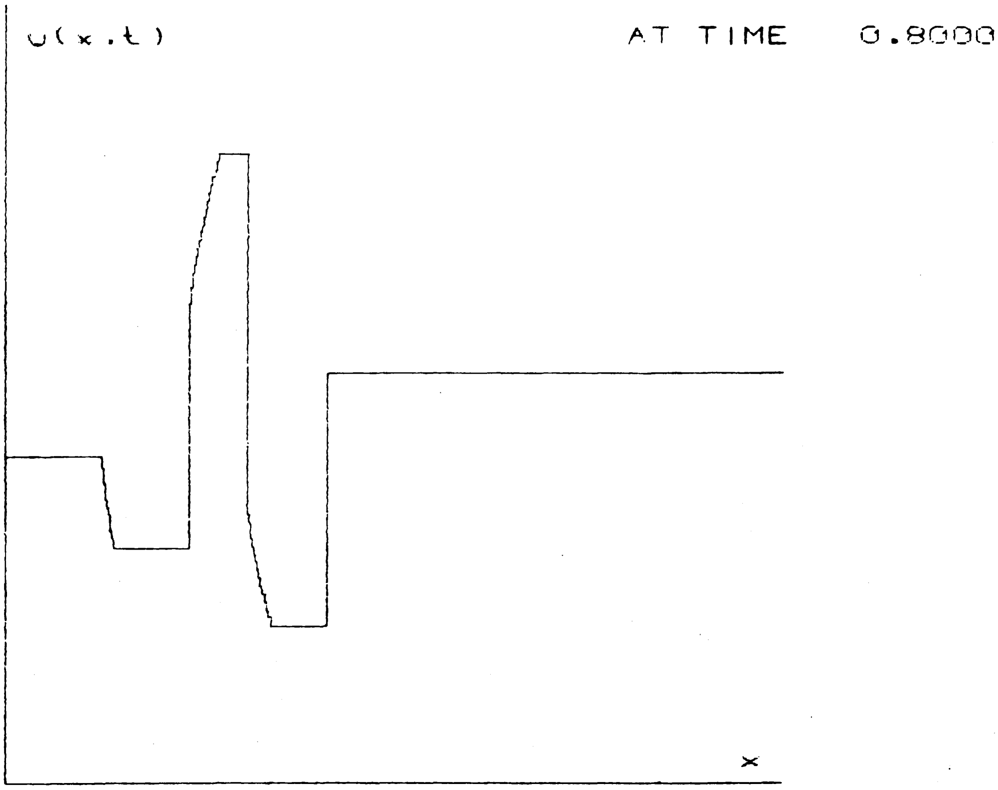

Figure 5.5

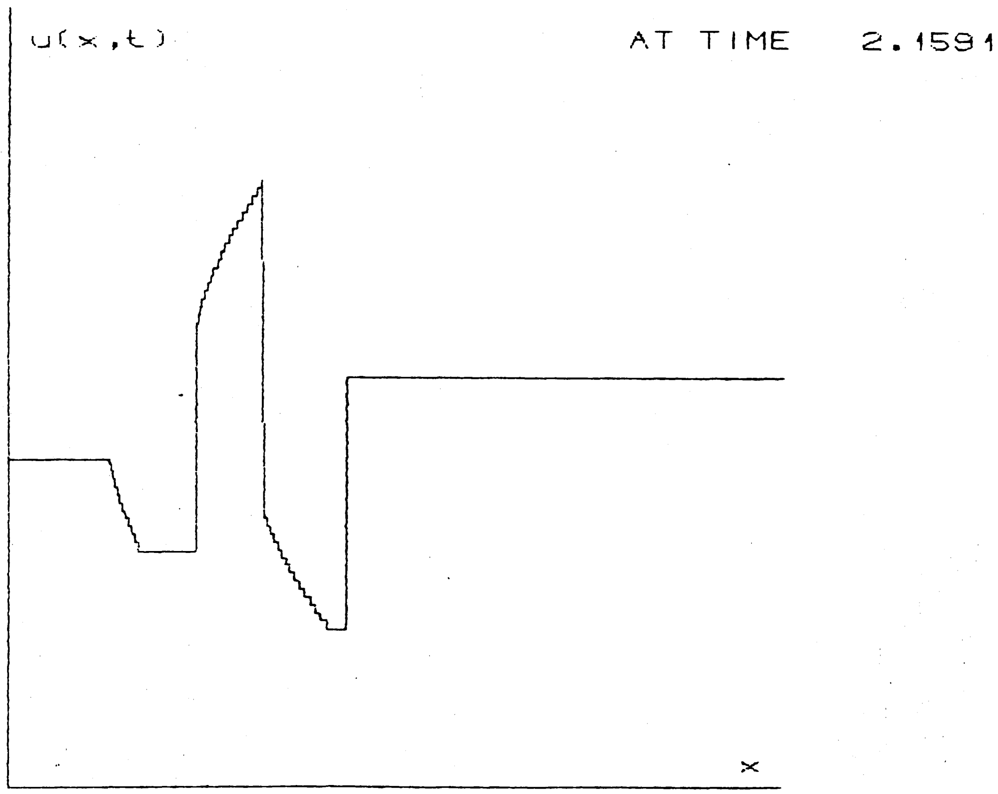

Figure 5.6

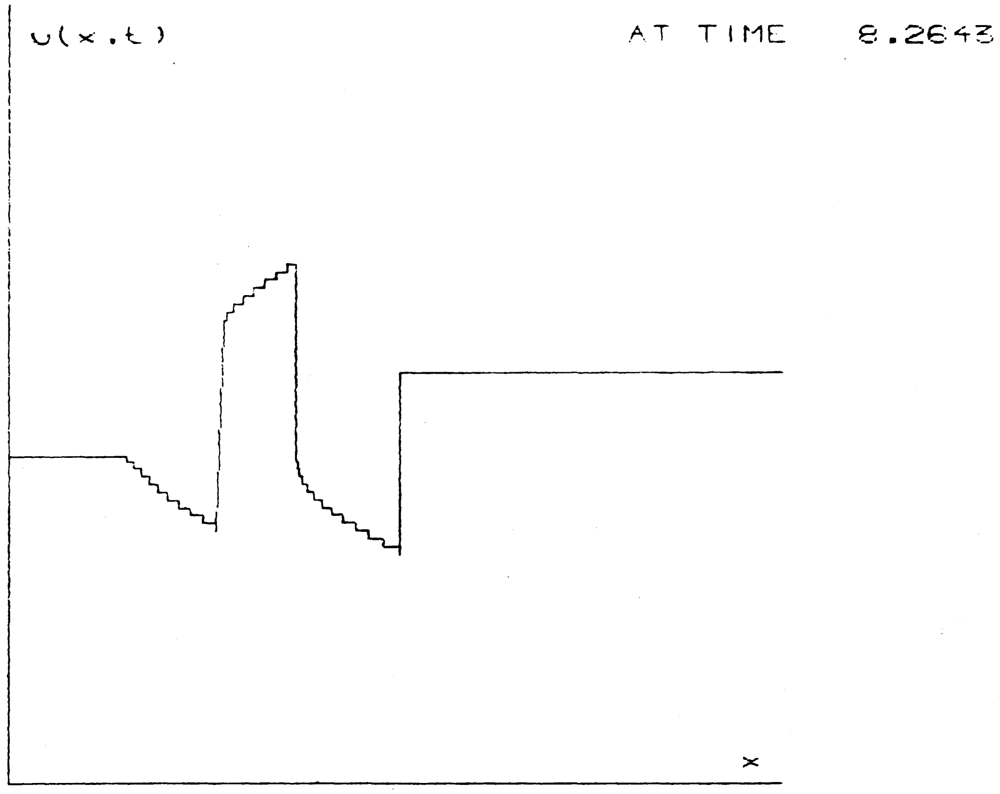

Figure 5.7

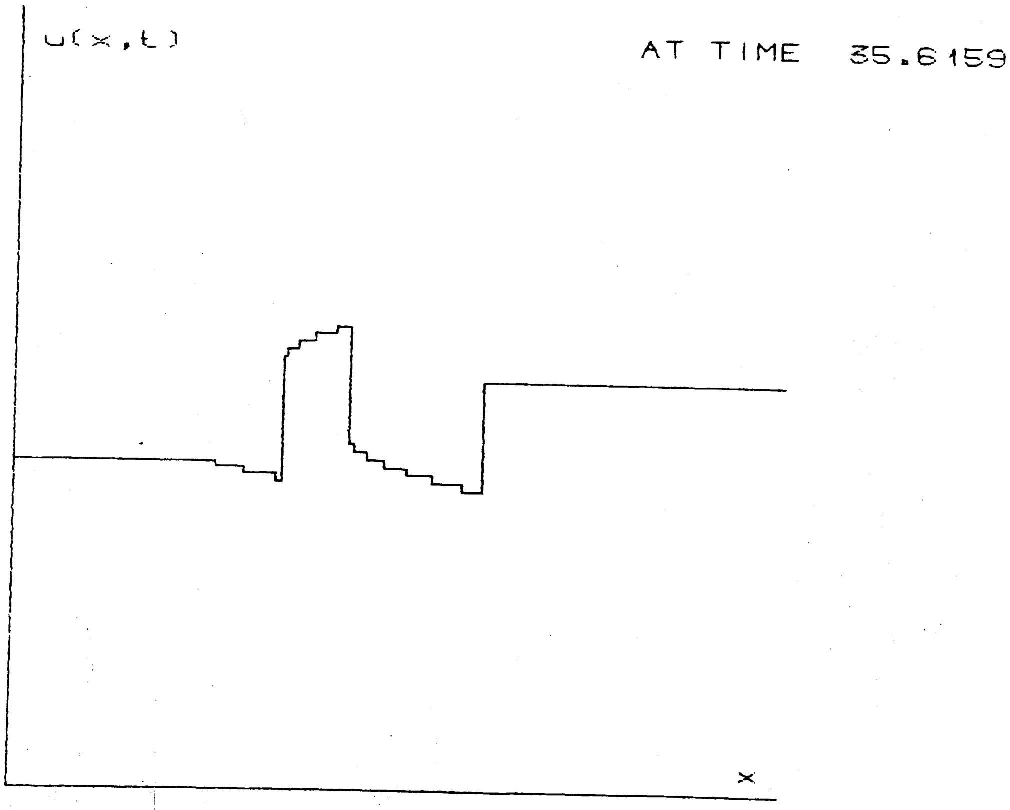

Figure 5.8

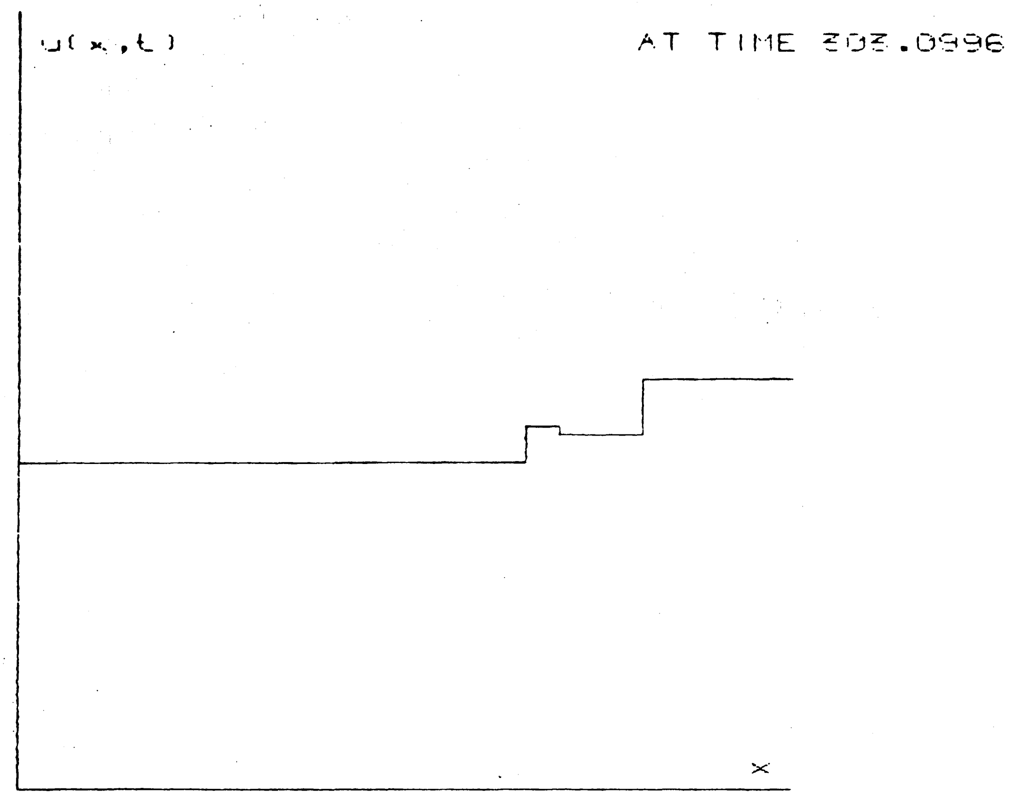

Figure 5.9

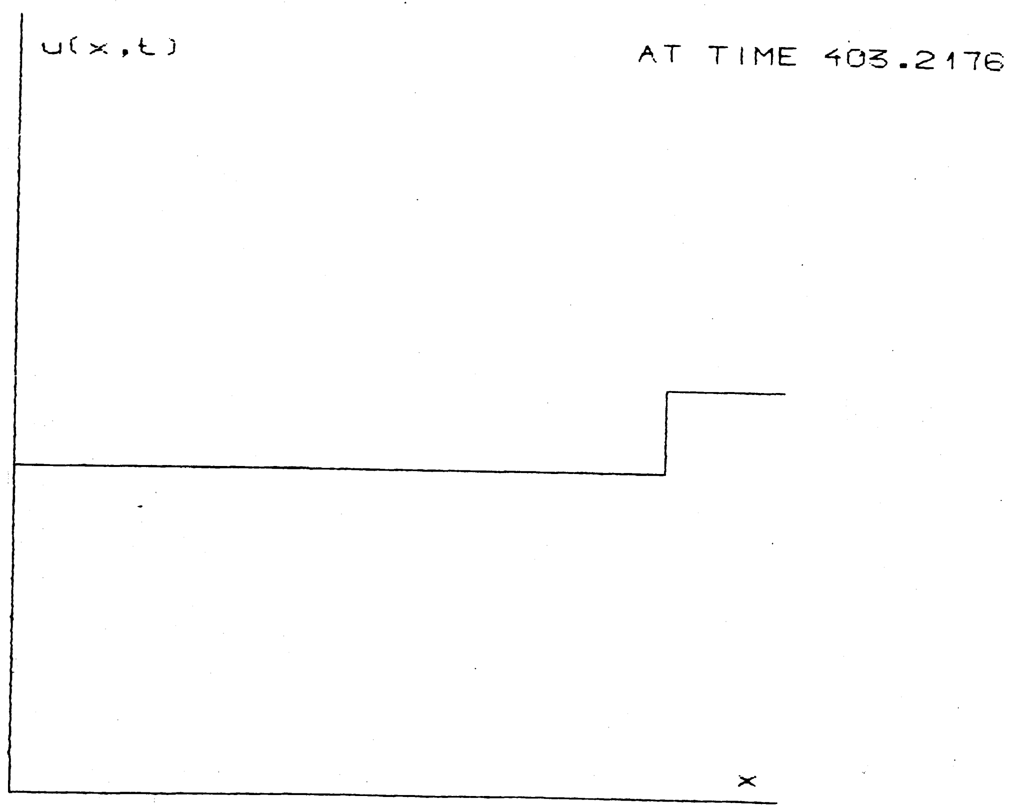

Figure 5.10