

UNIVERSITY OF READING  
DEPARTMENT OF MATHEMATICS

A HIGH ORDER MONOTONICITY PRESERVING ALGORITHM ON  
AN IRREGULAR GRID FOR NON-LINEAR CONSERVATION LAWS

---

P.K. SWEBY

Numerical Analysis Report 1/81

This work was carried out during the  
tenure of an S.R.C. CASE Studentship  
in conjunction with R.A.E. Farnborough,  
under the joint supervision of P.L. Roe  
(R.A.E. Bedford) and Dr. M.J. Baines  
(Mathematics Department, Reading)

1. Introduction
2. Roe's method
3. Monotonicity preservation
4. Irregular grids
5. Monotonicity preservation on an irregular grid
6. Numerical results for the equation  $u_t + (\frac{1}{2}u^2)_x = 0$
7. A three parameter scheme
8. A note on entropy violation
9. Summary

Appendices:

- A. An observation on stability regions
- B. A Fortran program to solve  $u_t + (\frac{1}{2}u^2)_x = 0$  and graphical results

Bibliography

## §1. Introduction

This report describes work carried out in developing various ideas, proposed by P. L. Roe of R.A.E. Bedford, concerning data-dependent, monotonicity preserving finite difference schemes for non-linear conservation laws, including in particular their extension to irregular grids.

Section 2 contains an outline of Roe's original scheme on a regular grid, along with his general approach to generating finite difference schemes.

Section 3 contains a new proof that Roe's scheme is monotonicity preserving.

In Section 4 a monotonicity preserving scheme on a restricted irregular grid is derived which is conservative and also second order in certain smooth regions of the flow. Section 5 contains a proof, along the lines of Section 3, showing that the scheme of Section 4 is monotonicity preserving.

Section 6 describes scalar test cases to which the scheme of Section 4 has been applied as well as the application to a Shock Tube problem, and references the graphs of Appendix B.

In Section 7 a possible extension to an arbitrary irregular grid, using 3 parameters, is presented along with various limitations. Section 8 comments on entropy violation and a possible method of avoiding this. These sections deal with topics still to be investigated fully.

Section 9 consists of a brief summary of previous sections.

Appendix A states and generalises an observation made by P. L. Roe on stability regions, whilst Appendix B contains a listing of the FORTRAN program used for the test cases along with graphical output.

## §2. Roe's Method

Consider first the simple linear scalar equation

$$u_t + au_x = 0 \quad (2.1)$$

where  $a$  is a constant. Discretise the region using a regular grid, with  $x$  values at the nodes  $x_k = x_{k-1} + \Delta x$ . The interval  $[x_{k-1}, x_k]$  will be referred to as the  $k$ -th cell.

Equation (2.1) integrated over the  $k$ -th cell gives

$$\frac{\partial}{\partial t} \left\{ \int_{x_{k-1}}^{x_k} u dx \right\} = -a(u_k - u_{k-1}) = -a\Delta u_k \quad (2.2)$$

where for notational convenience  $u_k - u_{k-1} = \Delta u_k$  is written as  $\Delta u_k$ .  $u_j$  is the restriction of  $u(x)$  to the grid. (i.e. point values)

A time discretization then gives

$$\Delta \left\{ \int_{x_{k-1}}^{x_k} u dx \right\} = -a\Delta t \Delta u_k \quad (2.3)$$

If  $u$  is regarded as a density then the l.h.s. can be thought of as the change of mass in cell  $k$  over time  $\Delta t$ . Hence if the r.h.s. is distributed with weights summing to unity conservation of mass is maintained. Once the mass has been allocated, it is converted back to a density by dividing by the length  $\Delta x$ , and added to  $u$  at appropriate points.



Thus we have

$$\frac{\Delta}{\Delta x} \left\{ \int_{x_{k-1}}^{x_k} u dx \right\} = -a \frac{\Delta t}{\Delta x} \Delta u_k = g_k \quad (2.4)$$

where  $g_k$  is the familiar quantity used to increment values of  $u$  in several

well known algorithms. For example, in first order upwind differencing it is used to increment  $u_j$  if  $a > 0$  and to increment  $u_{j-1}$  if  $a < 0$ . The notation  $v = a\Delta t/\Delta x$  will be used for the C.F.L. number.

If  $g_k$  is used to increment  $u_{k-1}$  and  $u_k$  with weights  $\alpha, \beta$  respectively then these weights can be chosen to give a second order accurate scheme, namely Lax Wendroff. To obtain second order accuracy it suffices to choose  $\alpha, \beta$  such that polynomials of degree zero, one and two are convected exactly, the exact increments being  $0, -a\Delta t, (a\Delta t)^2$  respectively.

i.e. scheme is:

$$u^k = u_k + \alpha g_{k+1} + \beta g_k \quad (2.5)$$

where subscripts denote time level  $t$ , superscripts denote time level  $t+\Delta t$ .  $u_k$  will be referred to as the data,  $u^k$  as the solution.

W.l.o.g. take  $x_k$  to be the origin of  $x$

zero-th order,  $u \equiv \text{const}$ ,  $g's \equiv 0$ . zero increment as required

first order,  $u \equiv x$ , increment =  $-av\Delta x - \beta v\Delta x = -(\alpha+\beta)a\Delta t$

exact increment =  $-a\Delta t$ .  $\alpha+\beta = 1$  (note same as for conservation)

second order,  $u \equiv x^2$ , increment =  $-av\Delta x^2 + \beta v\Delta x^2 = (\beta-\alpha)a\Delta t\Delta x$

exact increment =  $(a\Delta t)^2$ .  $\beta-\alpha = v$

giving  $\alpha = \frac{1}{2}(1-v)$   $\beta = \frac{1}{2}(1+v)$  (2.6)

Alternatively the  $g_k$  can be distributed to two upwind points. If  $v > 0$  this gives the scheme:

$$u^k = u_k + \gamma g_k + \delta g_{k-1} \quad (2.7)$$

Solving the second order weights using the same procedure as above gives

$$\gamma = \frac{1}{2}(3-v) \quad \delta = \frac{1}{2}(v-1) \quad (2.8)$$

the second order Warming and Beam Upwind Scheme. Similar weights are obtained for  $v < 0$ . Other scheme's may be obtained using 3 or more weights, e.g. Fromm, Roe, Lax-Courant (see [1], [2]).

So for all that has been done is to derive established schemes via a new approach. Roe now introduces the notion of compatibility [1], namely:

Definition

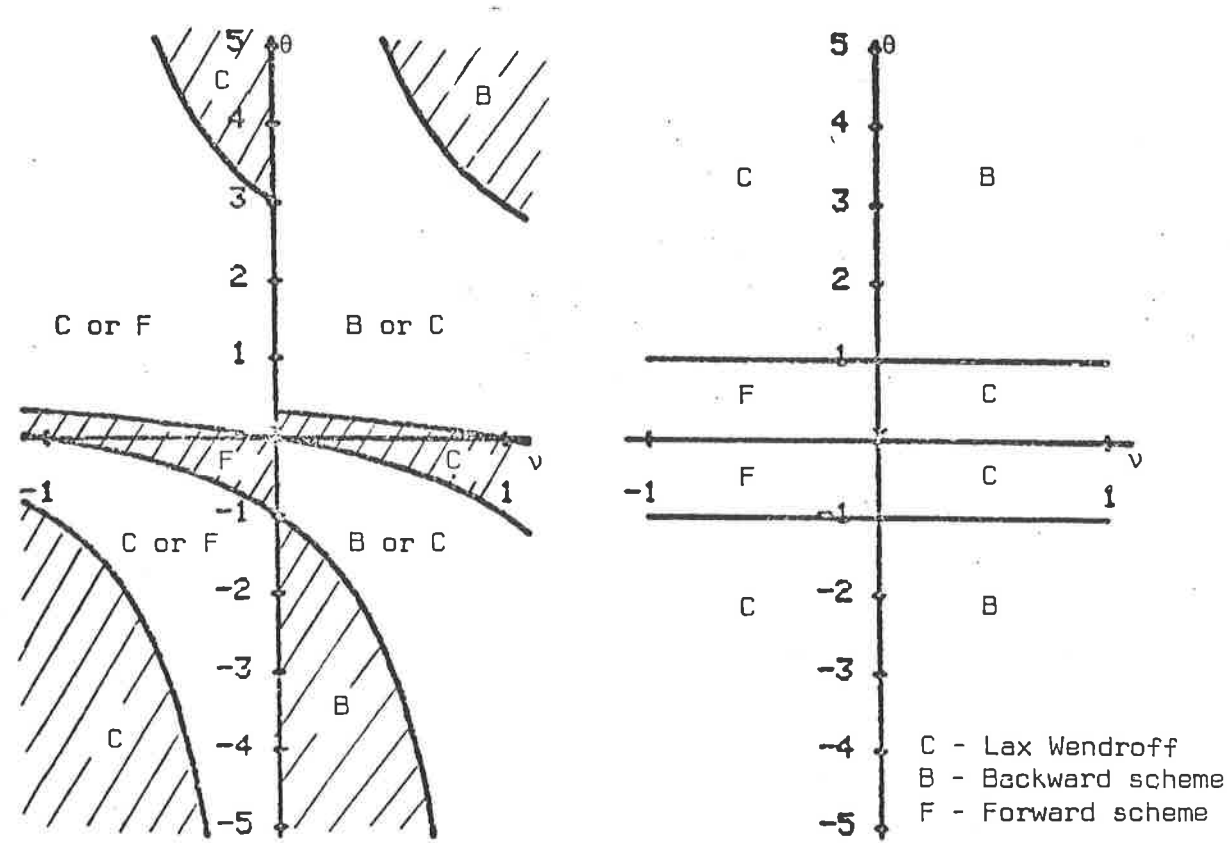
A given algorithm is compatible, in the sense of Roe, with a given set of data values  $\{u_n\}$  if the solution at each point within that set is bounded by the data values from the interval containing the characteristic through the solution point. That is

$$\min[u_{n-N-1}, u_{n-N}] \leq u^n \leq \max[u_{n-N-1}, u_{n-N}] \tag{2.9}$$

where  $N \leq v \leq N+1$ .

It is easily seen that given an algorithm compatible with monotonic data, the solution set is also monotonic. That is a compatible algorithm is monotonically preserving. This suggests (and it is borne out in practice) that serious oscillations introduced by Lax-Wendroff and other schemes near shocks will not be produced by such a compatible algorithm.

Roe now defines a data parameter  $\theta = (u_{n+1} - u_n) / (u_n - u_{n-1})$  (in fact a ratio of gradients) and shows that Lax-Wendroff and Warming and Beam (forward and backward) schemes between them provide compatibility over the part of  $\theta$ - $v$  plane bounded by the C.F.L. condition  $|v| \leq 1$ . See Figure 2.1



Compatibility regions

Switching strategy

Fig. 2.1

Fig. 2.2

Roe then proposes a switching strategy for switching between Lax-Wendroff and the appropriate upwind schemes which ensure compatibility. This is shown in Figure 2.2 and is implemented via the shifting of increments ([1]).

So far only the linear equation (2.1) has been considered, but the nonlinear equation

$$u_t + (f(u))_x = 0 \quad (2.10)$$

can be treated similarly. The equivalent of (2.4) is

$$\frac{\Delta}{\Delta x} \left\{ \int_{x_{k-1}}^{x_k} u dx \right\} = -\frac{\Delta t}{\Delta x} \Delta f_k = -\frac{\Delta t}{\Delta x} a_k \Delta u_k$$

$$= -v_k \Delta u_k = g_k \quad (2.11)$$

where

$$a_k = \frac{\Delta f_k}{\Delta u_k} \quad \text{and} \quad v_k = a_k \frac{\Delta t}{\Delta x}$$

the weights derived above are used, substituting  $v_k$  for  $v$  appropriately, and although it is not possible to prove the same degree of accuracy, empirical results suggest that this is so. Moreover, with the switching strategy, these results display an absence of spurious oscillation near shocks.

Roe's scheme can also be extended to systems of equations

$$\underline{u}_t + \underline{F}(\underline{u})_x = 0 \quad (2.12)$$

by decomposing into component waves in the directions of the eigenvectors of the system using the linear Riemann Problem (see [2]). Roe has used his scheme on the Euler equations, and has shown it to perform well on various problems. e.g. Sod's [3] Shock Tube problem.

### §3. Monotonicity Preservation

It should be noted that monotonicity as used here means monotonicity preserving in the sense that, given monotonic data  $\{u_k\}$  the solution set  $\{u^k\}$  is also monotonic. This is distinct from monotone schemes, i.e. ones which are non-decreasing functions of each of their arguments (see e.g. Harten, Hyman & Lax [4]).

Roe's scheme, incorporating his switching strategy can be written as

$$u^k = u_k + \alpha_{k+1} g_{k+1-s_{k+1}} + \bar{\alpha}_{k+1} (g_{k+1} - g_{k+1-s_{k+1}}) + \beta_k g_{k-s_k} + \bar{\beta}_k (g_k - g_{k-s_k})$$

where

$$\left. \begin{aligned} v_k &= \frac{\Delta f_k}{\Delta u_k} \frac{\Delta b}{\Delta x}, & s_k &= \text{sgn}(v_k) & g_k &= -v_k \Delta u_k \\ \alpha_k &= \frac{1}{2}(1 - v_k) & \beta_k &= \frac{1}{2}(1 + v_k) \\ \bar{\alpha}_k &= \begin{cases} \alpha_k & |g_k| < |g_{k-s_k}| \\ \frac{1}{2}(1 - s_k) & |g_k| \geq |g_{k-s_k}| \end{cases} \\ \bar{\beta}_k &= \begin{cases} \beta_k & |g_k| < |g_{k-s_k}| \\ \frac{1}{2}(1 + s_k) & |g_k| \geq |g_{k-s_k}| \end{cases} \end{aligned} \right\} \quad (3.1)$$

To simplify notation consider the linear case ( $f \equiv au$ ) for which  $v$  is constant and we may drop subscripts on  $v, \alpha, \beta$  and  $s$ , leaving

$$u^k = u_k + \alpha g_{k+1-s} + \bar{\alpha}_{k+1} (g_{k+1} - g_{k+1-s}) + \beta g_{k-s} + \bar{\beta}_k (g_k - g_{k-s}) \quad (3.2)$$

where  $g_k = v \Delta u_k$ .

We can now prove algebraically that this scheme is monotonicity preserving and also gives a lower bound on the gradients of the solution set.

#### Theorem 1

Given a subset of the data  $\{u_k\}$  which is monotonic then the solution set  $\{u^k\}$  produced by (3.2) is



(i) monotonic

(ii) such that  $|g^k| \geq \min\{|g_k|, |g_{k-s}|\}$

provided that the C.F.L. condition  $|v| \leq 1$  is satisfied.

Proof

First it should be noted that since  $v$  is of constant sign, monotonic  $\{u_k\}$  is equivalent to the  $\{g_k\}$  being of constant sign. To prove (i) it is therefore sufficient to show that if  $\{g_k\}$  are of constant sign, then  $\{g^k\}$  are of the same sign.

$$\text{Now } g^k = -v(u^k - u^{k-1})$$

$$= -v(u_k + i_k - u_{k-1} - i_{k-1})$$

where  $u_k + i_k$  is given by r.h.s. of (3.2)

$$= g_k - v\{\alpha g_{k+1-s} + \bar{\alpha}_{k+1}(g_{k+1} - g_{k+1-s}) + \beta g_{k-s} + \bar{\beta}_k(g_k - g_{k-s}) - \alpha g_{k-s} - \bar{\alpha}_k(g_k - g_{k-s}) - \beta g_{k-1-s} - \bar{\beta}_k(g_{k-1} - g_{k-1-s})\}$$

$$\text{i.e. } g^k = [1 + v(\bar{\alpha}_k - \bar{\beta}_k)]g_k + v\{(\bar{\alpha}_{k+1} - \alpha)(g_{k+1-s} - g_{k+1}) - \alpha g_{k+1} + (\bar{\beta}_k - \beta - \bar{\alpha}_k + \alpha)g_{k-s} + \beta g_{k-1} + (\bar{\beta}_{k-1} - \beta)(g_{k-1} - g_{k-1-s})\} \quad (3.3)$$

There are two distinct cases,  $v > 0$  and  $v < 0$

Case I  $v > 0, s = 1$

(3.3) becomes (remembering that  $\alpha + \beta = 1$ )

$$g^k = \{1 + v(\bar{\alpha}_k - \bar{\beta}_k - \alpha)\}g_k + v\{(\bar{\beta}_k - \bar{\alpha}_k + \alpha)g_{k-1} + \bar{\alpha}_{k+1}(g_k - g_{k+1}) + (\bar{\beta}_{k-1} - \beta)(g_{k-1} - g_{k-2})\} \quad (3.4)$$

Consider first the terms inside the square brackets.

$$(\bar{\beta}_k - \bar{\alpha}_k + \alpha)g_{k-1}$$

:- there are two possible cases, corresponding to switch and no switching

switched:  $\bar{\beta}_k = \beta, \bar{\alpha}_k = \alpha$  . . . whole term is of same sign as  $g_{k-1}$

unswitched:  $\bar{\beta}_k = \beta, \bar{\alpha}_k = \alpha$  and again term is of same sign as  $g_{k-1}$

$\bar{\alpha}_{k+1}(g_k - g_{k+1})$  if  $\bar{\alpha}_{k+1} = 0$  then whole term vanishes. If  $\bar{\alpha}_{k+1} = \alpha$  this is the case when  $|g_k| > |g_{k+1}|$  and hence whole term is of the same sign as  $g_k$

$(\bar{\beta}_{k-1} - \beta)(g_{k-1} - g_{k-2})$  If  $\bar{\beta}_{k-1} = \beta$  then the term is zero, if  $\bar{\beta}_{k-1} = 1$  this is caused by  $|g_{k-1}| > |g_{k-2}|$  and since  $\beta < 1$  whole term is of the same sign as  $g_{k-1}$

Hence expression inside square brackets is of the same sign as either  $g_k$  or  $g_{k-1}$ . This leaves the term

$$\{1 + v(\bar{\alpha}_k - \bar{\beta}_k - \alpha)\}g_k$$

If  $\bar{\alpha}_k = \alpha$ ,  $\bar{\beta}_k = \beta$  we have  $(1-v\beta)g_k$  which, since both  $\beta$  and  $v$  are less than 1, is of the same sign as  $g_k$ . Otherwise we have

$$\begin{aligned} & [1 - v(1+\alpha)]g_k \\ &= [1 - \frac{v}{2}(3 - v)]g_k \\ &= (1 - \frac{v}{2})(1 - v)g_k \text{ which is of the same sign as } g_k \text{ for } v \in [0, 1]. \end{aligned}$$

Putting the above results together we have that if  $g_k$  and  $g_{k-1}$  are of the same sign then  $g^k$  is also of that sign; hence proving (i) of the theorem for  $v > 0$ .

To show part (ii)

if  $g_k, g_{k-1} \leq C \leq 0$ , we have from (3.4)

$$g_k \leq [1 + v(\bar{\alpha}_k - \bar{\beta}_k - \alpha)]C + v(\bar{\beta}_k - \bar{\alpha}_k + \alpha)C + \text{two negative quantities}$$

i.e.  $g_k \leq C$ .

Similarly if  $g_k, g_{k-1} \geq C \geq 0$  then  $g^k > C$ .

These combined give

$|g^k| \geq \min\{|g_k|, |g_{k-1}|\}$  for  $g_k, g_{k-1}$  of same sign which is (ii) of theorem for  $v > 0$ .

For the case  $v < 0$ ,  $s = -1$ , the counterpart to (3.4) is

$$\begin{aligned} g^k &= [1 - v(\bar{\beta}_k - \bar{\alpha}_k - \beta)]g_k - v\{(\beta - \bar{\beta}_k + \bar{\alpha}_k)g_{k+1} + \bar{\beta}_{k-1}(g_k - g_{k-1}) \\ &\quad + (\alpha - \bar{\alpha}_{k+1})(g_{k+2} - g_{k+1})\} \end{aligned} \quad (3.5)$$

The monotonicity preserving proof is obtained with similar arguments as for  $v > 0$ , likewise as is the result

$$|g^k| > \min\{|g_{k+1}|, |g_k|\}.$$

Thus the proof is complete.

The theorem can be restated to include the nonlinear equation (2.10) with the added condition that  $v_k$  is of constant sign over both the time step in question and the spacial region. That is the data should be away from sonic points ( $f'(\bar{u}) = 0$ ) in both space and time. The proof of this will be demonstrated from the extended proof to cover irregular grids given below.

From the proof it can be seen that the sign of  $g^k$  depends on only the signs of  $g_k$  and  $g_{k-s}$ ; and hence if the data is piecewise monotonic, with each section more than two cells in length, then the solution will also be piecewise monotonic, the lengths of each section being no more than two cells different from the lengths of the corresponding sections of the data. This can best be seen by considering the data in Figure 3.1, with  $v > 0$ .

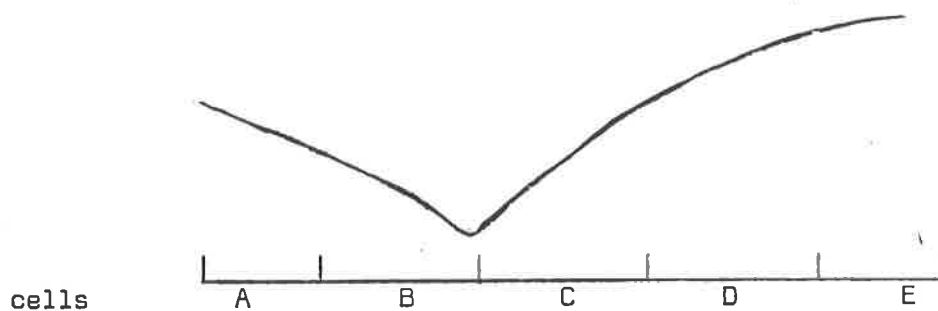


Figure 3.1

If the data continues monotonically at least one cell to the left of A, then by the theorem gradients of the solution will be non-positive in cells A and B, and non-negative in cells D and E. The theorem does not predict the sign of the gradient in C, but whichever sign it is, the solution is still piecewise monotonic.

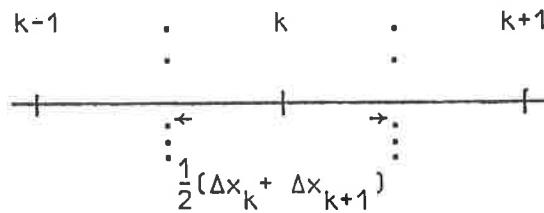
#### §4. Irregular grids

We now seek a monotonicity-preserving difference scheme on an irregular grid. Again we start with the linear equation (2.1) and then extend to the nonlinear equation (2.10).

For irregular grids we still have

$$\Delta \left\{ \int_{x_{k-1}}^{x_k} u dx \right\} = -a \Delta t \Delta u_k \quad (2.3)$$

but now the appropriate length, by which to divide to revert to a density at point  $k$ , is  $\frac{1}{2}(\Delta x_k + \Delta x_{k+1})$  since this is the length 'surrounding' point  $k$



Introducing weights  $\gamma_k$  and  $\delta_k$ , the increment at the point  $k$  is now

$$-\frac{2\gamma_{k+1} a \Delta t \Delta u_{k+1}}{\Delta x_{k+1} + \Delta x_k} - \frac{2\delta_k a \Delta t \Delta u_k}{\Delta x_{k+1} + \Delta x_k} \quad (4.1)$$

assuming that the mass from cell  $k$  is used to increment points  $k-1$  and  $k$  i.e. a centred scheme. The weights are also subscripted since it is expected that they will depend on the grid.

For conservation we need need the mass change given by (2.3) to be distributed exactly, i.e.

$$\gamma_k + \delta_k = 1 \quad (4.2)$$

It is convenient to define

$$\alpha_{k+1} = \frac{2\Delta x_{k+1} \gamma_{k+1}}{\Delta x_{k+1} + \Delta x_k}, \quad \beta_k = \frac{2\Delta x_k \delta_k}{\Delta x_{k+1} + \Delta x_k} \quad (4.3)$$

so that (4.1) becomes the familiar

$$\alpha_{k+1} g_{k+1} + \beta_k g_k \quad (4.4)$$

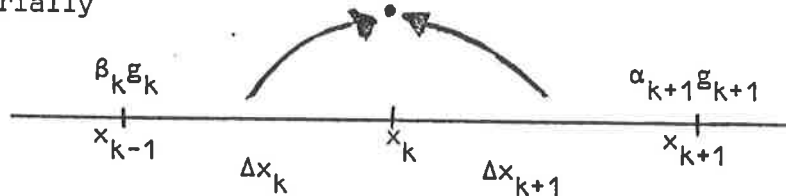
with  $g_k = -v_k \Delta u_k$ ,  $v_k = \frac{a \Delta t}{\Delta x_k}$  (4.5)

The conservation condition (4.2) now becomes

$$(\Delta x_{k-1} + \Delta x_k) \alpha_k + (\Delta x_k + \Delta x_{k+1}) \beta_k = 2 \Delta x_k \quad (4.6)$$

we have  $u^k = u_k + \alpha_{k+1} g_{k+1} + \beta_k g_k$  (4.7)

i.e. pictorially



Again we impose exactness when  $u$  is set equal to polynomials of degree 0, 1, 2, to obtain a second order scheme.

This gives

$$\alpha_k = \frac{\Delta x_{k-1} - a \Delta t}{\Delta x_{k-1} + \Delta x_k}, \quad \beta_k = \frac{\Delta x_{k+1} + a \Delta t}{\Delta x_k + \Delta x_{k+1}} \quad (4.8)$$

or, writing

$$R_k = \frac{\Delta x_k}{\Delta x_{k-1}}$$

$$\alpha_k = \frac{1 - v_{k-1}}{1 + R_k}, \quad \beta_k = \frac{R_{k+1} + v_k}{1 + R_{k+1}} \quad (4.9)$$

Note that

$$\alpha_{k+1} + \beta_k = 1 \quad (4.10)$$

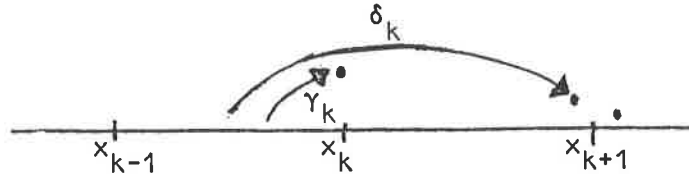
The conservation condition (4.6) is equivalent to

$$\Delta x_{k+1} - \Delta x_k = \Delta x_k - \Delta x_{k-1} \quad (4.11)$$

i.e. exact conservation is maintained on a grid where the lengths are linearly related.

An upwind second order scheme may also be derived, namely

$$u^k = u_k + \gamma_k g_k + \delta_{k-1} g_{k-1} \quad (4.12)$$



$\gamma_k, \delta_k$  are given by

$$\left. \begin{aligned} \gamma_k &= \frac{\Delta x_{k-1} + 2\Delta x_k - a\Delta t}{\Delta x_{k-1} + \Delta x_k} \\ \delta_k &= \frac{a\Delta t - \Delta x_{k+1}}{\Delta x_k + \Delta x_{k+1}} \end{aligned} \right\} \quad (4.13)$$

The conservation condition is

$$\gamma_k (\Delta x_k + \Delta x_{k+1}) + \delta_k (\Delta x_{k+1} + \Delta x_{k+2}) = 2\Delta x_k \quad (4.14)$$

$$\text{i.e. } (\Delta x_{k-1} + 2\Delta x_k - a\Delta t) \frac{(\Delta x_k + \Delta x_{k+1})}{(\Delta x_{k-1} + \Delta x_k)} + (a\Delta t - \Delta x_{k+1}) \frac{(\Delta x_{k+1} + \Delta x_{k+2})}{(\Delta x_k + \Delta x_{k+1})} = 2\Delta x_k \quad (4.15)$$

To ensure conservation the coefficient of  $a\Delta t$  in (4.15) must be zero, i.e.

$$\frac{\Delta x_{k+1} + \Delta x_{k+2}}{\Delta x_k + \Delta x_{k+1}} = \frac{\Delta x_k + \Delta x_{k+1}}{\Delta x_{k-1} + \Delta x_k}$$

$$\text{or } \frac{1 + R_{k+2}}{1/R_{k+1} + 1} = \frac{R_{k+1} + 1}{1/R_k + 1} \quad (4.16)$$

If we test this with the condition on the central scheme, namely (4.11)

which is equivalent to

$$R_{k+1} + 1/R_k = 2 \quad (4.17)$$

(4.16) becomes

$$\frac{1 + R_{k+2}}{3 - R_{k+2}} = \frac{1 + R_{k+1}}{3 - R_{k+1}} \quad \forall k$$

i.e.  $R_k = \text{const} = 1$  (from 4.17)  $\forall k$ , a regular grid, so the schemes can only both be conservative on a regular grid. This means that we cannot attempt to switch between these schemes to preserve monotonicity, and still have conservation

Consider now the scheme obtained by switching of the type given by (3.1), this is equivalent to in cell  $k$ , if  $v > 0$ , distribute  $g_{k-1}$  to points  $k-1, k$  with second order weights, then distribute  $g_k - g_{k-1}$  with first or second order weights depending on whether  $|g_{k+1}| > |g_k|$  or not.

If  $v > 0$ , and  $|g_{k+1}| > |g_k| > |g_{k-1}|$  point  $k$  gets total increment:

$$\underbrace{\alpha_{k+1} g_k + \beta_k g_{k-1}}_{\text{second order weights}} + \underbrace{(g_k - g_{k-1})}_{\text{first order weights}} = (\alpha_{k+1} + 1)g_k + (\beta_k - 1)g_{k-1} \quad (4.18)$$

We therefore have a scheme

$$u^k = u_k + \theta_k g_k + \phi_{k-1} g_{k-1}$$

with

$$\theta_k = \alpha_{k+1} + 1 \quad (4.19)$$

$$= \frac{\Delta x_{k+1} + 2\Delta x_k - a\Delta t}{\Delta x_k + \Delta x_{k+1}}$$

$$\phi_k = \beta_{k+1} - 1$$

$$= \frac{a\Delta t - \Delta x_{k+1}}{\Delta x_{k+1} + \Delta x_{k+2}}$$

Note that  $\theta_k + \phi_{k-1} = \alpha_{k+1} + \beta_k = 1$ , this first order accuracy condition, so this scheme is first order, but not second order since the unique 2 parameter upwind scheme is given by (4.12), (4.13).

The conservation condition is (4.14) with  $\gamma, \delta$  replaced by  $\theta, \phi$ . This gives

$$a\Delta t - \Delta x_{k+1} + \Delta x_{k+1} + 2\Delta x_k - a\Delta t = 2\Delta x_k$$

which is satisfied without any restrictions on the grid, so the fully switched scheme is conservative on any grid, but second order only on a regular grid.

Since this type of switching is monotonicity preserving on a regular grid it is plausible to suspect that it will also do so on an irregular grid. This is in fact so, as will be proved in the next section, but if compatibility is

investigated, there are certain values of  $v$  (within linear stability region) such that the switched scheme is not compatible. However, compatibility is a stronger condition than monotonicity preservation, and in some cases is undesirable, e.g.

Consider data such that  $u_{k-1} = u_k$  but varies inbetween. Compatibility will ensure that  $u^k$  is bounded by  $u_{k-1}$  and  $u_k$ , i.e. here  $u_k = u_{k-1}$ , hence clipping the peaks of any maxima or minima. (Figure 4.1)

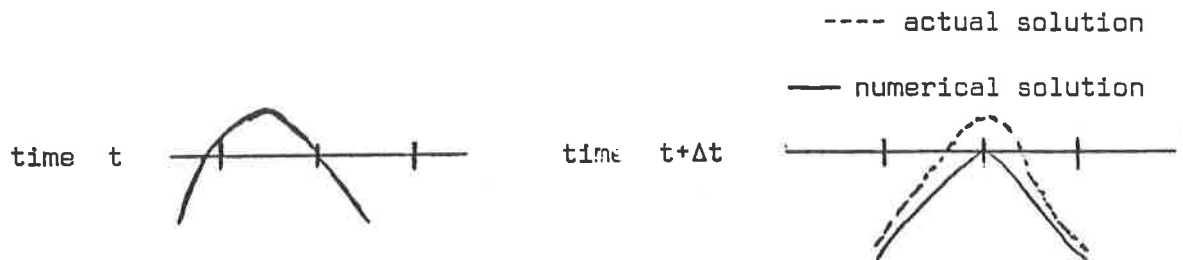


Figure 4.1

As shown above, the scheme is only conservative on a particular type of irregular grid, but, if some prior knowledge of the solution is known, this need not be as restrictive as it at first seems. For instance, if it is known that over a certain region of the grid  $\Delta f$  is zero for the time period in question, then two grids may be joined here without loss of conservation, since the mass increment is zero. Similarly if it is known that  $|g_{k-s}| > |g_k|$  over a region of the grid for the whole time period, then again two grids may be joined without loss of conservation since the scheme will be in effect the upwind (4.18), on this part of the grid, which is conservative on all grids.



### §5. Monotonicity Preservation on an Irregular grid

The switched scheme on an irregular grid can be stated explicitly as

$$u^k = u_k + \alpha_{k+1} g_{k+1-s_{k+1}} + \bar{\alpha}_{k+1} (g_{k+1} - g_{k+1-s_{k+1}}) + \beta_k g_{k-s_k} + \bar{\beta}_k (g_k - g_{k-s_k}) \quad (5.1)$$

where

$$s_k = \text{sgn}(v_k), \quad v_k = \frac{\Delta t}{\Delta x_k}$$

$$g_k = -v_k \Delta u_k, \quad R_k = \frac{\Delta x_k}{\Delta x_{k-1}}$$

$$\alpha_k = \frac{1 - R_k v_k}{1 + R_k}, \quad \beta_k = \frac{R_{k+1} + v_k}{1 + R_{k+1}} \quad (5.1a)$$

$$\bar{\alpha}_k = \begin{cases} \alpha_k & |g_k| < |g_{k-s_k}| \\ \frac{1}{2}(1 - s_k) & |g_k| \geq |g_{k-s_k}| \end{cases}$$

$$\bar{\beta}_k = \begin{cases} \beta_k & |g_k| < |g_{k-s_k}| \\ \frac{1}{2}(1 + s_k) & |g_k| \geq |g_{k-s_k}| \end{cases}$$

(Note  $v_{k-1} = R_k v_k$ )

i.e.

$$\alpha_k g_{k-s_k} + \bar{\alpha}_k (g_k - g_{k-s_k}) \quad \leftarrow \quad \beta_k g_{k-s_k} + \bar{\beta}_k (g_k - g_{k-s_k})$$

We can extend the scheme to the nonlinear equation by writing

$$v_k = \frac{\Delta t}{\Delta x_k} \frac{\Delta f_k}{\Delta u_k}, \quad \Delta u_k \neq 0 \quad \text{i.e. approximate } \frac{\partial f}{\partial u} \text{ by } \frac{\Delta f}{\Delta u}$$

$$g_k = -\frac{\Delta t}{\Delta x_k} \Delta f_k \quad (5.2)$$

If  $\Delta u_k = 0$ ,  $g_k$  is simply zero (assuming  $f(x)$  is single valued), and The previous known value of  $v$  is used in the weights. If  $\Delta f = 0$  but  $\Delta u \neq 0$  then a new estimate for  $v$  will need to be taken if followed by  $\Delta u = 0$ .

We now state and prove the following theorem

Theorem 2

Given a region of the grid on which the data  $\{u_k\}$  is monotonic and the local C.F.L.  $v$  does not change in sign over the time step  $\Delta t$ , then the solution set  $\{u^k\}$  produced by (5.1) is

(i) monotonic

(ii) such that  $|g^k| \geq \min\{|g_k|, |g_{k-s_k}|\}$

provided the C.F.L. condition  $|v| \leq 1$  is satisfied.

Proof

On such a region of the grid, since  $v$  is of constant sign we can drop the subscript on  $s$ .

$$\begin{aligned}
 g^k &= -v_k(u^k - u^{k-1}) \\
 &= g_k - v_k\{\alpha_{k+1}g_{k+1-s} + \bar{\alpha}_{k+1}(g_{k+1} - g_{k+1-s}) + \beta_k g_{k-s} \\
 &\quad + \bar{\beta}_k(g_k - g_{k-s}) - \alpha_k g_{k-s} - \bar{\alpha}_k(g_k - g_{k-s}) - \beta_{k-1}g_{k-1-s} \\
 &\quad - \bar{\beta}_k(g_{k-1} - g_{k-1-s})\} \\
 &\quad \text{(using 5.1)} \\
 &= [1 + v_k(\bar{\alpha}_k - \bar{\beta}_k)]g_k + v_k\{(\bar{\alpha}_{k+1} - \alpha_{k+1})(g_{k+1-s} - g_{k+1}) \\
 &\quad - \alpha_{k+1}g_{k+1} + (\bar{\beta}_k - \beta_k - \bar{\alpha}_k + \alpha_k)g_{k-s} + \beta_{k-1}g_{k-1} \\
 &\quad + (\bar{\beta}_{k-1} - \beta_{k-1})(g_{k-1} - g_{k-1-s})\}. \tag{5.2}
 \end{aligned}$$

Take first the case  $v > 0$ , i.e.  $s = 1$ , then

$$\begin{aligned}
 g^k &= \{1 + v_k(\bar{\alpha}_k - \bar{\beta}_k - \alpha_{k+1})\}g_k + v_k\{(\bar{\beta}_k - \bar{\alpha}_k + \alpha_{k+1})g_{k-1} \\
 &\quad + \bar{\alpha}_{k+1}(g_k - g_{k+1}) + (\bar{\beta}_{k-1} - \beta_{k-1})(g_{k-1} - g_{k-2})\} \tag{5.3} \\
 &\quad \text{(using } \alpha_{k+1} + \beta_k = 1)
 \end{aligned}$$

As before we use the fact that for monotonic  $u$ , the  $g$ 's are of constant sign. The last two terms in the square brackets are of the same sign as of  $k$

and  $g_{k-1}$  by the definition of  $\bar{\alpha}_{k+1}$  and  $\bar{\beta}_{k-1}$ . The first term in the square brackets is either  $(1 - \alpha_k)g_{k-1}$  if there is no switch or  $(1 + \alpha_{k+1})g_{k-1}$  if there is. Since  $R_k v_k = v_{k-1} \leq 1$  (to satisfy C.F.L. condition) we have  $0 \leq \alpha_k \leq 1$  and hence the whole of [.....] is of the same sign as either  $g_k$  or  $g_{k-1}$ .

Consider now the term in curly brackets. If there is no switch this becomes  $1 - v_k(1 - \alpha_k)$ , and since both  $v_k$  and  $1 - \alpha_k$  are  $< 1$  the term is positive. If there is a switch then we have

$$\begin{aligned} & 1 - v_k(1 + \alpha_{k+1}) \\ &= 1 - v_k(2 - \beta_k) \\ &= 1 - v_k \frac{(2 + R_{k+1} - v_k)}{1 + R_{k+1}} \end{aligned}$$

$$= F(v_k) \text{ say.}$$

$$\text{Now } F(0) = 1, F(1) = 0 \text{ and also } F'(v_k) = \frac{2(v_k - 1) - R_{k+1}}{1 + R_{k+1}} \neq 0$$

in interval  $[0, 1]$ ; hence  $F(v_k) \geq 0$  for  $v_k \in [0, 1]$  and thus the whole r.h.s. of (5.3) is of the same sign as  $g_k, g_{k-1}$  - proving (i) for  $v > 0$ .

Now consider the case where  $g_k, g_{k-1} \leq K \leq 0$ . (5.3) gives

$$g^k < [1 + v_k(\bar{\alpha}_k - \bar{\beta}_k - \alpha_{k+1})]K + v_k(\bar{\beta}_k - \bar{\alpha}_k + \alpha_{k+1})K + \text{-ve terms}$$

$$\text{i.e. } g^k < K$$

Similarly if  $g_k, g_{k-1} \geq L \geq 0$  we get  $g^k > L$ .

These two combined give

$$|g^k| > \min\{|g_k|, |g_{k-1}|\}$$

for  $v > 0$ .

If  $v < 0$ , then  $s = -1$  and (5.2) becomes

$$\begin{aligned} g^k &= [1 - v_k(\bar{\beta}_k - \bar{\alpha}_k - \beta_{k-1})]g_k - v_k\{(\beta_{k-1} - \bar{\beta}_k + \bar{\alpha}_k)g_{k+1} \\ &\quad + \bar{\beta}_{k-1}(g_k - g_{k-1}) + (\alpha_{k+1} - \bar{\alpha}_{k+1})(g_{k+2} - g_{k+1})\} \end{aligned} \quad (5.4)$$

Both (i) and (ii) follow by similar arguments as for  $v > 0$  thus completing the proof.

§6. Numerical results for the equation  $u_t + (\frac{1}{2}u^2)_x = 0$

In this section various initial-value test problems are used to illustrate the performance of the scheme (5.1). These test problems are the same as those used by P. L. Roe to demonstrate his modified version of MacCormack's Algorithm [8]. The results are presented graphically in Appendix B. For each example both grids with increasing and decreasing intervals are used. The grid points  $X(I)$  and grid intervals  $DX(I)$  are generated by:

$$\left. \begin{aligned} DX(I) &= DX(I-1) + R \\ X(I) &= X(I-1) + DX(I) \end{aligned} \right\} I = 3, 4, \dots \quad (6.1)$$

where  $R$ ,  $DX(2)$  and  $X(1)$  are given. In the examples chosen  $x$  is taken in the range  $-2 \leq x \leq 2$ , i.e.  $X(1) = -2$ , with

$DX(2) = 0.01$  and  $R = 0.005$  for the increasing grid  
and  $DX(2) = 0.2$  and  $R = -0.005$  for the decreasing grid.

The problems were as follows:

Problem 1 - Shock wave.

The exact solution of the problem with initial data

$$\begin{aligned} u(x, 0) &= 1 & x < 0 \\ u(x, 0) &= 0 & x > 0 \end{aligned} \quad (6.2)$$

$$\text{is } \begin{aligned} u(x, t) &= 1 & x < \frac{1}{2}t \\ u(x, t) &= 0 & x > \frac{1}{2}t \end{aligned} \quad (6.3)$$

The problem was run for  $\Delta t = 0.005$  and  $0.009$  giving a maximum C.F.L. number of  $0.5$  and  $0.9$  respectively. The results are shown in graphs B1-B4. As can be seen comparing B1 with B2 and B3 with B4 the irregular grid appears not to affect the solution. In each case the exact solution is superimposed on the final timestep for easy comparison, and the profile is reproduced correctly apart from a one point discrepancy. In graphs B5 and B6 the unswitched scheme is used, and as can be seen spurious oscillations soon develop behind the shock, these oscillations being absent from the switched scheme.

Problem 2 - An abrupt expansion.

The initial data

$$\begin{aligned} u(x, 0) &= 0 & x < 0 \\ u(x, 0) &= 1 & x > 0 \end{aligned} \quad (6.4)$$

has the exact solution

$$\begin{aligned} u(x, t) &= 0 & x < 0 \\ u(x, t) &= x/t & 0 < x < t \\ u(x, t) &= 1 & x > t \end{aligned} \quad (6.5)$$

The results are shown in graphs B7 and B8. The scheme gives satisfactory results, the greatest deviation being a slight rounding of the right hand 'corner'. Again the irregularity of the grid appears not to affect the solution, and other runs with different  $\Delta t$ 's gave similar results. The unswitched scheme is shown in B9 and B10 where it can be seen that it is again having difficulties with spurious oscillations which affect the profile of the expansion.

Problem 3 - A shock collision.

Consider the initial data

$$\begin{aligned} u(x, 0) &= 1 & x < 0 \\ u(x, 0) &= 0 & 0 < x < 1 \\ u(x, 0) &= -2 & x > 1 \end{aligned} \quad (6.6)$$

The solution for  $0 < t < 2/3$  is

$$\begin{aligned} u(x, t) &= 1 & x < t/2 \\ u(x, t) &= 0 & t/2 < x < 1-t \\ u(x, t) &= -2 & x > 1-t \end{aligned} \quad (6.7)$$

while for  $t > 2/3$  it is

$$\begin{aligned} u(x, t) &= 1 & x < 2/3 - t/2 \\ u(x, t) &= -2 & x > 2/3 - t/2 \end{aligned} \quad (6.8)$$

(6.7) represents two shock waves approaching each other while (6.8) is the solution when the shock waves have merged into one.

Graphs B 11 and B 12 show the results, with the exact solution superimposed

at times  $t = 0.5$  and  $t = 1.75$ . The method using the increasing grid, B11, resolves the shocks better, with a one point discrepancy at each shock whereas the decreasing grid does not perform quite as well.

Problem 4.

We take now the less simple problem with initial data

$$\begin{aligned} u(x, 0) &= -2x & -1 < x < 0 \\ u(x, 0) &= 2x & 0 < x < 1 \\ u(x, 0) &= 2 & x < -1; x > 1 \end{aligned} \quad (6.9)$$

The compression between  $-1$  and  $0$  steepens until at  $t = \frac{1}{2}$  it becomes a shock wave. This interacts with the expansion and weakens, and then follows a curved path in the  $x-t$  plane.

i.e. the solution for  $0 < t < \frac{1}{2}$  is

$$\begin{aligned} u(x, t) &= 2x/(2t-1) & 2t-1 < x < 0 \\ u(x, t) &= 2x/(2t+1) & 0 < x < 2t+1 \\ u(x, t) &= 2 & x < 2t-1, x > 2t+1 \end{aligned} \quad (6.10)$$

and for  $t > \frac{1}{2}$

$$\begin{aligned} u(x, t) &= 2x/(2t+1) & x_s(t) < x < (2t+1) \\ u(x, t) &= 2 & x < x_s(t); x > (2t+1) \end{aligned} \quad (6.11)$$

where  $x_s(t) = 1 + 2t - (2(2t+1))^{1/2}$ .

As can be seen from B13 and B14 results on the decreasing grid follow the solution more accurately, although neither is exact in the region of the shock.

Problem 5.

The final example is a combination of Problem 1 and Problem 2. The initial data is

$$\begin{aligned} u(x, 0) &= 1 & x < -0.5 \\ u(x, 0) &= 0 & -0.5 < x < 0.5 \\ u(x, 0) &= 1 & x > 0.5 \end{aligned} \quad (6.12)$$

B15 shows the results, the one-point deviation at the shock and the rounding of the expansion, as seen in Problems 1 and 2, being apparent. The results were

the same for both grids here and so only one is reproduced here.

It is not claimed that an irregular grid has advantages where a regular grid would suffice, but that should an irregular grid be desirable then the scheme (5.1) gives good results without any spurious oscillations around the discontinuities of the solution.

Also reproduced in B16 and B17 are the results obtained when Sod's [3] well known Shock Tube problem when solved using Roe's second order scheme on a regular grid and the scheme of Section 4 on an irregular grid respectively. This demonstrates the straightforward extension to systems of conservation laws of these schemes.

### §7. A Three Parameter Scheme

As mentioned in Section 4, the centred 2nd order scheme (4.7), (4.8) is conservative, only on arithmetic grids, i.e. grids with a linear relation between their cell lengths. This is not surprising, since there are three non-trivial conditions to be satisfied, conservation and first and second order accuracy, but only two weights. Therefore to achieve second order accuracy and conservation on a general grid, a minimum of three weights will be needed.

Such a scheme with three parameters will have the form

$$u^k = u_k + \alpha_{k+1}g_{k+1} + \beta_k g_k + \gamma_{k-1}g_{k-1} \quad (7.1)$$

The conditions to be satisfied are

first order accuracy

$$\alpha_{k+1} + \beta_k + \gamma_{k-1} = 1 \quad (7.2)$$

second order accuracy

$$-\Delta x_{k+1}\alpha_{k+1} + \Delta x_k\beta_k + (\Delta x_{k-1} + 2\Delta x_k)\gamma_{k-1} \quad (7.3)$$

conservation

$$(\Delta x_{k-1} + \Delta x_k)\alpha_k + (\Delta x_k + \Delta x_{k+1})\beta_k + (\Delta x_{k+1} + \Delta x_{k+2})\gamma_k \quad (7.4)$$

If we solve (7.2) - (7.4) we obtain for the  $\alpha$ s and  $\beta$ s

$$\alpha_k = \frac{\Delta x_{k-1} - a\Delta t + (\Delta x_{k-2} + \Delta x_k)\gamma_{k-2}}{(\Delta x_{k-1} + \Delta x_k)} \quad (7.5)$$

$$\beta_k = \frac{\Delta x_{k+1} + a\Delta t - (\Delta x_{k+1} + 2\Delta x_k + \Delta x_{k-1})\gamma_{k-1}}{(\Delta x_k + \Delta x_{k+1})}$$

while for the  $\gamma$ s we have a recurrence relation,

$$\begin{aligned} (\Delta x_{k+1} + \Delta x_{k+2})\gamma_k - (\Delta x_{k+1} + 2\Delta x_k + \Delta x_{k-1})\gamma_{k-1} + (\Delta x_{k-2} + \Delta x_{k-1})\gamma_{k-2} \\ = -\Delta x_{k-1} + 2\Delta x_k - \Delta x_{k+1} \end{aligned} \quad (7.6)$$

where  $\gamma_0, \gamma_1$  are so far arbitrary.

If this scheme is tested numerically it is soon found that for most values of  $\gamma_0, \gamma_1$  the solution very quickly becomes unstable. This can be explained



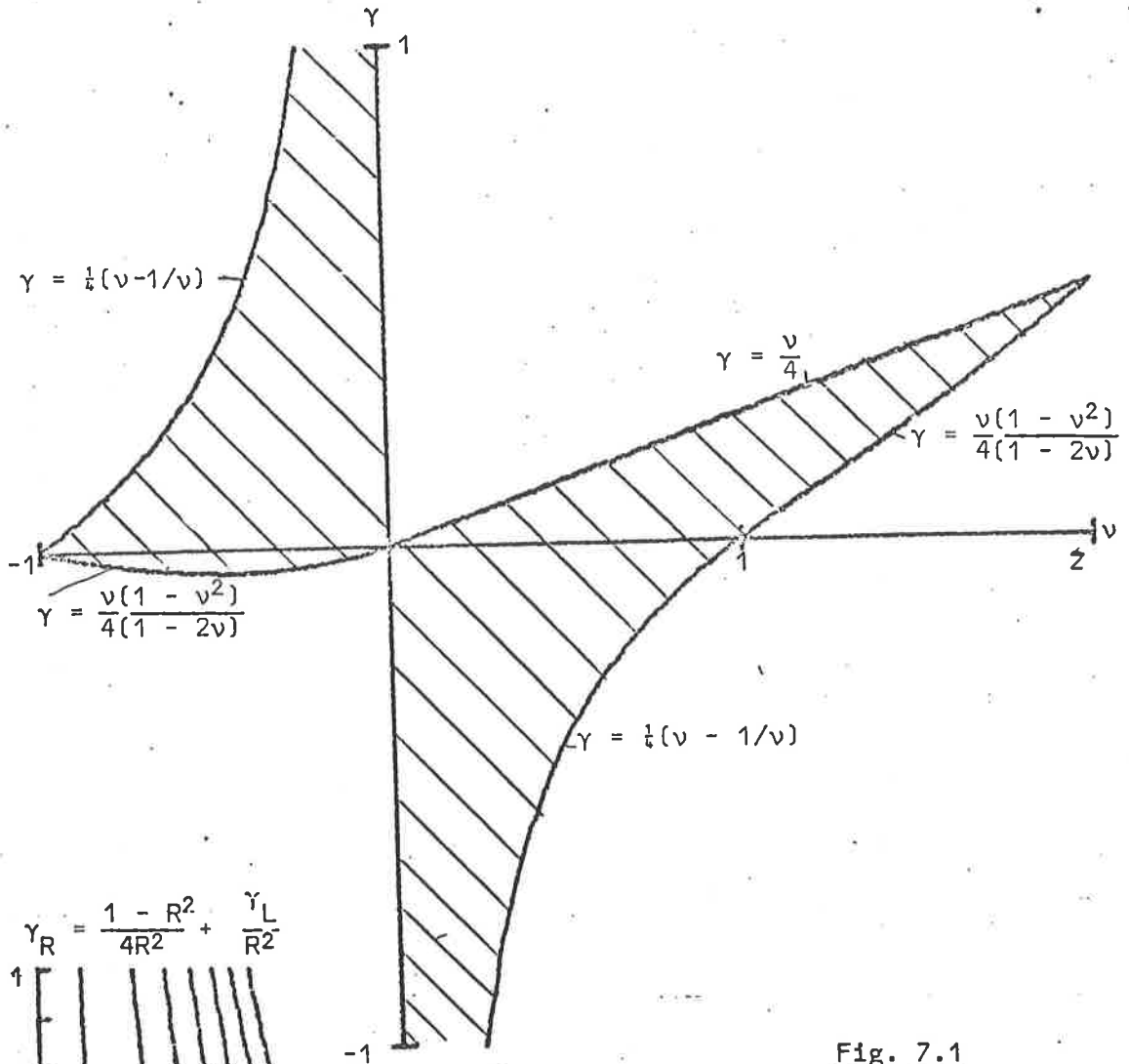


Fig. 7.1

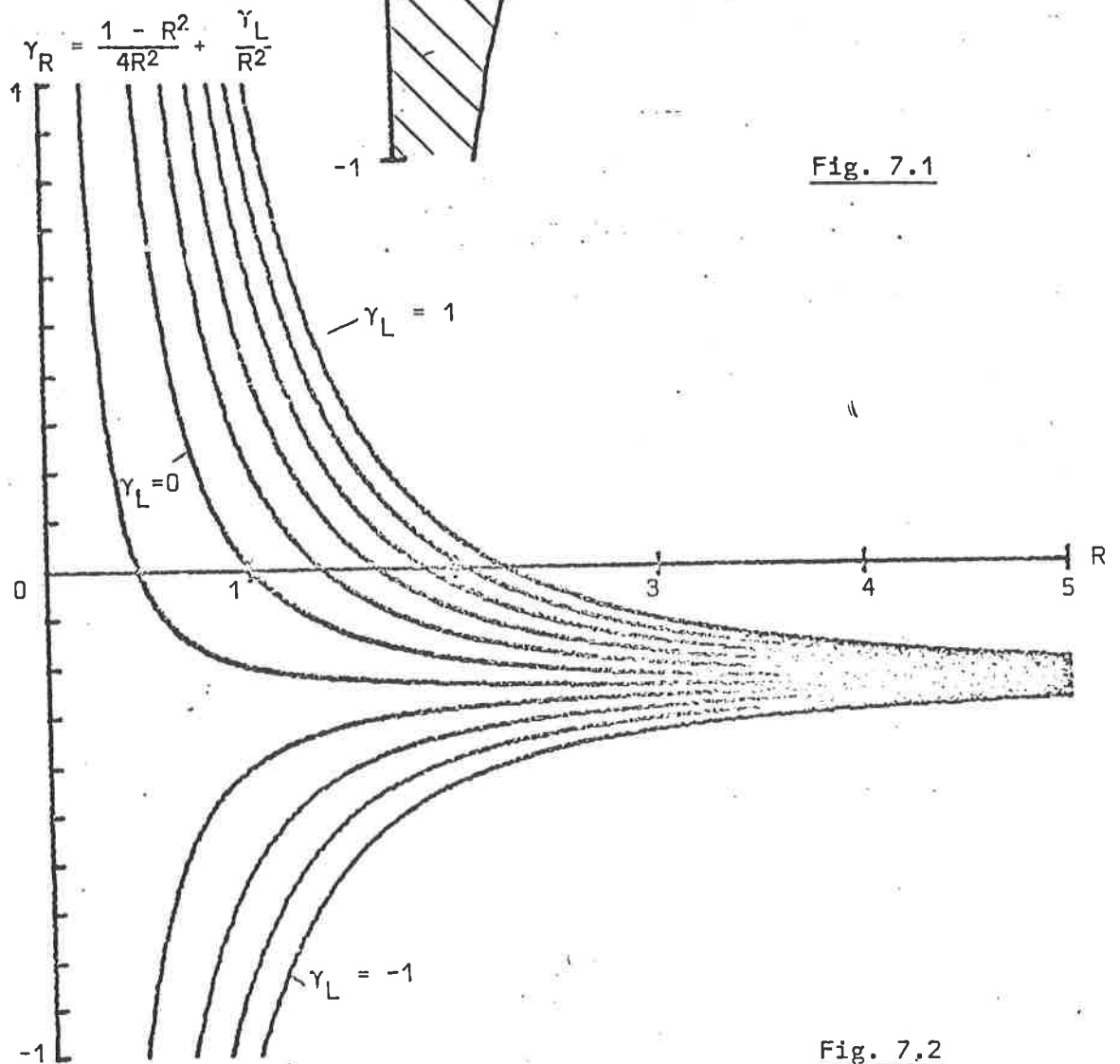


Fig. 7.2

by looking at the linear stability of the scheme. The amplification factor is given by

$$|K|^2 = 1 + 4v^2s^4(v^2 - 1) + 16vs^4\gamma\{2(2\gamma + 1 - v)vs^2 + 1 - 2v\} \quad (7.7)$$

where  $s \equiv \sin \frac{\xi \Delta x}{2}$ .

The stability region in the  $\gamma v$ -plane is given in Figure 7.1 and it can be easily seen that for a general grid it will be extremely difficult to choose  $\gamma_0, \gamma_1$  so that the recurrence relation (7.6) keeps  $\gamma_k$  in the stability region  $\forall k$ . So again the grid is in effect restricted.

One possible application of this scheme is where it may be desired to interface two constant grids. To do this using the two parameter scheme would not ensure conservation unless the join were in a region in which the data/solution had small or zero gradients.

The small infringement of the conservation condition (4.6) would be amplified at each stage, giving inaccurate shock locations. However it can easily be shown that if  $\gamma$  is constant on one of the grids, then except at the join,  $\gamma$  is also constant (different) on the second grid.

The relation between these constants is

$$\gamma_R = \frac{1 - R^2}{4R^2} + \frac{1}{R^2} \gamma_L \quad (7.8)$$

where  $\gamma_L$  is the constant on lefthand grid,  $\gamma_R$  the constant on the right hand grid and  $R = \frac{\Delta x_R}{\Delta x_L}$  is the ratio between the grids.

Hence using a graph, such as Figure 7.2, it is possible to choose  $\gamma_0 = \gamma_1$  appropriately so as to ensure conservation and linear stability on both grids.

Unfortunately, however, this scheme is not monotonicity preserving, and it is not apparent how to extend the method used in the above sections to achieve monotonicity preservation in this case. The scheme will hence suffer from the spurious oscillations often found with second order schemes like Lax-Wendroff.

§8. A Note on Entropy Violation

The equation

$$u_t + f(u)_x = 0, \quad u(x, 0) = \phi(x) \quad (8.1)$$

can be written in the form

$$u_t + a(u)u_x = 0, \quad a(u) = \frac{df}{du} \quad (8.2)$$

This asserts that  $u$  is constant along the characteristics  $x = x(t)$  where

$$\frac{dx}{dt} = a(u) \quad (8.3)$$

Since  $u$  is constant along these characteristics so is  $a(u)$  and hence from (8.3) the characteristics are straight lines, whose slopes depend on the solution. These lines may intersect for nonlinear  $f$ , and where this happens the continuous solution breaks down. Therefore to get existence for all time, weak solutions, which satisfy

$$\int_0^\infty \int_{-\infty}^\infty [w_t u + w_x f(u)] dx dt + \int_{-\infty}^\infty w(x, 0) \phi(x) dx = 0 \quad (8.4)$$

for all continuous  $w(x, t)$ , of compact support (i.e. only non-zero in a finite interval), are admitted.

However, the class of all weak solutions is such that there is no uniqueness for the initial value problem (8.1) and hence an extra condition is required to determine a physically relevant solution. This can be taken as the limit solutions of the parabolic equation

$$u_t + f(u)_x = \epsilon [\beta(u)u_x]_x \quad \beta(u) > 0 \quad \text{as } \epsilon \rightarrow 0+ \quad (8.5)$$

Oleinik [10] showed, that discontinuities of such solutions can be characterized by

$$\frac{f(u) - f(u_L)}{u - u_L} \geq S \geq \frac{f(u) - f(u_R)}{u - u_R} \quad (8.6)$$

for all  $u$  between  $u_L$  and  $u_R$ , where  $S$  is the speed of propagation of the discontinuity and  $u_L, u_R$  are the states on the left and right of the discontinuity. This is called the entropy condition, and weak solutions satisfying

(8.6) are uniquely determined by their initial data, see [4], [5], [6] and others.

Consider now the data (see [7])

$$\begin{aligned} u(x) &= -1 & x < 0 \\ u(x) &= 1 & x > 0 \end{aligned} \tag{8.7}$$

for the equation  $u_t + (\frac{1}{2}u)_x = 0$

$$f(u_L) = f(u_R) = \frac{1}{2}, \quad 0 \leq f(u) \leq \frac{1}{2} \quad \forall u \in (u_L, u_R)$$

The left hand side of (8.6) is negative while the right hand side is positive, so clearly the entropy condition (8.6) is violated. (Here  $S = 0$  by Rankine-Hugoniot relation  $f(u_R) - f(u_L) = S(u_R - u_L)$ )

Now consider any of the schemes mentioned in the previous sections,  $\Delta f_k = 0 \quad \forall k$  and hence at each time step all the increments are zero. Thus (8.7) is produced as a steady state solution and the schemes do not always produce entropy satisfying solutions.

This is due to the fact that for the cell in which the discontinuity lies, the nett increment is zero, whereas if the cell were subdivided so that the division was at the point where  $f'(u) = 0$  (the sonic point) then the 2 increments derived would be equal and opposite and the steady state solution would be disturbed.

This suggests a possible way of overcoming this entropy violation. We proceed as follows.

If  $f_k = 0$  but  $u_k \neq 0$  divide the cell  $k$  into two at the sonic point (obtained by interpolation), calculate new mass increments for each of these cells and distribute these increments with first order weights.

If the discontinuity is entropy satisfying (characteristics pointing towards it) then the increments will go to the imaginary point and cancel. If however the discontinuity is entropy violating then the increments will go to points  $k$  and  $k - 1$  and thus disturb the discontinuity.

It should be noted that this procedure is purely speculative at present and has not been tested.

## §9. Summary

A monotonicity preserving algorithm on an irregular grid has been developed which is conservative <sup>almost everywhere</sup> and, for parts of the flow, second order accurate. There are restrictions as to the grid, but as is pointed out in Section 4 these are not as severe as might at first seem. This scheme may easily be extended to systems of conservation laws, using the Riemann problem approach adopted by P. Roe [2], as demonstrated by graph B17.

The scheme has been run on various test problems and has performed well suppressing oscillations. It may transpire that a pertinent choice of grid will improve the results further.

Two particular lines of approach to be investigated more fully are the three parameter scheme (Section 7) and entropy violation (Section 8).

## Acknowledgement

It is with pleasure that I acknowledge the help given to me by both my supervisors P. L. Roe and Dr. M. J. Baines during this work, and the invaluable help given by Dr. Baines in writing this report.

Appendix A

An Observation on Stability Regions

In a private communication, P.L. Roe made the following observation. On regular grids, any centred 2 parameter first order scheme can be written as

$$u^k = u_k + \lambda g_k + (1 - \lambda)g_{k+1} \quad (\text{A-1})$$

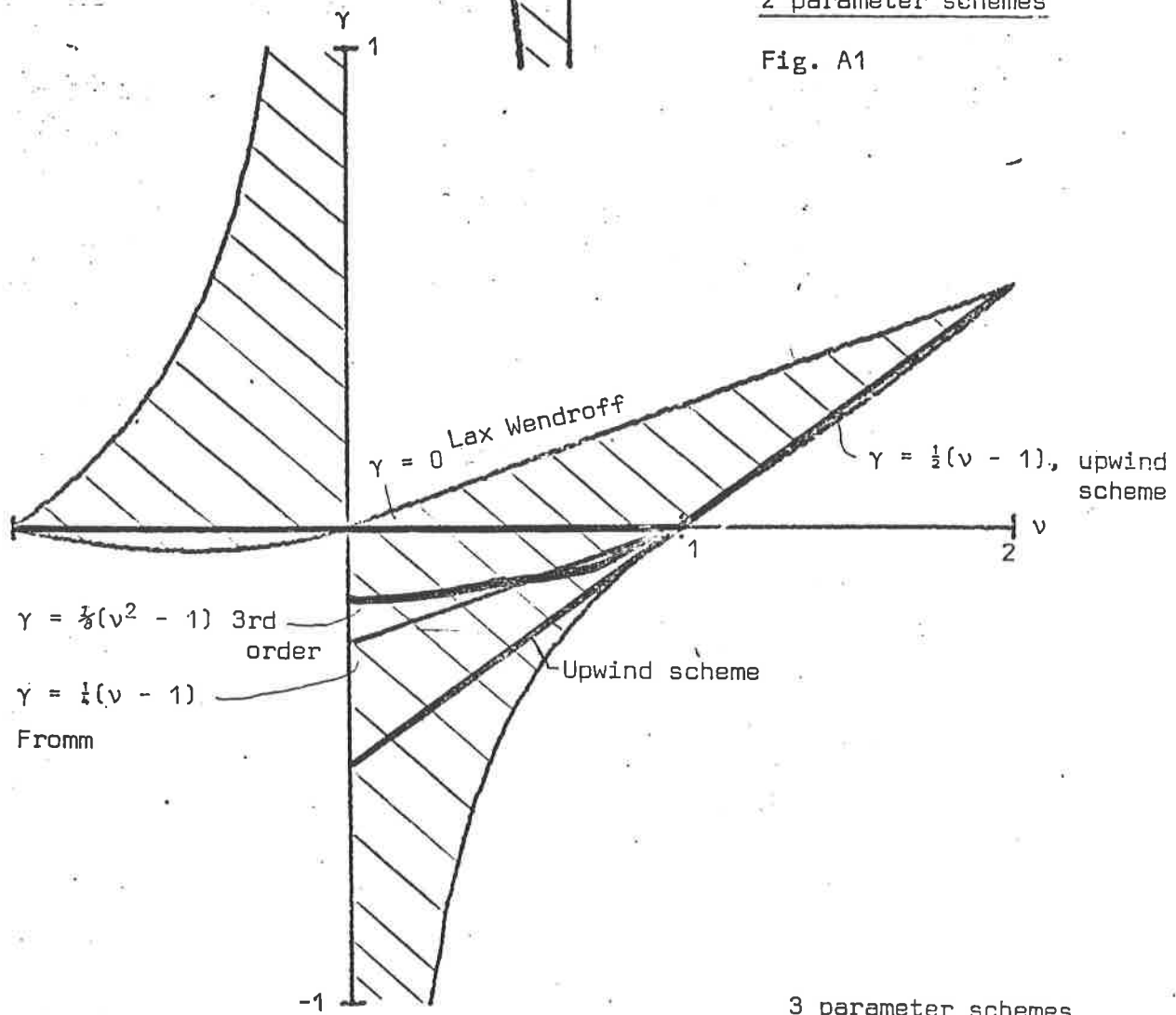
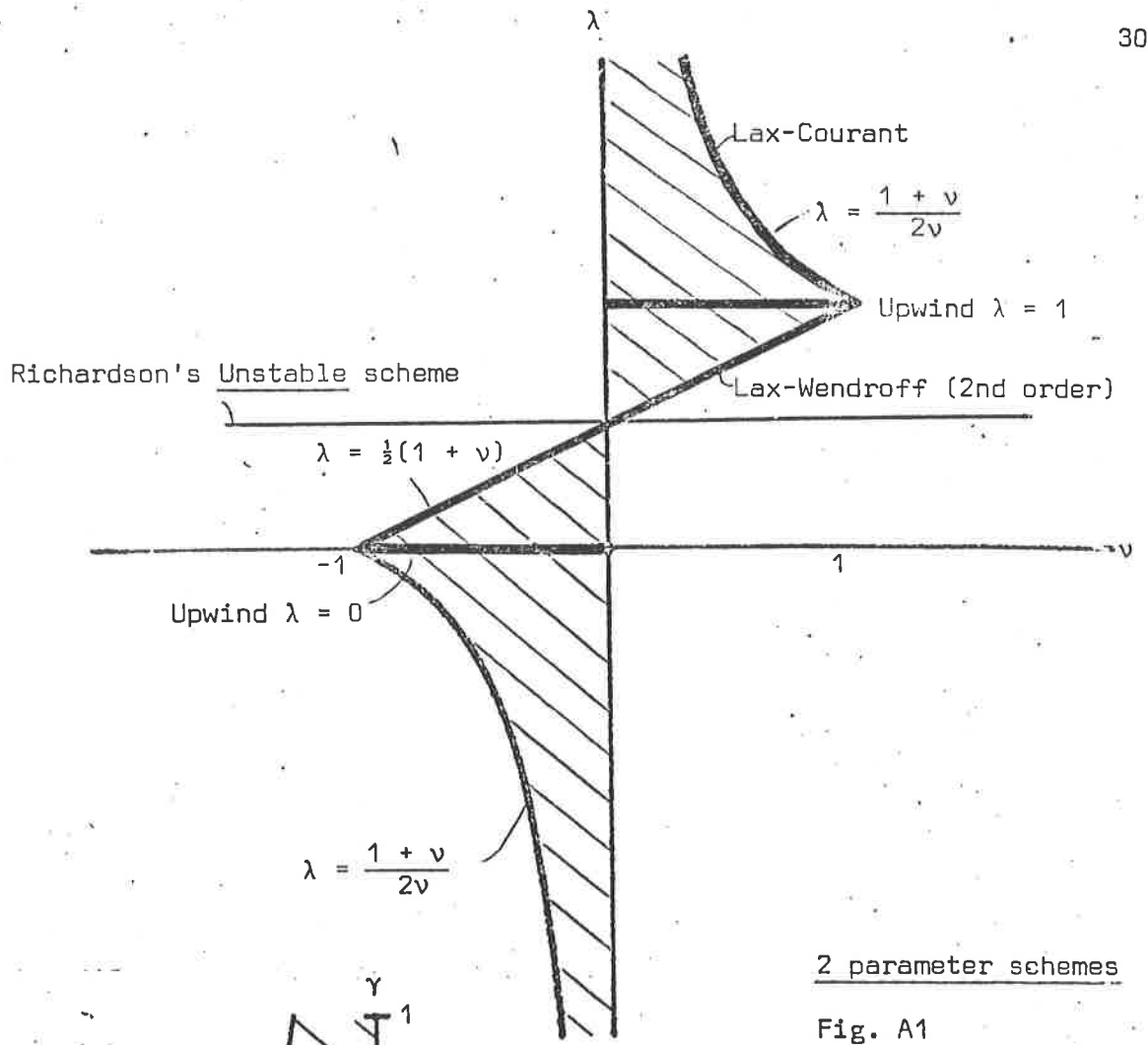
where  $\lambda$  is a parameter.

On analysing the stability, the amplification factor  $K$  satisfies

$$\begin{aligned} |K|^2 = 1 - 2v(2\lambda - 1)(1 - \cos m) + v^2(2\lambda - 1)^2(1 - \cos m)^2 \\ + v^2(1 - \cos^2 m) \end{aligned} \quad (\text{A-2})$$

This gives the stability region shown in Figure A1. Roe goes on to observe that many well-known 2 parameter schemes can be shown as lines on this region, including Lax Wendroff second order scheme and Richardson's unstable scheme.

Following this lead, it is possible to show 3 parameter schemes on the stability region shown in Figure (701) reproduced here in Figure (A2).





## Appendix B

A Fortran program to solve  $u_t + (\frac{1}{2}u^2)_x = 0$  and graphical results

This program has been included to demonstrate how the somewhat complicated looking scheme (5.1) can be easily programmed.

Subroutine STEP implements the scheme (5.1) by first allocating the downward increment  $g_{k-s}$  with second order weights, then testing the gradients and finally allocating the difference  $g_k - g_{k-s}$  with first or second order weights, whichever are appropriate.

The scheme works in such a way that the effective boundary conditions are either transparent (i.e. "no boundary") if the wave is moving into the boundary or constant if the wave is moving away from the boundary (for a discussion on boundary conditions see [9]). There is also a compiling option to have the unswitched scheme.

Subroutine MESH sets up the mesh and determines the number of points, ILIM, used. Subroutine INITIAL initialises  $u$  with the initial solution  $u(x, 0)$ . This subroutine is changed for each problem.

OUTPUT and the other external subroutines are all concerned with graphical output and in no way affect the operation of the scheme.

```

PROGRAM TEST
DIMENSION DX(105),X(105),U(0:105)
COMMON/REGP/ILIM,DT
COMMON/OUT/N,DX2,R

```

```

C PROGRAM TO SOLVE  $U_t + (U*U/2)x = C$ 

```

```

CALL BASH(X,DX)
CALL INITIAL(X,U)
PRINT *, 'OUTPUT EVERY ? STEPS'
PRINT *
READ(1,*)N
CALL FILNAM('TSTPLOT',7)
CALL PAPER(1)
CALL HPOCHR(1)
CALL OUTPUT(X,U,0)

```

```

DO 10 M=1,7*N
CALL STEP(DX,U)
IF(M/N,ER,(M-1)/N) GOTO 10
CALL OUTPUT(X,U,M*DT)
10 CONTINUE

```

```

CALL OREND
STOP
END

```

```

SUBROUTINE STEP(DX,U)
DIMENSION DX(105),U(0:105),U1(0:105),F(0:105)
COMMON/REGP/ILIM,DT
REAL NU

```

```

C ADVANCES SOLUTION ONE TIME STEP.
C CONDITIONALLY COMPILE SYMBOL U FOR UNSWITCHED VERSION

```

```

DO 10 I=0,ILIM+1
U1(I)=U(I)
F(I)=U(I)*U(I)/2
10 CONTINUE

```

```

DO 20 K=2,ILIM
NU=DT*(U1(K)+U1(K-1))/(2*DX(K))
IS=1
IF(NU,LT,0)IS=-1
GKMS=DT*(F(K-1-IS)-F(K-IS))/DX(K-IS)
GK=DT*(F(K-1)-F(K))/DX(K)
ALF=(DX(K-1)-DX(K)*NU)/(DX(K-1)+DX(K))
BET=(DX(K+1)+DX(K)*NU)/(DX(K)+DX(K+1))
U(K-1)=U(K-1)+ALF*GKMS
U(K)=U(K)+BET*GKMS

```

```

CU GOTO 15
IF(ABS(GK),GT,ABS(GKMS)) THEN
ALF=(1-IS)/2.0
BET=(1+IS)/2.0
ENDIF

```

```

15 U(K-1)=U(K-1)+ALF*(GK-GKMS)
U(K)=U(K)+BET*(GK-GKMS)
20 CONTINUE

```

```

U(0)=U(1)
U(ILIM+1)=U(ILIM)

```

```

RETURN
END

```

```

SUBROUTINE OUTPUT(X,U,T)
DIMENSION X(105),U(0:105)
COMMON/REGP/ILIM,DT
COMMON/OUT/N,DX2,R
PARAMETER IO=46

```

```

IF(T.EQ.0.0) THEN
CALL PSPACE(0.0,1.0,0.9,1.0)
CALL MAP(0.0,2.0,0.0,1.0)
CALL PLOTCS(0.2,0.5,'DT=',3)
CALL TYPENF(DT,3)
CALL PLOTNI(0.7,0.5,ILIM,3)
CALL TYPECS('POINTS',6)
CALL PLOTCS(1.0,0.5,'DX(2)=',6)
CALL TYPENF(DX2,3)
CALL PLOTCS(1.6,0.5,'R=',2)
CALL TYPENF(R,4)
CALL PSPACE(0.3,0.7,0.0,0.1)
CALL MAP(-2.0,2.0,-1.0,0.1)
CALL XAXISI(0.5)
ENDIF

```

```

CALL PSPACE(0.0,0.7,0.8-T*0.1/(DT*N),0.9-T*0.1/(DT*N))
CALL MAP(-3.0,4.0,-2.5,2.5)
CALL YAXISI(1.0)
CALL PLOTCS(-2.9,0.5,'T=',2)
CALL TYPENF(T,3)
CALL MAP(-5.0,2.0,-2.5,2.5)
CALL PTPLOT(X,U,1,ILIM,IO)
RETURN
END

```

```

SUBROUTINE MESH(X,DX)
DIMENSION X(105),DX(105)
COMMON/REGP/ILIM,DT
COMMON/OUT/N,DX2,R

```

C      SETS UP MESH

```

PRINT *, 'INPUT DX(2) AND MESH PARAMETER'
PRINT *,
READ(1,*)DX2,R
DX(1)=DX2-R
X(1)=-2.0
I=1

```

```

DO WHILE(X(I).LT.2.0)
I=I+1
DX(I)=DX(I-1)+R
X(I)=X(I-1)+DX(I)
ENDDO
ILIM=I
DX(ILIM+1)=DX(ILIM)+R
X(ILIM+1)=X(ILIM)+DX(ILIM+1)

RETURN
END

```

```
SUBROUTINE INITIAL(X,U)
DIMENSION X(105),U(0:105)
COMMON/REGF/ILIM,DT
```

```
C   SETS UP INITIAL DATA
```

```
PRINT *, 'INPUT DT'
PRINT *,
READ(1,*)DT
```

```
DO 10 I=1,ILIM+1
U(I)=0.0
IF(X(I).LT.-0.5) U(I)=1.0
IF(X(I).GT.0.5) U(I)=1.0
CONTINUE
10  U(0)=U(1)
    U(ILIM+1)=U(ILIM)
```

```
RETURN
END
```

```
END OF FILE
```

```
@
```

DT = 0.005

40POINTS

DX (2) = 0.010

T = 0.000

T = 0.050

T = 0.100

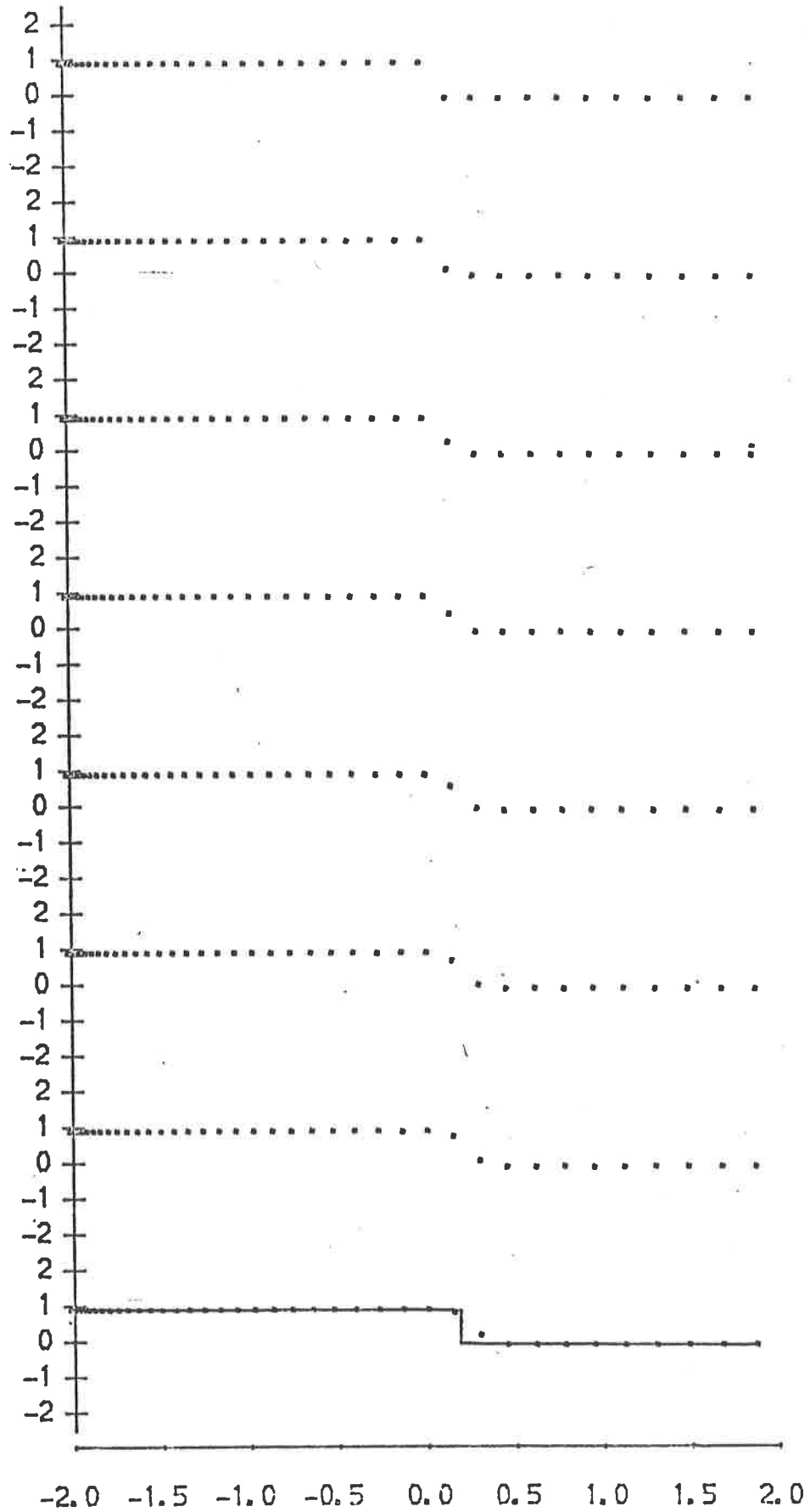
T = 0.150

T = 0.200

T = 0.250

T = 0.300

T = 0.350



Shock problem

B1

DT= 0.005

36POINTS

DX (2) = 0.200

T= 0.000



T= 0.050



T= 0.100



T= 0.150



T= 0.200



T= 0.250



T= 0.300

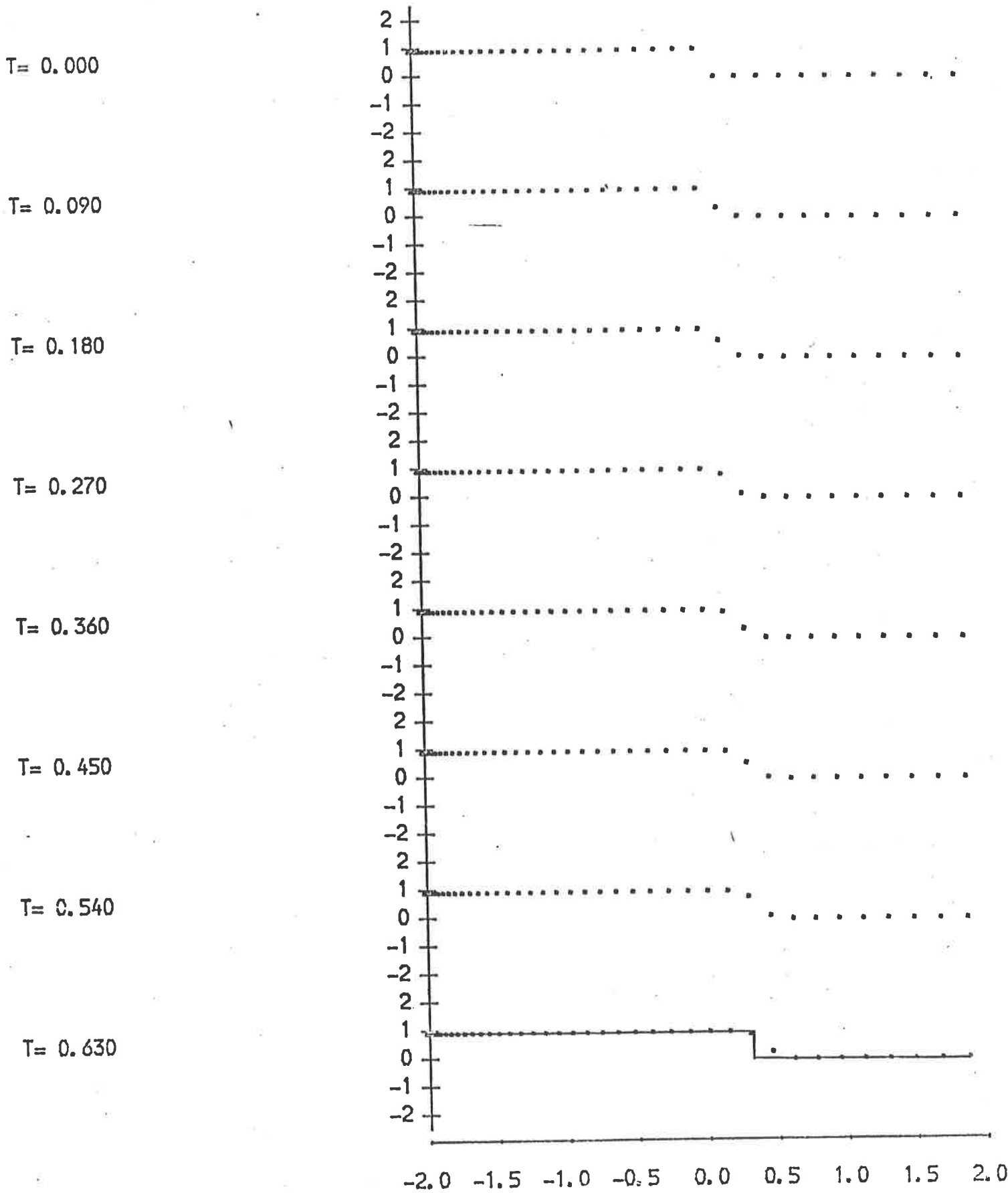


T= 0.350



Shock problem

B2



Shock problem

B3

T= 0.000

T= 0.090

T= 0.180

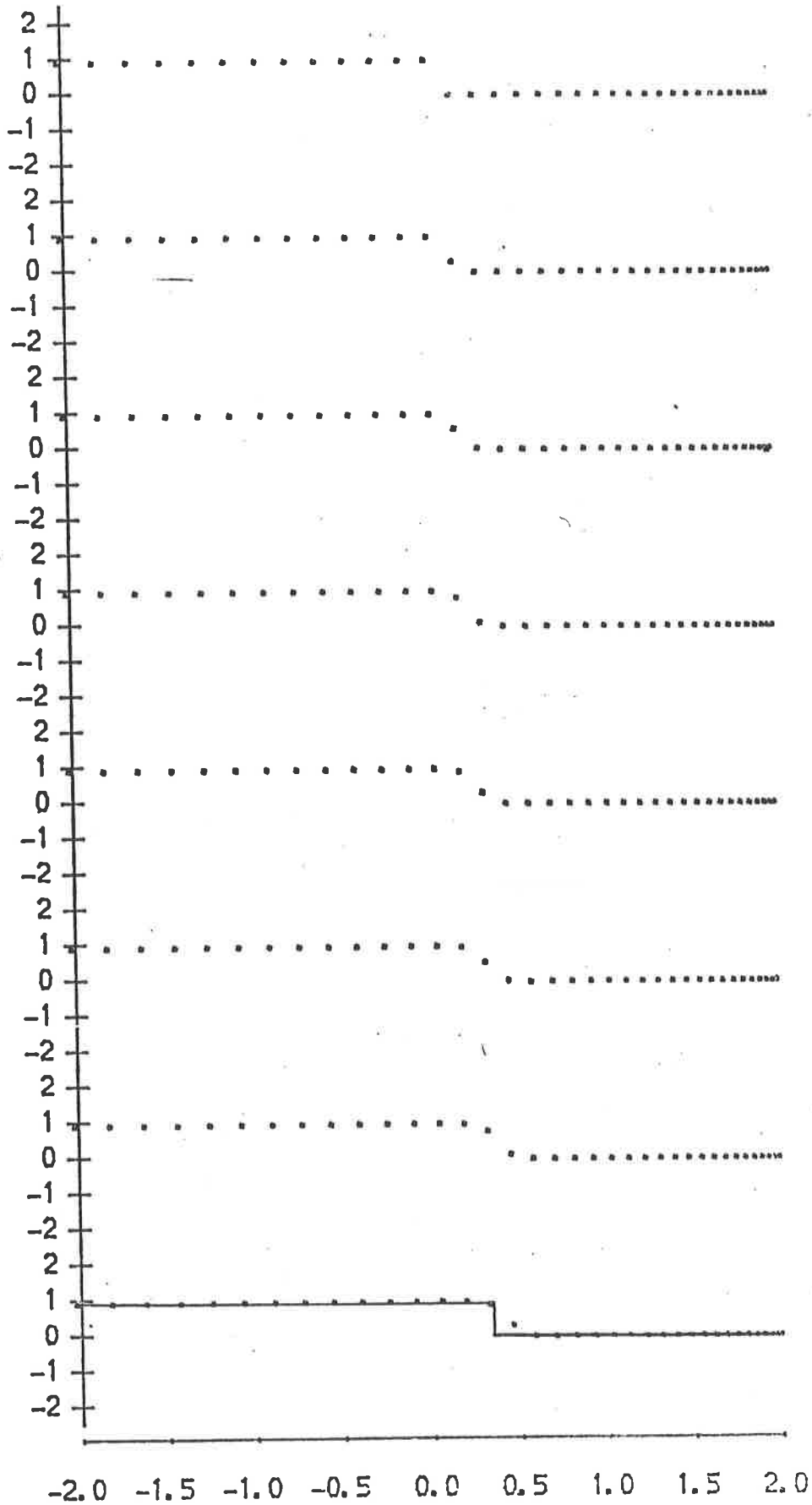
T= 0.270

T= 0.360

T= 0.450

T= 0.540

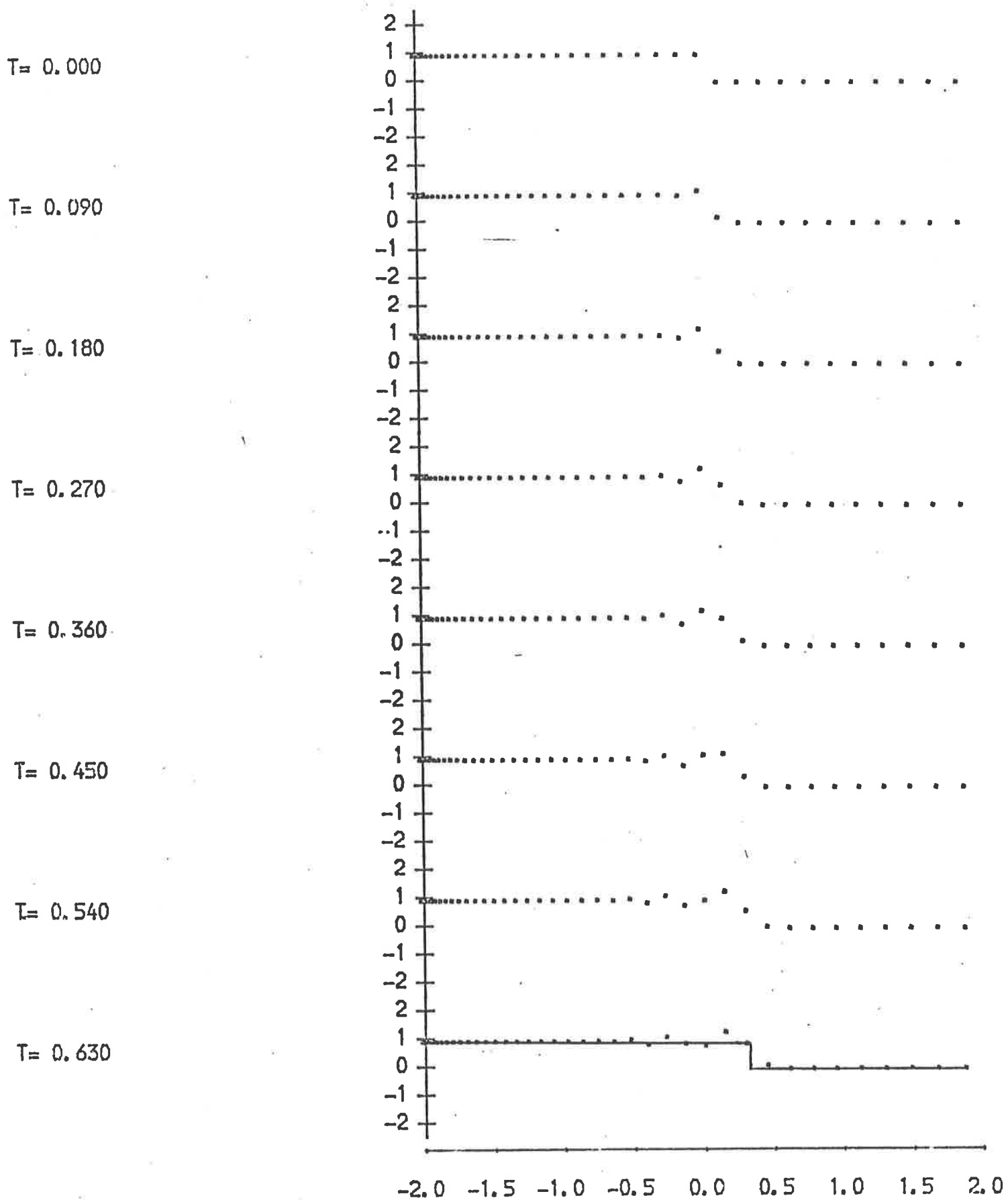
T= 0.630



Shock problem

B4



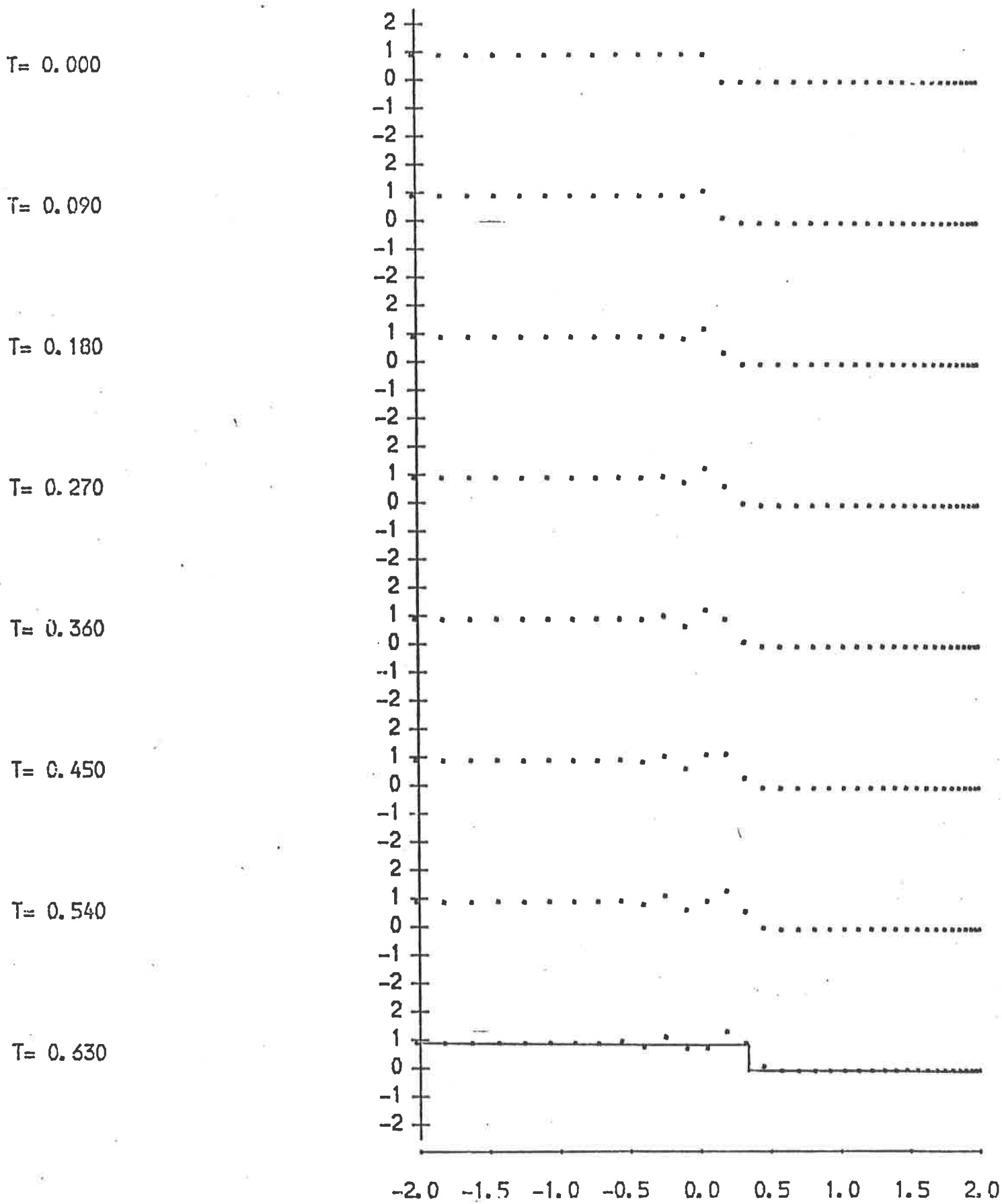


Shock problem - unswitched scheme

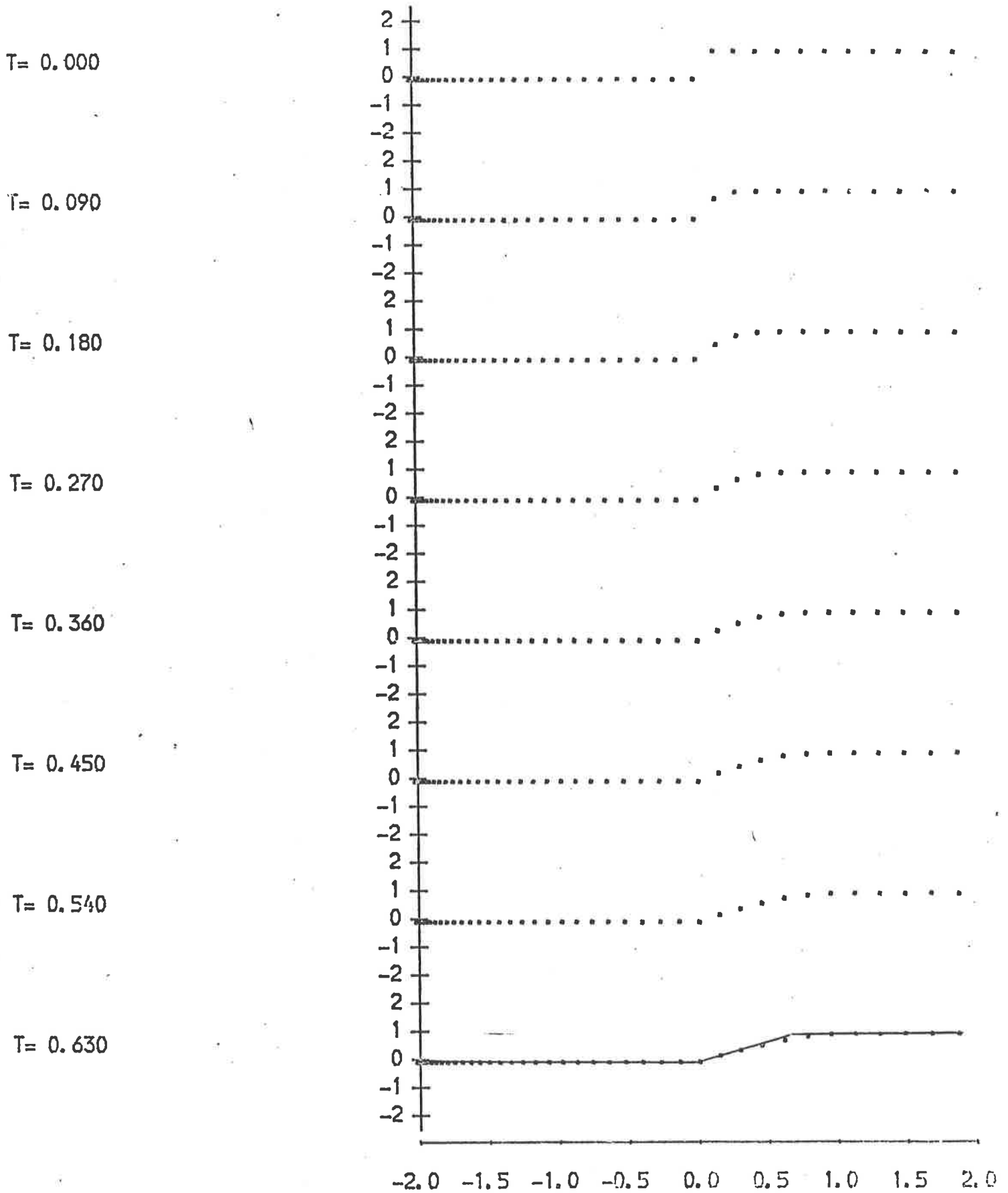
DT= 0.009

36POINTS

DX (2) = 0.200 .



Shock problem - unswitched scheme



Expansion problem

B7

DT= 0.009

36POINTS

DX (2) = 0.200

T= 0.000



T= 0.090



T= 0.180



T= 0.270



T= 0.360



T= 0.450



T= 0.540

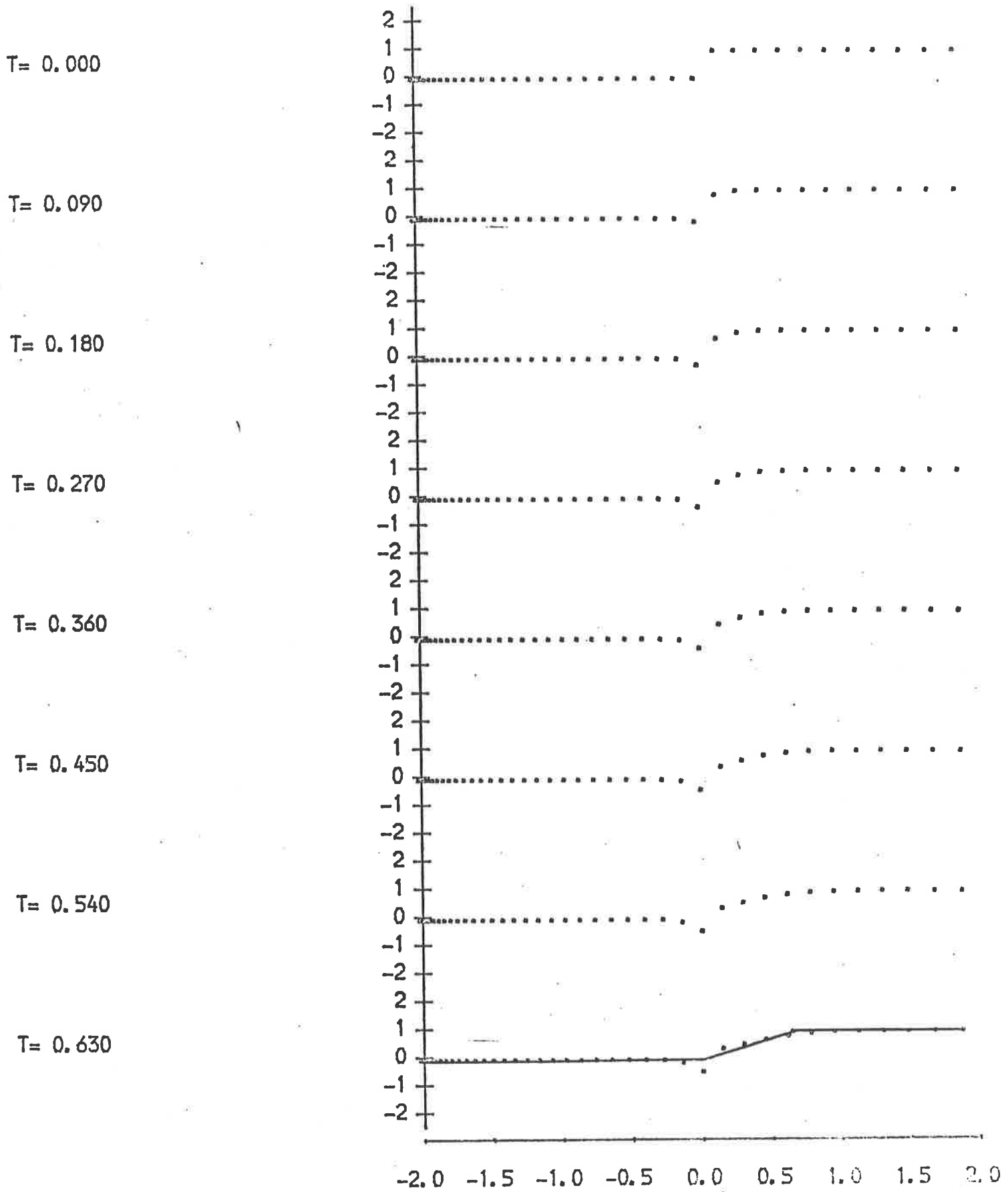


T= 0.630

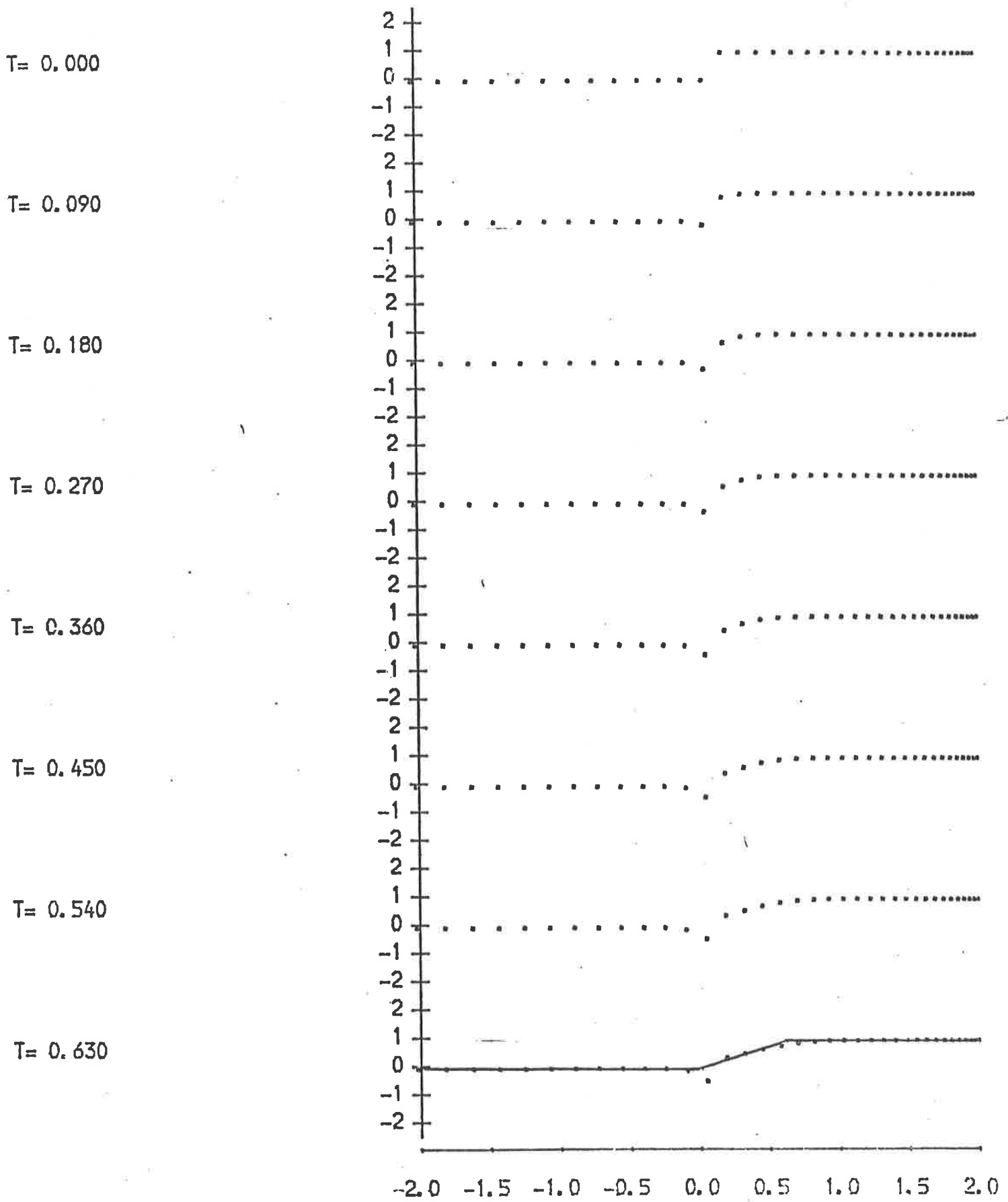


-2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0

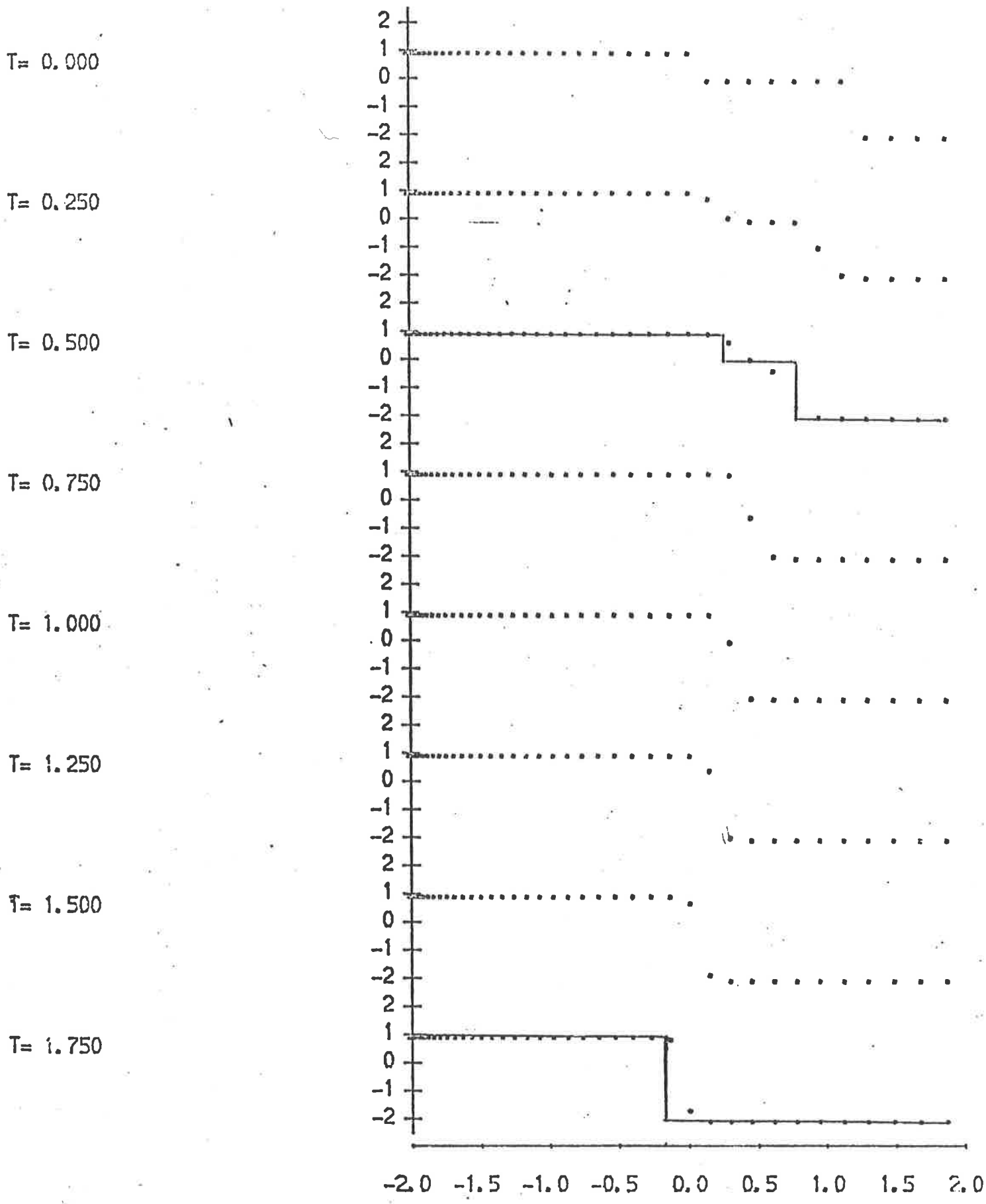
Expansion problem



Expansion problem - unswitched scheme



Expansion problem - unswitched scheme



Shock collision

DT= 0.005

36POINTS

DX (2) = 0.200

T= 0.000



T= 0.250



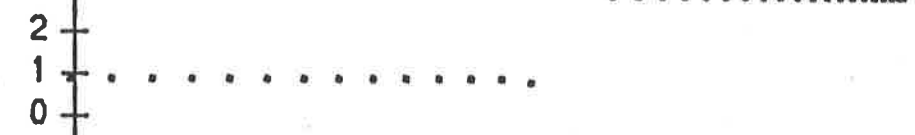
T= 0.500



T= 0.750



T= 1.000



T= 1.250



T= 1.500



T= 1.750

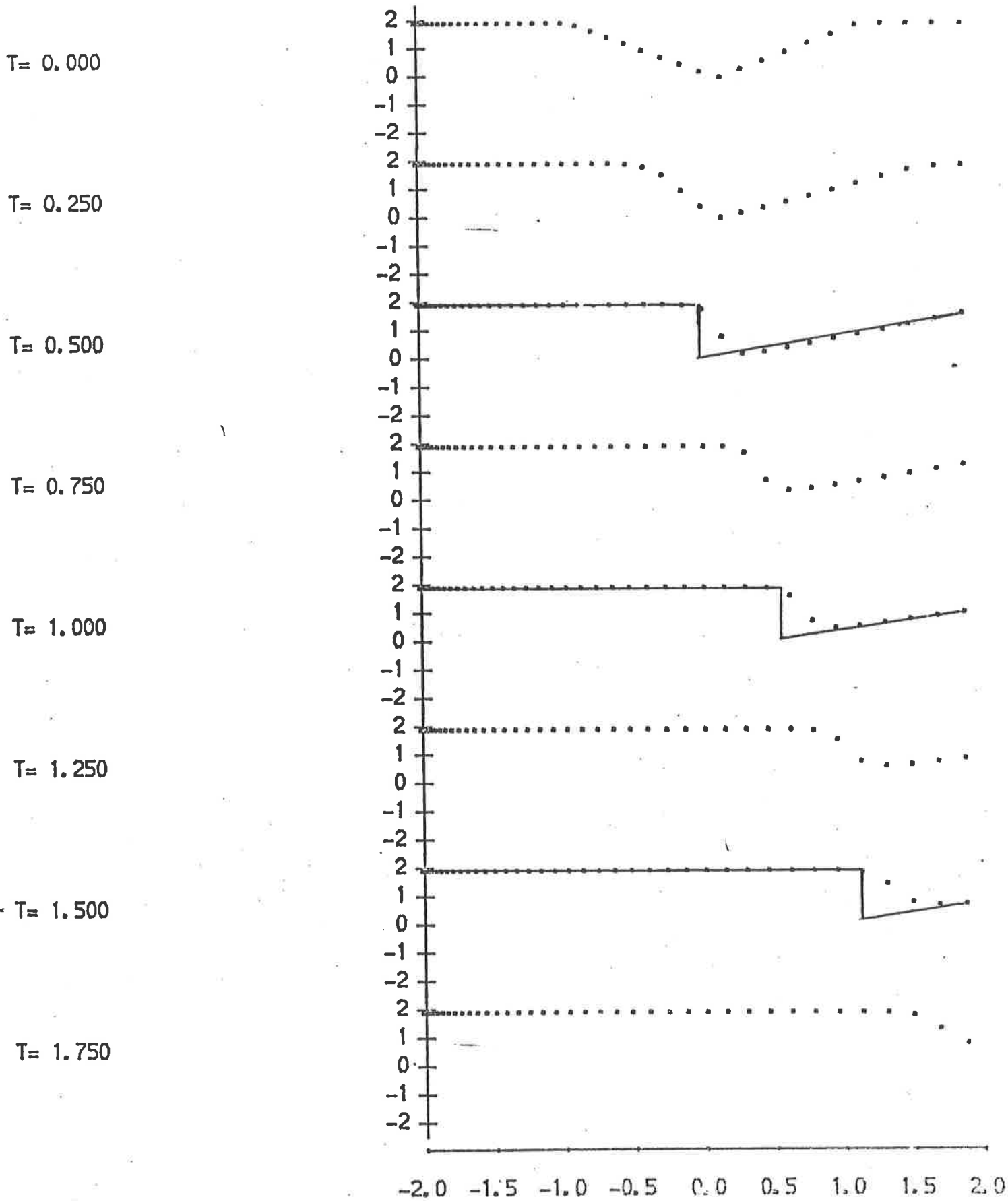


-2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0

Shock collision

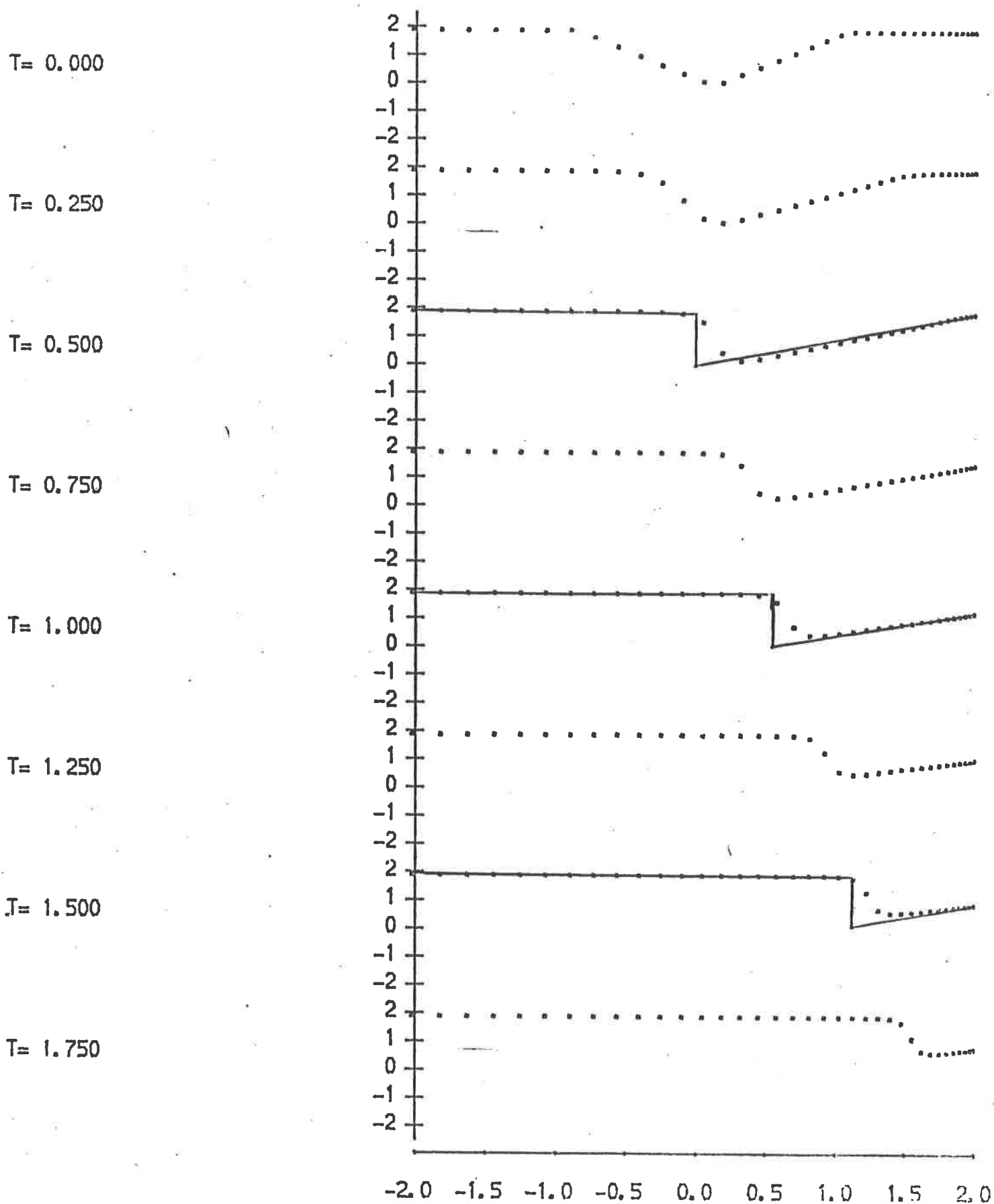
B12





Problem 4

B13



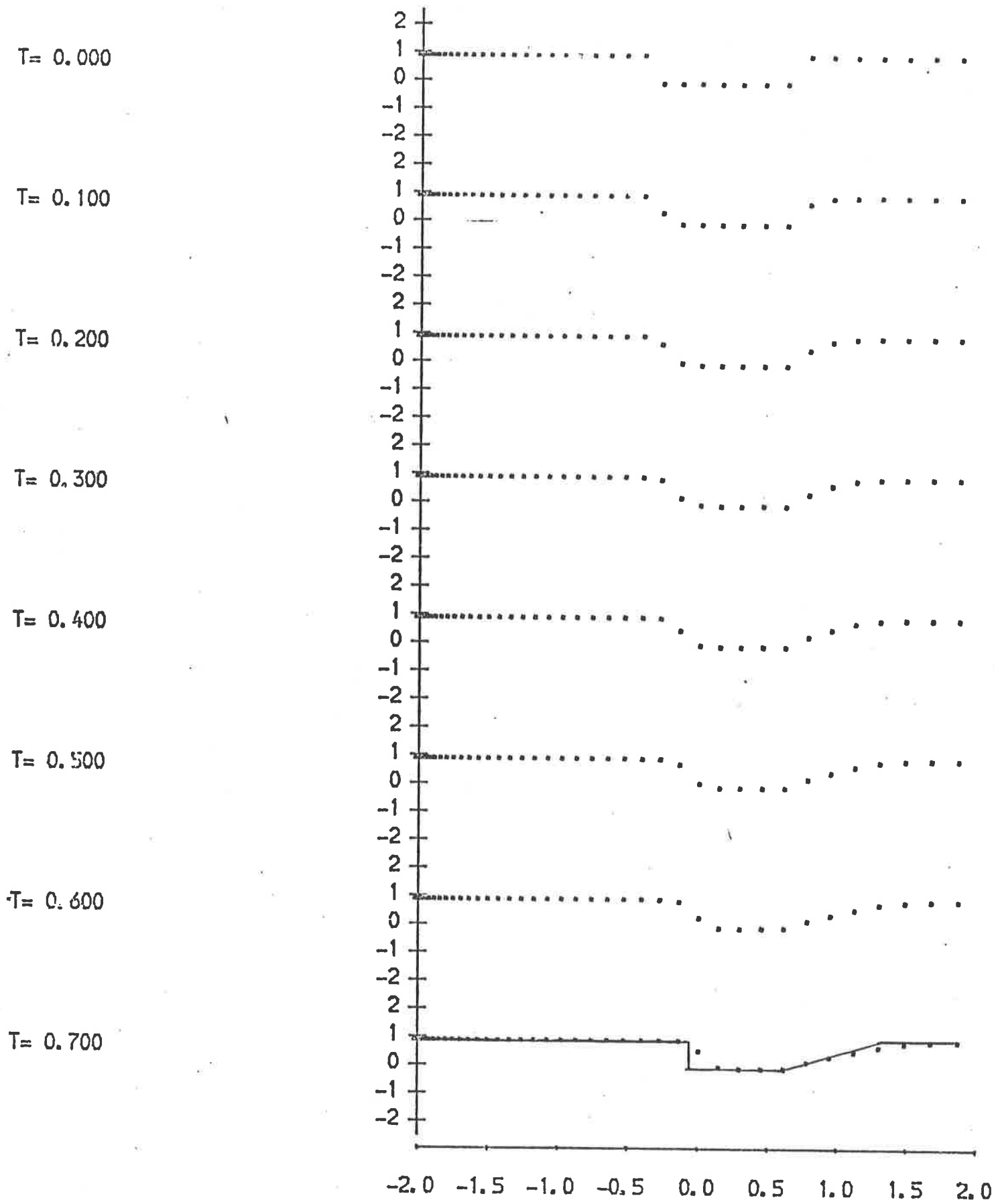
Problem 4

B14

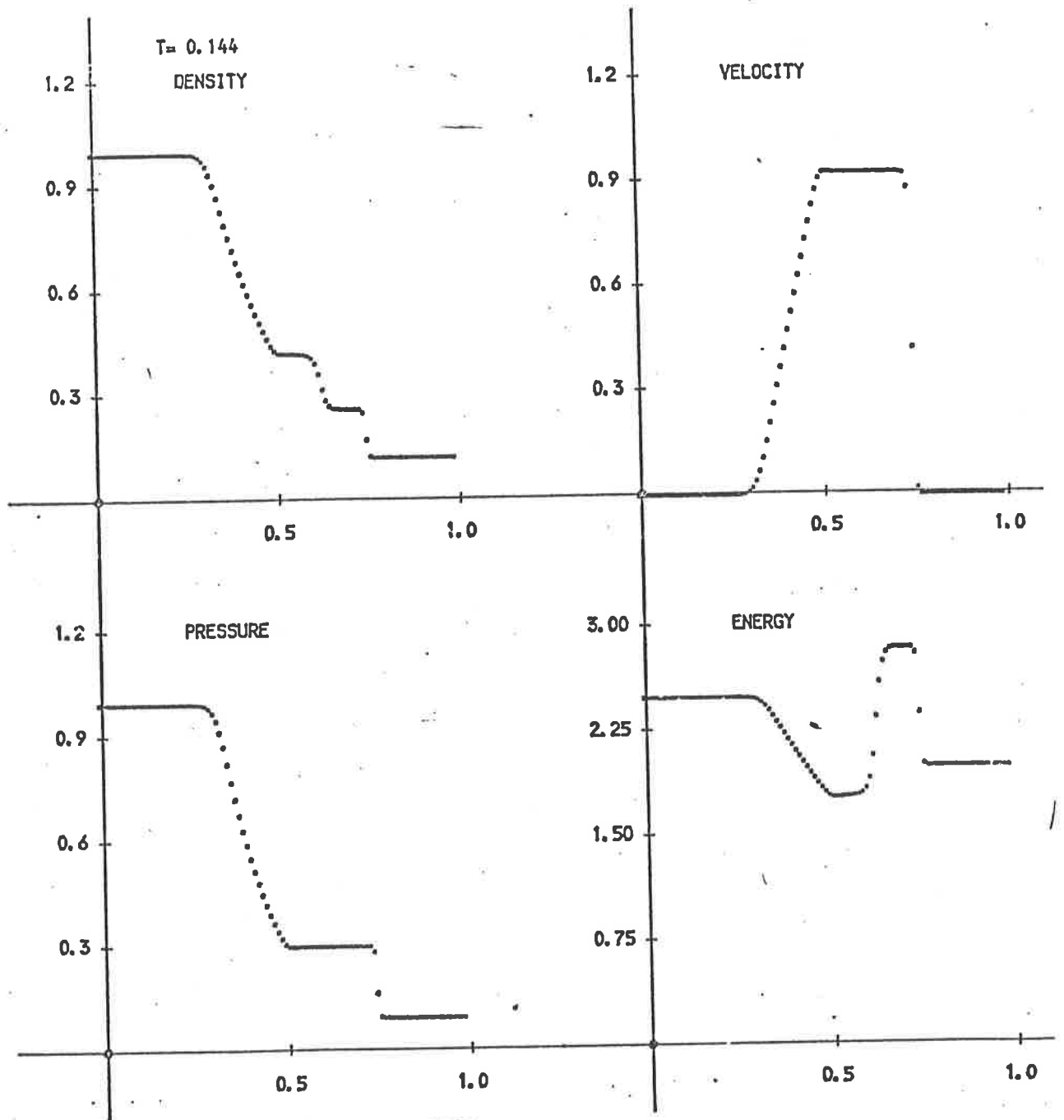
DT= 0.005

40POINTS

DX (2) = 0.010



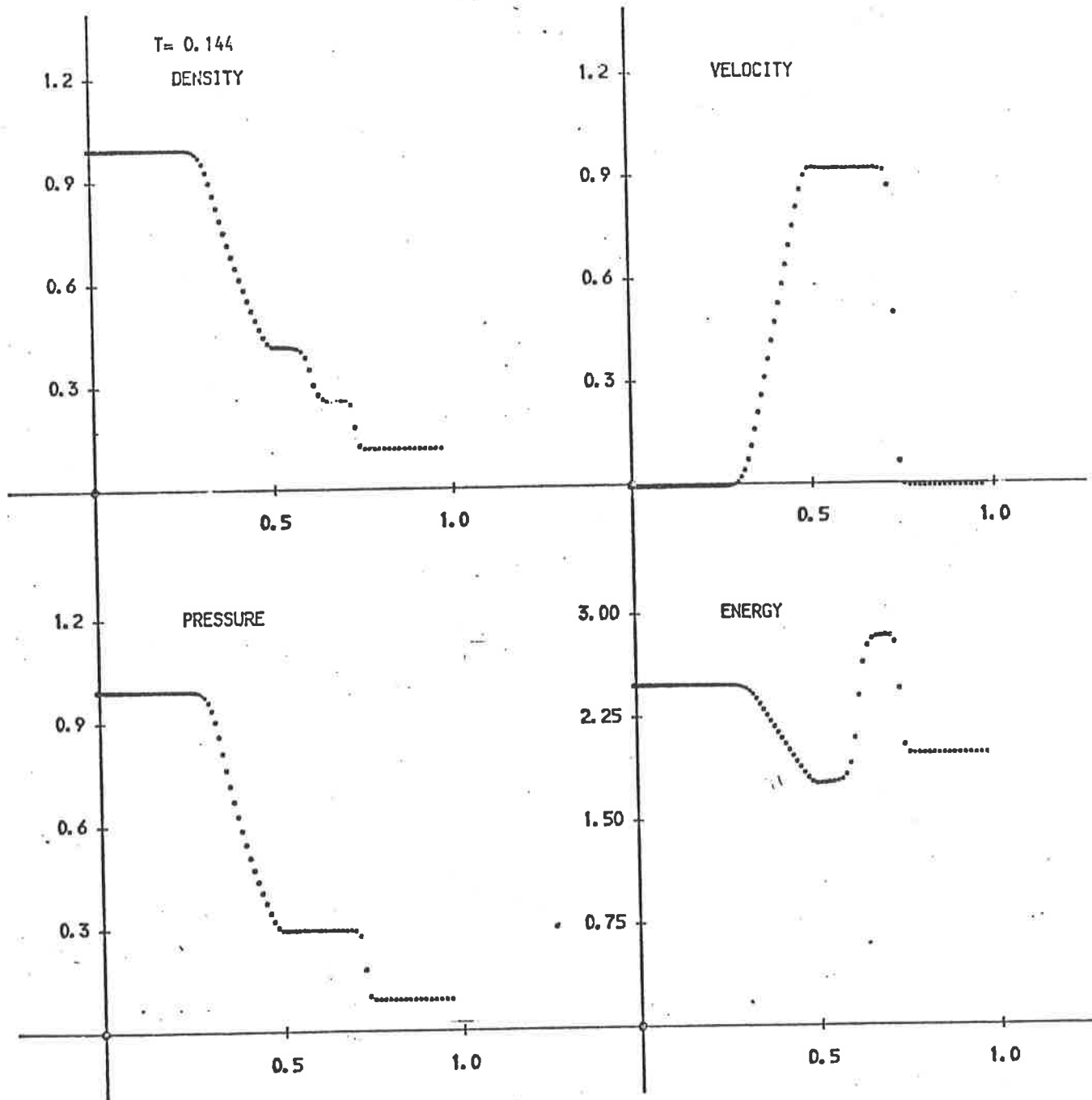
A shock wave and an expansion wave



Sod's shock tube problem using Roe's scheme (regular grid)

$\Delta x = 0.01$ ,  $t = 0.004$

B16



Sod's shock tube problem using scheme of Section 4 (irregular grid)

$DX(2) = 0.005$ , MESH PARAMETER = 0.001,  $\Delta t = 0.004$

B17

Bibliography

- [1] P.L. Roe Numerical algorithms for the linear wave equation. Private communication.
- [2] P.L. Roe The use of the Riemann problem in finite difference schemes. Proc. 7th Int. Conf. Num. Meth. in Fluid Dynmaics.
- [3] G.A. Sod A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. J.C.P. 27 1978, 1-31.
- [4] Harten, Hyman & Lax On finite-difference approximations and entropy conditions for shocks. Comm. Pure Appl. Maths. 29 1976, 297-322.
- [5] P.D. Lax Hyperbolic systems of conservation laws and the mathematical theory of shock waves. SIAM Regional Conference Series in Appl. Maths. 11.
- [6] A.Y. Le Roux A numerical conception of entropy for quasi-linear equations.
- [7] B. Enquist & S. Osher Stable and entropy satisfying approximations for transonic flow calculations. Maths. Comp. 34 149 Jan. 1980, 45-75.
- [8] P.L. Roe An improved version of MacCormack's shock-capturing algorithm. R.A.E. Technical Report 79041 April 1979.
- [9] M.J. Baines A numerical algorithm for the solution of systems of conservation laws in two dimensions. Univ. of Reading Numerical Analysis Report 2/80 1980.
- [10] O.A. Oleinik Discontinuous solutions of nonlinear differential equations. Amer. Math. Soc. Transl. Ser. 2 26, 95-172.