TVB RUNGE–KUTTA LOCAL PROJECTION DISCONTINUOUS GALERKIN FINITE ELEMENT METHOD FOR CONSERVATION LAWS III: ONE DIMENSIONAL SYSTEMS

By

Bernardo Cockburn

San–Yih Lin

and

Chi–Wang Shu

IMA Preprint Series # 415

April 1988

TVB RUNGE-KUTTA LOCAL PROJECTION DISCONTINUOUS GALERKIN FINITE ELEMENT METHOD FOR CONSERVATION LAWS III: ONE DIMENSIONAL SYSTEMS

BERNARDO COCKBURN†, SAN-YIH LIN† AND CHI-WANG SHU‡

Abstract. This is the third paper in a series in which we construct and analyze a class of TVB (total variation bounded) discontinuous Galerkin finite element methods for solving conservation laws $\mathbf{u}_t + \sum_{i=1}^{d}(\mathbf{f}_i(\mathbf{u}))_{x_i} = 0$. In this paper we present the method in system of equations, stressing the point of how to use the weak form in the component spaces but to use the local projection limiting in the characteristic fields, and how to implement boundary conditions. One dimensional system is thus chosen as a model. Different implementation techniques are discussed, theory analogous to scalar cases are proven for linear systems, and numerical results are given illustrating the method on nonlinear systems. Discussions of handling complicated geometries via adaptive triangle elements will appear in future papers.

Key words. TVD, TVB, finite element, conservation law, Euler's equation

AMS(MOS) subject classifications. 65M60, 65N30, 35L65

1. Introduction. In [3, 4] we constructed and analyzed a new class of finite element methods—we call them RK$\Lambda\Pi\mathrm{P}^k$, or $(k+1)$-th order TVB Runge-Kutta local projection discontinuous Galerkin finite element method—for solving the hyperbolic conservation law

$$\mathbf{u}_t + \sum_{i=1}^{d}(\mathbf{f}_i(\mathbf{u}))_{x_i} = 0, \tag{1.1}$$

with suitable initial or initial-boundary conditions. In (1.1) $\mathbf{u} = (u_1, \ldots, u_m)^T$, $\mathbf{x} = (x_1, \ldots, x_d)$, and $\sum_{i=1}^{d} \xi_i \frac{\partial \mathbf{f}_i}{\partial \mathbf{u}}$ always has m real eigenvalues and a complete set of eigenvectors, with real ξ_i. In [4] we presented the general framework in the case $d = m = 1$, keeping in mind about the possibility of natural extensions for $d > 1$ and/or $m > 1$. For details, history and related work, see [1–4, 8, 10, 11, 13–15]. As indicated in [4], the main advantages of these methods over most other finite element methods are their time explicitness, hence they can be equipped with high order TVD Runge-Kutta type time discretizations in [15], and their TVB provableness in 1-D, scalar nonlinear case; the main advantage over finite difference methods is the flexibility in handling complicated geometries and boundary conditions. In this paper we carry out the generalization to one dimensional system $m > 1$, $d = 1$. The multidimensional case $d > 1$ needs adaptive triangle elements to handle complicated geometries, and will be discussed in future papers.

†School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455.
‡Division of Applied Mathematics, Brown University, Providence, Rhode Island 02912.

Theories about (1.1) as well as about numerical methods solving (1.1) are far less advanced for the system case $m > 1$ than for the scalar case $m = 1$. A numerical method will be considered acceptable for practical purpose if it verifies a convergence theory for the linear system case $f(u) = Au$ (the scheme in this case is usually still nonlinear) as well as for scalar nonlinear case, if it is easily implementable to nonlinear systems, and if it gives good results for solving nonlinear systems with discontinuities. We will follow this conventional approach in this paper. In section 2 we will consider initial value problems. We present the weak form and the general framework of the scheme, and consider different possible avenues to apply monotone fluxes at the interfaces and the local projection limiters. A total variation bounded estimate, similar to the scalar case $m = 1$, is proven for linear systems. In section 3 boundary condition implementation is considered. A similar total variation bounded estimate is proven, again for linear systems. Section 4 includes some numerical results, mainly for the (nonlinear) Euler's equation in gas dynamics, to illustrate the behavior of our schemes for nonlinear systems. The test problems chosen are standard. For comparison with non-oscillatory finite difference schemes and other finite element methods, we refer the readers to [5, 7, 14].

2. Initial Value Problems. We consider in this section the equation (1.1) with $m > 1$, $d = 1$, and with a pure initial condition (periodic or compact supported):

$$\mathbf{u}(x, 0) = \mathbf{u}^0(x). \tag{2.1}$$

As in the scalar case $m = 1$, we shall first discretize (1.1)-(2.1) in the spatial variable x. Let $I_j = (x_{j-1/2}, x_{j+1/2})$, $I = \cup_j I_j$ be a partition of the real line. Denote $\Delta x_j = x_{j+1/2} - x_{j-1/2}$ and $h = sup_j \Delta x_j$. The finite element method we are going to use is a Galerkin method for which the finite dimensional space $\mathbf{V}_h$ to which the approximate solution $\mathbf{u}^h(t)$ belongs for $t \in [0, T]$ is taken as

$$\begin{aligned} \mathbf{V}_h = \mathbf{V}_h^k = \{\mathbf{p} :\ & \text{each of its components } p_i \in BV \cap L^1 \ : \\ & p_i|_{I_j} \ \text{ is a polynomial of degree } \le k\}, \end{aligned} \tag{2.2}$$

All the general framework developed in [4] for scalar case, except the choice of monotone flux and the local projection limiting, can be applied here, just component-wisely. We have

$$\mathbf{u}^h(x, t) = \sum_{\ell=0}^{k} a_\ell \mathbf{u}_j^{(\ell)}(t) v_\ell^{(j)}(x) \text{ for } x \in I_j, \tag{2.3}$$

where $v_\ell^{(j)}(x)$ form a local orthogonal basis over I_j:

$$v_0^{(j)}(x) = 1, \quad v_1^{(j)}(x) = x - x_j, \quad v_2^{(j)}(x) = (x - x_j)^2 - \frac{1}{12}\Delta x_j^2, \ldots, \tag{2.4}$$

the coefficients a_ℓ are

$$a_\ell = \frac{\Delta x_j^{\ell+1}}{\int_{I_j} \left(v_\ell^{(j)}(x)\right)^2 dx}; \quad \text{i.e., } a_0 = 1,\ a_1 = \frac{12}{\Delta x_j},\ a_2 = \frac{180}{\Delta x_j^2}, \cdots, \tag{2.5}$$

and the degrees of freedom, $\mathbf{u}_j^{(\ell)}(t)$, are defined by

$$\mathbf{u}_j^{(\ell)} = \mathbf{u}_j^{(\ell)}(t) = \frac{1}{\Delta x_j^{\ell+1}} \int_{I_j} \mathbf{u}(x,t) v_\ell^{(j)}(x) dx, \qquad \ell = 0, 1, \ldots, k. \tag{2.6}$$

In order to determine the degrees of freedom of $\mathbf{u}^h$ we proceed as in [4]. We multiply (1.1) by $\mathbf{v}^h \in \mathbf{V}_h^k$, integrate over I_j, and integrate by parts formally to obtain

$$\begin{aligned} \frac{d}{dt} \int_{I_j} \mathbf{u}(x,t)\, \mathbf{v}^h(x) dx + [\Delta_+(\mathbf{v}^h(x_{j-\frac{1}{2}}) \mathbf{f}(\mathbf{u}(x_{j-\frac{1}{2}}, t))] \\ - \int_{I_j} \mathbf{f}(\mathbf{u}(x,t)) \frac{d}{dx} \mathbf{v}^h(x)\, dx = 0, \forall \mathbf{v}^h \in \mathbf{V}_h^k, \end{aligned} \tag{2.7}$$

where $\Delta_\pm$ are the usual difference operators $\Delta_\pm a_j = \pm(a_{j\pm1} - a_j)$.

Next, we replace the exact solution $\mathbf{u}$ by its approximation $\mathbf{u}^h$, and $\mathbf{f}(\mathbf{u}(x_{j-\frac{1}{2}}, t))$ by some monotone flux (whose choice in the current system case will be determined later) $\mathbf{h}_{j-\frac{1}{2}} = \mathbf{h}(\mathbf{u}_{j-\frac{1}{2}}^-, \mathbf{u}_{j-\frac{1}{2}}^+)$, where $\mathbf{u}_{j-\frac{1}{2}}^\pm = \mathbf{u}^h(x_{j-\frac{1}{2}}^\pm, t)$ are defined by (2.3), subject to some local projection limiting to be discussed later. We obtain, after some simple algebraic manipulations,

$$\begin{aligned} \frac{d}{dt}\, \mathbf{u}_j^{(\ell)} + \frac{1}{\Delta x_j^{\ell+1}} [\Delta_+(v_\ell^{(j)}(x_{j-\frac{1}{2}}) \mathbf{h}_{j-\frac{1}{2}}] \\ - \frac{1}{\Delta x_j^{\ell+1}} \int_{I_j} \mathbf{f}(\mathbf{u}^h(x,t)) \frac{d}{dx} v_\ell^{(j)}(x)\, dx = 0, \quad \ell = 0, 1, \ldots, k. \end{aligned} \tag{2.8}$$

The integration in (2.8) can be approximated by a suitable quadrature whose error is at most $0(h^{k+\ell+2})$, and (2.8) is solved in time by a TVD Runge-Kutta type method [4, 15]:

(2.9a)

$$(\mathbf{u}^h)^{(i)} = \sum_{\ell=0}^{i-1} \left[\alpha_{i\ell} (\mathbf{u}^h)^{(\ell)} + \beta_{i\ell} \Delta t \mathbf{L}_h((\mathbf{u}^h)^{(\ell)}, t^n + d_\ell \Delta t)\right], \quad i = 1, \cdots, r,$$

(2.9b)

$$(\mathbf{u}^h)^{(0)} = (\mathbf{u}^h)^n, \qquad (\mathbf{u}^h)^{(r)} = (\mathbf{u}^h)^{n+1}.$$

For example

(2.10a)

$$\text{second order } (r=2): \alpha_{10} = \beta_{10} = 1,\ \alpha_{20} = \alpha_{21} = \beta_{21} = \tfrac{1}{2},\ \beta_{20} = 0;\ d_0 = 0,\ d_1 = 1;\ \text{CFL} = 1,$$

(2.10b)

$$\text{third order } (r=3): \alpha_{10} = \beta_{10} = 1,\ \alpha_{20} = \tfrac{3}{4},\ \beta_{20} = 0,\ \alpha_{21} = \beta_{21} = \tfrac{1}{4},\ \alpha_{30} = \tfrac{1}{3},\ \beta_{30} = \alpha_{31} = \beta_{31} = 0,\ \alpha_{32} = \beta_{32} = \tfrac{2}{3},\ d_0 = 0,\ d_1 = 1,\ d_2 = \tfrac{1}{2};\ \text{CFL} = 1.$$

etc. The starting point of (2.9)–(2.10) is (2.8) written in a concise ODE form:

$$\frac{d}{dt}\mathbf{u}^h = \mathbf{L}_h(\mathbf{u}^h, t), \tag{2.11}$$

where we include the time variable t in case there are time dependent forcing terms or boundary conditions.

We now turn to the problem of choosing $\mathbf{h}_{j+\frac{1}{2}}$ in (2.8), and applying the local projection limiting. For this purpose we write (see (2.3)):

$$\mathbf{u}^-_{j+\frac{1}{2}} = \mathbf{u}^{(0)}_j + \tilde{\mathbf{u}}_j\ , \qquad \mathbf{u}^+_{j-\frac{1}{2}} = \mathbf{u}^{(0)}_j - \tilde{\tilde{\mathbf{u}}}_j, \tag{2.12}$$

and apply the local projection limiting on $\tilde{\mathbf{u}}_j$, $\tilde{\tilde{\mathbf{u}}}_j$, as in the scalar case [4].

One simple way is to do everything just component-wisely. We define

$$\tilde{\mathbf{u}}^{(\text{mod})}_j = \mathbf{m}\left(\tilde{\mathbf{u}}_j, \Delta_+\mathbf{u}^{(0)}_j, \Delta - \mathbf{u}^{(0)}_j\right), \tilde{\tilde{\mathbf{u}}}^{(\text{mod})}_j = \mathbf{m}\left(\tilde{\tilde{\mathbf{u}}}_j, \Delta + \mathbf{u}^{(0)}_j, \Delta - \mathbf{u}^{(0)}_j\right), \tag{2.13}$$

where $\mathbf{m}$ is a vector minmod function with TVB correction:

$$\mathbf{m}(\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n) = \begin{pmatrix} m\left((a_1)_1, (a_2)_1, \cdots, (a_n)_1\right) \\ \vdots \\ m\left((a_1)_m, (a_2)_m, \cdots, (a_n)_m\right) \end{pmatrix}, \tag{2.14}$$

where $\mathbf{a}_i = ((a_i)_1, (a_i)_2, \cdots, (a_i)_m)^T$, and

(2.15a)

$$m(b_1, \cdots, b_n) = \begin{cases} b_1 & , \text{ if } |b_1| \le Mh^2, \\ s \cdot \min_{1\le i\le n} |b_i| & , \text{ if } |b_1| > Mh^2, \text{ and } \operatorname{sign}(b_1) = \cdots = \operatorname{sign}(b_n) = s, \\ 0 & , \text{ otherwise,} \end{cases}$$

with

(2.15b)

$$M = \frac{2}{3}M_2,$$

or

(2.15c)

$$M = M_{j,\ell} = \frac{2}{9}(3 + 10M_2)M_2 \frac{h^2}{h^2 + |(\Delta_+ \mathbf{u}_j^{(0)})_\ell| + |(\Delta_- (\mathbf{u}_j^{(0)}))_\ell|},$$

and $M_2 = \max_{i,J} |\frac{\partial^2}{\partial x^2}(\mathbf{u}^0(x))_i|$, the maximum being taken over a neighborhood J of smooth critical points of $\mathbf{u}^0(x)$ in (2.1).

We then use the (local) Lax-Friedrichs flux

$$\text{(2.16a)} \quad \mathbf{h}_{j+\frac{1}{2}} = \mathbf{h}\left(\mathbf{u}^-_{j+\frac{1}{2}}, \mathbf{u}^+_{j+\frac{1}{2}}\right) = \frac{1}{2}\left[\mathbf{f}\left(\mathbf{u}^-_{j+\frac{1}{2}}\right) + \mathbf{f}\left(\mathbf{u}^+_{j+\frac{1}{2}}\right) - \alpha_{j+\frac{1}{2}}\left(\mathbf{u}^+_{j+\frac{1}{2}} - \mathbf{u}^-_{j+\frac{1}{2}}\right)\right],$$

with

$$\text{(2.16b)} \qquad \alpha_{j+\frac{1}{2}} = \max_{1 \le p \le m}\left(|\lambda^{(p)^+}_{j+\frac{1}{2}}|, |\lambda^{(p)^-}_{j+\frac{1}{2}}|\right) \qquad \text{(local Lax-Friedrichs)},$$

or

$$\text{(2.16c)} \qquad \alpha_{j+\frac{1}{2}} \equiv \alpha = \max_{j,p}\left|\lambda^{(p)^\pm}_{j+\frac{1}{2}}\right| \qquad \text{(Lax-Friedrichs)},$$

where $\lambda^{(p)^\pm}_{j+\frac{1}{2}}$, $p = 1, \cdots, m$, are the m real eigenvalues of the Jacobian $\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\big|_{\mathbf{u}=\mathbf{u}^\pm_{j+\frac{1}{2}}}$.

Notice that we do not need to evaluate the Jacobian or its eigenvectors, just its eigenvalues. Thus computationally it is very simple. Unfortunately this simple version does not have a TVB theory, even for linear systems. Computationally we do observe wriggles (see section 4), although these wriggles are usually rather small for second order schemes.

To achieve better qualities at the price of more complicated computations, we use characteristic field decompositions. We denote by $A_{j+\frac{1}{2}} = \left(\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right)_{\mathbf{u}=\mathbf{u}_{j+\frac{1}{2}}}$ some "average" Jacobian, *e.g.*, $\mathbf{u}_{j+\frac{1}{2}} = \frac{\mathbf{u}^{(0)}_j + \mathbf{u}^{(0)}_{j+1}}{2}$ (simple arithmetic mean), or (for Euler equations of gas dynamics) $\mathbf{u}_{j+\frac{1}{2}} = \mathbf{R}\left(\mathbf{u}^{(0)}_j, \mathbf{u}^{(0)}_{j+1}\right)$ where $\mathbf{R}$ is the Roe average [12]. We denote the eigenvalues and left and right eigenvectors of $A_{j+\frac{1}{2}}$ by $\lambda^{(p)}_{j+\frac{1}{2}}$, $\mathbf{l}^{(p)}_{j+\frac{1}{2}}$, $\mathbf{r}^{(p)}_{j+\frac{1}{2}}$, $p = 1, \cdots, m$, normalized so that $\mathbf{l}^{(p)}_{j+\frac{1}{2}} \cdot \mathbf{r}^{(q)}_{j+\frac{1}{2}} = \delta_{pq}$. Then, in computing $\mathbf{h}_{j+\frac{1}{2}}$, we project everything to the eigenspace of $A_{j+\frac{1}{2}}$:

$$\text{(2.17)} \qquad a^{(p)} = \mathbf{l}^{(p)}_{j+\frac{1}{2}} \cdot \mathbf{a},$$

where we take $\mathbf{a} = \tilde{\mathbf{u}}_j, \tilde{\tilde{\mathbf{u}}}_{j+1}, \mathbf{u}^{(0)}_j, \mathbf{u}^{(0)}_{j+1}, \Delta_- \mathbf{u}^{(0)}_j, \Delta_+ \mathbf{u}^{(0)}_j, \Delta_+ \mathbf{u}^{(0)}_{j+1}$. We then apply the local projection limiting in each characteristic field:

$$\text{(2.18a)} \qquad (\tilde{u}_j)^{(p)(\text{mod})} = m\left((\tilde{u}_j)^{(p)}, \left(\Delta_+ u^{(0)}_j\right)^{(p)}, \left(\Delta_- u^{(0)}_j\right)^{(p)}\right),$$

(2.18b)

$$\left(\tilde{\tilde{u}}_{j+1}\right)^{(p)(\text{mod})} = m\left(\left(\tilde{\tilde{u}}_{j+1}\right)^{(p)}, \left(\Delta_+ u^{(0)}_j\right)^{(p)}, \left(\Delta_+ u^{(0)}_{j+1}\right)^{(p)}\right),$$

(We will drop the superscript "mod" in the following.)

We form

$$\left(u^-_{j+\frac{1}{2}}\right)^{(p)} = \left(u^{(0)}_j\right)^{(p)} + (\tilde{u}_j)^{(p)}, \quad \left(u^+_{j+\frac{1}{2}}\right)^{(p)} = \left(u^{(0)}_{j+1}\right)^{(p)} - \left(\tilde{\tilde{u}}_{j+1}\right)^{(p)}, \tag{2.19}$$

return to the component space

$$\mathbf{a} = \sum_{p=1}^{m} a^{(p)} \mathbf{r}^{(p)}_{j+\frac{1}{2}}, \tag{2.20}$$

by taking $\mathbf{a} = \mathbf{u}^{\pm}_{j+\frac{1}{2}}$, compute $\mathbf{f}^{\pm}_{j+\frac{1}{2}} = \mathbf{f}\left(\mathbf{u}^{\pm}_{j+\frac{1}{2}}\right)$, compute $\left(f^{\pm}_{j+\frac{1}{2}}\right)^{(p)}$ by (2.17), then use any scalar monotone flux or E-flux [11] in the p-th characteristic field, $p = 1, \cdots, m$. For example, we may use the local Lax-Friedrichs flux

$$h^{(p)}_{j+\frac{1}{2}} = \frac{1}{2}\left[\left(f^+_{j+\frac{1}{2}}\right)^{(p)} + \left(f^-_{j+\frac{1}{2}}\right)^{(p)} - \alpha^{(p)}_{j+\frac{1}{2}}\left(\left(u^+_{j+\frac{1}{2}}\right)^{(p)} - \left(u^-_{j+\frac{1}{2}}\right)^{(p)}\right)\right], \tag{2.21a}$$

with

$$\alpha^{(p)}_{j+\frac{1}{2}} = \max\left(|\lambda^{(p)}_j|, |\lambda^{(p)}_{j+1}|\right), \tag{2.21b}$$

(for convex case only, otherwise the maximum should be taken in the whole interval) or the Roe flux with entropy correction

$$h^{(p)}_{j+\frac{1}{2}} = \begin{cases} \left(f^+_{j+\frac{1}{2}}\right)^{(p)} & \text{, no sonic point and } \lambda^{(p)}_{j+\frac{1}{2}} < 0, \\ \left(f^-_{j+\frac{1}{2}}\right)^{(p)} & \text{, no sonic point and } \lambda^{(p)}_{j+\frac{1}{2}} \geq 0, \\ \text{same as in (2.21)} & \text{, otherwise.} \end{cases} \tag{2.22}$$

We finally get $\mathbf{h}_{j+\frac{1}{2}}$ by (2.20) with $\mathbf{a} = \mathbf{h}_{j+\frac{1}{2}}$:

$$\mathbf{h}_{j+\frac{1}{2}} = \sum_{p=1}^{m} h^{(p)}_{j+\frac{1}{2}} \mathbf{r}^{(p)}_{j+\frac{1}{2}}. \tag{2.23}$$

Computationally this approach needs much more work. However it works better both theoretically and numerically. Theoretically we have the following proposition, similar to the results for scalar case [4], for linear systems

PROPOSITION 2.1. *Scheme (2.8)-(2.9)-(2.23) is TVBM (total variation bounded in the means* $\mathbf{u}^{(0)}$*) and TVB, under the total variation definition (see (2.17) for notations)*

(2.24)
$$TV(\mathbf{u}^{(0)}) = \sum_j \sum_{p=1}^{m} \left| \left(u_{j+1}^{(0)}\right)^{(p)} - \left(u_j^{(0)}\right)^{(p)} \right|, \; TV(\mathbf{u}) = \sum_j \sum_{p=1}^{m} \left| (u_{j+1})^{(p)} - (u_j)^{(p)} \right|,$$

for linear system $\mathbf{f}(\mathbf{u}) = A\mathbf{u}$ *where* A *is a constant matrix, hence has a convergent subsequence in this case.*

Proof. Since $\lambda_{j+\frac{1}{2}}^{(p)}$, $\mathbf{l}_{j+\frac{1}{2}}^{(p)}$, $\mathbf{r}_{j+\frac{1}{2}}^{(p)}$ do not vary with j, in each characteristic field, $\left(u^{(0)}\right)^{(p)}$ will satisfy a scalar TVB scheme in [4]. The same argument as in the scalar case [4] now leads to TVBM and TVB. □

Notice that the scheme in this case is still nonlinear. The result can be generalized to $A = A(x)$.

For numerical results see Section 4.

3. Initial Boundary Value Problems. We now turn our attention to the initial boundary value problems. For simplicity we take the interval $(0, +\infty)$ and consider one boundary at $x = 0$ only. The general case of two boundaries can be handled similarly. if at the boundary $x = 0$

$$\lambda^{(1)} \leq \cdots \leq \lambda^{(s)} < 0 < \lambda^{(s+1)} \leq \cdots \leq \lambda^{(m)}, \tag{3.1}$$

where $\lambda^{(p)}$ are the eigenvalues of $\frac{\partial \mathbf{f}}{\partial \mathbf{u}} \mid_{x=0}$, then a well-posed boundary condition takes the form (see (2.17) for notations):

$$\begin{pmatrix} (u)^{(s+1)}(0,t) \\ \vdots \\ (u)^{(m)}(0,t) \end{pmatrix} = B(t) \begin{pmatrix} (u)^{(1)}(0,t) \\ \vdots \\ (u)^{(s)}(0,t) \end{pmatrix} + g(t), \tag{3.2}$$

where $B(t)$ is a $(m-s)\times s$ matrix with Lipschitz continuous components, $g(t)$ is a $(m-s)\times 1$ vector with bounded variation.

We put the boundary at $x_{-\frac{1}{2}} = 0$, and implement the boundary condition (3.2) as follows:

$$\left(u_{-\frac{1}{2}}^{-}\right)^{(p)} = \left(u_{-\frac{1}{2}}^{+}\right)^{(p)}, (\tilde{u}_0)^{(p)(\mathrm{mod})} = m\left((\tilde{u}_0)^{(p)}, \left(\Delta_+ u_0^{(0)}\right)^{(p)}\right),$$

$$\left(\tilde{\tilde{u}}_0\right)^{(p)(\mathrm{mod})} = m\left(\left(\tilde{\tilde{u}}_0\right)^{(p)}, \left(\Delta_+ u_0^{(0)}\right)^{(p)}\right), \tag{3.3a}$$

for $p = 1, \cdots, s$;

$$\text{(3.3b)} \qquad \begin{pmatrix} \left(u^-_{-\frac{1}{2}}\right)^{(s+1)} \\ \vdots \\ \left(u^-_{-\frac{1}{2}}\right)^{(m)} \end{pmatrix} = B(t) \begin{pmatrix} \left(u^-_{-\frac{1}{2}}\right)^{(1)} \\ \vdots \\ \left(u^-_{-\frac{1}{2}}\right)^{(s)} \end{pmatrix} + g(t),$$

$$(\tilde{u}_0)^{(p)(\text{mod})} = m\left((\tilde{u}_0)^{(p)}, \left(\Delta_+ u_0^{(0)}\right)^{(p)}, 2\left(\left(u_0^{(0)}\right) - \left(\sum_{\ell=1}^{s} B_{p-s,\ell}(t)\left(u^-_{-\frac{1}{2}}\right)^{(\ell)} + (g(t))_{p-s}\right)\right)\right),$$

$$\left(\tilde{\tilde{u}}_0\right)^{(p)(\text{mod})} = m\left((\tilde{u}_0)^{(p)}, \left(\Delta_+ u_0^{(0)}\right)^{(p)}\right),$$

for $p = s+1, \cdots, m$.

Notice that, as in the scalar case [4], accuracy is not affected by the limiters.

We then have the following proposition

PROPOSITION 3.1. *Scheme (2.8)-(2.9)-(2.23)-(3.3) is TVBM, under the total variation definition for the mean* $\mathbf{u}^{(0)}$*:*

$$\text{(3.4)} \qquad TV\left(\mathbf{u}^{(0)}\right) = \sum_{j \geq -1} \left(Q \sum_{p=1}^{s} + \sum_{p=s+1}^{m}\right) \left| \left(u_{j+1}^{(0)}\right)^{(p)} - \left(u_j^{(0)}\right)^{(p)} \right|,$$

with $Q = \max(1, 2\|B\|_1) = \max\left(1, 2\max_{\substack{0 \leq t \leq T \\ 1 \leq p \leq s}} \sum_{\ell=1}^{m-s} |B_{\ell p}(t)|\right)$, *and*

$$\text{(3.5)} \qquad \left(u_{-1}^{(0)}\right)^{(p)} \equiv \begin{cases} \left(u_0^{(0)}\right)^{(p)} & , \ p = 1, \cdots, s \\ \sum_{\ell=1}^{s} B_{p-s,\ell}(t)\left(u_0^{(0)}\right)^{(\ell)} + (g(t))_{p-s} & , \ p = s+1, \cdots, m, \end{cases}$$

and TVB under the total variation definition (2.24), for linear constant coefficiented system $\mathbf{f}(\mathbf{u}) = A\mathbf{u}$.

Proof. We only need to prove the result for the Euler forward version of (2.8). Following the lines of proofs in [4, Proposition 3.1] we have

$$\left(\left(u_j^{(0)}\right)^{(p)}\right)^{n+1} = \left(\left(u_j^{(0)}\right)^{(p)}\right)^{n} + C_{j+\frac{1}{2}}^{(p)} \Delta_+ \left(\left(u_j^{(0)}\right)^{(p)}\right)^{n} - D_{j-\frac{1}{2}}^{(p)} \Delta_- \left(\left(u_j^{(0)}\right)^{(p)}\right)^{n},$$

for $j \geq 1$, $p = 1, \cdots, m$ and $j = 0$, $p = s+1, \cdots, m$, and

$$\left(\left(u_0^{(0)}\right)^{(p)}\right)^{n+1} = \left(\left(u_j^{(0)}\right)^{(p)}\right)^{n} + C_{\frac{1}{2}}^{(p)} \Delta_+ \left(\left(u_0^{(0)}\right)^{(p)}\right)^{n},$$

for $p = 1, \cdots, s$, where all the C's and D's are non-negative, module $0(h^2)$.

Following the lines of proof in [14, Theorem 3.1], we then have, for $p = s+1, \cdots, m$,

$$\begin{aligned}\Delta_+\left(\left(u_{-1}^{(0)}\right)^{(p)}\right)^{n+1} &= \left(1 - D_{-\frac{1}{2}}^{(p)}\right)\Delta_+\left(\left(u_{-1}^{(0)}\right)^{(p)}\right)^{n} \\ &\quad + C_{\frac{1}{2}}^{(p)}\Delta_+\left(\left(u_0^{(0)}\right)^{(p)}\right)^{n} - \left(g(t^{n+1}) - g(t^n)\right)_{p-s} \\ &\quad - \sum_{\ell=1}^{s}\left(B_{p-s,\ell}(t^{n+1})\left(\left(u_0^{(0)}\right)^{(\ell)}\right)^{n+1} - B_{p-s,\ell}(t^n)\left(\left(u_0^{(0)}\right)^{(\ell)}\right)^{n}\right).\end{aligned}$$

Hence, after some technical manipulations similar to [14, Theorem 3.1], we arrive at

$$\begin{aligned}TV\left((\mathbf{u}^{(0)})^{n+1}\right) &= \sum_{j\geq -1}\left(Q\sum_{p=1}^{s} + \sum_{p=s+1}^{m}\right)\left|\Delta_+\left(\left(u_j^{(0)}\right)^{(p)}\right)^{n+1}\right| \\ &\leq TV\left((\mathbf{u}^{(0)})^{n}\right) - \sum_{p=1}^{s}\left(Q - 2\sum_{\ell=1}^{m-s}\left|B_{\ell p}^{n+1}\right|\right)C_{\frac{1}{2}}^{(p)}\left|\Delta_+(u_0^{(0)})^{(p)}\right| \\ &\quad + L\Delta t\mathrm{TV}\left(\left(\mathbf{u}^{(0)}\right)^{n}\right) + \sum_{p=1}^{m-s}\left|\left(g\left(t^{n+1}\right) - g(t^n)\right)_p\right| + H\Delta t,\end{aligned}$$

where the $L\Delta t\mathrm{TV}(\mathbf{u}^{(0)})$ term is related to the Lipschitz condition of $B(t)$, it does not appear if $B(t)$ is a constant matrix. The $H\Delta t$ term is related to the TVB correction constant M in (2.15); it does not appear if $M \equiv 0$.

The remaining of the proof is straight forward. □

Remark 3.2. As in [14], the total variation definition in (3.4) has different weights for incoming and outgoing components. This is motivated by the differential equation theory, and guarantees total variation diminishing of the boundary treatment (3.3) under this definition of total variation, *i.e.*, if $B(t)$ is a constant matrix, $g \equiv 0$, $M = 0$ in (2.15), then $\mathrm{TV}((\mathbf{u}^{(0)})^{n+1}) \leq \mathrm{TV}((\mathbf{u}^{(0)})^{n})$. □

4. Numerical Results.

Example 1. We consider the Riemann problems of the Euler equation of gas dynamics for a polytropic gas:

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0, \mathbf{u}(x,0) = \mathbf{u}^0(x) = \begin{cases}\mathbf{u}_L, x < 0, \\ \mathbf{u}_R, x > 0,\end{cases} \tag{4.1a}$$

$$\mathbf{u} = (\rho, m, E)^T, \mathbf{f}(\mathbf{u}) = q\mathbf{u} + (0, p, qp)^T, \tag{4.1b}$$

with

(4.1c) $$p = (\gamma - 1)\left(E - \frac{1}{2}\rho q^2\right), m = \rho q,$$

$\gamma = 1.4$ is used in the following computation. For details of the Jacobian, its eigenvalues, eigenvectors, etc., see [5, 12].

Two sets of initial conditions are considered. One is proposed by Sod [17]:

(4.2a) $$(\rho_L, q_L, p_L) = (1, 0, 1); (\rho_R, q_R, p_R) = (0.125, 0, 0.10).$$

The other is used by Lax [9]

(4.2b) $$(\rho_L, q_L, p_L) = (0.445, 0.698, 3.528); (\rho_R, q_R, P_R) = (0.5, 0, 0.571).$$

We test our second order and third order schemes, *i.e.*, $k = 2$, and 3 in (2.8), $r = 2$ and 3 in (2.10). Both component-wise limiters (2.13)-(2.16a,b) and characteristic-wise limiters (2.18)-(2.22) are tested. Local Lax-Friedrichs flux (2.16a,b) and (2.21) are used. The results are in Figures 1–24. As expected, we see some wriggles in the component-wise version (and the wriggles become more severe with the third order scheme), and good behaviors in the characteristic-wise version.

We remark that the contact discontinuities and the corners of rarefaction waves are smeared more than the shocks. Some artificial compression or "sub-cell resolution" [6, 16] should help.

Example 2. We consider the interaction of blast waves of the Euler equation (4.1) with

(4.3) $$\mathbf{u}(x,0) = \begin{cases} \mathbf{u}_L, 0 \le x < 0.1, \\ \mathbf{u}_M, 0.1 \le x < 0.9, \\ \mathbf{u}_R, 0.9 \le x < 1, \end{cases}$$

where $\rho_L = \rho_M = \rho_R = 1$, $q_L = q_M = q_R = 0$, $p_L = 10^3$, $p_M = 10^{-2}$, $p_R = 10^2$. A reflecting boundary condition is applied to both ends. See [18, 5].

The results are in Figures 25–26. We see that the pictures are satisfactory, except for the above mentioned smearing of contact discontinuities.

Example 3. We use our second order and third order methods (2.8)-(2.10), with the boundary treatment (3.3), to solve the equation

(4.4) $$\begin{pmatrix} u \\ v \end{pmatrix}_t = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}_x,$$

with the initial boundary conditions

(4.5a) $$\begin{cases} u(x,0) = v(x,0) = \sin 2\pi x, \ \ 0 \le x \le 1, \\ u(0,t) = -v(0,t), v(1,t) = -u(1,t), \end{cases}$$

and

$$
\text{(4.5b)} \qquad \begin{cases} v(x,0) & = \begin{cases} 1, \text{if } \frac{1}{3} \le x \le \frac{2}{3}, \\ 0, \text{ otherwise,} \end{cases} \\ u(x,0) \equiv 0, x \ge 0, & \\ u(0,t) & = v(0,t). \end{cases}
$$

The numerical errors at $t = 2$ for (4.4)-(4.5a) are listed in Table 1. We can see that the boundary treatment works very well for smooth problems.

The solutions of (4.4)-(4.5b), at $t = 0.5$ and $t = 1.0$, are in Figures 27–32. We can see that the boundary treatment is total variation stable. The smearing is again due to the fact that the equation is linear, and we expect improvements by artificial compressions or sub-cell corrections [6, 16].

Table 4.1 . L_∞ : L_∞-error; L_1 : L_1-error; r = numerical order of convergence.

	Δx	second order				third order			
		L_∞	r	L_1	r	L_∞	r	L_1	r
u	$\frac{1}{20}$	0.21×10^{-1}		0.13×10^{-1}		0.42×10^{-3}		0.17×10^{-3}	
	$\frac{1}{40}$	0.49×10^{-2}	2.08	0.32×10^{-2}	2.07	0.37×10^{-4}	3.51	0.17×10^{-4}	3.28
	$\frac{1}{80}$	0.12×10^{-2}	2.03	0.77×10^{-3}	2.03	0.38×10^{-5}	3.31	0.20×10^{-5}	3.15
v	$\frac{1}{20}$	0.21×10^{-1}		0.13×10^{-1}		0.29×10^{-3}		0.13×10^{-3}	
	$\frac{1}{40}$	0.49×10^{-2}	2.08	0.32×10^{-2}	2.07	0.35×10^{-4}	3.06	0.15×10^{-4}	3.14
	$\frac{1}{80}$	0.12×10^{-2}	2.03	0.77×10^{-3}	2.03	0.42×10^{-5}	3.07	0.18×10^{-5}	3.06

ACKNOWLEDGEMENT. The work of Cockburn and Shu was initiated when both were in residence at the Institute of Mathematics and its Applications, University of Minnesota. We thank IMA and its staff for support and hospitality.

REFERENCES

[1] G. CHAVENT AND B. COCKBURN, *The Local Projection P^0P^1-Discontinuous-Galerkin Finite Element Method for Scalar Conservation Laws*, IMA Preprint Series No. 341, University of Minnesota, September 1987.

[2] G. CHAVENT AND G. SALZANO, J. Comput. Phys., 45, 307 (1982).

[3] B. COCKBURN AND C.-W. SHU, *The Runge Kutta Local Projection P^1-Discontinuous-Galerkin Finite Element Method for Scalar Conservation Laws*, IMA Preprint Series No. 388, University of Minnesota, January 1988.

[4] B. COCKBURN AND C.-W. SHU, *TVB Runge-Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws II: General Framework*, IMA Preprint Series No. 392, University of Minnesota, March 1988.

[5] A. HARTEN, B. ENGQUIST, S. OSHER AND S. CHAKRAVARTHY, J. Comput. Phys., 71, 231 (1987).

[6] A. HARTEN, *ENO Schemes with Subcell Resolutions*, ICASE Report 87-56, August 1987.

[7] T. HUGHES AND M. MALLET, *A High-Precision Finite Element Method for Shock-Tube Calculations*, Finite Element in Fluids, 6, 339 (1985).

[8] C. JOHNSON AND J. PITKARANTA, Math. Comput., 46, 1 (1986).

[9] P. LAX, Comm. Pure Appl. Math., 7, 159 (1954).

[10] P. LESAINT AND P. RAVIART, in Mathematical Aspects of Finite Element in Partial DifferentialEquations, edited by C. DeBoor, Academic Press, p. 89.

[11] S. OSHER, SIAM J. Numer. Anal., 22, 947 (1985).

[12] P. ROE, J. Comput. Phys., 43, 357 (1981).

[13] C.-W. SHU, Math. Comput., 49, 105 (1987).

[14] C.-W. SHU, Math. Comput., 49, 123 (1987).

[15] C.-W. SHU AND S. OSHER, *Efficient Implementation of Essentially Non-Oscillatory Shock Capturing Schemes*, ICASE Report 87-33, J. Comput. Phys., to appear.

[16] C.-W. SHU AND S. OSHER, *Efficient Implementation of Essentially Non-Oscillatory Shock Capturing Schemes, II*, in preparation.

[17] G. SOD, J. Comput. Phys., 27, 1 (1978).

[18] P. WOODWARD AND P. COLELLA, J. Comput. Phys., 54, 115 (1984).

NOTES ABOUT THE FIGURES:

(1) *Except in Figures 25-26, solid lines are for the exact solutions, and '+' are for the numerical solutions (just one point per cell is printed);*

(2) *In Figures 1-24, we solve (4.1)-(4.2a) to $t = 2.0$, and (4.1)-4.2b) to $t = 1.3$, using 100 cells. Density, velocity and pressure are pictured for each case;*

(3) *In figures 25a, 26a, the solid lines are for numerical results with 400 cells, and '+' for 200 cells.*

(4) *In figures 25b, 26b, the solid lines are for numerical results with 800 cells, and '+' for 400 cells.*

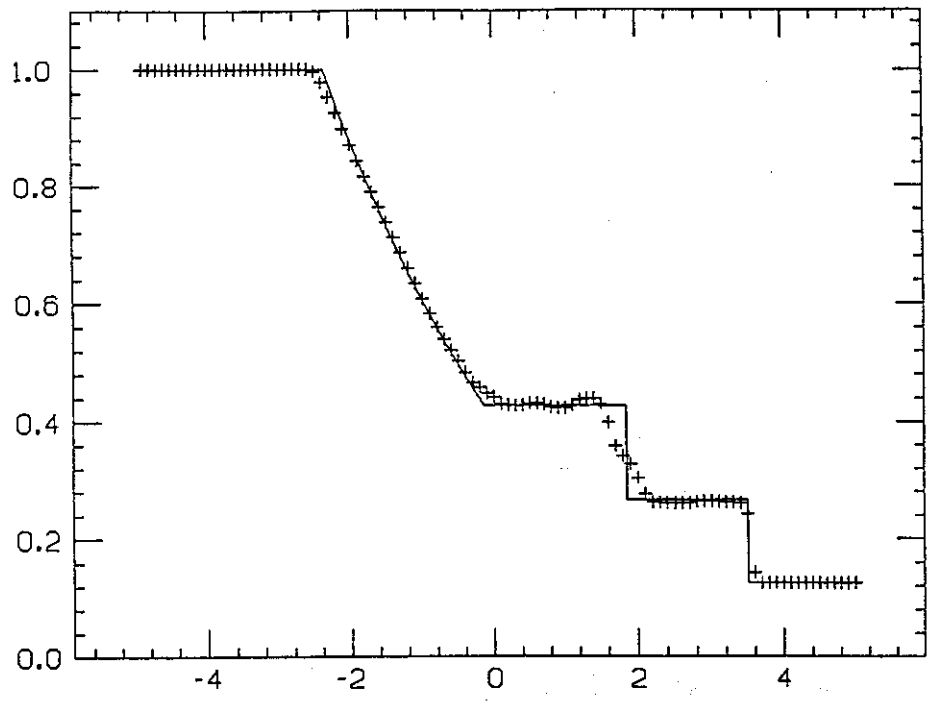

Figure 1: second order, component-wise limiter, (4.2a), density.

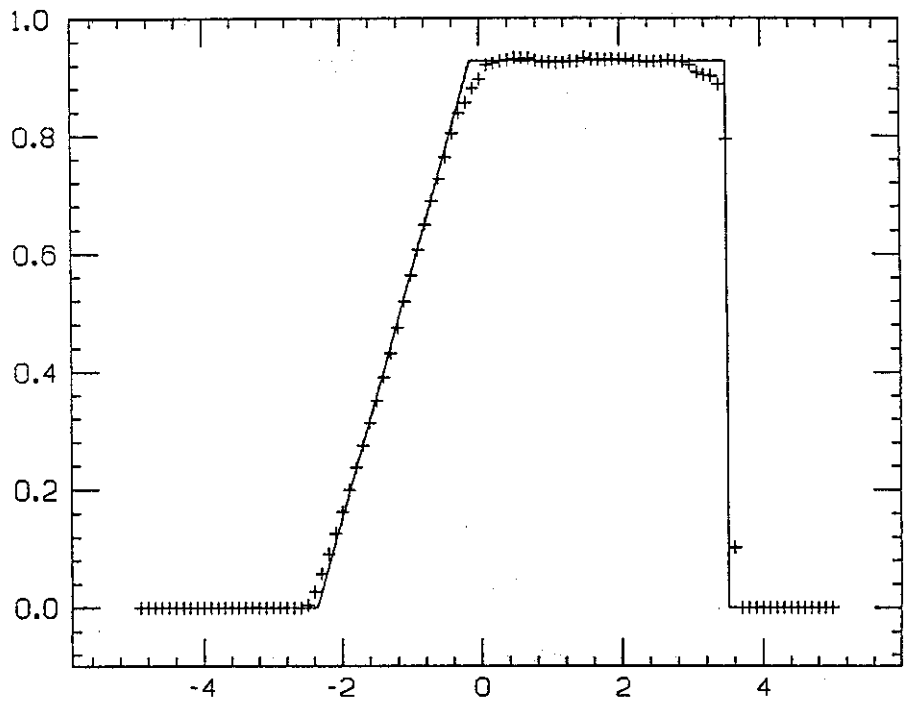

Figure 2: second order, component-wise limiter, (4.2a), velocity.

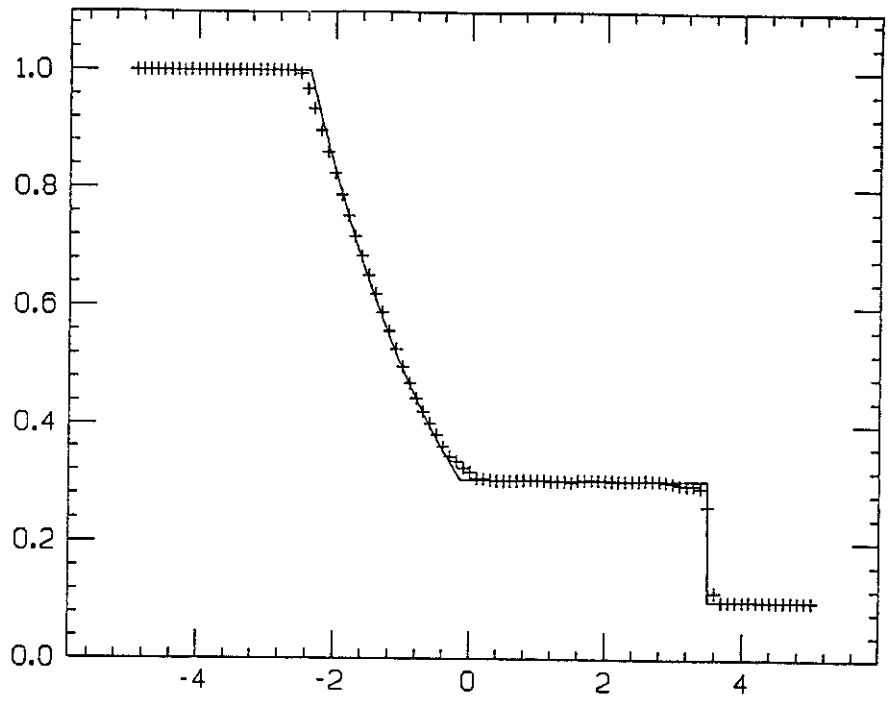

Figure 3: second order, component-wise limiter, (4.2a), pressure.

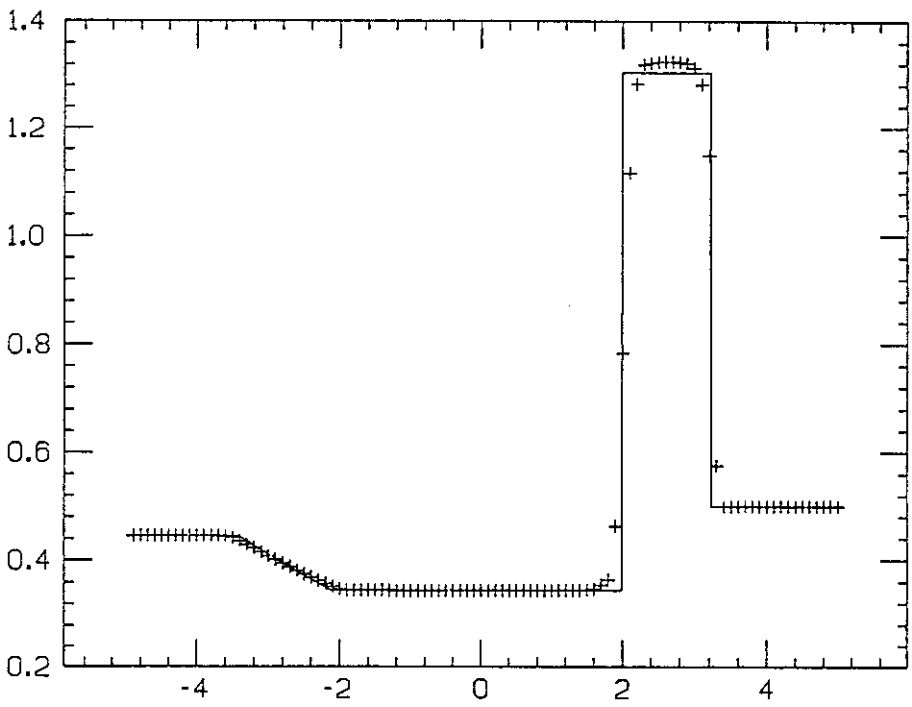

Figure 4: second order, component-wise limiter, (4.2b), density.

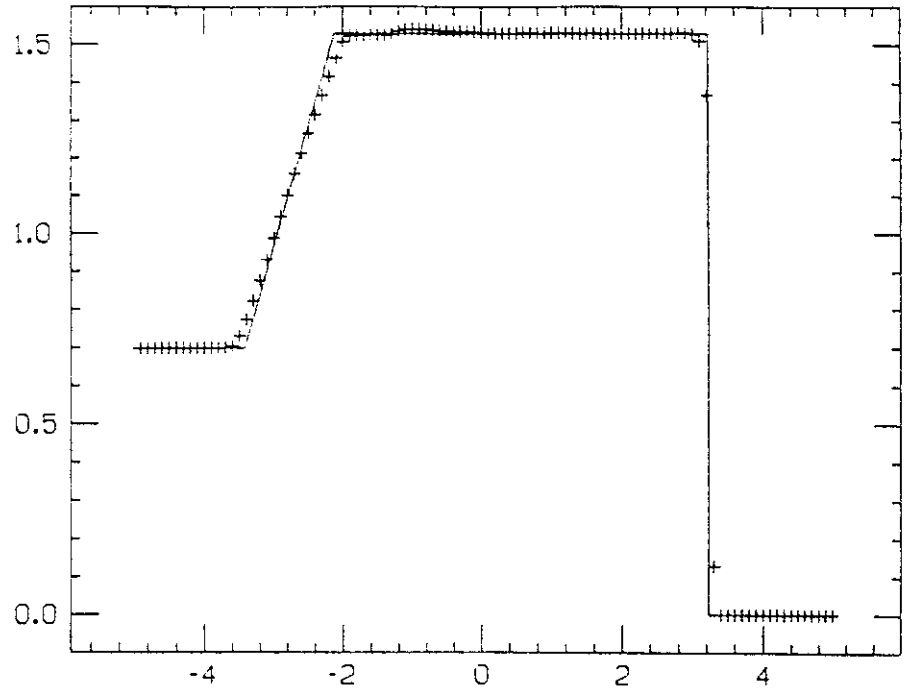

Figure 5: second order, component-wise limiter, (4.2b), velocity.

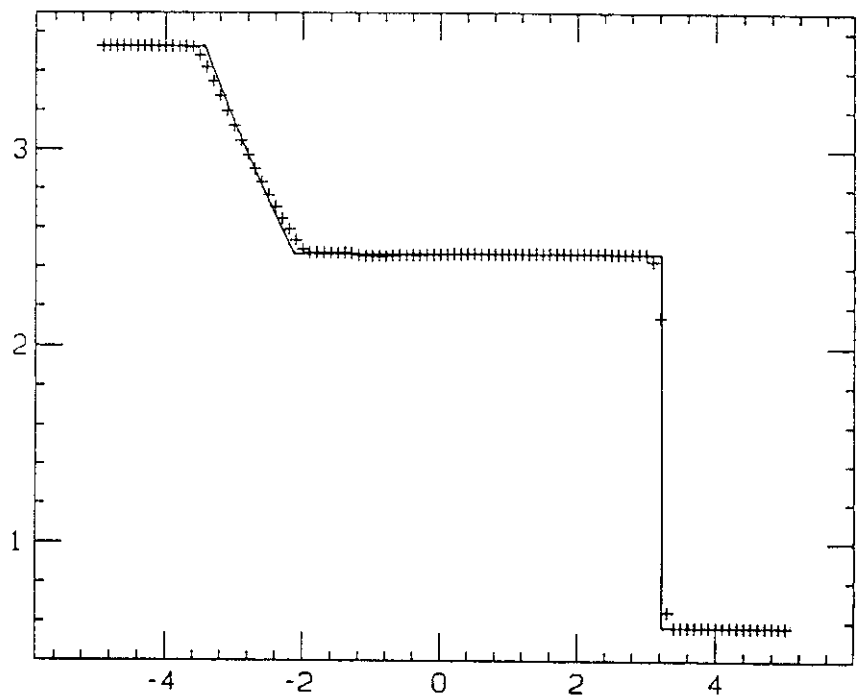

Figure 6: second order, component-wise limiter, (4.2b), pressure.

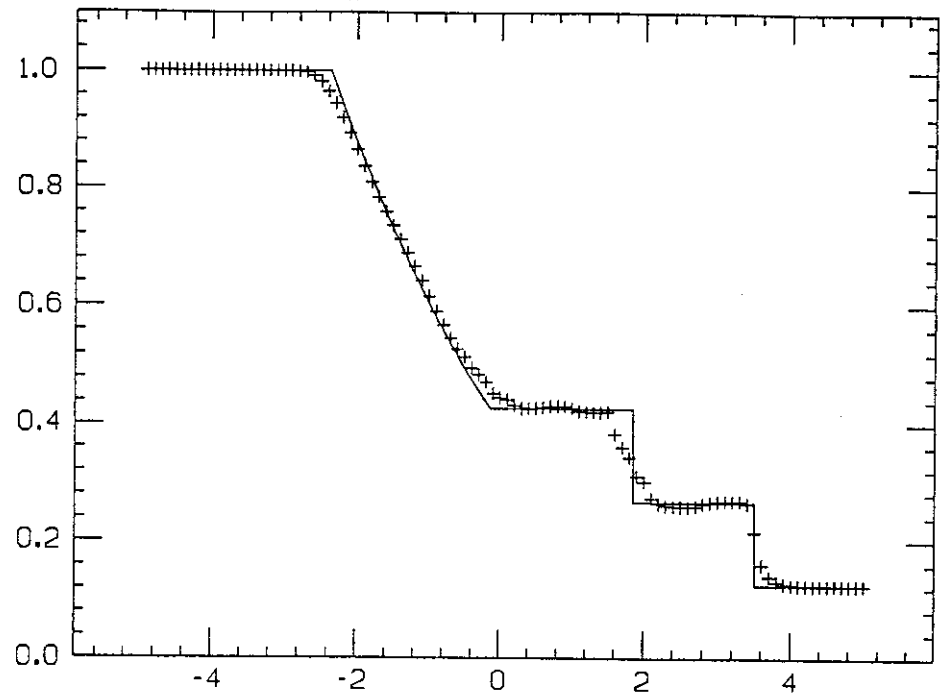

Figure 7: third order, component-wise limiter, (4.2a), density.

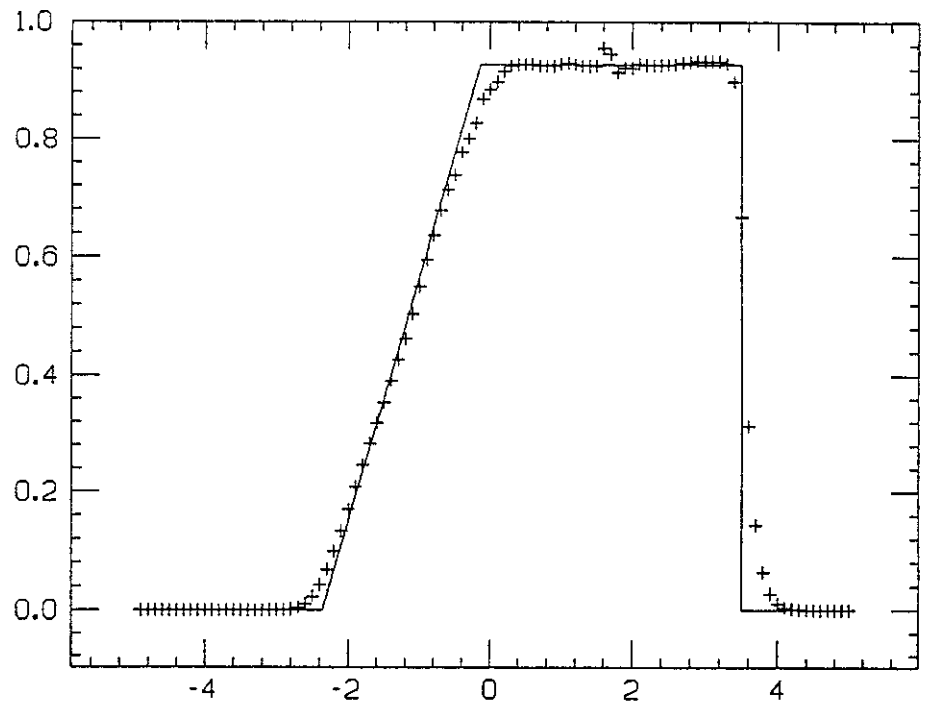

Figure 8: third order, component-wise limiter, (4.2a), velocity.

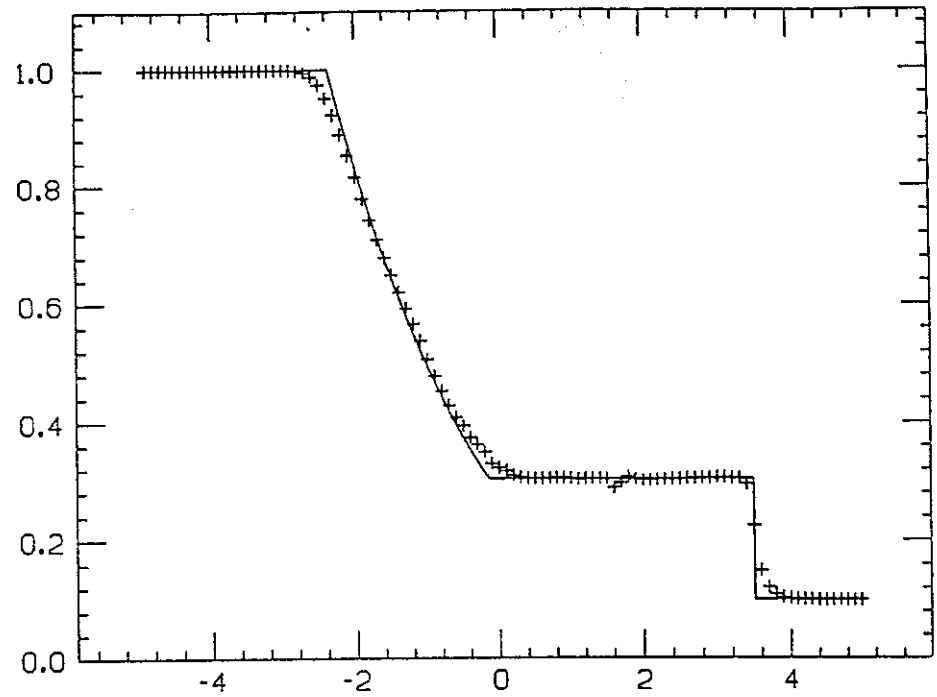

Figure 9: third order, component-wise limiter, (4.2a), pressure.

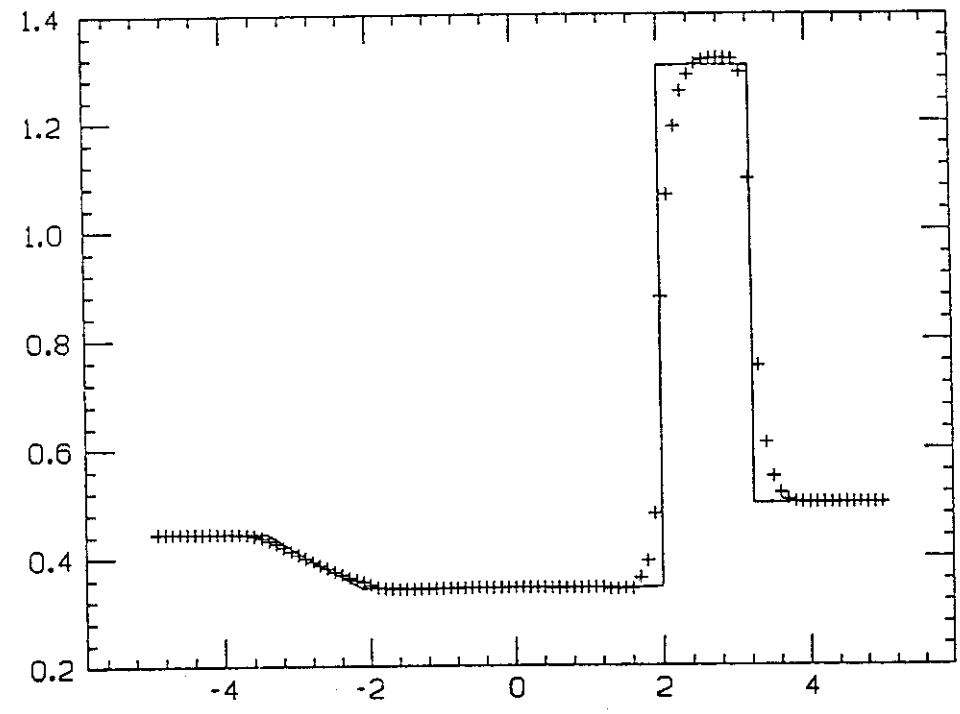

Figure 10: third order, component-wise limiter, (4.2b), density.

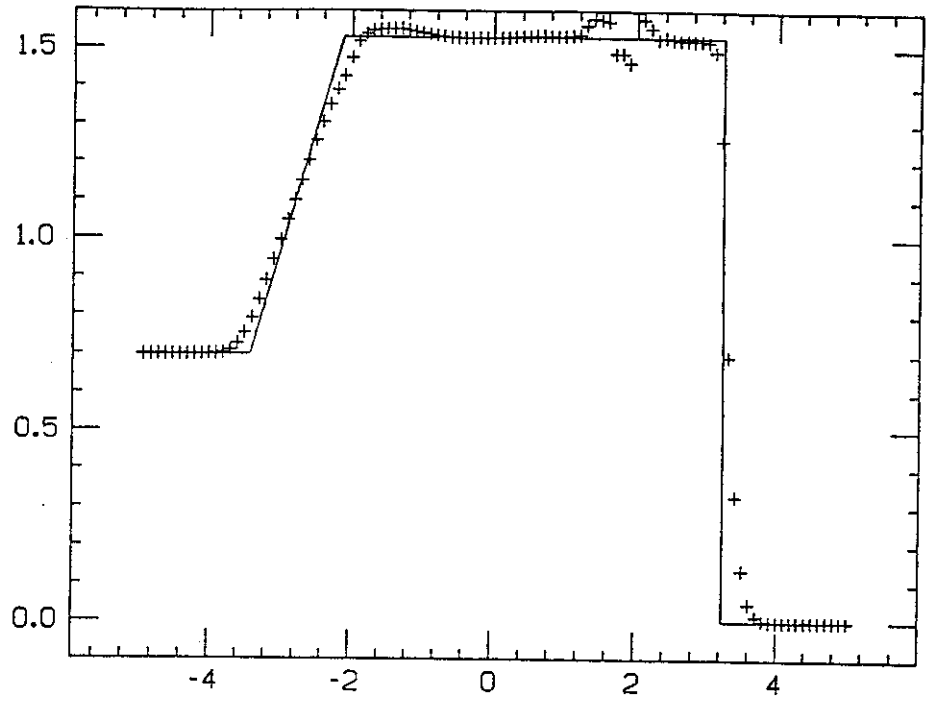

Figure 11: third order, component-wise limiter, (4.2b), velocity.

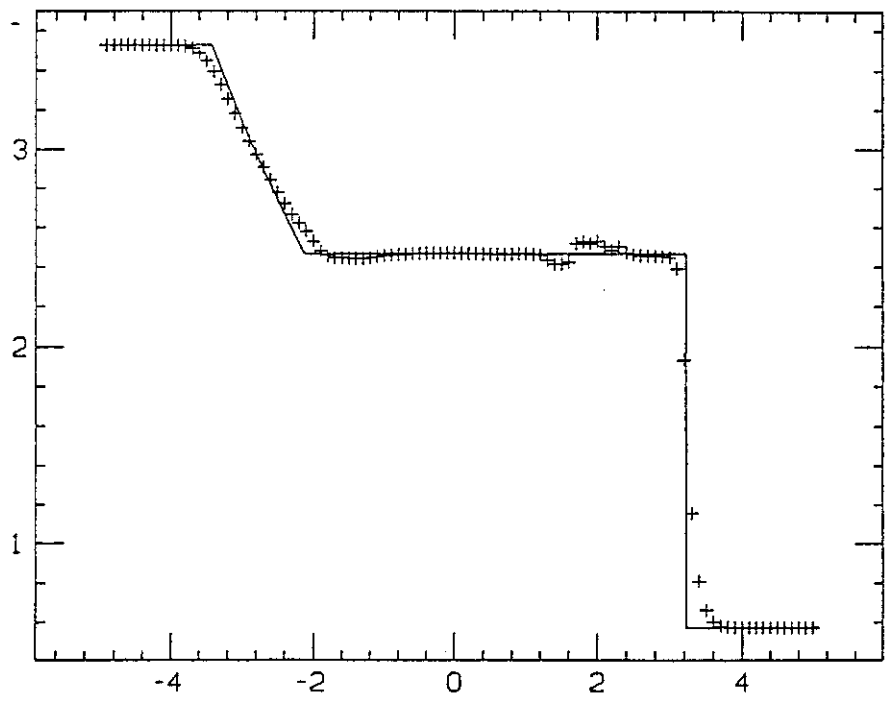

Figure 12: third order, component-wise limiter, (4.2b), pressure.

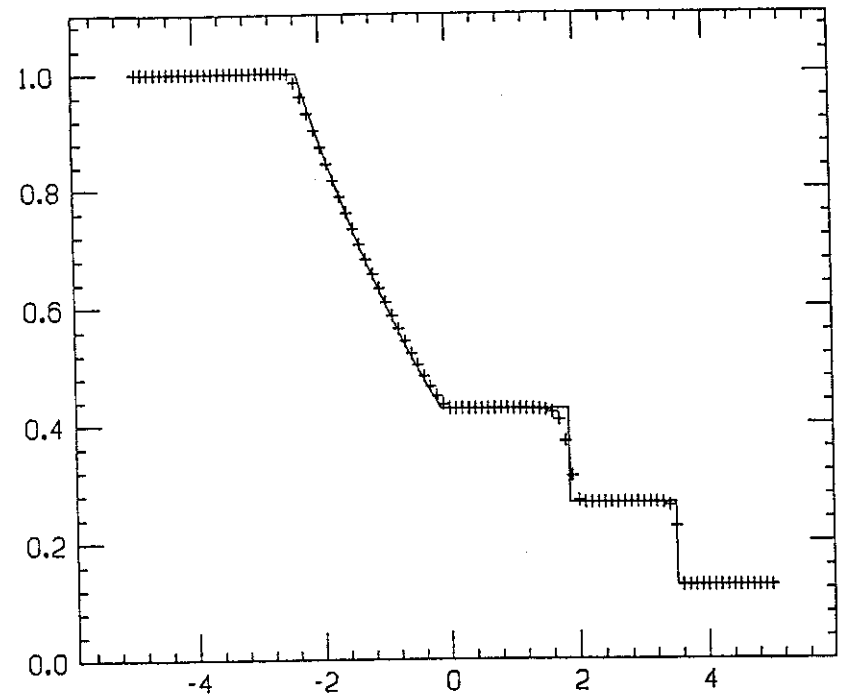

Figure 13: second order, characteristic-wise limiter, (4.2a), density.

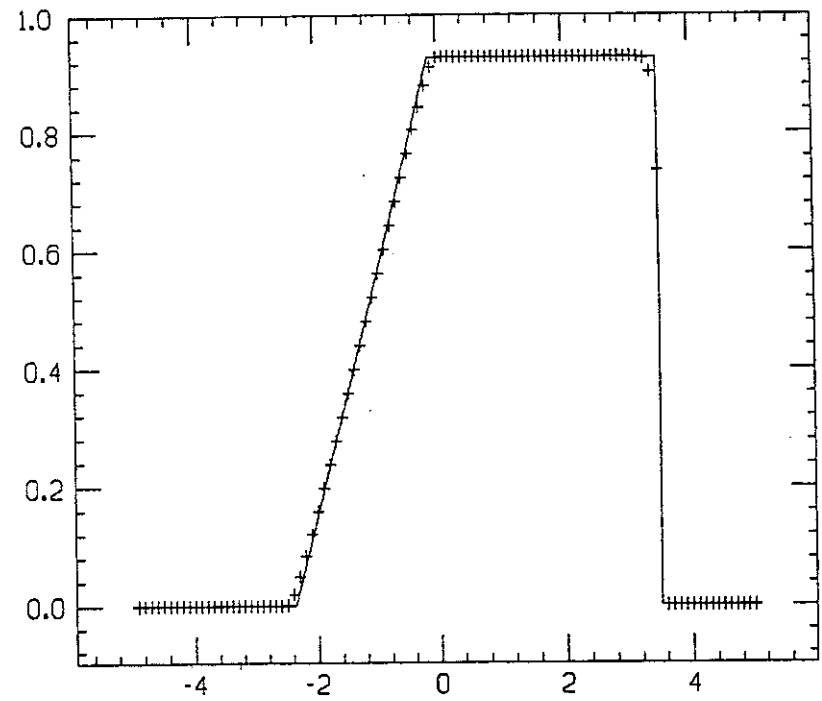

Figure 14: second order, characteristic-wise limiter, (4.2a), velocity.

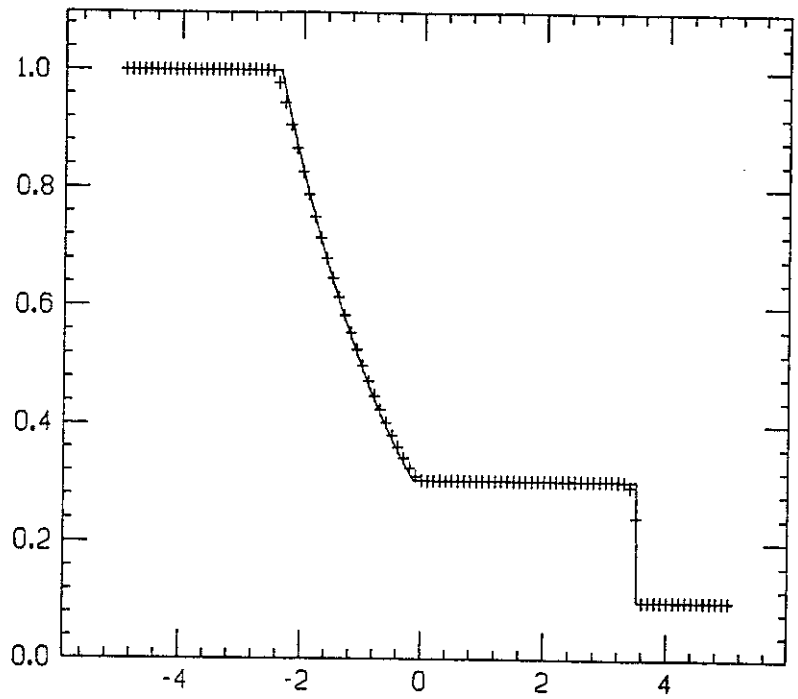

Figure 15: second order, characteristic-wise limiter, (4.2a), pressure.

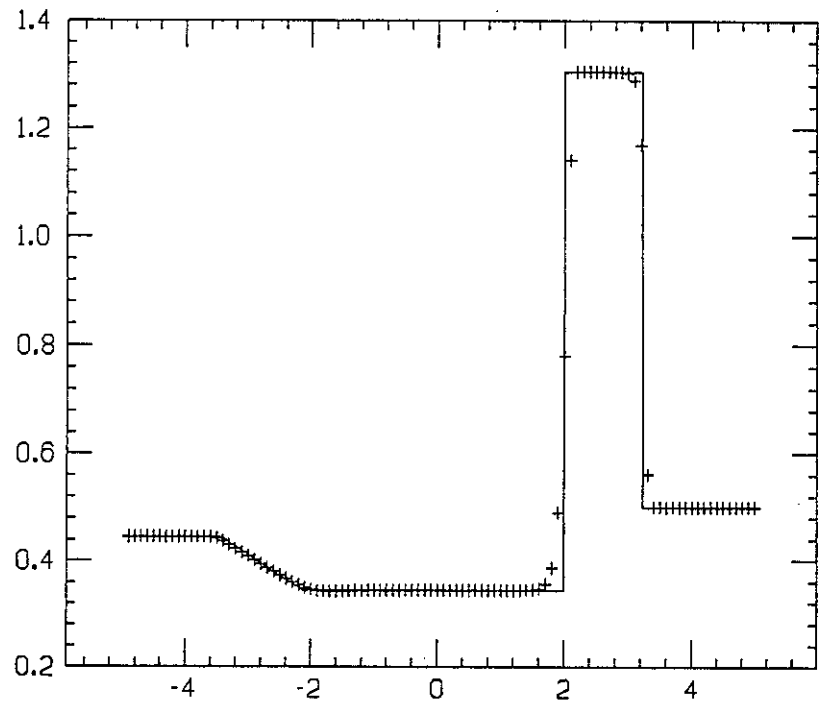

Figure 16: second order, characteristic-wise limiter, (4.2b), density.

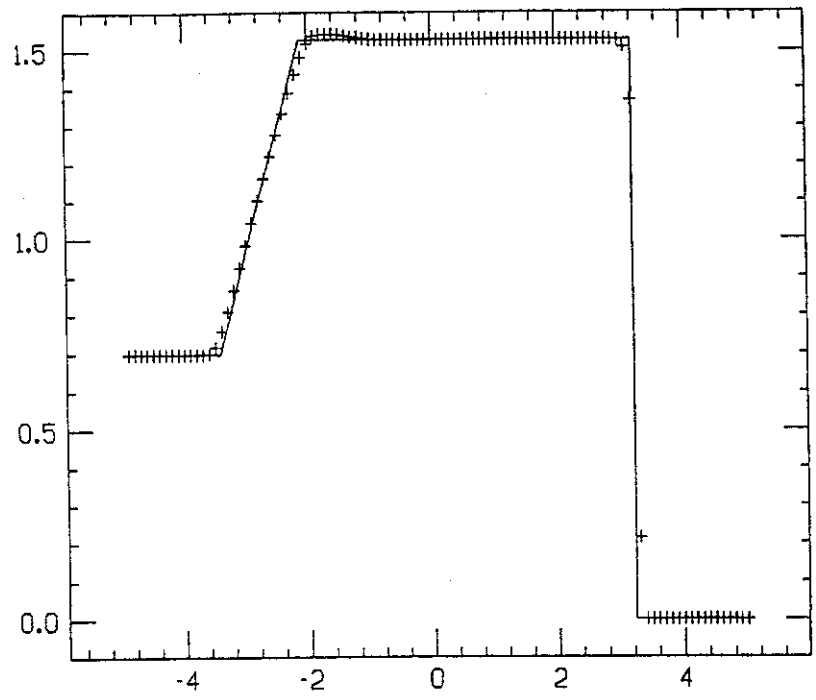

Figure 17: second order, characteristic-wise limiter, (4.2b), velocity.

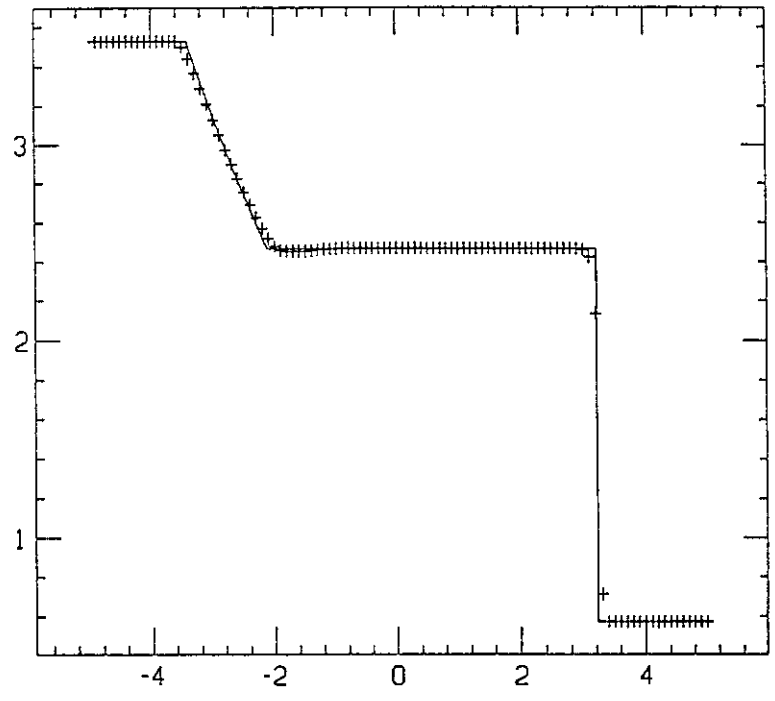

Figure 18: second order, characteristic-wise limiter, (4.2b), pressure.

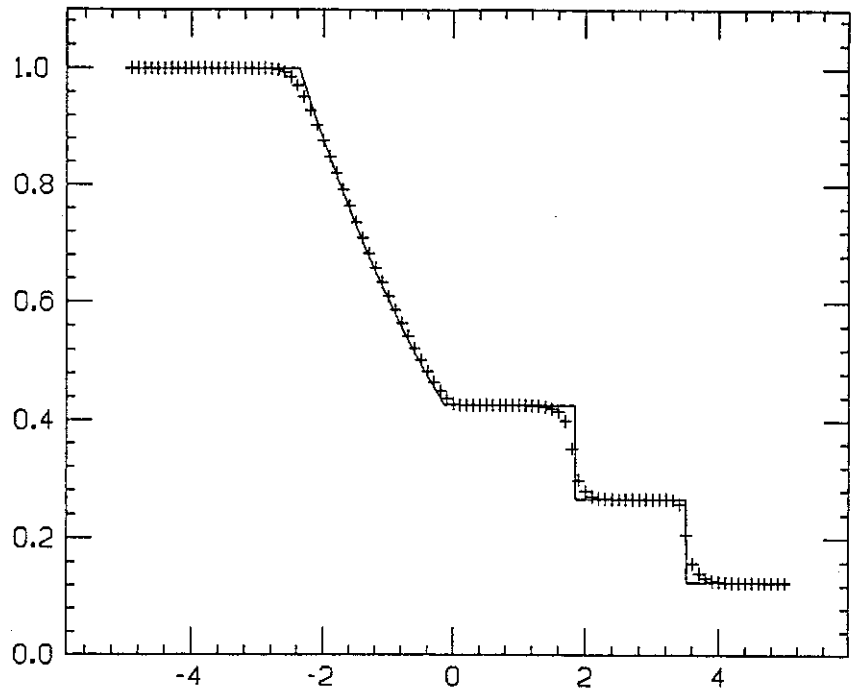

Figure 19: third order, characteristic-wise limiter, (4.2a), density.

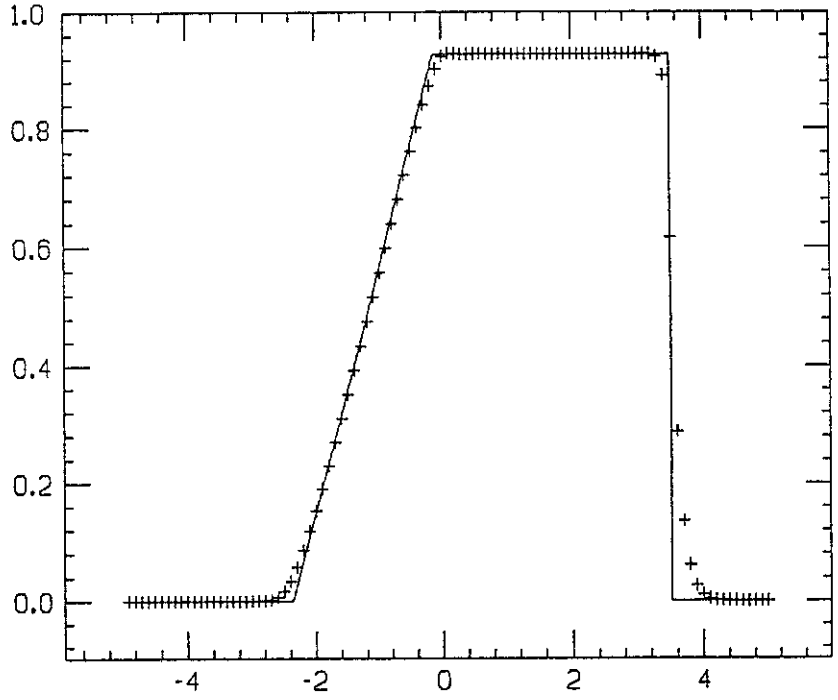

Figure 20: third order, characteristic-wise limiter, (4.2a), velocity.

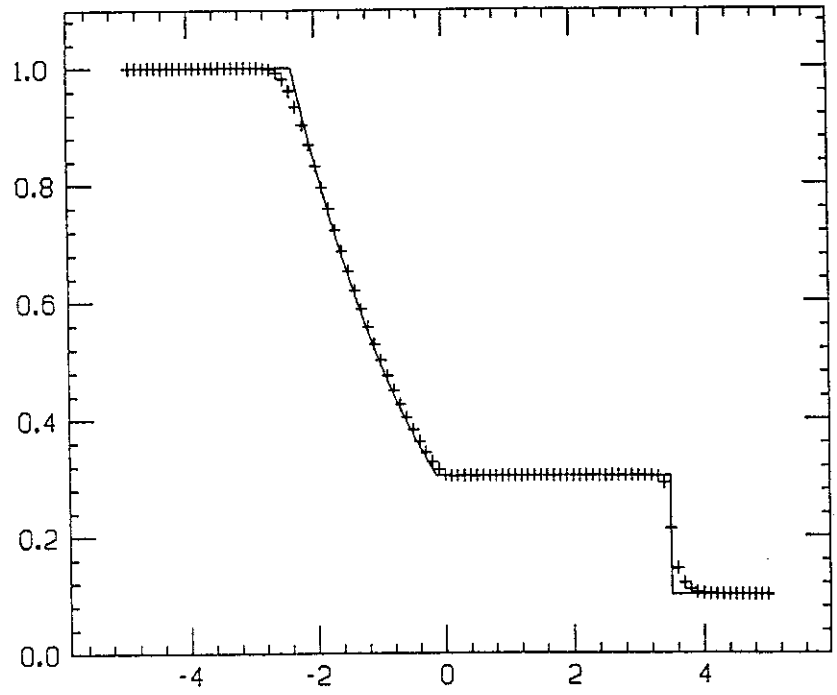

Figure 21: third order, characteristic-wise limiter, (4.2a), pressure.

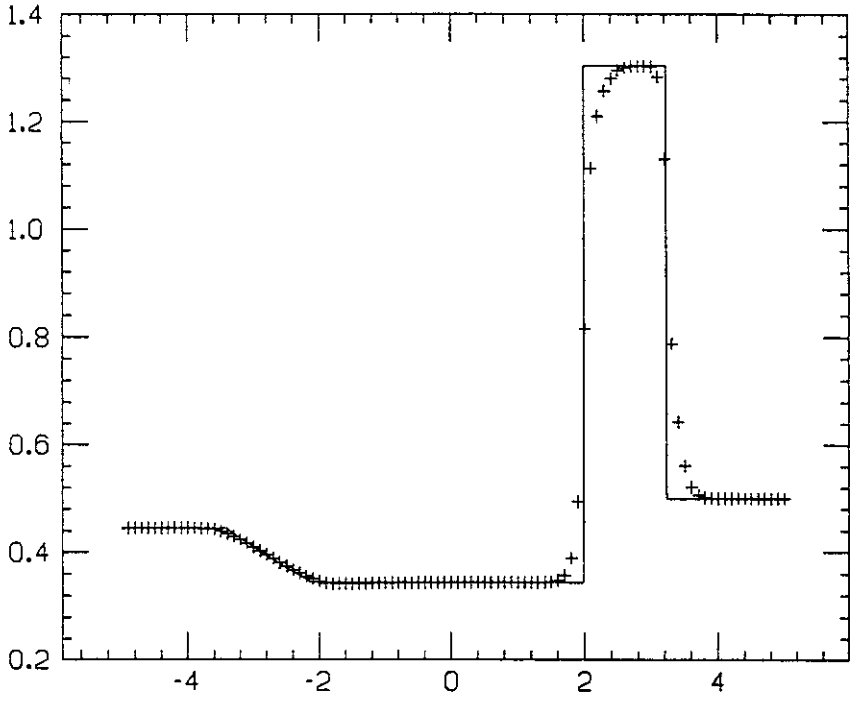

Figure 22: third order, characteristic-wise limiter, (4.2b), density.

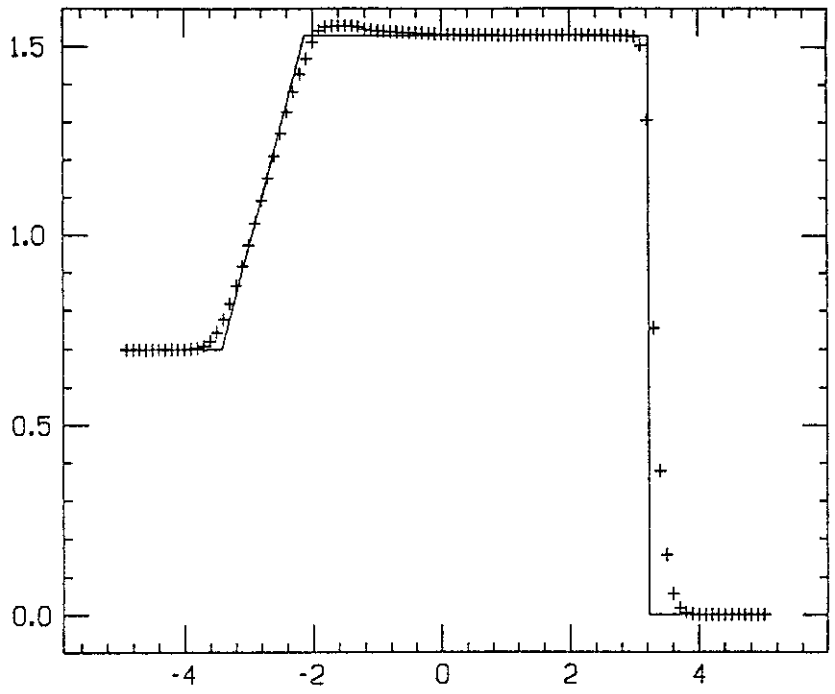

Figure 23: third order, characteristic-wise limiter, (4.2b), velocity.

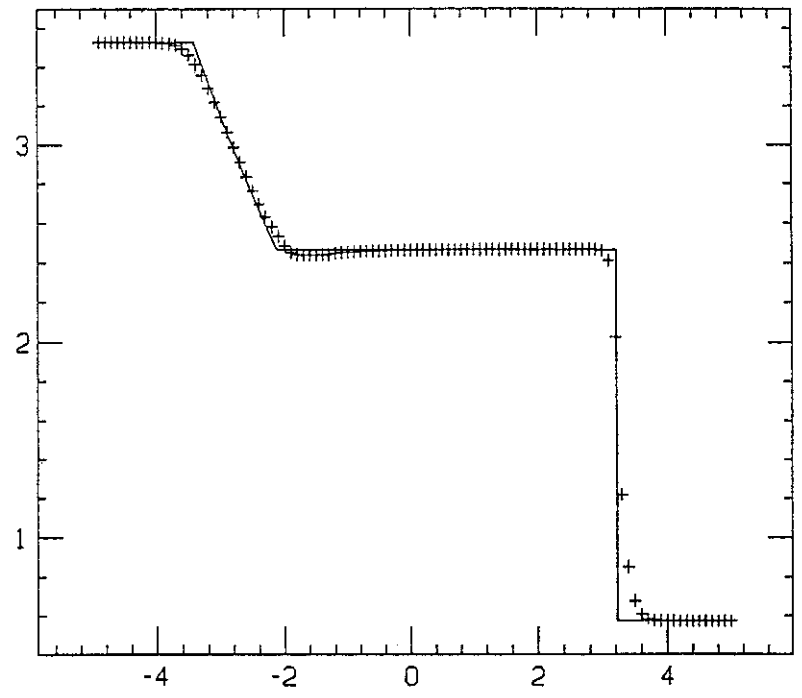

Figure 24: third order, characteristic-wise limiter, (4.2b), pressure.

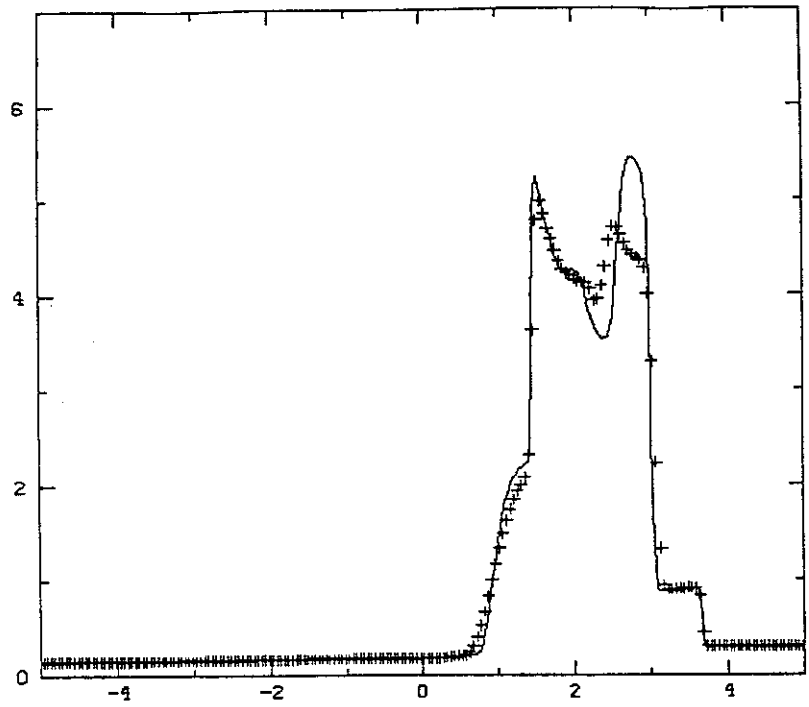

Figure 25a: second order, characteristic-wise limiter, 200 and 400 points, density.

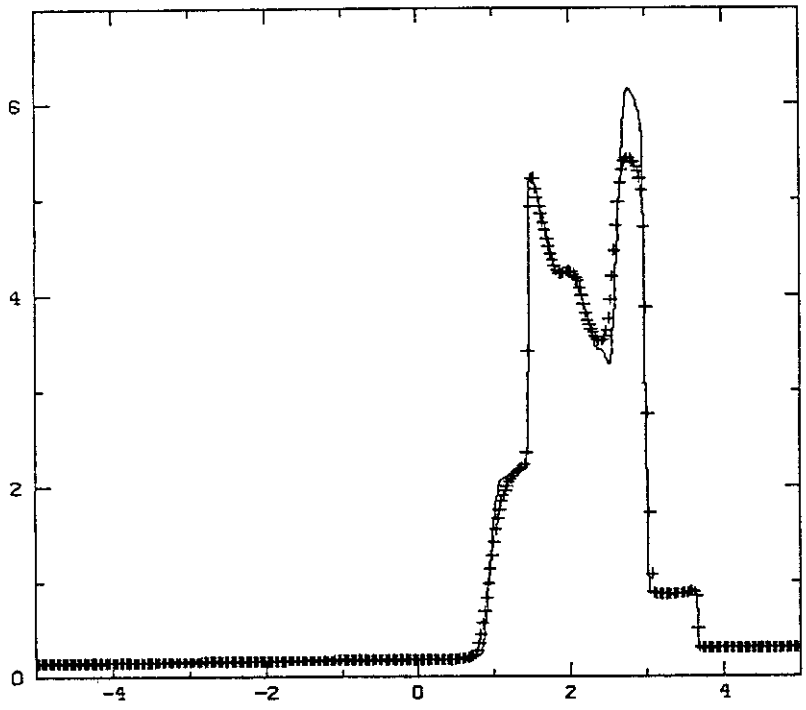

Figure 25b: second order, characteristic-wise limiter, 400 and 800 points, density.

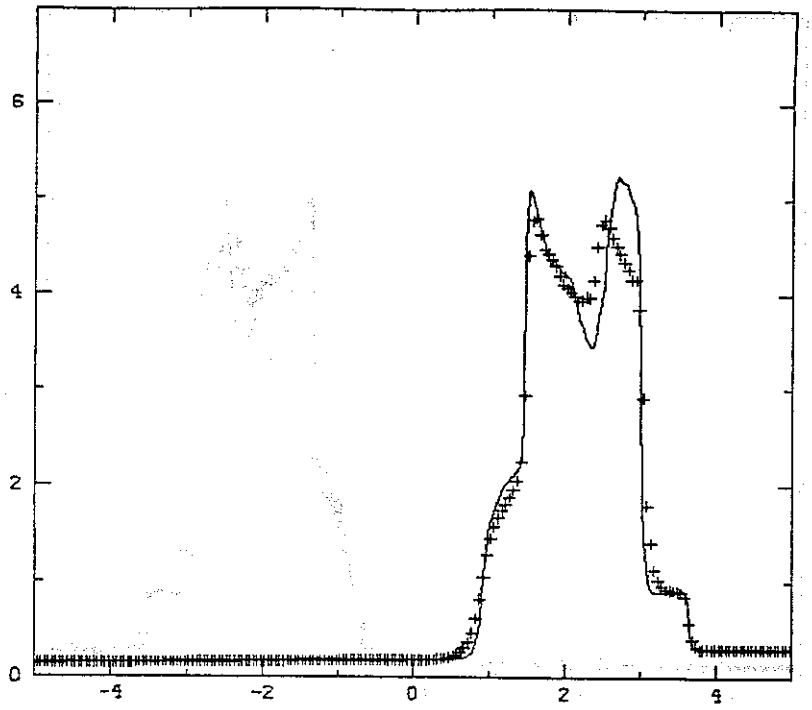

Figure 26a: third order, characteristic-wise limiter, 200 and 400 points, density.

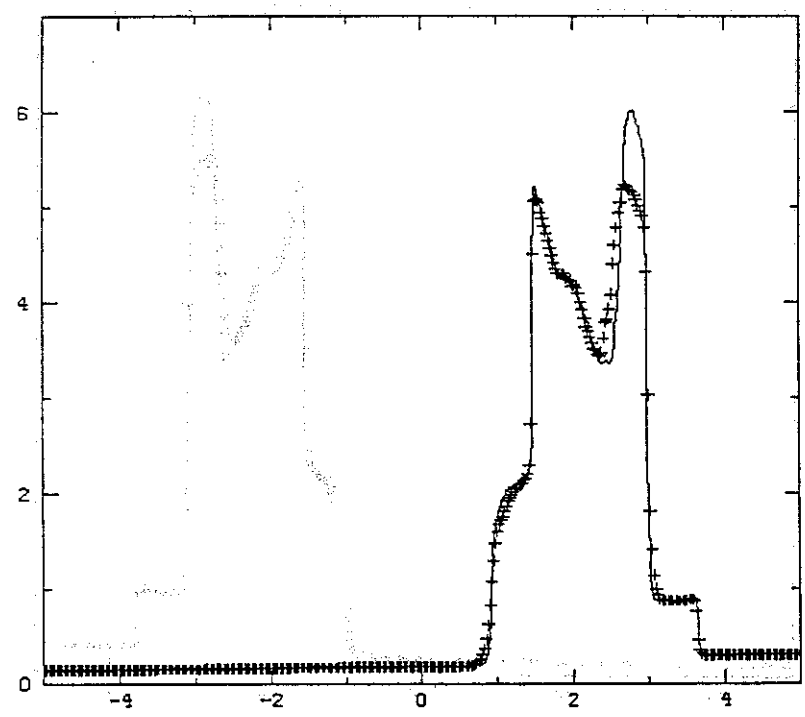

Figure 26b: third order, characteristic-wise limiter, 400 and 800 points, density.

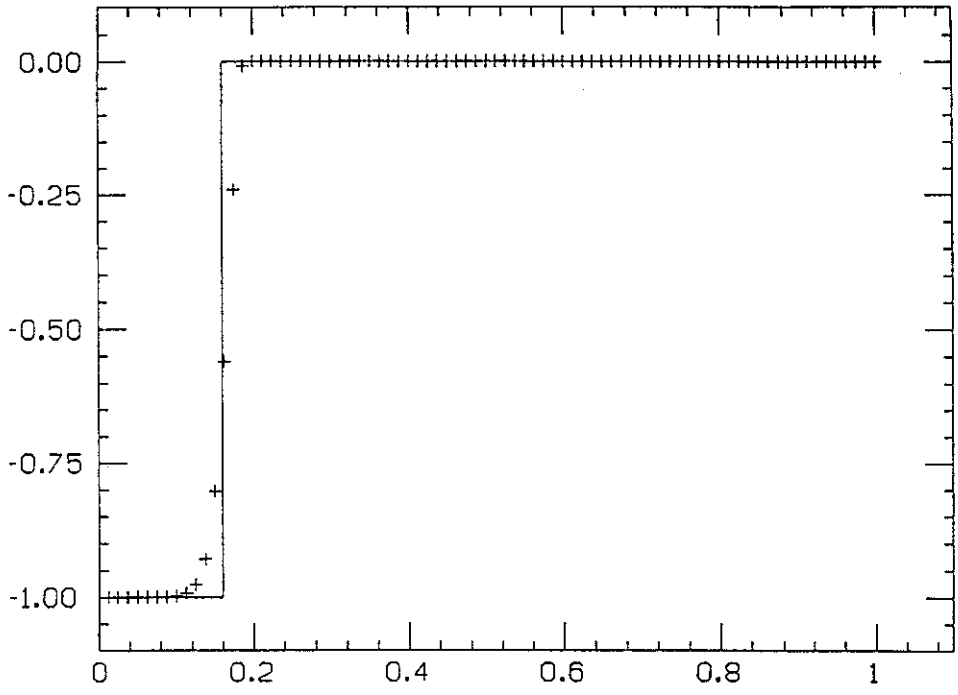

Figure 27: second order, $t = 0.5$, u.

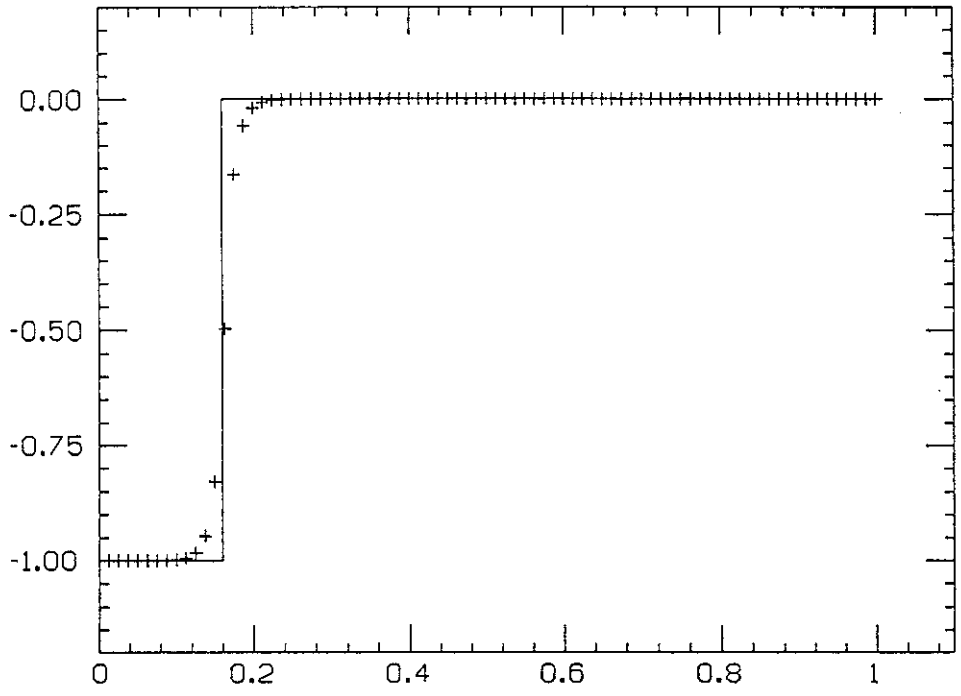

Figure 28: third order, $t = 0.5$, u.

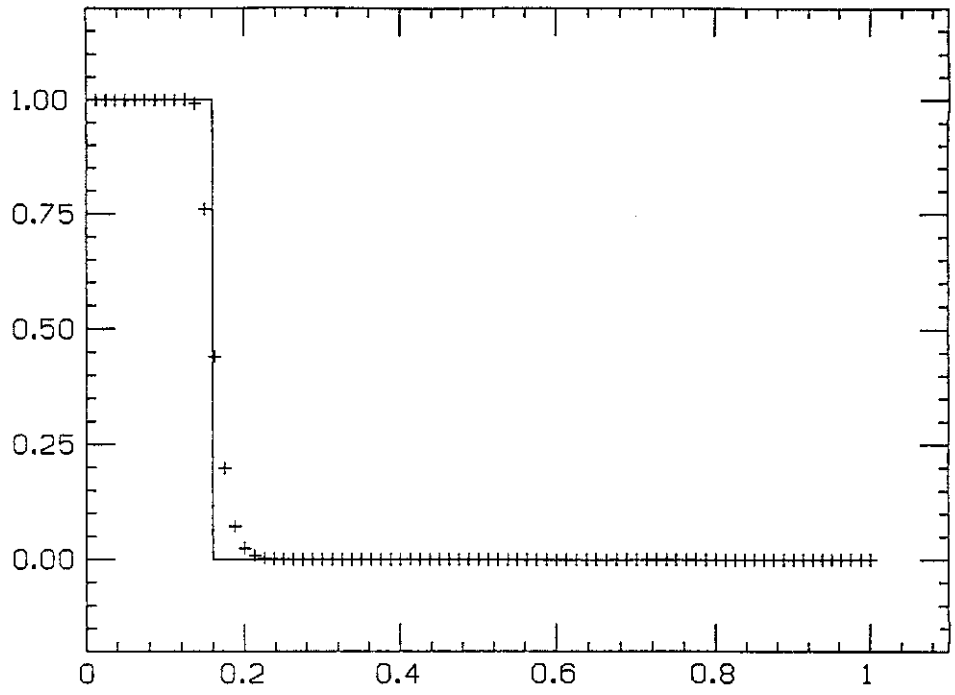

Figure 29: second order, $t = 0.5$, v.

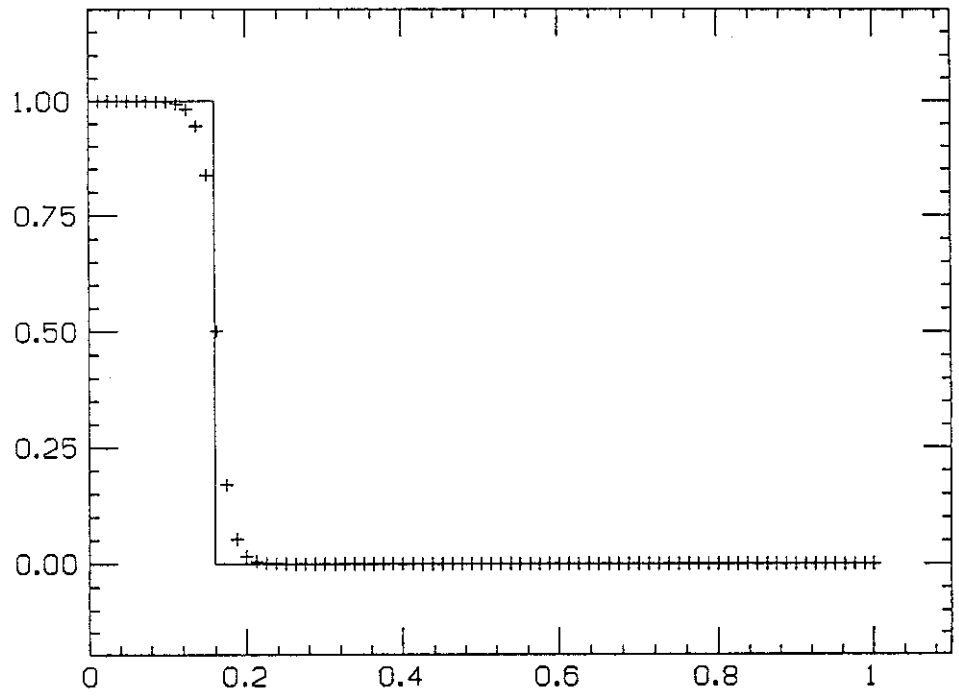

Figure 30: third order, $t = 0.5$, v.

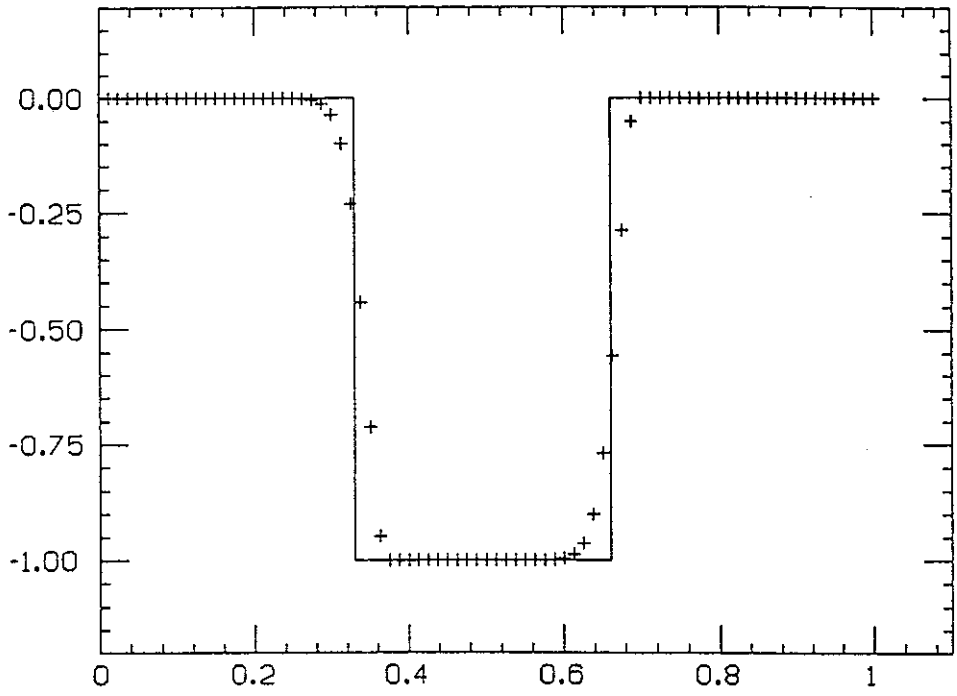

Figure 31: second order, $t = 1$, u.

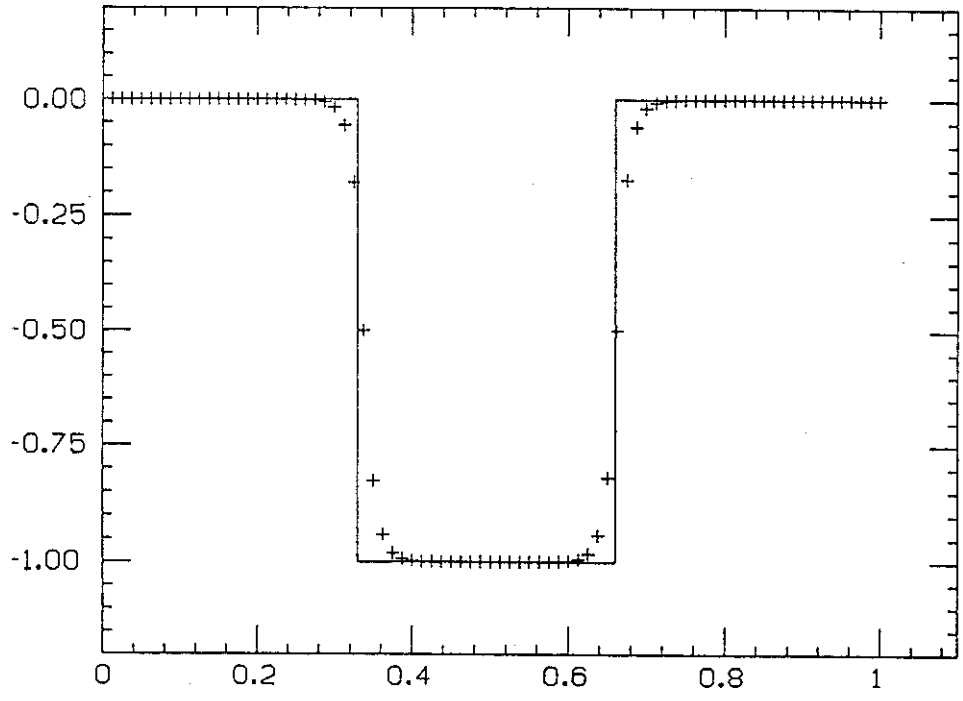

Figure 32: third order, $t = 1$, u.