A Discontinuous finite element method for scalar nonlinear conservation laws

Veerappa Gowda, Jérôme Jaffré

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A discontinuous finite element method for scalar nonlinear conservation laws

Veerappa GOWDA
Jérôme JAFFRÉ

N° 1848
Février 1993

PROGRAMME 6

Calcul Scientifique,
Modélisation et
Logiciels numériques

A Discontinuous Finite Element Method for Scalar Nonlinear Conservation Laws

Veerappa Gowda
Tata Institute of Fundamental Research, P.B. 1234, Bangalore 560012, India

Jérôme Jaffré
INRIA, BP 105, 78153 Le Chesnay Cédex, France

February 9, 1993

Abstract

Solutions of scalar nonlinear conservation laws are calculated by using discontinuous finite elements in one or in several dimensions. The standard first order finite difference scheme is obtained with piecewise constant approximations, while higher degree piecewise polynomial approximations give more accurate schemes. At discontinuities of the approximate solution, numerical fluxes are calculated by one-dimensional approximate Riemann solvers. The method is stabilized with truely multidimensional slope limiters. Special attention is given to piecewise linear approximation which is shown to be total variation diminishing and convergent. In two dimensions numerical experiments are presented on structural as well as on unstructured meshes.

Keywords : conservation laws, nonlinear equations, finite elements, finite volumes.

Une Méthode d'Eléments Finis Discontinus pour les Lois de Conservation Scalaires Non Linéaires

Résumé

Les solutions des lois de conservation scalaires non linéaires sont calculées à l'aide d'éléments finis discontinus en dimension un ou supérieure à un. Le schéma standard aux différences finis d'ordre un est obtenu avec une approximation constante par morceaux, tandis que des approximations polynômiales de degré plus élevé fournissent des schémas plus précis. Aux points de discontinuité de la solution approchée, les flux numériques sont calculés par des solveurs de Riemann unidimensionnels. La méthode est stabilisée avec des limiteurs de pente multidimensionnels. Une attention particulière est accordée au cas de l'approximation linéaire dont on montre en dimension un qu'elle est à variation totale diminuante et convergente. En dimension 2 des résultats numériques sont présentés pour des maillages structurés et non structurés.

Mots clés : Lois de conservation, équations hyperboliques non linéaires, éléments finis, volumes finis.

1 Introduction

Several finite element approaches have already been used to calculate solutions of hyperbolic equations with high resolution : a conforming approach using control volumes associated to the vertices of the mesh [23], the streamline diffusion Petrov-Galerkin method [14],[15], and characteristics Galerkin methods [10],[20],[2]. The method described below belongs to a different class, that of cell-centered finite volume methods.

In this paper, we construct accurate Godunov methods for scalar conservation laws in several dimensions by using discontinuous finite elements. We obtain a method which is truely multidimensional and can be used on unstructured meshes as well as on rectangular meshes.

The origine of discontinuous finite element methods can be traced back to the nuclear engineering litterature [21] where they have been used to solve linear first order hyperbolic equations. Still in the linear case, discontinuous finite elements have been analyzed in [19] and this analysis has been improved later in [16],[22]. The method was extended to solve nonlinear scalar conservation laws arising in reservoir simulation, first without slope limitation [7],[4] and then with a multidimensional slope limiter [3],[6],[5]. Another version of the discontinuous finite element method with a different time- stepping and a different slope limiter has been developped and analyzed in [8].

On one hand, one advantage of discontinuous finite element methods is that they can be easily interpreted as cell-centered finite volume methods, methods which are very popular in many areas of physics and enginneering. On the other hand, higher order Godunov finite difference schemes are now generally constructed through a discontinuous piecewise polynomial representation of the solution, linear in the MUSCL scheme [25], parabolic in the PPM method [9]. Therefore it seems natural to try to interpret these methods in terms of finite element methods.

In our method, the solution is approximated by discontinuous piecewise linear or bilinear polynomials. Numerical fluxes are calculated on the edges at integration points through one-dimensional Riemann solvers. To prevent spurious oscillations, we extend Van Leer's slope limiter [24] to the multidimensional case. Though the numerical scheme is presented in one and two dimensions its extension to three dimensions is straightforward. The method was also extended to the case of systems of conservation laws [17].

A finite volume method in two dimensions on rectangular meshes using discontinuous bilinears and a similar slope limiter can be found in [1], but the values inside the cells are calculated by tracing back the characteristics instead of using a variational formula.

The equation to solve is

$$\frac{\partial u}{\partial t} + div\vec{f}(u) = 0\,, x \in \mathbf{R}^n\,, \tag{1.1}$$

$$u(x,0) = u_0(x)\,, x \in \mathbf{R}^n\,, \tag{1.2}$$

given $\vec{f} : [0,1] \rightarrow \mathbf{R}^n$ and $u_0 : \mathbf{R}^n \rightarrow [0,1]$. In the following section we consider one-dimensional semi-discretization in space while in section 3, still in one dimension, equation (1.1) is discretized in time and space. In the last section we describe the two-dimensional method and present numerical experiments.

2 One-dimensional semi-discretization in space

2.1 Discontinuous finite elements

Let us denote by $\ldots < x_{i-1/2} < x_{i+1/2} < \ldots$ and by $K(i) =]\, x_{i-1/2}, x_{i+1/2}\, [$, $i \in \mathbf{Z}$ the points and the intervals of the discretization of $\mathbf{R}$. The measure of the intervals is h and we introduce the approximation space

$$V^k = \{v \mid_{K(i)} \in P^k(x_{i-1/2}, x_{i+1/2})\}$$

where $P^k(x_{i-1/2}, x_{i+1/2})$ denotes the set of polynomials of degree k defined on the interval $(x_{i-1/2}, x_{i+1/2})$. Note that functions of V^k are in general discontinuous at the points of discretization so that we denote by $v^L_{i+1/2}$ and $v^R_{i+1/2}$ respectively the lefthand and righthand limits of a function in V^k. Also we shall need in the following the averages of a function v over the discretization intervals that we denote by

$$v_i = \frac{1}{h} \int_{K(i)} v \, dx, \quad i \in \mathbf{Z}.$$

The approximation equation is obtained by multiplying equation (1.1) by test functions in V^k, by integrating over the intervals and by integrating by parts the term containing the derivative with respect to space. Thus the approximate problem consists of seeking $u_h \in V^k$ solution of

$$(2.1) \int_{K(i)} \frac{\partial u_h(t)}{\partial t} v \, dx - \int_{K(i)} f(u_h(t)) \frac{dv}{dx} \, dx + F_{i+1/2}(u_h(t)) v^L_{i+1/2} - F_{i-1/2}(u_h(t)) v^R_{i-1/2} = 0, \qquad \text{for } v \in V^k, \quad i \in \mathbf{Z}, \quad t > 0.$$

The numbers $F_{i+1/2}(u_h(t)), i \in \mathbf{Z}$ are called numerical fluxes since they approximate the boundary terms resulting from the integration by parts, i.e. the fluxes of f across the extremities of the interval of discretization. Therefore to preserve mass balance, they have to be uniquely defined at these points despite of the discontinuity of u_h. The next paragraph is devoted to their calculation.

Remark 1 *Note that for $k = 0$ scheme (2.1) reduces to the finite difference scheme :*

$$\frac{du_i}{dt} + \frac{F_{i+1/2} - F_{i-1/2}}{h} = 0, \quad \in \mathbf{Z}, \quad t > 0,$$

and that, for $k > 0$, when taking for v in (2.1) the characteristic functions of the intervals, we obtain the conservation equations

$$(2.2) \qquad \int_{K(i)} \frac{\partial u_h}{\partial t} \, dx + \frac{F_{i+1/2} - F_{i-1/2}}{h} = 0 \quad i \in \mathbf{Z}, \quad t > 0.$$

2.2 Numerical flux

The numerical flux $F_{i+1/2}(t)$ should be calculated by solving a generalized Riemann problem at point $x_{i+1/2}$ with initial data $u_h(t)$. However, due to its complexity, such a problem is solved only approximately by replacing it by a standard Riemann problem with initial data $u^L_{i+1/2}(t)$ for $x < x_{i+1/2}$, $u^R_{i+1/2}(t)$ for $x > x_{i+1/2}$. Godunov ([12]), Engquist and Osher ([11]) among others have given formulas for $F_{i+1/2}(t)$ as a function of $u^L_{i+1/2}(t)$ and $u^R_{i+1/2}(t)$:

$$F_{i+1/2}(u_h(t)) = F(u^L_{i+1/2}(t), u^R_{i+1/2}(t)). \tag{2.3}$$

For Godunov numerical flux the function F is

$$F(u,v) = \begin{cases} \min\limits_{W\in[u,v]} f(w) & \text{if} \quad u \le v, \\ \max\limits_{W\in[v,u]} f(w) & \text{if} \quad u > v, \end{cases}$$

and for Engquist-Osher's it is

$$F(u,v) = [f(u) + f(v) + \int_v^u \mid f'(w) \mid \, dw]/2.$$

In these formulas, the numerical flux function F has the following properties :

(2.4) F is increasing (resp. decreasing) with respect to its first (resp. second argument),

(2.5) F is consistent, i.e. $F(v,v) = f(v)$,

$$\begin{array}{rcl} 0 \le (F(u_1,v) - F(u_2,v))/(u_1-u_2) & \le & \left(\int_{u_2}^{u_1} (f'(u))^+ \, du\right)/(u_1-u_2), \\ 0 \ge (F(u,v_1) - F(u,v_2))/(v_1-v_2) & \ge & \left(\int_{u_2}^{u_1} (f'(u))^- \, du\right)/(v_1-v_2), \end{array} \tag{2.6}$$

Inequalities (2.6) implies that F is Lipschitz continuous with respect to each argument.

If the function f is increasing (resp. decreasing) in the interval $(u^L_{i+1/2}(t), u^R_{i+1/2}(t))$, then $F_{i+1/2} = f(u^L_{i+1/2}(t))$ (resp. $= f(u^R_{i+1/2}(t))$). In other words the numerical flux is calculated with the upstream value of u_h and the scheme is said to be upstream weighted.

In the linear case $f(u) = u$, the space discretization reduces to the one analyzed by Lesaint-Raviart ([19]). They have shown that the L^2 approximation error is of order h^{k+1} for smooth enough functions.

2.3 Slope limiter

For $k \ge 1$ scheme (2.1) do not have good stability properties and the calculated solution oscillates. To stabilize them, we extend to the discontinuous finite element method the notion of slope limiters already introduced by Van Leer ([24]) for finite difference schemes. We introduce a slope limitation operator L which associates to each function $u_h \in V^k$ a function $L(u_h) = w_h \in V^k$ satisfying

$$(\int_{K(i)} w_h \, dx)/h = w_i(t) = u_i(t), \tag{2.7}$$

in order to preserve mass balance, and

$$(2.8)\qquad (1-\alpha)u_i(t)+\alpha\min(u_{i-1}(t),u_i(t)) \le w^R_{i-1/2}(t) \le \\ \le (1-\alpha)u_i(t)+\alpha\max(u_{i-1}(t),u_i(t)),$$

$$(2.9)\qquad (1-\alpha)u_i(t)+\alpha\min(u_i(t),u_{i+1}(t)) \le w^L_{i+1/2}(t) \le \\ \le (1-\alpha)u_i(t)+\alpha\max(u_i(t),u_{i+1}(t)),$$

$$(2.10)\qquad 0 \le \alpha \le 1,$$

$$(2.11)\qquad w_h(x) \equiv u_i \text{ for } x \in K(i) \text{ if } u_i \ge \max(u_{i-1},u_{i+1}) \text{ or } u_i \le \min(u_{i-1},u_{i+1}),$$

in order to limit the slope.

Note that $w_h(t)$ is not uniquely defined by (2.7) (2.8) (2.9) (2.10) (2.11). To define it uniquely, one can also impose for instance that $w_h(t)$ be as close as possible to $u_h(t)$ with respect to the L^2 norm.

Then the slope limited version of scheme (2.1) can be written as

$$(2.12)\int_{K(i)} \frac{\partial u_h(t)}{\partial t} v\,dx - \int_{K(i)} f(L(u_h(t)))\frac{dv}{dx}\,dx + \\ F_{i+1/2}(Lu_h(t)))v^L_{i+1/2} - F_{i-1/2}(L(u_h(t)))v^R_{i-1/2} = 0, \text{ for } v \in V^k,\ i \in \mathbf{Z},\ t>0.$$

The parameter α controls the slope limitations and the corresponding added numerical diffusion. In the case of a piecewise linear approximation ($k=1$), when $\alpha=0$ the slope limited solution is piecewise constant so it is the strongest slope limitation possible ; then we are taken back to a first order finite difference scheme and the added numerical diffusion is maximum. On the other hand, the larger α is the looser is the slope limitation and the smaller is the added numerical diffusion. However, as we will see below, $\alpha \le 1$ is necessary for the method to be total variation diminishing.

2.4 Total variation

Theorem 2.1 *Assume that the numerical flux function is monotone (i.e. (2.4) holds). Then the solution u_h calculated by equation (2.12) is such that its averages over the discretization intervals are total variation diminishing, i.e.*

$$\sum_{i\in\mathbf{Z}} |\, u_i(t_1) - u_{i-1}(t_1)\,| \le \sum_{i\in\mathbf{Z}} |\, u_i(t_2) - u_{i-1}(t_2)\,|, \quad t_1 \ge t_2 \ge 0.$$

Proof :

Taking for v in (2.12) the characteristic functions of the intervals of discretization, using (2.7) and denoting $w_h = L(u_h)$, we obtain

$$du_i/dt + (F_{i+1/2}(w_h) - F_{i-1/2}(w_h))/h = 0, \quad i \in \mathbf{Z}, \quad t>0,$$

that we write in the incremental form

$$du_i/dt = d_1(i)(u_{i+1} - u_i) + d_{-1}(i)(u_i - u_{i-1})$$

where

$$d_1(i) = (1/h)[F(w^L_{i+1/2}, w^R_{i-1/2}) - F(w^L_{i+1/2}, w^R_{i+1/2})]/(u_{i+1} - u_i),$$
$$d_{-1}(i) = (1/h)[F(w^L_{i-1/2}, w^R_{i-1/2}) - F(w^L_{i+1/2}, w^R_{i-1/2})]/(u_i - u_{i-1}).$$

Using the continuous time version of Harten's lemma ([13]), it is sufficient to show that $d_i(i) \geq 0$ and $d_{-1}(i) \leq 0$ for $i \in \mathbf{Z}$. Slope limiting relations (2.8),(2.9), (2.11) imply that we may define two real numbers $\sigma^R_{i-1/2}, \sigma^L_{i+1/2}$ such that

$$(2.13) \quad \begin{cases} w^R_{i-1/2} = u_i + \alpha\sigma^R_{i-1/2}(u_{i-1} - u_i), & 0 \leq \sigma^R_{i-1/2} \leq 1, \\ w^L_{i+1/2} = u_i + \alpha\sigma^L_{i+1/2}(u_{i+1} - u_i), & 0 \leq \sigma^L_{i+1/2} \leq 1, \\ \sigma^R_{i-1/2} = \sigma^L_{i+1/2} = 0 \text{ if } u_i \geq \max(u_{i-1}, u_{i+1}) \text{ or } u_i \leq \min(u_{i-1}, u_{i+1}). \end{cases}$$

Now $d_1(i)$ and $d_{-1}(i)$ can be written as

$$d_1(i) = (1/h)\left[-\frac{F(w^L_{i+1/2}, w^R_{i+1/2}) - F(w^L_{i+1/2}, w^R_{i-1/2})}{w^R_{i+1/2} - w^R_{i-1/2}}\right]\left[\frac{w^R_{i+1/2} - w^R_{i-1/2}}{u_{i+1} - u_i}\right]$$

$$d_{-1}(i) = (1/h)\left[-\frac{F(w^L_{i+1/2}, w^R_{i-1/2}) - F(w^L_{i-1/2}, w^R_{i-1/2})}{w^L_{i+1/2} - w^L_{i-1/2}}\right]\left[\frac{w^L_{i+1/2} - w^L_{i-1/2}}{u_i - u_{i-1}}\right].$$

Hence :

$$d_1(i) = (1/h)\left[-\frac{F(w^L_{i+1/2}, w^R_{i+1/2}) - F(w^L_{i+1/2}, w^R_{i-1/2})}{w^R_{i+1/2} - w^R_{i-1/2}}\right]\left[1 - \alpha\sigma^R_{i+1/2} + \alpha\sigma^R_{i-1/2}\frac{u_i - u_{i-1}}{u_{i+1} - u_i}\right],$$

$$d_{-1}(i) = (1/h)\left[-\frac{F(w^L_{i+1/2}, w^R_{i-1/2}) - F(w^L_{i-1/2}, w^R_{i-1/2})}{w^L_{i+1/2} - w^L_{i-1/2}}\right]\left[1 - \alpha\sigma^L_{i-1/2} + \alpha\sigma^L_{i+1/2}\frac{u_{i+1} - u_i}{u_i - u_{i-1}}\right].$$

Let us consider $d_1(i)$. The term in the first brackets in the right-hand side is positive since F is a monotone numerical flux function. For the term in the second brackets, we have $0 \leq \alpha\sigma^R_{i+1/2} \leq 1$ and, because of (2.13), we obtain

$$\alpha\sigma^R_{i-1/2}\frac{u_i - u_{i-1}}{u_{i+1}u_i} \begin{cases} = 0 & \text{if } u_i \geq \max(u_{i-1}, u_{i+1}) \text{ or } u_i \leq \min(u_{i-1}, u_{i+1}), \\ \geq 0 & \text{otherwise.} \end{cases}$$

Therefore $d_1(i) \geq 0$ and a similar argument holds to prove that $d_{-1}(i) \leq 0$.

3 Piecewise linears in space and explicit time discretization

3.1 Formulation of the scheme

Denote by $0 = t_0 \leq t_1 \leq \ldots \leq t_n \ldots, n \in \mathbf{N}$ with $\Delta t_n = t_{n+1} - t_n$ a time discretization. First we discretize in time using forward differencing. This gives a scheme which is first order in time. Considering space discretization, as mentionned in (1) we obtain for $k = 0$ the standard first order finite difference scheme. Using the finite element formulation, the natural way to obtain higher order in space is to increase the degree of the approximation polynomials, say $k = 1$. Then we calculate $u^{n+1}_h \in V^1$ from $u^n_h \in V^1$ in two steps :

Step 1 : finite element calculation predicting u_h at $n+1$

Given $u_h^n \in V^1$ calculate $u_h^* \in V^1$ satisfying

$$(3.1)\quad \int_{K(i)} \frac{u_h^* - u_h^n}{\Delta t_n} v\, dx - \int_{K(i)} f(u_h^n(t)) \frac{dv}{dx}\, dx + F_{i+1/2}^n v_{i+1/2}^L - F_{i-1/2}^n v_{i-1/2}^R = 0, \quad \text{for } v \in V^1, \quad i \in \mathbf{Z}, \quad n \in \mathbf{N}.$$

Step 2 : slope limitation

From u_h^* we calculate u_h^{n+1} as close as possible to u_h^* with respect to the L^2 norm and satisfying the discrete analogue to (2.7) (2.8) (2.9) (2.10) (2.11) :

$$(3.2)\quad (u_{i+1/2}^{L,n+1} + u_{i-1/2}^{R,n+1})/2 = u_i^{n+1} = u_i^*,$$

$$(3.3)\quad (1-\alpha)u_i^* + \alpha \min(u_{i-1}^*, u_i^*) \le u_{i-1/2}^{R,n+1} \le (1-\alpha)u_i^* + \alpha \max(u_{i-1}^*, u_i^*),$$

$$(3.4)\quad (1-\alpha)u_i^* + \alpha \min(u_i^*, u_{i+1}^*) \le u_{i+1/2}^{L,n+1} \le (1-\alpha)u_i^* + \alpha \max(u_i^*, u_{i+1}^*),$$

$$(3.5)\quad 0 \le \alpha \le 1,$$

$$(3.6)\quad u_{i+1/2}^{L,n+1} = u_{i-1/2}^{R,n+1} = u_i^* \text{ if } u_i^* \ge \max(u_{i-1}^*, u_{i+1}^*) \text{ or } u_i^* \le \min(u_{i-1}^*, u_{i+1}^*).$$

Step 1 amounts to solving a series of linear systems of dimension 2 and step 2 to solving a series of minimization problems of dimension 2.

Since u_h is slope limited, we may define $\sigma_{i-1/2}^R, \sigma_{i+1/2}^L$ such that

$$(3.7)\quad \begin{cases} u_{i-1/2}^R = u_i + \alpha \sigma_{i-1/2}^R (u_{i-1} - u_i), & 0 \le \sigma_{i-1/2}^R \le 1, \\ u_{i+1/2}^L = u_i + \alpha \sigma_{i+1/2}^L (u_{i+1} - u_i), & 0 \le \sigma_{i+1/2}^L \le 1, \\ \sigma_{i-1/2}^R = \sigma_{i+1/2}^L = 0 \text{ if } u_i \ge \max(u_{i-1}, u_{i+1}) \text{ or } u_i \le \min(u_{i-1}, u_{i+1}). \end{cases}$$

$$(3.8)\quad \sigma_{i-1/2}^R (u_{i-1} - u_i) + \sigma_{i+1/2}^L (u_{i+1} - u_i) = 0.$$

This last equality is just a rewritting of (3.2) using (3.7).

3.2 Total variation

In this section, we prove the following theorem.

Theorem 3.1 *Assume that the numerical flux function satisfies (2.4), (2.6). Then the solution u_h calculated by equation (3.1) associated with a slope limiter satisfying (3.2) (3.3) (3.4) (3.5) (3.6) is such that that its averages over the discretization intervals are total variation diminishing, i.e.*

$$\sum_{i \in \mathbf{Z}} | u_i^{n+1} - u_{i-1}^{n+1} | \le \sum_{i \in \mathbf{Z}} | u_i^n - u_{i-1}^n |, \quad n \in \mathbf{N},$$

provided that the following stability condition is satisfied :

$$(3.9)\quad \sup | f' | (\Delta t / h) \le \max(1/(1+2\alpha), 1/2).$$

Proof :

Taking for v in (3.1) the characteristic functions of the intervals of discretization and using (3.2), we obtain

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + \frac{F_{i+1/2}^n - F_{i-1/2}^n}{h} = 0, \quad i \in \mathbf{Z}, \quad n \in \mathbf{N}, \tag{3.10}$$

that we write in the incremental form

$$u_i^{n+1} = u_i^n + d_1^n(i)(u_{i+1}^n - u_i^n) + d_{-1}^n(i)(u_i^n - u_{i-1}^n) \tag{3.11}$$

where $d_1(i)$ and $d_{-1}(i)$ are now defined by :

$$\begin{aligned} d_1(i) &= (\Delta t/h)[F(u_{i+1/2,}^L, u_{i-1/2}^R) - F(u_{i+1/2}^L, u_{i+1/2}^R)]/(u_{i+1} - u_i), \\ d_{-1}(i) &= (\Delta t/h)[F(u_{i-1/2,}^L, u_{i-1/2}^R) - F(u_{i+1/2}^L, u_{i-1/2}^R)]/(u_i - u_{i-1}). \end{aligned}$$

Following Harten'lemma ([13]), the scheme is TVD if :

$$d_{-1}(i) \leq 0 \leq d_1(i), \tag{3.12}$$

$$1 + d_{-1}(i) - d_1(i-1) \geq 0. \tag{3.13}$$

A similar argument to that used in the proof of theorem 2.1 shows that (3.12) is satisfied, and we are left to interpret (3.13) as a stability condition. We have :

$$d_1(i-1) - d_{-1}(i) = \frac{(\Delta t/h)}{u_i - u_{i-1}} \left[F(u_{i-1/2}^L, u_{i-3/2}^R) + F(u_{i+1/2}^L, u_{i-1/2}^R) - 2F(u_{i-1/2}^L, u_{i-1/2}^R)\right]$$

Using (2.6) we may write

$$\begin{aligned} &d_1(i-1) - d_{-1}(i) \leq \\ &\quad \frac{(\Delta t/h)}{u_i - u_{i-1}} \left[\int_{u_{i-1}}^{u_i} |f'| + \int_{u_{i-3/2}^R}^{u_{i-1}} (f')^- + \int_{u_i}^{u_{i-1/2}^R} (f') + \int_{u_{i-1/2}^L}^{u_{i-1}} (f')^+ + \int_{u_i}^{u_{i+1/2}^L} (f')^+\right]. \end{aligned}$$

On one hand, the third and the fourth term of the sum in the righthand side are negative and therefore can be dropped when calculating an upper bound. By using the slope limiter, the three remaining terms can be bounded to obtain

$$d_1(i-1) - d_{-1}(i) \leq (\Delta t/h) \sup |f'| [1 + \alpha\sigma_{i-1/2}^L + \alpha\sigma_{i-1/2}^R] \leq (\Delta t/h) \sup |f'| (1 + 2\alpha).$$

On the other hand, if we do not drop any term, we have :

$$\begin{aligned} d_1(i-1) - d_{-1}(i) &\leq \frac{(\Delta t/h)}{u_i - u_{i-1}} \left[\int_{u_{i-3/2}^R}^{u_{i-1/2}^R} (f')^- + \int_{u_{i-1/2}^L}^{u_{i+1/2}^L} (f')^+\right] \\ &\leq \sup |f'| \frac{(\Delta t/h)}{u_i - u_{i-1}} \left[u_{i-1/2}^R - u_{i-3/2}^R + u_{i+1/2}^L - u_{i-1/2}^L\right] \\ &\leq (\Delta t/h) \sup |f'| 2. \end{aligned}$$

This completes the proof of theorem 3.1.

Remark 2 *When the function f is monotone, the stability condition can be improved. Assuming for instance that f is increasing, i.e. $f' \geq 0$, then we can write :*

$$\begin{aligned} d_1(i-1) - d_{-1}(i) &\leq \frac{(\Delta t/h)}{u_i - u_{i-1}} \left[\int_{u_{i-1}}^{u_i} | f' | + \int_{u_{i-1/2}^L}^{u_{i-1}} (f')^+ + \int_{u_i}^{u_{i+1/2}^L} (f')^+ \right] \\ &\leq (\Delta t/h) \sup | f' | [1 + \alpha \sigma_{i-1/2}^R] \leq (\Delta t/h) \sup | f' | (1+\alpha). \end{aligned}$$

Therefore, in the case of monotone functions, stability condition (3.9) can be replaced by

$$\sup | f' | (\Delta t/h) \leq 1/(1+\alpha).$$

3.3 L^∞-stability

Now we prove the following result concerning L^∞-stability :

Theorem 3.2 *Assume that the numerical flux function satisfies (2.4),(2.5), (2.6). Then the solution u_h calculated by equation (3.1) associated with a slope limiter satisfying (3.2), (3.3), (3.4), (3.5), (3.6), is L^∞-stable provided that the following stability condition is satisfied :*

$$\sup | f' | (\Delta t/h) \leq 1/(1+\alpha). \tag{3.14}$$

Proof

We use the incremental form (3.11) of the scheme and we suppose first that u_i^n lies between u_{i-1}^n and u_{i+1}^n :

$$u_i^n = \theta u_{i-1}^n + (1-\theta) u_{i+1}^n, \quad 0 \leq \theta \leq 1.$$

Therefore we can write

$$u_i^{n+1} = \lambda_1 u_{i+1}^n + \lambda_2 u_{i-1}^n$$

with

$$\lambda_1 = \theta d_1^n(i) + (1-\theta)(1 + d_{-1}^n(i)), \quad \lambda_2 = (1-\theta)(-d_{-1}^n(i)) + \theta(1 - d_1^n(i)).$$

Obviously we have $\lambda_1 + \lambda_2 = 1$. Moreover one can easily check that, from (2.6),(3.2),(3.3), (3.4),(3.5) and the stability condition (3.14) we have

$$0 \leq d_1(i) \leq \sup | f' | (\Delta t/h)(1+\alpha) \leq 1, \quad 0 \leq -d_{-1}(i) \leq \sup | f' | (\Delta t/h)(1+\alpha) \leq 1,$$

and consequently $0 \leq \lambda_1 \leq 1$, $0 \leq \lambda_2 \leq 1$. Therefore u_i^{n+1} lies between u_{i+1}^n and u_{i-1}^n.

In the case where u_i^n is a maximum or a minimum, say for instance u_i^n greater than u_{i+1}^n and u_{i-1}^n, since $d_1(i)$ is positive and $d_{-1}(i)$ is negative, we obtain immediately from the incremental form (3.11) of the scheme that

$$u_i^{n+1} \leq u_i^n.$$

On the other hand, by using (3.6) we have :

$$d_1(i) = (\Delta t/h)[F(u_i, u_i) - F(u_i, u_{i+1/2}^R)]/(u_{i+1} - u_i),$$
$$d_{-1}(i) = (\Delta t/h)[F(u_{i-1/2}^L, u_i) - F(u_i, u_i)]/(u_i - u_{i-1}),$$

and it follows from (3.3),(3.4) that

$$d_1(i) \leq \bar{d}_1^n(i) = (\Delta t/h)[F(u_i, u_i) - F(u_i, u_{i+1})]/(u_{i+1} - u_i),$$
$$d_{-1}(i) \geq \bar{d}_{-1}^n(i) = (\Delta t/h)[F(u_{i-1}, u_i) - F(u_i, u_i)]/(u_i - u_{i-1}).$$

Therefore, we can write :

$$u_i^{n+1} \geq u_i^n + \bar{d}_1^n(i)(u_{i+1}^n - u_i^n) + \bar{d}_{-1}^n(i)(u_i^n - u_{i-1}^n)$$

where $\bar{d}_1^n(i)$ and $\bar{d}_{-1}^n(i)$ are the coefficients of the incremental form of the first order scheme. Since the latter is L^∞-stable, we obtain :

$$u_i^{n+1} \geq \min(u_{i+1}^n, u_{i-1}^n).$$

When u_i^n is a minimum, a similar argument gives :

$$u_i^n \leq u_i^{n+1} \leq \max(u_{i+1}^n, u_{i-1}^n).$$

This terminates the proof of (3.2).

3.4 Convergence

To prove convergence to the entropy solution we shall use the same ideas as in ([27]). These ideas can be put together in the following theorem whose proof can be found in ([27]) :

Theorem 3.3 *Suppose that a scheme can be written in the following form :*

(3.15) $$u_i^{n+1} = \bar{u}_i^{n+1} - a_{i+1/2}^n + a_{i-1/2}^n$$

where

(i) $\bar{u}_i^{n+1}$ *is obtained from* u_i^n *by using a monotone scheme which yields an approximate sequence converging in* L^1_{loc} *to the entropy satisfying solution of (1.1),*

(ii) *the higher order corrections satisfy* $| a_{i+1/2}^n | \leq Ch^\delta$ *for some constant* C *independent of* h *and some* $\delta \in]0,1[$,

(iii) *the scheme is TVD and* L^∞*-stable.*

Then the scheme yields an approximate sequence converging in L^1_{loc} *to the entropy satisfying solution of (1.1).*

In order to satisfy hypothesis (ii) to apply this theorem to our scheme it is necessary to add to the slope limiter (3.2), (3.3), (3.4), (3.5), (3.6) the following condition on u_h :

(3.16) $| u_{i+1/2}^L - u_{i-1/2}^R | \leq Ch^\delta$ for some C independent of h and some $\delta \in]0,1[$.

Condition (3.16) is not new and has already been used in [18], [27]. In practice it is usually not implemented.

Theorem 3.4 *Assume that the numerical flux function satisfies (2.4) (2.5) (2.6). Then the solution u_h calculated by scheme (3.1) associated with a slope limiter satisfying (3.2), (3.3),(3.4),(3.5),(3.6) and (3.16) converges to the entropy satisfying solution of (1.1) provided that the following stability condition is satisfied :*

$$\sup \mid f' \mid (\Delta t/h) \leq \max(1/(1+2\alpha), 1/2).$$

Proof :

We apply (3.3). Hypothesis (iii) is satisfied from the stability condition and theorems 3.1 and 3.2. Taking for v in (3.1) the characteristic functions of the intervals of discretization and using (3.2), the scheme can be written in the form (3.15) with

$$\begin{aligned} \bar{u}_i^{n+1} &= u_i^n - (\Delta t/h)[F(u_i^n, u_{i+1}^n) - F(u_{i-1}^n, u_i^n)], \\ a_{i+1/2}^n &= (\Delta t/h)[F(u_{i+1/2}^{L,n}, u_{i+1/2}^{R,n}) - F(u_i^n, u_{i+1}^n)]. \end{aligned}$$

Therefore hypothesis (i) is satisfied and concerning (ii) we have from (2.6), (3.16) and the stability condition :

$$\begin{aligned} a_{i+1/2}^n &= (\Delta t/h)[F(u_{i+1/2}^{L,n}, u_{i+1/2}^{R,n}) - F(u_i^{L,n}, u_{i+1}^n) + F(u_i^{L,n}, u_{i+1}^n) - F(u_i^n, u_{i+1}^n)] \\ &\leq C_1 \mid u_{i+1/2}^{R,n} - u_{i+1}^n \mid + C_2 \mid u_{i+1/2}^{L,n} - u_i^n \mid \\ &\leq (C_1/2) \mid u_{i+1/2}^{R,n} - u_{i+3/2}^{L,n} \mid + (C_2/2) \mid u_{i+1/2}^{L,n} - u_{i-1/2}^{R,n} \mid \leq Ch^\delta. \end{aligned}$$

This terminates the proof of theorem 3.4.

3.5 Numerical experiments

In all one-dimensional experiments, we consider two examples. The first one is the shock produced by the following data :

$$f(u) = u^2/(u^2 + 5(1-u)^2),$$
$$u_0(x) = \begin{cases} 1 \text{ for } x \leq 0, \\ .1/(x+.1) \text{ for } x \in [0,1], \\ 1/11 \text{ for } x \geq 1, \end{cases}$$

and the second one is the rarefaction wave produced by :

$$f(u) = u(1-u),$$
$$u_0(x) = \begin{cases} 1 \text{ for } x \leq 1/2, \\ 1 \text{ for } x \geq 1/2. \end{cases}$$

In all following figures the space discretization step is shown on the x-axis.

Since the proofs of theorems 3.1 and 3.2 rely only upon (3.10 and the slope limiter, there are several possibilities to calculate the integrals in (3.1), as long as equation (3.10) holds. Considering the integral with the derivative with respect to time, one can choose to calculate it exactly or using the trapezoidal rule, and for the integral with the derivative with respect to space, we tried trapezoidal, Simpson and mid-point rules. These choices are not at all indifferent with respect to the quality of numerical results as shown on figures 3.1 and 3.2. In these figures, we can see that the best results are obtained using the trapezoidal rule for the integral containing the derivative with respect to time, and the mid-point rule for the integral containing the derivative with respect to space.

At this point, in one dimension, the discontinuous finite element method is very similar to a higher order in space finite difference method as invented by Van Leer [24]. The only difference is that in the discontinuous finite element method, the slopes of the solution are predicted by the first step of the scheme consisting in the finite element calculation (3.1) and then modified by the slope limiter, while in the finite difference method the slopes are derived directly from the slope limiter. Thus the finite element method is slightly more expensive in one dimension, but as we shall see in section 4, it will extend in a truely multidimensional scheme and also it will enable us to devise several implicit higher order methods [26].

In figures 3.3 and 3.4, we analyze the effect of the parameter α. In the calculations, the choice of Δt is optimal with respect to the stability condition (3.9) for $\alpha = 1$. For the given discretization, when $\alpha = 0$ the method reduces to the first order finite difference scheme and is diffusive, when $\alpha = 1$ it is antidiffusive, and the best results for the shock are obtained with $\alpha = 0.3$ while for the rarefaction wave they are obtained with $\alpha = 0.5$. This last value of α is usually used though in case of a shock the corresponding scheme is still antidiffusive.

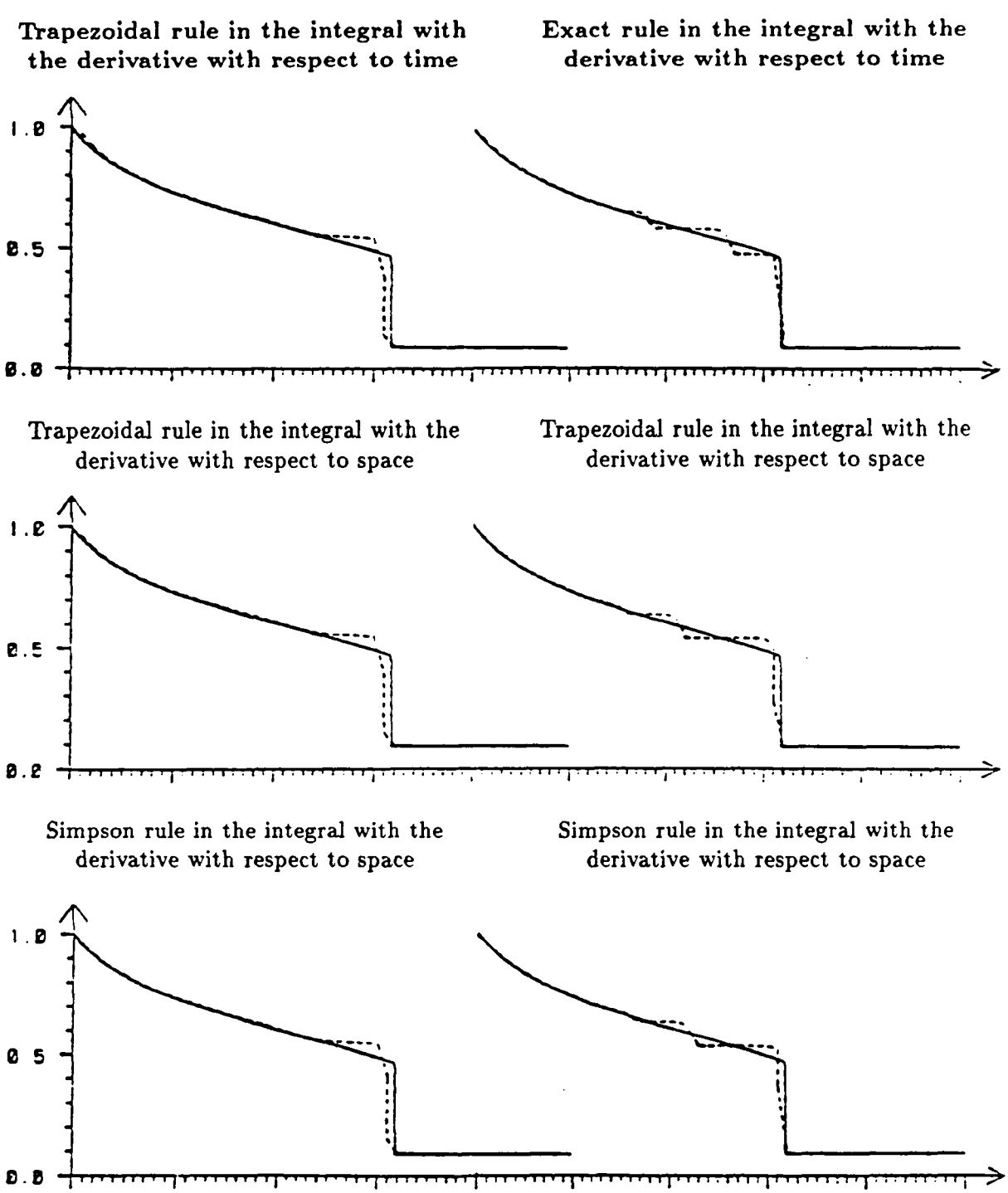
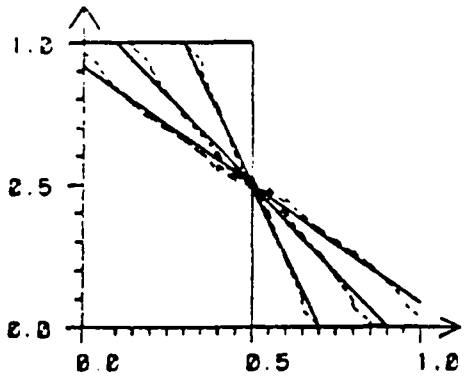


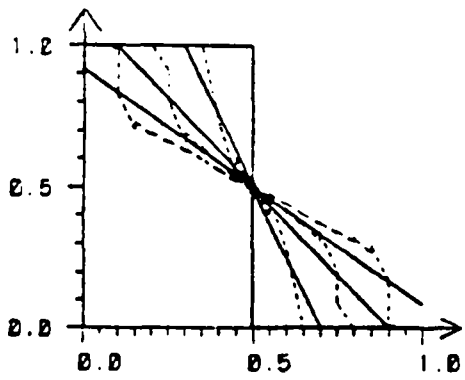
Figure 3.1: Comparisons of different choices of integration formulas in equation (3.1) when $\alpha = 1$ and for a shock wave

Trapezoidal rule in the integral with the derivative with respect to time

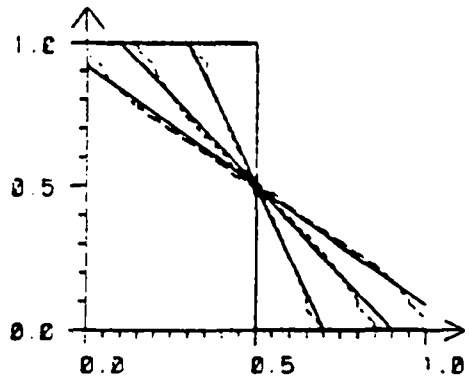


Trapezoidal rule in the integral with the derivative with respect to space

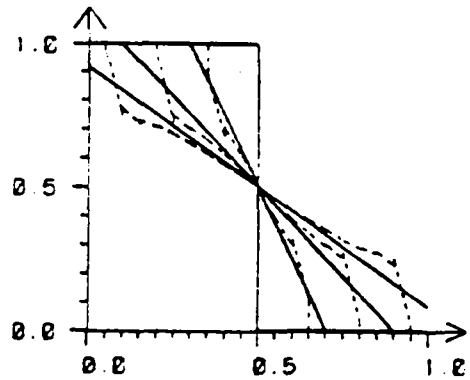
Exact rule in the integral with the derivative with respect to time



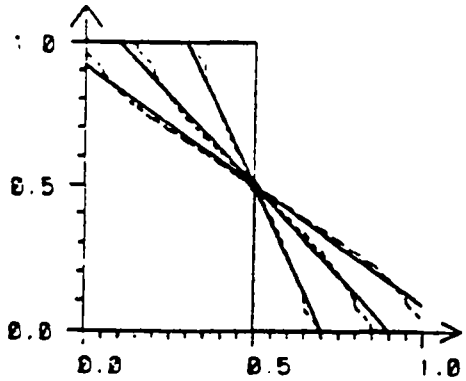
Trapezoidal rule in the integral with the derivative with respect to space



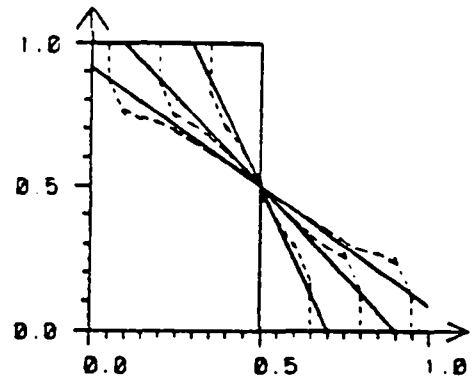
Simpson rule in the integral with the derivative with respect to space



Simpson rule in the integral with the derivative with respect to space



Mid point rule in the integral with the derivative with respect to space



Mid-point rule in the integral with the derivative with respect to space

Figure 3.2: Comparisons of different choices of integration formulas in equation (3.1) when $\alpha = 1$ and for a rarefaction wave

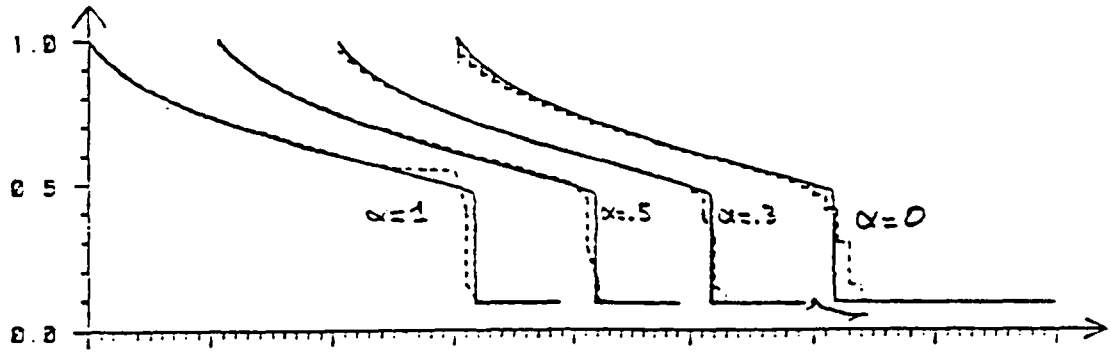


Figure 3.3: The effect of the slope limiting parameter α for a shock wave

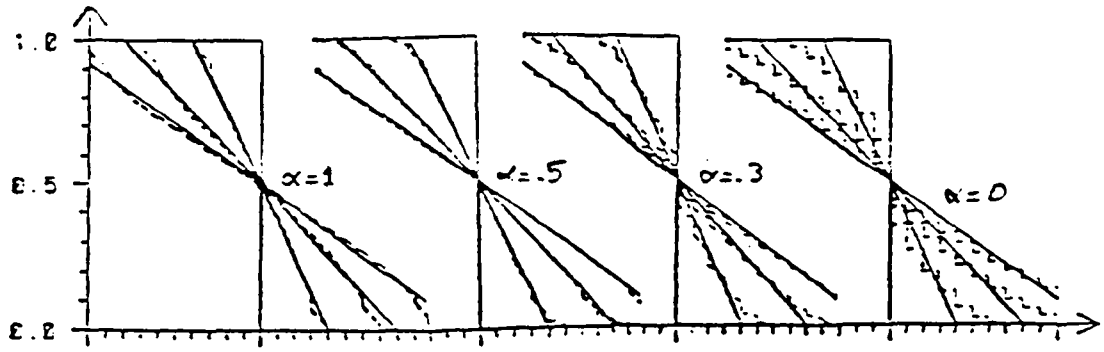


Figure 3.4: The effect of the slope limiting parameter α for a rarefaction wave

Tables (3.1) and (3.2) show some L^1 errors and their convergence rates δ calculated as follows :

$$e_1 = \| u - u_{h_1} \|_{L^1} \text{ with } h_1 = h, \quad e_2 = \| u - u_{h_2} \|_{L^1} \text{ with } h_2 = h/2, \quad \delta = \frac{\ln(e_1/e_2)}{\ln 2} \tag{3.17}$$

These numbers have been obtained when using the best integration formulas, i.e. the trapezoidal rule in the integral with the derivative with respect to time and the midpoint rule in the integral with the derivative with respect to space.

We observe again that the errors obtained with the piecewise linear approximation are significantly smaller than with piecewise constants. However the rates of convergence is not improved, which can be easily explained by the fact that the scheme is only first order in time. For the rarefaction wave it appears that the errors are not yet in the asymptotic range for the given discretizations.

		$\alpha = 0$		$\alpha = 0.3$		$\alpha = 0.5$	
h	Δt	$\|\|u - u_h\|\|_{L^1}$	δ	$\|\|u - u_h\|\|_{L^1}$	δ	$\|\|u - u_h\|\|_{L^1}$	δ
1/25	1/105	0.02612		0.01242		0.01698	
1/50	1/210	0.01351	0.947	0.00618	1.005	0.00867	0.970
1/100	1/420	0.00639	1.084	0.00240	1.365	0.00436	0.989

Table 3.1: L^1-error and rate of convergence for a shock wave for three values of the slope limiting parameter α (first order in time)

		$\alpha = 0$		$\alpha = 0.5$	
h	Δt	$\|\|u - u_h\|\|_{L^1}$	δ	$\|\|u - u_h\|\|_{L^1}$	δ
1/20	1/40	0.02724		0.00518	
1/40	1/80	0.01695	0.684	0.00467	0.149
1/80	1/160	0.01026	0.724	0.00370	0.335

Table 3.2: L^1-error and rate of convergence for a rarefaction wave for two values of the slope limiting parameter α (first order in time)

3.6 Second order time discretization

In this section, we present a second order in time version of the finite element scheme that we described in the previous sections. This scheme is obtained through a Lax-Wendroff type modification of the first order in time scheme in a similar manner to what is done for some finite difference schemes. It contains an intermediate step in which the solution is calculated at the time $t_{n+1/2}$ by means of a local calculation. However to respect our framework, this calculation is formulated in terms of finite elements so it can be extended to multidimensional calculations on nonstructured meshes.

We now calculate n_h^{n+1} from u_h^n in three steps :

Step 1 : Finite element calculation calculating u_h at $t_{n+1/2} = (t_n + t_{n+1})/2$

Given $u_h^n \in V^1$ calculate $u_h^{n+1/2} \in V^1$ satisfying

$$\int_{K(i)} \frac{u_h^{n+1/2} - u_h^n}{1/2\Delta t_n} v \, dx - \int_{K(i)} f(u_h^n) \frac{dv}{dx} \, dx + f(u_{i+1/2}^{L,n}) v_{i+1/2}^L - f(u_{i-1/2}^{R,n}) v_{i-1/2}^R = 0, \quad \text{for } v \in V^1, \quad i \in \mathbf{Z}. \tag{3.18}$$

Step 2 : Finite element calculation predicting u_h at t_{n+1}

Calculate $u_h^* \in V^1$ satisfying :

$$(3.19) \qquad \int_{K(i)} \frac{u_h^* - u_h^n}{\Delta t_n} v \, dx - \int_{K(i)} f(u_h^{n+1/2}) \frac{dv}{dx} \, dx + F_{i+1/2}^{n+1/2} v_{i+1/2}^L - F_{i-1/2}^{n+1/2} v_{i-1/2}^R = 0, \quad \text{for } v \in V^1, \quad i \in \mathbf{Z}.$$

Step 3 : Slope limitation

From u_h^* we calculate u_h^{n+1} as before, i.e. it is as close as possible to u_h^* with respect to the L^2 norm and satisfying (3.2) (3.3) (3.4) (3.5) (3.6).

In figures 3.5 and 3.6 we show numerical experiments for the same data as in section 3.5. In the two first steps of the scheme we used the trapezoidal rule for the integral with the derivatives with respect to time and the mid-point rule for the integral with the derivatives with respect to space. Again we chose an optimal time step with respect to the stability condition (3.9) with $\alpha = 1$. One can see that the second order in time scheme is less diffusive than the first order one and the best results are now obtained for values of α close to 1.

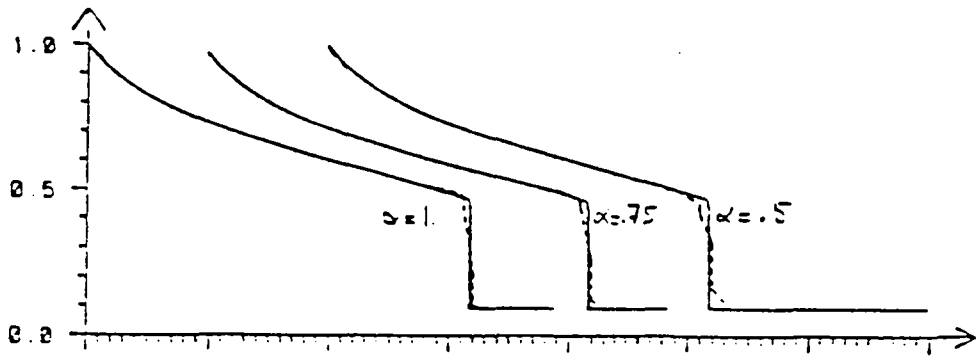


Figure 3.5: The effect of the slope limiting parameter α for a shock wave

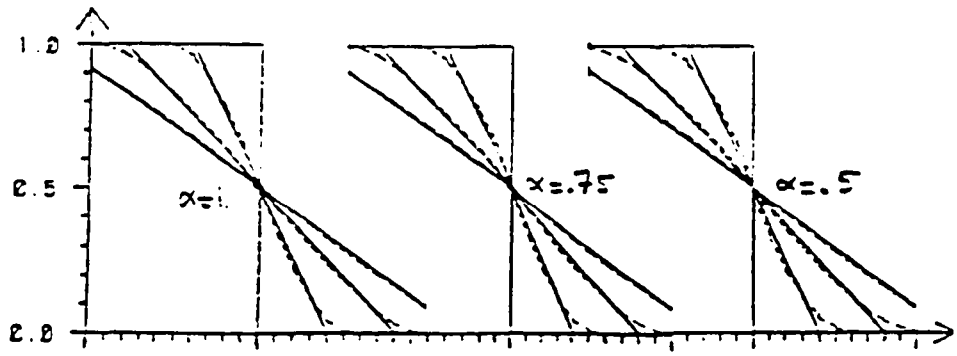


Figure 3.6: The effect of the slope limiting parameter α for a rarefaction wave

Tables 3.3 and 3.4 give some L^1 errors and convergence rates to be compared with those in tables 3.1 and 3.2. The convergence rate δ is again calculated using formulas (3.17). Now that the time stepping is second order in time we observe a significant improvement in the errors and in the convergence rates, for the shock wave as well as for the rarefaction wave.

		$\alpha = 0.5$		$\alpha = 1$	
h	Δt	$\|\|u - u_h\|\|_{L^1}$	δ	$\|\|u - u_h\|\|_{L^1}$	δ
1/25	1/105	0.00918		0.00614	
1/50	1/210	0.00447	1.039	0.00274	1.164
1/100	1/420	0.00185	1.271	0.00119	1.193

Table 3.3: L^1-error and rate of convergence for a shock wave for two values of the slope limiting parameter α (second order in time)

		$\alpha = 0.5$		$\alpha = 1$	
h	Δt	$\|\|u - u_h\|\|_{L^1}$	δ	$\|\|u - u_h\|\|_{L^1}$	δ
1/20	1/40	0.00919		0.00733	
1/40	1/80	0.00460	0.997	0.00372	0.979
1/80	1/160	0.00230	0.999	0.00190	0.967

Table 3.4: L^1-error and rate of convergence for a rarefaction wave for two values of the slope limiting parameter α (second order in time)

4 Two-dimensional space approximation

4.1 Discontinuous finite elements

We consider a regular discretization of a domain Ω with triangles and quadrangles $K \in T_h$ of diameter less than or equal to h and we define the approximation space

$$V^1 = \left\{ v \mid v_{|K} \in P^1 \text{ (resp. } Q^1 \text{) if } K \text{ is a triangle (resp. quadrangle), } K \in T_h \right\}.$$

The degrees of freedom of functions of V^1 are, element by element, their values at the vertices (see figure 4.1). We denote them $v_{K,A}$ with $K \in T_h$ and A a vertex of K. Another

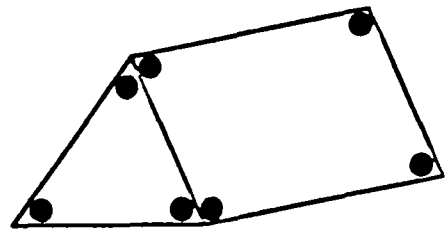


Figure 4.1: Degrees of freedom of functions of the approximation space V^1

choice is possible in case of a structured mesh of rectangles : one can still take P^1 polynomials (instead of Q^1) in the definition of V^1. A convenient choice of the three degrees of freedom is the average value of the function in the element and its slopes in the directions parallel to the axes. Then a two-dimensional scheme is obtained by writing the one-dimensional scheme in directions parallel to the axes. This is what is done usually in higher order finite difference schemes but this is not what we can call a truely multidimensional scheme.

As in one dimension, given $u_h^n \in V^1$ we calculate $u_h^{n+1} \in V^1$ in two steps : a finite element calculation giving a predicted solution $u_h^* \in V^1$ followed by a slope limitation yielding u_h^{n+1}. The finite element calculation consists in solving the following equation :

$$(4.1) \quad \int_K \frac{u_h^* - u_h^n}{\Delta t_n} v \, dx - \int_K \vec{f}(u_h^n) \cdot \vec{grad} \, dx + \int_{\partial K} F^n v \, d\gamma = 0, \text{ for } v \in V^1, K \in T_h,$$

where the numerical flux F^n defined on the edges of the mesh is calculated as follows.

First we notice that the integral over ∂K is the sum of integrals over three or four edges. Any integral over an edge E will be calculated by means of an integration formula

$$(4.2) \quad \int_E F^n v \, d\gamma \simeq \sum_{i=1,npi} \beta_i F^n(P_i) v(P_i),$$

where npi, β_i, P_i denote respectively the number of integration points, the weights and the points of the integration formula.

Then we note that the numerical flux F^n is an approximation on the edge E of the quantity $\vec{f} \cdot \vec{\nu}$ where $\vec{\nu}$ is a unit normal to the edge E. Therefore it is legitimate to calculate $F^n(P_i)$ by solving the one-dimensional Riemann problem in the direction of $\vec{\nu}$, relative to the function $\vec{f}(P_i).\vec{\nu}$ and as initial data the two limit values $u(P_i)^+$, $u(P_i)^-$ of u_h at the points P_i. Thus $F^n(P_i)$ is calculated by the same formulas as described in section 2.2 for the one-dimensional case with $u(P_i)^-$, $u(P_i)^+$, $\vec{f}.\vec{\nu}$ replacing respectively $u^L_{i+1/2}$, $u^R_{i+1/2}$, f.

In practice we need integration formulas to calculate (4.1). Extending to two dimensions the conclusions of the one-dimensional experiments described in section 3.5, we are led to use mass lumping for the integral with the derivative with respect to time and the mid-point rule for the integral with the derivative with respect to space. For the integrals over the edges we used the two-point Gauss formula in the calculations described below, but other experiments have shown that using instead the mid-point rule does not alter significantly the numerical results [17].

4.2 Slope limiter

We formulate now a multidimensional extension of the one-dimensional slope limiter (3.2), (3.3), (3.4), (3.5), (3.6). For any element $K \in T_h$ and any $v \in V^1$ we introduce the following notations :

$$
\begin{aligned}
& nv(K) = \text{ number of vertices of } K, \\
& T(A) = \{K \in T_h \mid A \text{ is a vertex of } K\}, \\
& \bar{v}_K = \frac{1}{nv(K)} \sum_{i=1}^{nv(K)} v_{K,A_i} = \text{ average of } v \text{ over } K, \\
& V_K = (v_{K,A_i})_{i=1,nv(K)}, \\
& J_K(V_K) = \frac{1}{2} \sum_{i=1}^{nv(K)} (v_{K,A_i} - u^*_{K,A_i})^2
\end{aligned}
$$

J_K measures the distance between v and u_h^* inside the element K.

The slope limited function u_h^{n+1} must have the same cell averages as u_h^* to preserve mass balance and its degrees of freedom will satisfy inequalities similar to (3.3),(3.4), (3.5), (3.6). This can be achieved in the following way. Given u_h^* obtained from (4.1), we calculate for any vertex A of the discretization the minimum and the maximum of the averages in the cells surrounding A :

$$
u_{min}(A) = \min_{K \in T(A)} u_K^*, \quad u_{max}(A) = \max_{K \in T(A)} u_K^*.
$$

Then u_h^{n+1} is obtained by solving the series of minimization problems :

$$
\begin{cases} \text{Find } u_h^{n+1} \in V^1 \text{ such that} \\ U_K^{n+1} \in P_K \cap Q_K, \quad , J_K(U_K^{n+1}) = \min_{V_K \in P_K \cap Q_K} J_K(V_K), \quad \text{for all } K \in T_h, \end{cases}
$$

where P_K and Q_K are respectively the following hyperplane and hypercube in $\mathbf{R}^{nv(K)}$:

$$
\begin{aligned}
& P_K = \left\{ x \in \mathbf{R}^{nv(K)} \mid \sum_{i=1}^{nv(K)} x_i = nv(K) U_K^* \right\}, \\
& Q_K = \prod_{i=1}^{nv(K)} [(1-\alpha)u_K^* + \alpha u_{min}(A_i), (1-\alpha)u_K^* + \alpha u_{max}(A_i)], \; 0 \le \alpha \le 1.
\end{aligned}
$$

It is easy to check that each minimization problem in K has a unique solution which can be calculated by dualizing the constraint $V_K \in P_K$ and solving the associated saddle point problem ([6]) (see appendix).

4.3 Numerical experiments

We first present two-dimensional calculations for a one-dimensional problem and we study the mesh effects. The example that we consider is the same shock wave as that in section (3.5) though the initial data is now $u_0 = 0$ for $x > 0$. We ran our calculation with five different meshes. The two first ones are rectangular grids, one perpendicular and the other one diagonal with respect to the propagation of the shock. Next we used two structured triangular meshes and the last one is an unstructured triangular mesh. All the meshes have been chosen in order to have roughly the same h. The corresponding numerical results are shown in figures 4.2, 4.3, 4.4, 4.5, 4.6. In each figure, the mesh is shown as well as three pictures of the solution. The first picture represent the solution obtained with a standard scheme first order in time and space. The last two pictures represent the same solution calculated with our method (with the parameter α of the slope limiter set to .5), but the last picture is a piecewise constant representation of this solution obtained by using only the mean values over the elements. Of course the results look best in figure 4.2 since in this case the two-dimensional calculation reduces actually to a one-dimensional calculation. In figures 4.5 and 4.6 one can notice in the second pictures small oscillations on the upper part of the shock, in the direction perpendicular to the propagation of the shock. This shows that the scheme is not anymore oscillation free as it was in one dimension. However in the last pictures one can see that the mean values are oscillation free. Of course one can observe mesh effects in the sense that the discontinuity line is perturbed by the elements when the edges are not parallel to it. However in all cases there is a good representation of the shock.

For a two-dimensional example, we consider the problem of a two-phase incompressible displacements in porous media. Water (subscript w) is injected at an injection well in order to displace oil (subscript o) towards a production well. The equations governing the flow in the domain are

$$div\ \vec{v} = 0, \quad x \in \mathbf{R}^2, \tag{4.3}$$

$$\vec{v} = -(\lambda_0 + \lambda_w) gr\vec{a}dp \tag{4.4}$$

$$s_t + \ div(f(s)\vec{v}) = 0 \tag{4.5}$$

where $f = \lambda_w/(\lambda_o + \lambda_w)$. Equation (4.3) expresses the incompressibility of the flow and conservation of the total (oil + water) while equation (4.5) expresses the conservation of water. Equation (4.4) is Darcy's law. To equations (4.3) (4.4) (4.5), boundary condition or point sources in case of wells are associated. We consider the case where $\lambda_o = k(1-s)^2/\mu_o$, $\lambda_w = ks^2/\mu_w$ and the ratio μ_o/μ_w is now equal to 4 as in [1]. We studied again mesh effects for this viscosity problem considering a diagonal grid (mesh 1 in fig. 4.2), a parallel grid (mesh 2 in fig. 4.3) and an unstructured mesh (mesh 5 in fig. 4.6). We also considered a refined version of these meshes by dividing each element into four elements. In figure 4.7 we show results for all these meshes.

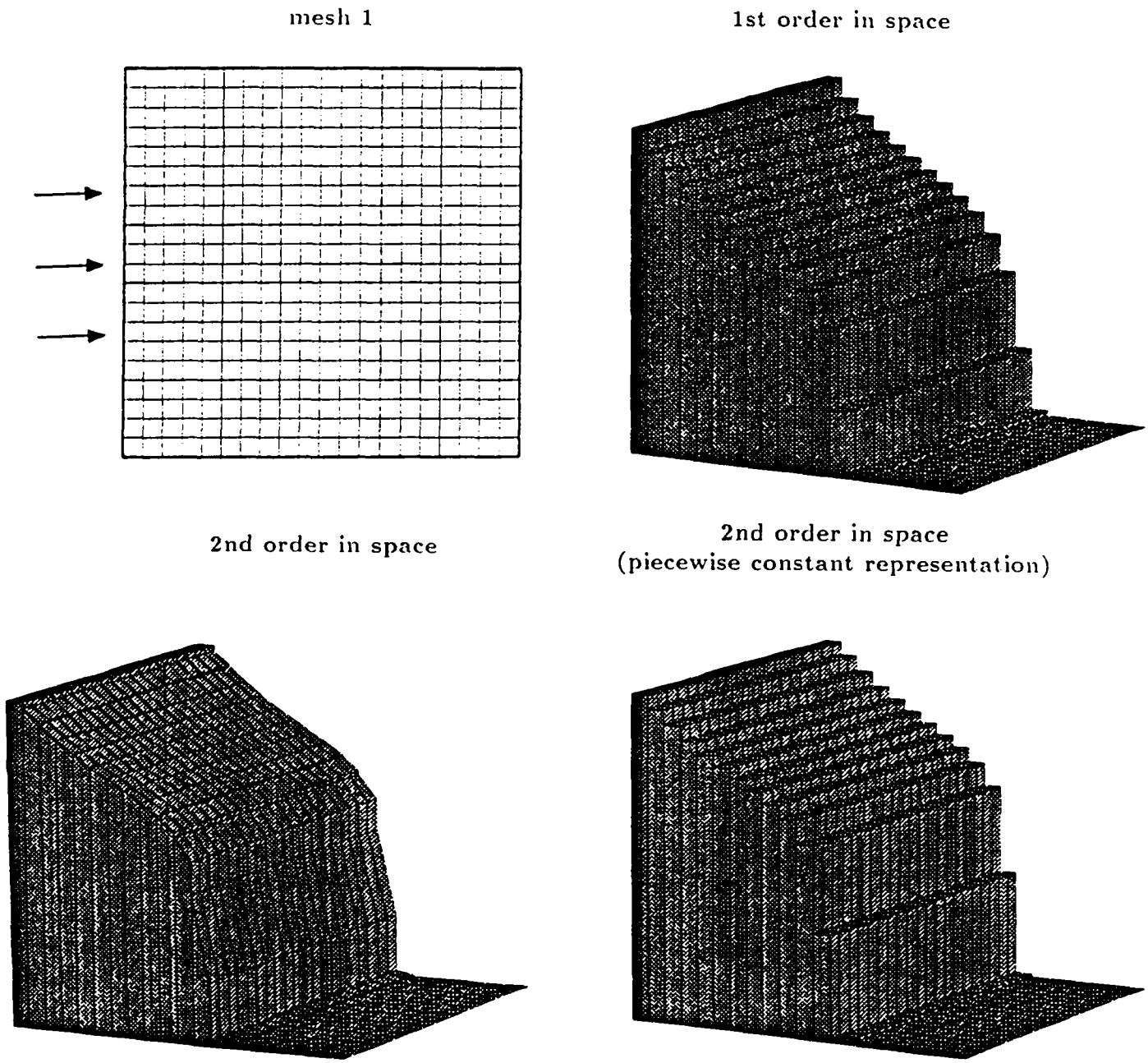


Figure 4.2: A shock wave calculated with mesh 1

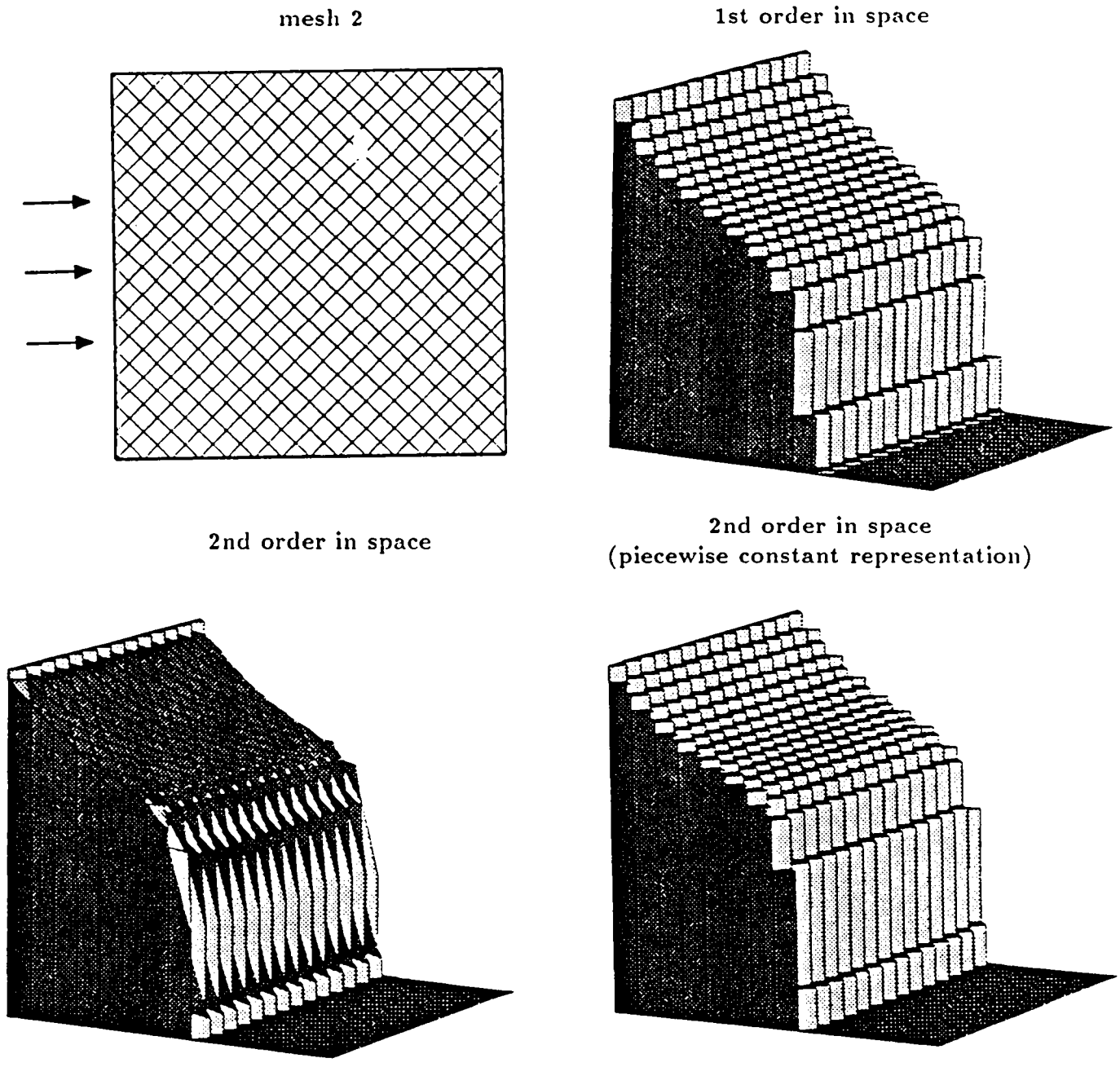


Figure 4.3: A shock wave calculated with mesh 2

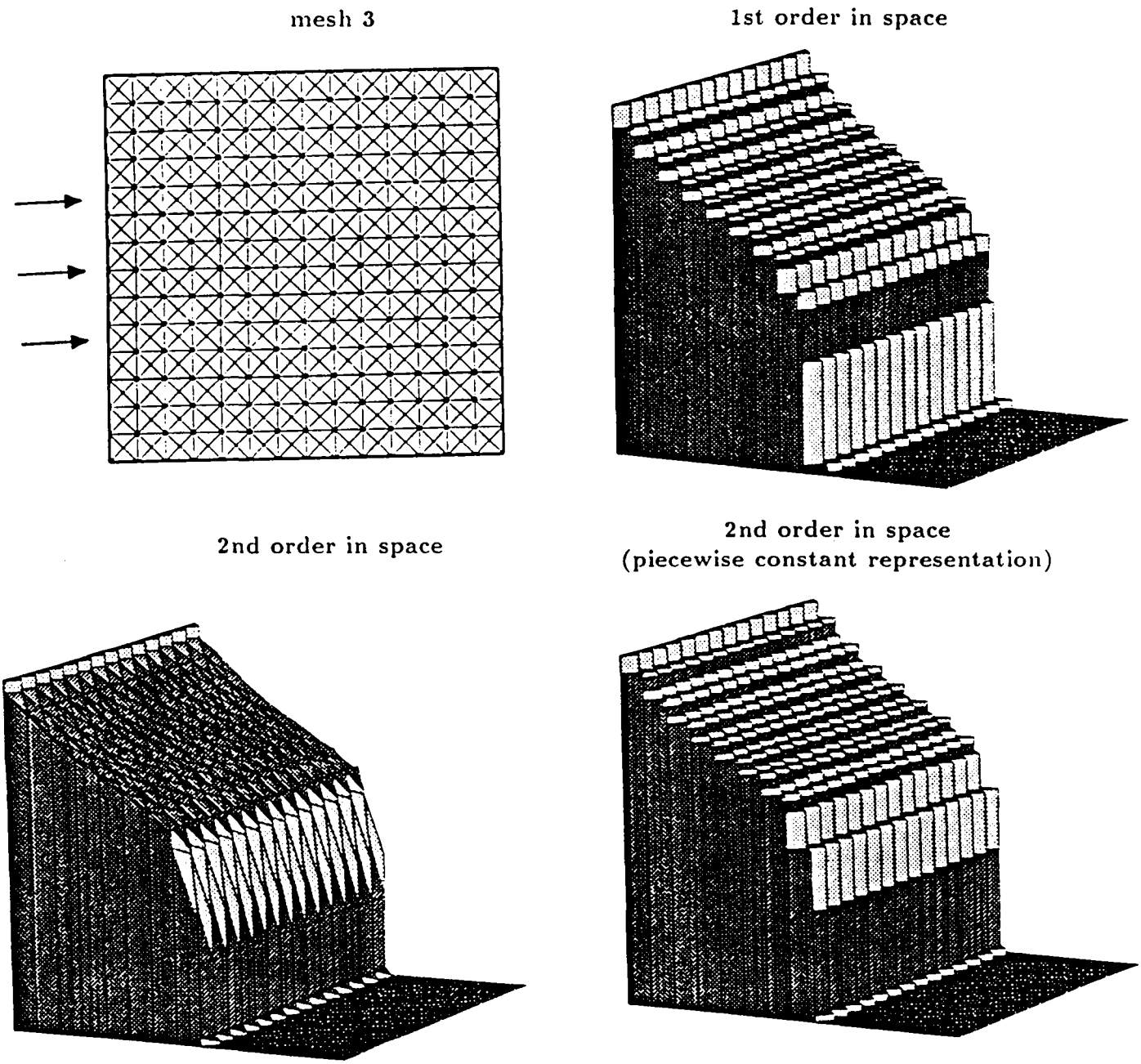


Figure 4.4: A shock wave calculated with mesh 3

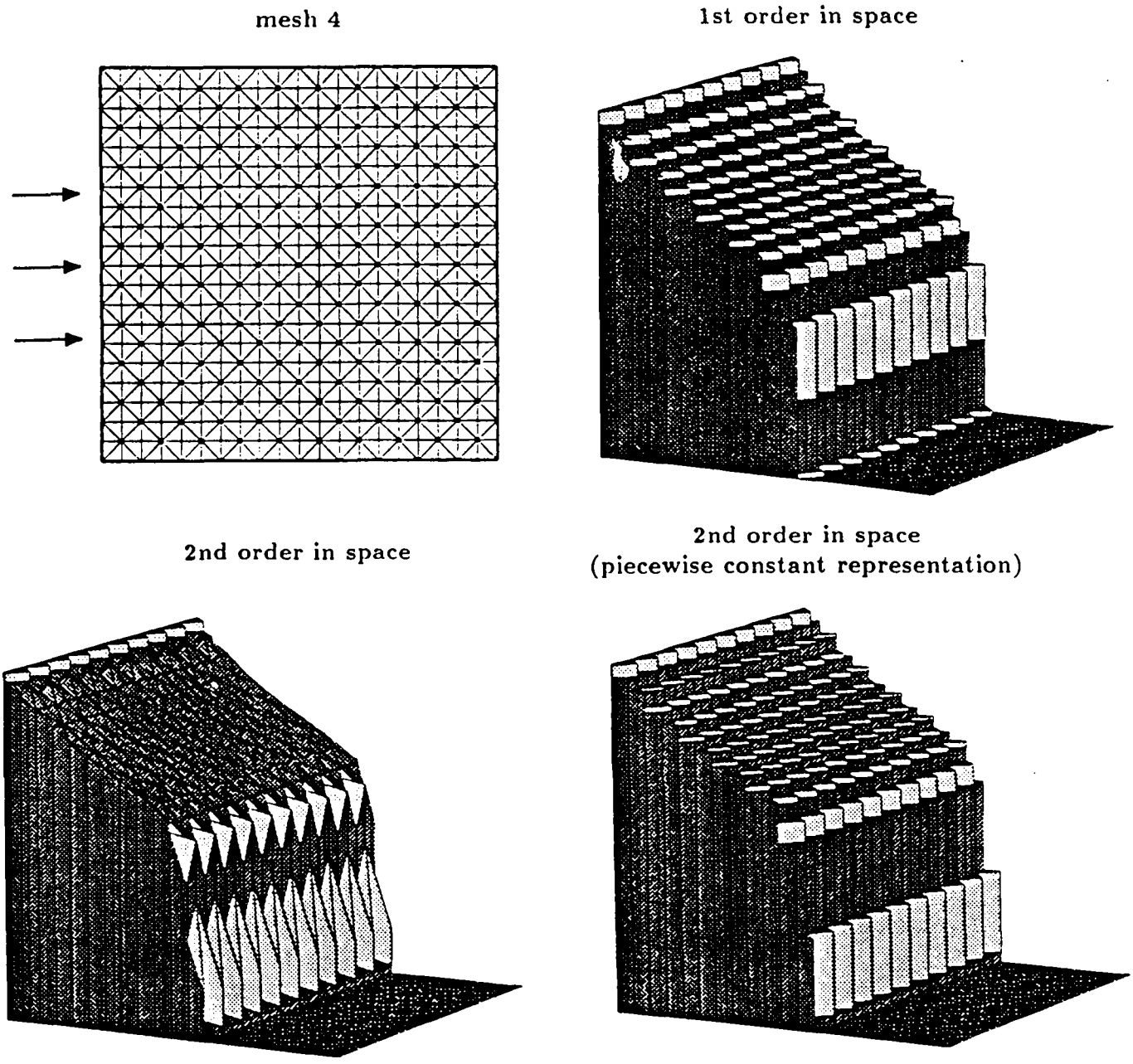


Figure 4.5: A shock wave calculated with mesh 4

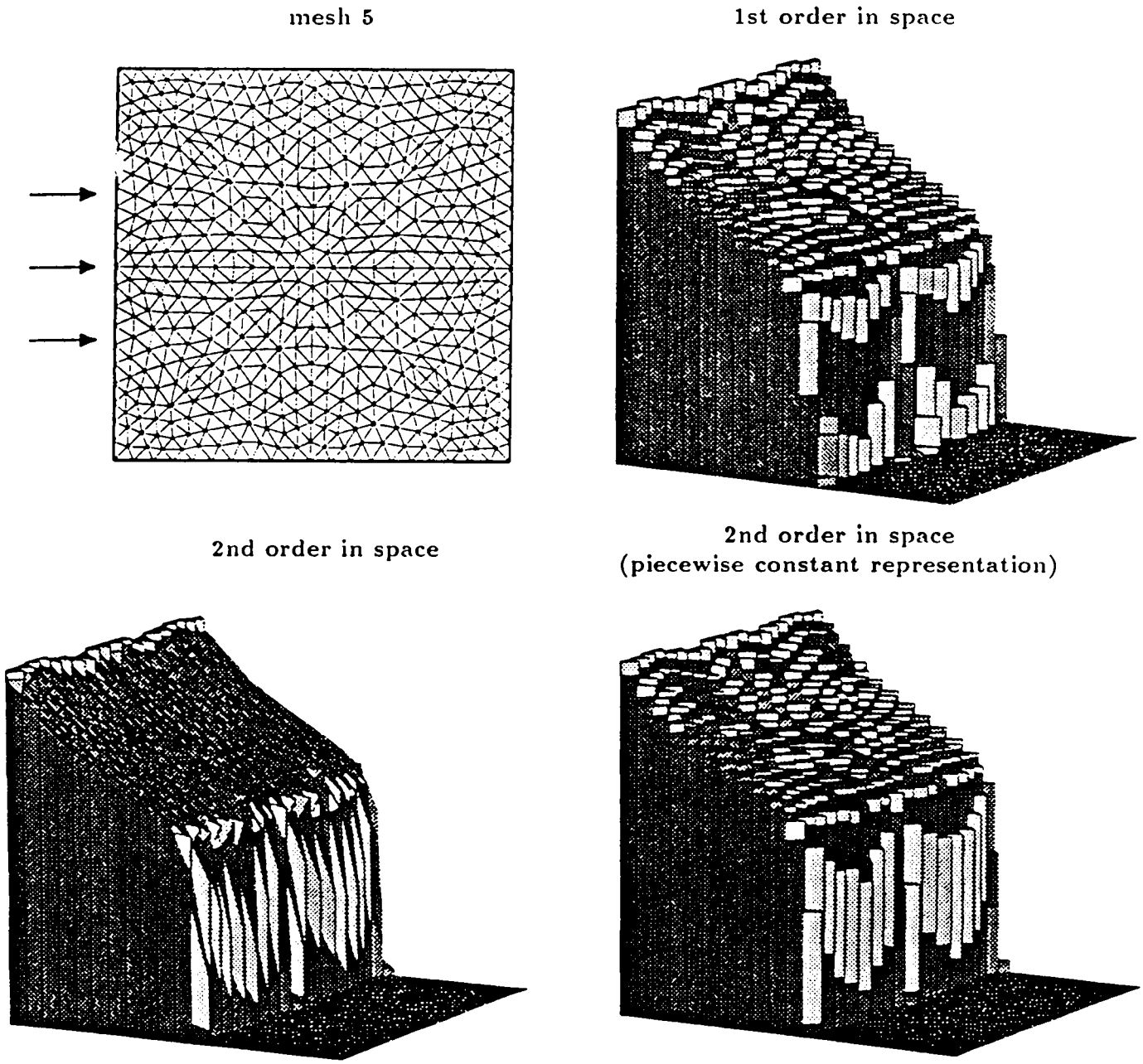


Figure 4.6: A shock wave calculated with mesh 5

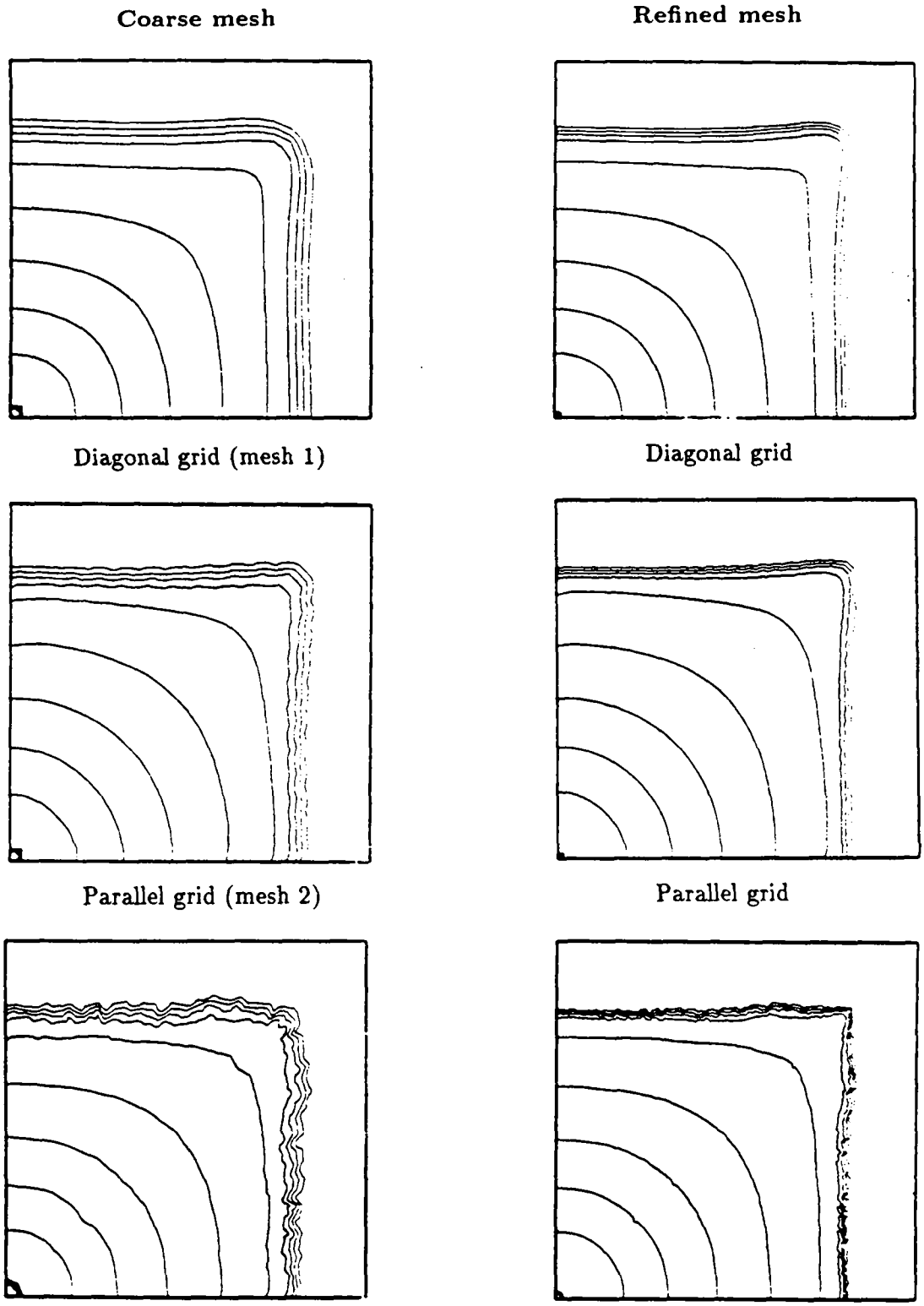


Figure 4.7: Mesh effects on a quarter of five-spot simulation

One can observe no mesh effect in the sense that the shock is always positioned at the same place though in the case of unstructured mesh the discontinuity line is perturbed by elements. The quality of these results is similar to those obtained in [1]. The method are similar in that the slope limitation is the same though we solve completly the minimization problems shown in the appendix whereas in [1] a suboptimal solution is obtained heuristically. However the two methods differ through the flux calculation which is simpler in our method and make it easy to extend to the three-dimensional case.

Conclusion

The discontinuous finite element method is associated to a slope limitation to construct an accurate method for nonlinear hyperbolic conservation laws. For each time step the slope limitation is introduced as a separate step performed after a finite element calculation. In one-dimension, the method has been shown to be total variation diminishing and convergent. However numerical experiments showed the importance of the choice of the integration formulas.

In several dimensions, the slope limitation is truely multidimensional. The method has been experimented on structured or unstructured meshes, trangular or quadrangular. We studied the mesh effects. In all cases the method performs well.

A Appendix

In this appendix we explain how to solve the minimization problem stated in section (4.2), by dualizing the constraint $V_K \in P_K$ and solving the associated saddle point problem. Dualize the constraint by introducing the Lagrangian in $\mathbf{R}^{nv(K)} \times \mathbf{R}$:

$$L(V_K, \mu) = J_K(V_K) - \mu[(\sum_{i=1}^{nv(K)} v_{K,A_i}) - nv(K)\bar{u}_K^*]. \tag{A.1}$$

Then the problem

$$\begin{cases} \text{Find } U_K^{n+1} = (u_{K,A_i}^{n+1})_{i=1,\ldots,nv(K)} \in P_K \cap Q_K \text{ such that} \\ J_K(U_K^{n+1}) = \min\limits_{V_K \in P_K \cap Q_K} J_K(V_K) \end{cases} \tag{A.2}$$

is equivalent to the saddle point problem

$$\begin{cases} \text{Find } (U_K^{n+1}, \lambda) \in Q_K \times \mathbf{R} \text{ such that} \\ L(U_K^{n+1}, \lambda) = \max\limits_{\mu \in \mathbf{R}} \min\limits_{V_K \in Q_K} L(V_K, \mu). \end{cases} \tag{A.3}$$

Therefore we first solve, for a given $\mu \in \mathbf{R}$, the minimization problem

$$\begin{cases} \text{Find } V(\mu) \in Q_K \text{ such that} \\ L(V(\mu), \mu) = \min\limits_{V_K \in Q_K} L(V_K, \mu). \end{cases} \tag{A.4}$$

Then we solve the maximization problem

$$\begin{array}{l} \text{Find } \lambda) \in \mathbf{R} \text{ such that} \\ L(V(\lambda), \lambda) = \min\limits_{\mu \in \mathbf{R}} L(V(\mu), \mu). \end{array} \tag{A.5}$$

and U_K^{n+1} satisfies

$$U_K^{n+1} = V(\lambda).$$

Since $U_K^* = (u_{K,A_i}^*)_{i=1,\ldots,nv(K)} \in P_K$, the Lagrangian L defined in (A.1) can be rewritten as

$$L(V_K, \mu) = (1/2) \parallel V_K - (U_K^* - \mu U) \parallel^2 - (\mu^2/2) \parallel U \parallel^2 \tag{A.6}$$

where $\parallel \parallel$ denotes the Eucledian norm in $\mathbf{R}^{nv(K)}$ and U is a vector such that $U = (u_i = 1)_{i=1,\ldots,nv(K)}$, which is normal to P_K. Expression (A.6) shows that $V(\mu)$, the solution to problem (A.4) is the projection of $(U_K^* - \mu U)$ on the hypercube Q_K, so that $V(\mu)$ is simply obtained by truncation of the components of $(U_K^* - \mu U)$.

Thus the function $\mu \to F(\mu) = L(V(\mu), \mu)$ is easy to calculate, and finding U_K^{n+1} reduces to solving the one-dimensional maximization problem (A.5). One can check that the derivatives of F are

$$\begin{aligned} F'(\mu) &= \sum_{i=1}^{nv(K)} V(\mu)_i \\ F''(\mu) &= -card\left[i \in [0, 1, \ldots, nv(K)] \mid (u_{K,A_i}^* - \mu) \in Q_{K_i}\right] \end{aligned}$$

where $Q_{K_i} = [(1-\alpha)u_K^* + \alpha u_{min}(A_i), (1-\alpha)u_K^* + \alpha u_{max}(A_i)]$. The slope limiting the saturation reduces to maximizing, for each element of the mesh, a one dimensional concave function which has piecewise constant second derivatives.

References

[1] J. Bell, C. Dawson, and G. Shubin, *An unsplit higher order Godunov methods for scalar conservation laws in multiple dimensions*, J. Comp. Phys., 74 (1988), pp. 1–24.

[2] M. Celia, T. Russel, I. Herrera, and R. Ewing, *An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation*, Adv. Water Resources, 13 (1991), pp. 187–206.

[3] G. Chavent, B. Cockburn, G. Cohen, and J. Jaffré, *A discontinuous finite element method for nonlinear hyperbolic equations*, in Innovative Numerical Methods in Engineering, R.P. Schaw and J. Periaux and A. Chaudouet and J. Wu and C. Marino and C.A. Bredia, Berlin, 1986.

[4] G. Chavent, G. Cohen, and J. Jaffré, *Discontinuous upwinding and mixed finite elements for two-phase flow in reservoir simulation*, Comp. Meth. Appl. Mech. Eng., 47 (1984), pp. 93–118.

[5] G. Chavent, G. Cohen, J. Jaffré, R. Eymard, D. Guérillot, and L. Weill, *Discontinuous and mixed finite elements for two-phase incompressible flow*, SPE Resevoir Engineering, 5 (1990), pp. 567–575.

[6] G. Chavent and J. Jaffré, *Mathematical Models and Finite Elements for Reservoir Simulation*, North Holland, Amsterdam, 1986.

[7] G. Chavent and G. Salzano, *A finite element method for the 1-D water flooding problem with gravity*, J. Comp. Phys., 45 (1982), pp. 1–21.

[8] B. Cockburn, *Quasimonotone schemes for scalar conservation laws. Part III*, SIAM J. Numer. Anal., 27 (1990), pp. 259–276.

[9] P. Collela and P. Woodward, *The piecewise parabolic method (PPM) for gas dynamics*, J. Comp. Phys., 54 (1984), pp. 174–201.

[10] J. Douglas Jr and T. Russel, *Numerical methods for convection-dominated problems based on combining the method of characteristics with finite element or finite difference procedures*, SIAM J. Num. Anal., 19 (1982), pp. 871–885.

[11] B. Engquist and S. Osher, *One sided difference approximations for scalar conservation laws*, Math. Comp., 36 (1981), pp. 321–351.

[12] S. Godunov, *Finite difference methods for numerical computation of discontinuous solutions of the equations of fluid dynamics*, Math. Sbornik, 47 (1959), pp. 271–306.

[13] A. Harten, *High resolution schemes for hyperbolic conservation laws*, J. Comp. Phys., 49 (1983), pp. 357–393.

[14] T. Hugues and al, *A new finite element formulation for computational fluid dynamics III, IV*, Comp. Math. Appli. Mech. Eng., 58 (1986), pp. 305–336.

[15] C. Johnson, *Streamline diffusion methods for problems in fluid dynamics*, in Finite Elements in Fluids, Vol.6, G. et al., ed., Wiley, New York, 1985.

[16] C. JOHNSON AND J. PITKARANTA, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation*, Math. Comp., 46 (1986), pp. 1–26.

[17] L. KADDOURI, *Une méthode d'éléments finis discontinus pour les équations d'Euler des fluides compressibles.* Thèse de Doctorat de lÚniversité Paris 6, 1993.

[18] A. LEROUX, *Convergence of an accurate scheme for first order quasi linear equations*, RAIRO Anal. Numér., 15 (1981), pp. 151–170.

[19] P. LESAINT AND P. RAVIART, *On a finite element method for solving the neutron transport equations*, in Mathematical Aspects of Finite Elements in Partial Differentia l Equations, Academic Press, New York, 1974.

[20] K. MORTON AND A. STOKES, *Generalized Galerkin methods for hyperbolic problems*, in Proc. MAFELAP, Academic Press, London, 1982, pp. 421–431.

[21] W. REED AND T. HILL, *Triangular mesh methods for the neutron transport equation*, Tech. Rep. LA-UR-73-479, Los Alamos Scientific Laboratory, 1973.

[22] G. RICHTER, *An optimal order error estimate for the discontinuous Galerkin method*, Math. Comp., 50 (1988), pp. 75–88.

[23] B. STOUFFLET, J. PERIAUX, F. FEZOUI, AND A. DERVIEUX, *Numerical simulation of 3-D hypersonic Euler flows around space vehicles using adapted finite elements.* AIAA paper 87-0560, 1987.

[24] B. VAN LEER, *Towards the ultimate conservative scheme : IV. A new approach to numerical convection*, J. Comp. Phys., 23 (1977), pp. 276–299.

[25] ——, *Towards the ultimate conservative scheme : V. A second order sequel to Godunov's method*, J. Comp. Phys., 32 (1979), pp. 101–136.

[26] G. VEERAPPA GOWDA, *Discontinuous finite elements for non linear scalar conservation laws.* Thèse de Doctorat de l'Université Paris 9, 1988.

[27] J. VILA, *Sur la théorie et l'approximation numérique de problèmes hyperboliques non linéaires ; applications aux equations de saint venant et à la modélisation des avalanches de neige dense.* Thèse de Doctorat de l'Université Paris 6, 1988.

Unité de Recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

Unité de Recherche INRIA Lorraine Technopôle de Nancy-Brabois - Campus Scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 VILLERS LES NANCY Cedex (France)
Unité de Recherche INRIA Rennes IRISA, Campus Universitaire de Beaulieu 35042 RENNES Cedex (France)
Unité de Recherche INRIA Rhône-Alpes 46, avenue Félix Viallet - 38031 GRENOBLE Cedex (France)
Unité de Recherche INRIA Sophia Antipolis 2004, route des Lucioles - B.P. 93 - 06902 SOPHIA ANTIPOLIS Cedex (France)

EDITEUR
INRIA - Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

ISSN 0249 - 6399