Chapitre 5

Discontinuous finite elements and finite volume versions of the Lax-Friedrichs and Nessyahu-Tadmor schemes for compressible flows on unstructured grids

by

Paul ARMINJON[†] , Marie-Claude VIALLON[⋆],

Aziz MADRANE[†] and Lahcen KADDOURI[◦]

[†]Dép. de mathématiques et de statistique, Université de Montréal

C.P. 6128, Succ. centre-ville, Montréal, Québec, Canada, H3C 3J7

[⋆]Équipe d'analyse numérique, Université de Saint-Étienne,

23 rue Paul Michelon, 42023 Saint-Étienne Cedex 2, France

[◦] Dép. de mathématiques appliquées Université du Littoral

50, Rue F. Buisson , B.P. 699 F-62228 Calais Cedex France

ABSTRACT – We present a new finite volume version ([1],[2],[3]) of the 1-dimensional Lax-Friedrichs and Nessyahu-Tadmor schemes ([5]) for nonlinear hyperbolic equations on unstructured grids, and compare it to a recent discontinuous finite element method ([6],[23]) in the computation of some typical test problems for compressible flows.
The non-oscillatory central difference scheme of Nessyahu and Tadmor, in which the resolution of Riemann problems at the cell interfaces is by-passed thanks to the use of the staggered Lax-Friedrichs scheme, is extended here to a two-step, two-dimensional non-oscillatory centered scheme in finite volume formulation. The construction of the scheme rests on a finite volume extension of the Lax-Friedrichs scheme, in which the finite volume cells are the barycentric cells constructed around the nodes of an FEM triangulation, for even time steps, and some quadrilateral cells associated with this triangulation, for odd time steps. Piecewise linear cell interpolants using least-squares gradients combined with a van Leer-type slope limiting allow for an oscillation-free second-order resolution.
The discontinuous finite element method consists of two steps. We first perform a finite element computation which includes calculation of the fluxes across the edges of the triangular elements using 1-D Riemann solvers with a modification to satisfy the entropy condition. We then proceed to a truly multidimensional slope limitation performed on the physical variables.

Numerical applications to several test problems show the accuracy and stability of the finite volume method.

1. Introduction

We present a new finite volume method for nonlinear hyperbolic systems on unstructured triangular grids inspired by the Lax-Friedrichs and Nessyahu-Tadmor one-dimensional difference schemes [5] , and compare it to a recent discontinuous finite element method introduced by Jaffré, Kaddouri and Gowda ([6],[23]) in the computation of some typical two-dimensional test problems for compressible flows.

Discontinuous finite elements were first introduced independently, for the neutron transport equation and applications to nuclear engineering, by Reed and Hill [7] and Lesaint and Raviart [8], and then adapted to the equations of hydrodynamics and elasticity ([9],[10], [11],[12]) or even to such applications as the motion of a load on an ice layer ([13], in joint work with P.Jamet), reservoir simulation both without slope limiters ([14],[15]) and with limitation ([16],[17],[18]), and many other applications where they proved to be very successful.

Numerical analysis of discontinuous finite element methods can be found e.g. in [8], with improvements in [19] for scalar hyperbolic equations, and in [20].

In the finite element method used in this paper, the solution is approximated by discontinuous piecewise linear polynomials; numerical fluxes are calculated, on the triangle edges, at appropriate integration points through the use of one-dimensional Riemann

solvers. To prevent spurious oscillations, we use a multidimensional version of van Leer's limiter borrowed from ([6],[23]).

Our finite volume method ([1],[2],[3],[4]) is a two-dimensional finite volume scheme inspired by an elegant difference scheme proposed, for one-dimensional problems, by Nessyahu and Tadmor [5], which is itself a high resolution non-oscillatory Godunov-type method for hyperbolic systems of conservation laws constructed on the principle of the staggered Lax-Friedrichs scheme.

It is decomposed in two time steps for second order accuracy in time, performed on alternate, staggered grids, thus allowing a complete by-pass of the (usually expensive) detailed exact or approximate solution of the local Riemann problems generated at the cell interfaces, thanks to the use of staggered form of the Lax-Friedrichs scheme.

In [4], we have presented an extension of the Nessyahu-Tadmor difference scheme to the simpler case of rectangular grids; several applications (linear advection, Burgers' equation, diffraction of a planar shock wave around a 90^o corner for the Euler equations [21] , Mach 3 wind tunnel with forward facing step [22]) showed the feasibility of the method, which led to good results on regular grids without any mesh adaptation.

In ([1],[2],[3]), we had constructed a finite volume scheme for unstructured triangular grids inspired by the Lax-Friedrichs and Nessyahu-Tadmor one-dimensional difference schemes; some numerical experiments (Supersonic flow around a NACA 0012 airfoil, supersonic flow around a double ellipse) showed the high accuracy of the method, but did not provide a systematic comparison with an already well established method.

In this paper we attempt to present such a comparison, and introduce for that purpose in section 3 the discontinuous finite element, second order accurate method recently proposed, and successfully tested, by Jaffré, Kaddouri and Gowda ([6],[23],[25]).

Notice that in this paper, we used Roe's Riemann solver with an entropy correction recently proposed by Dubois and Mehlman, while the solvers used in references ([6],[23],[25]) are the Osher and Osher-Solomon solvers, respectively.

In section 2 we describe the mathematical modelling of the problem; in section 3 we give a detailed presentation of the discontinuous finite element method of Jaffré, Kaddouri and Gowda.

In section 4, we present the first two authors' construction [1], for unstructured triangular grids, of a finite volume version of the staggered Lax-Friedrichs scheme, and then describe their second order accurate non-oscillatory finite volume method [1] inspired by the Nessyahu-Tadmor one-dimensional central difference scheme [5].
In section 5,we present applications to some typical test problems for supersonic flows (supersonic flow past a blunt body;supersonic flow around a double ellipse [35]).
Finally, section 6 is an appendix describing the solution of the projection problems generated by the slope limitation process.

2. Mathematical modelling

2.1. Governing equations.

We consider the two-dimensional Euler equations for compressible flows, written in conservation form as:

$$\frac{\partial}{\partial t}U(x,y;t)+\frac{\partial}{\partial x}F(U(x,y;t))+\frac{\partial}{\partial y}G(U(x,y;t))=0 \ \text{ for } (x,y;t)\in\Omega\times\mathbb{R}_+ \quad (2.1)$$

where Ω is a closed bounded domain of the plane,

$$U=\begin{pmatrix}\rho\\ \rho u\\ \rho v\\ \rho E\end{pmatrix},\quad F(U)=\begin{pmatrix}\rho u\\ \rho u^2+p\\ \rho uv\\ (\rho E+p)u\end{pmatrix},\quad G(U)=\begin{pmatrix}\rho v\\ \rho uv\\ \rho v^2+p\\ (\rho E+p)v\end{pmatrix} \quad (2.2)$$

ρ is the density, $\vec{V}=(u,v)$ is the velocity vector, E is the total energy by unit mass, and p is the pressure. We assume that the fluid satisfies the perfect gas law :

$$p=(\gamma-1)(\rho E-\frac{1}{2}\rho(u^2+v^2)) \quad (2.3)$$

where γ,the ratio of specific heats, is taken equal to 1.4 for air.

2.2. Boundary conditions :

In the sequel, we consider domains of computation related to external flows around bodies; in fig.5.53 the body is represented by a double-ellipse [40] which limits the

FIGURE 5.53. Boundary of the computational domain

domain of computation by its wall Γ_B . In order to deal with a bounded computational domain, a second (artificial) farfield boundary $\Gamma_\infty \cup \Gamma_S$ is introduced, with $\Gamma_S = \Gamma_S^1 \cup \Gamma_S^2$. The flow is assumed to be uniform at infinity, and we prescribe

$$\rho_\infty = 1, \quad V_\infty = \begin{pmatrix} cos\alpha \\ sin\alpha \end{pmatrix}, \quad p_\infty = \frac{1}{\gamma M_\infty^2} \tag{2.4}$$

where α is the angle of attack and M_∞ denotes the free-stream Mach number.
On the wall Γ_B we use the usual "no normal velocity" condition: $\vec{V}.\vec{n} = 0$, where $\vec{n} \in R^2$ is the outer normal vector to Γ_B.
Finally, for unsteady calculations, an initial flow is prescribed :

$$U(x, y; 0) = U_0(x, y) \quad \forall (x, y) \in \Omega. \tag{2.5}$$

3. A two-dimensional discontinuous finite element method

In this section, we give a detailed description of the discontinuous finite element method proposed by Jaffré,Kaddouri and Gowda ([6], [23],[25]).

We consider here a (P_1)-piecewise linear spatial discretization based on unstructured triangular elements (the number of elements sharing a common node is not

constant); the computational domain Ω is subdivided into triangles by a triangulation $\mathcal{T}_h$:

$$\Omega = \bigcup_{i=1}^{N_e} K_i \ ; \ K_i \in \mathcal{T}_h,$$

where the K_i's are the elements of the triangulation, N_e is the total number of triangles, and h is the largest diameter of all elements.

Let $U_{j,i} = U(A_{j,i}; t)$ denote the value of the dependent variable vector U, at time t and at the j^{th} vertex $A_{j,i}$ of element $K_i \in \mathcal{T}_h$ $(i = 1, \ldots, N_e; j = 1, 2, 3)$.

The degrees of freedom are the components of the vector U at the vertices of all elements of the triangulation (fig. 5.54). Let W denote the approximation space formed by

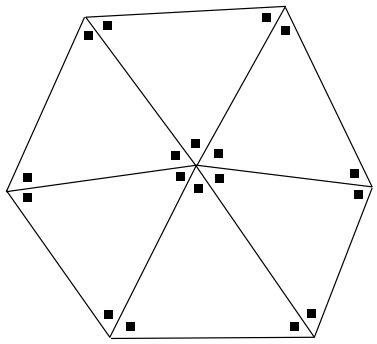

FIGURE 5.54. Degrees of freedom of the triangulation

the piecewise continuous functions which are linear on each triangle $K_i \in \mathcal{T}_h$ (P_1-approximation).

The vector U of conservative variables ρ, ρu, ρv, ρE is approximated by functions in the product space W^4 and the corresponding approximations will again be denoted by ρ, ρu, ρv, ρE, elements of W, for simplicity.

For each element $K_i (i = 1, \ldots, N_e)$ and each node $A_{j,i} \in K_i$ there exists a unique basis or shape function $(\varphi_{j,i})$ with the property

$$\forall\ i = 1, \ldots, N_e\ , \quad \varphi_{j,i}(A_{k,i}) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

The functions $(\varphi_{j,i})_{i=1,\ldots,N_e}^{j=1,\ldots,3}$ form a basis of the approximation space W.

In the present formulation we shall use an explicit Euler time discretization.

We now consider a Galerkin discontinuous finite element approximation, which proceeds from a variational formulation of the Euler equations.

Multiplying (2.1) by a shape function $\varphi_{j,i}$ and integrating by parts the terms with spatial derivatives, we obtain the following system for the piecewise linear vector U^{n+1} of approximate dependent variables to be computed at time t^{n+1}:

$$\left\{ \begin{array}{l} \text{Find } \; U^{n+1} \in W^4 \\ \\ \displaystyle\int_{K_i} \frac{U^{n+1}-U^n}{\Delta t^n}\varphi_{j,i}\, dS \;\; = \;\; \int_{K_i} \left(F(U^n)\frac{\partial \varphi_{j,i}}{\partial x} + G(U^n)\frac{\partial \varphi_{j,i}}{\partial y} \right) dS \\ \\ \qquad\qquad\qquad\qquad\qquad\qquad - \;\; \displaystyle\int_{\partial K_i} \left(\vec{\mathbb{F}}(U^n)\cdot \boldsymbol{n} \right) \varphi_{j,i}\, d\sigma \end{array} \right. \tag{3.1}$$

where $\Delta t^n = t^{n+1} - t^n$ is the time step, $U^n = U(x,y;t^n) \in W^4$, and $\vec{\boldsymbol{n}} = \begin{pmatrix} n_x \\ n_y \end{pmatrix}$ is the unit outer normal (directed towards the exterior of K_i).

In (3.1), $\vec{\mathbb{F}}(U^n)\cdot\vec{\boldsymbol{n}} = F(U^n)\cdot n_x + G(U^n)\cdot n_y \equiv F_n$ is the (outward) numerical flux across an edge $\mathcal{A}$ of ∂K_i; it will be computed with the help of an approximate Riemann solver for the Riemann problem generated, in the direction normal to the edge $\mathcal{A}$, by the limits of the values of the dependent variables on both sides of $\mathcal{A}$ as one tends to $\mathcal{A}$ along $\vec{\boldsymbol{n}}$.

In this paper, we have used Roe's Riemann solver ([28],[24],[34]) with an entropic correction due to Dubois and Mehlman [24].

REMARQUE 5.1. *Replacing the above approximation space by the space of piecewise constant functions (constant on each triangle $K_i \in \mathcal{T}_h$) reduces (3.1) to the equation governing the standard finite volume scheme*

$$\int_{K_i} \frac{U_i^{n+1}-U_i^n}{\Delta t^n}\, dS \; + \; \int_{\partial K_i} \left(\vec{\mathbb{F}}(U^n)\cdot \boldsymbol{n} \right) d\sigma = 0, \;\; K_i \in \mathcal{T}_h.$$

3.1. Numerical integration.

To complete the description of the spatial discretization, we must specify the quadrature formulas which will be used to compute the integrals appearing in (3.1). Numerical experiments for scalar equation ([25],[18],[6]) have suggested the following quadrature techniques.

In equation (3.1) the terms containing the spatial derivatives are computed with the help of the values at the centroid M of mesh K_i :

$$\int_{K_i}\left(F(U^n)\frac{\partial\varphi_{j,i}}{\partial x}+G(U^n)\frac{\partial\varphi_{j,i}}{\partial y}\right)dS\simeq \mathrm{Area}(K_i)\left(F(\overline{U}_i^n)\frac{\partial\varphi_{j,i}}{\partial x}(M_i)+G(\overline{U}_i^n)\frac{\partial\varphi_{j,i}}{\partial y}(M_i)\right)$$

where M_i is the centroid of K_i and $\overline{U}_i^n$ is the average value of U^n on K_i :

$$\overline{U}_i^n=\frac{1}{3}\sum_{j=1}^{3}U(A_{j,i})$$

since U^n is linear on K_i.

For the integral associated with the outward flux

$$\int_{\partial K_i}\left(\vec{\mathbb{F}}(U^n)\cdot\boldsymbol{n}\right)\varphi_{j,i}\,d\sigma=\sum_{\mathcal{A}\in\partial K_i}\int_{\mathcal{A}}\left(\vec{\mathbb{F}}(U^n)\cdot\boldsymbol{n}\right)\varphi_{j,i}\,d\sigma$$

we use either the values at the midpoints of the edges $\mathcal{A}$, or the values at both Gauss points of each edge

<u>1st choice :</u>

$$\int_{\mathcal{A}}\left(\vec{\mathbb{F}}(U^n)\cdot\boldsymbol{n}\right)\varphi_{j,i}\,d\sigma\;\simeq\;\mathrm{l}(\mathcal{A})\cdot\mathcal{F}^n_{|\mathcal{A}}(U_l,U_r)\cdot\varphi_{j,i}(M)$$

where $\mathrm{l}(\mathcal{A})$ and M denote the length and midpoint of edge $\mathcal{A}$, respectively, and

$\mathcal{F}^n_{|\mathcal{A}}(U_l,U_r)$ is the <u>numerical</u> <u>flux</u> across edge $\mathcal{A}$ which separates the states U_l and U_r, obtained by taking the limits of U along the normal to $\mathcal{A}$ at M; this numerical flux will be computed with Roe's Riemann solver, as described below.

<u>2nd choice :</u>

$$\int_{\mathcal{A}}\left(\vec{\mathbb{F}}(U^n)\cdot\boldsymbol{n}\right)\varphi_{j,i}\,d\sigma\;\simeq\;\frac{\mathrm{l}(\mathcal{A})}{2}\left(\mathcal{F}^n_{|\mathcal{A}}(U_l,U_r)(G_1)\varphi_{j,i}(G_1)+\mathcal{F}^n_{|\mathcal{A}}(U_l,U_r)(G_2)\varphi_{j,i}(G_2)\right)$$

where $G_1,\ G_2$ are the Gauss points of edge $\mathcal{A}$.

Numerical experiments have shown that the computation at the edge midpoint (1st choice) is sufficient.

For the first integral in (3.1) (time derivative) we use the values at the three vertices:

$$\int_{K_i}\frac{U^{n+1}-U^n}{\Delta t^n}\varphi_{j,i}\,dS\;\simeq\;\frac{\mathrm{Area}(K_i)}{3}\frac{U^{n+1}_{K_i,A_j}-U^n_{K_i,A_j}}{\Delta t^n}$$

by the properties of $\varphi_{j,i}$.

3.2. Multidimensional slope limitation.

We will describe a multidimensional extension of a slope limitation procedure which has been successfully used for scalar equations ([17],[6],[23]).

When dealing with the Euler equations, it has been widely recognized that one should limit the physical variables ρ, u, v, p rather than the conservative variables

ρ, ρu, ρv, ρE.

Let $U^n \in W^4$ denote the solution previously computed at time t^n, and $U^\star \in W^4$ the solution predicted at t^{n+1} by solving system (3.1).

We want to modify $U^\star$ and obtain a corrected vector of conservative variables U^{n+1}, by the following procedure.

For each triangle $K_i \in \mathcal{T}_h$, let

- $w_{K,A_i} = w_{|_K}(A_i) : \quad i = 1, \ldots, 3$ be the value of $w_{|_K}$ at node i,
- $\overline{w}_K = \dfrac{1}{3}\displaystyle\sum_{i=1}^{3} w_{K,A_i}$, the mean value of $w_{|_K}$ in element K,
- $T(A)$ be the set of element $K \in \mathcal{T}_h$ such that vertex $A \in K$ (Fig. 5.55).

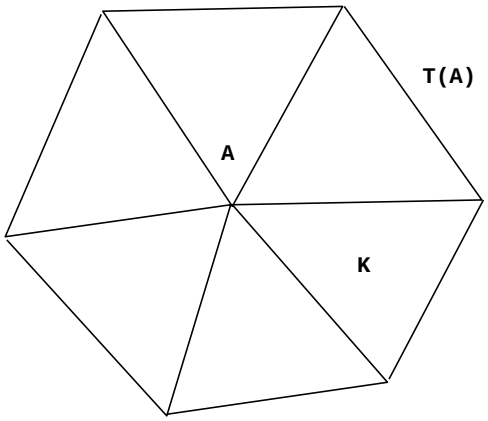

FIGURE 5.55. The set $T(A)$

For each element K, we compute the mean values of the conservative variables, noted $\overline{\rho}^\star_K$, $\overline{(\rho u)}^\star_K$, $\overline{(\rho v)}^\star_K$, $\overline{(\rho E)}^\star_K$,which are simply the arithmetic means of these variables at the three vertices of K.

In order to obtain a conservative scheme, the vectors U^{n+1} and $U^\star$ must have the same mean value on each element.

We then compute the mean values of the physical variables, $\overline{\rho}^\star_K, \overline{u}^\star_K, \overline{v}^\star_K, \overline{p}^\star_K$ (the mean value $\overline{\rho}^\star_K$ of the density has already been calculated).

For pressure we also take the arithmetic mean of p at the vertices of K.

In contrast, the mean value of the velocity components u, v are defined by

$$\overline{u}_K^\star = \frac{\overline{(\rho u)}_K^\star}{\overline{\rho}_K^\star} \ , \ \overline{v}_K^\star = \frac{\overline{(\rho v)}_K^\star}{\overline{\rho}_K^\star} \, . \tag{3.2}$$

For the components of momentum at time t^{n+1}, we will use the following value

$$\overline{(\rho u)}_K^{n+1} = \frac{1}{3} \sum_{i=1}^{3} \rho_{K,A_i}^{n+1} u_{K,A_i}^{n+1} \ , \quad \overline{(\rho v)}_K^{n+1} = \frac{1}{3} \sum_{i=1}^{3} \rho_{K,A_i}^{n+1} v_{K,A_i}^{n+1} \, . \tag{3.3}$$

Observe that these values are different from $\overline{u}_K^\star$, $\overline{v}_K^\star$, as they use the nodal values of density and velocity instead of those of the momentum.

Formulas (3.2)-(3.3) have been chosen to ensure existence and uniqueness of the solution of the minimisation problems to be defined below.

For every node A of the grid we compute the minimum and maximum of the mean values of the physical variables in the elements sharing node A :

$$w_{min}(A) = \min_{K \in T(A)} \overline{w}_K^\star, \quad w_{max}(A) = \max_{K \in T(A)} \overline{w}_K^\star, \qquad \text{for} \qquad w = \rho, u, v, p \tag{3.4}$$

The slopes of the physical variables ρ, u, v, p will be limited, in this order, in the following way.

Let V denote the vector $(\rho, u, v, p)^T$ of physical variables.

In each element K with vertices A_i $(i = 1, \ldots, 3)$, $V_{|K}^{n+1}$ is defined by :

$$\begin{cases} (i) \quad \overline{w}_K^{n+1} = \overline{w}_K^\star, \qquad \text{for} \quad w = \rho, \ p, \ \text{ and } \ \overline{(\rho w)}_K^{n+1} = \overline{(\rho w)}_K^\star, \ \text{for} \ \ w = u, \ v, \\ \\ (ii) \quad \text{For } i = 1, \ldots, 3, \ w = \rho, \ u, \ v, \ p : \\ \\ \qquad (1-\alpha)\overline{w}_K^\star + \alpha w_{min}(A) \le w_{K,A_i}^{n+1} \le (1-\alpha)\overline{w}_K^\star + \alpha w_{max}(A) \ , \ \ 0 \le \alpha \le 1, \\ \\ (iii) \quad \text{For } w = \rho, \ u, \ v, \ p \text{ the distance in } \mathbb{R}^3 \text{ between } w^{n+1} = (w_{K,A_i}^{n+1})_{i=1,\ldots,3} \\ \\ \qquad \text{and } w^\star = (w_{K,A_i}^\star)_{i=1,\ldots,3} \text{ is minimum .} \end{cases}$$

The computation of $V_{|K}^{n+1}$ from $V_{|K}^\star$ thus amounts to four projection problems in $\mathbb{R}^3$ (one for each physical variable); as (i) defines a plane, and (ii) a cube, we look for the projection, on their intersection, of the corresponding variable $w^\star = (w_{K,A_i}^\star)_{i=1,\ldots,3}$.

Condition (i) allows for mass conservation, (ii) limits the variation of $\rho,\ u,\ v,\ p$ (in that order), and (iii) guarantees uniqueness of the solution.

After computing the vector of physical variables $V^{n+1} = \begin{pmatrix} \rho^{n+1} \\ u^{n+1} \\ v^{n+1} \\ p^{n+1} \end{pmatrix}$ we return to the conservative variables according to

$$U^{n+1} = \begin{pmatrix} \rho^{n+1} \\ \rho^{n+1} u^{n+1} \\ \rho^{n+1} v^{n+1} \\ \dfrac{p^{n+1}}{(\gamma-1)} + \dfrac{1}{2}\rho^{n+1}\left[(u^{n+1})^2 + (v^{n+1})^2\right] \end{pmatrix}.$$

The slope limitation therefore requires the solution of a series of local minimization problems in 3 dimensional space, with the constraints (i) and (ii).

These projection problems can easily be solved by duality, as will be shown in the **appendix.**

In order to ensure the existence of a solution for the projection problems, we have to make sure that the intersection of the corresponding plane and cube is not empty. For density and pressure, it is easily seen that if we let $\rho_{K,A_i}^{n+1} = \overline{\rho}_K^\star$ and $p_{K,A_i}^{n+1} = \overline{p}_K^\star$ for $i = 1, 2, 3$, conditions (i) and (ii) are then satisfied, so that the relevant intersection is not empty.

As regards the velocity components, we can easily check, applying definitions (3.2) and (3.3), that if we let $u_{K,A_i}^{n+1} = \overline{u}_K^\star$ $(i = 1, 2, 3)$, then

$$\begin{aligned} \overline{(\rho u)}_K^{n+1} = \frac{1}{3}\sum_{i=1}^{3} \rho_{K,A_i}^{n+1}\,\overline{u}_K^\star &= \frac{1}{3}\left(\sum_{i=1}^{3}\rho_{K,A_i}^{n+1}\right)\frac{\overline{(\rho u)}_K^\star}{\overline{\rho}_K^\star} \\ &= \overline{(\rho u)}_K^\star\ . \end{aligned}$$

with a similar result for the second velocity component.

The parameter α controls the extent of the slope limitation process. For $\alpha = 0$, we get the most stringent limitation: the solution V^{n+1} (and therefore U^{n+1}) is piecewise constant, thus reducing the method to the usual (spatially) first order accurate scheme. In our numerical experiments, we have usually chosen $\alpha = 0.5$, a value which led to optimal results in the scalar case (cf. [25],[23],[6]).

For one-dimensional problems, the limitation procedure reduces to the usual van Leer solpe limitation ([26],[41]).

3.3. The numerical scheme.

The above description of the Jaffré-Kaddouri-Gowda discontinuous finite element method can be summarized within the frame of a two-step scheme.

Assuming for simplicity that we use an explicit Euler time discretization, let $U^n \in W^4$ be the solution obtained at time $t = t^n$. In the first step (predictor), we compute an approximation $U^\star \in W^4$ of the solution at time t^{n+1}. This predictor step consists in a finite element calculation, but features the use of Riemann solvers.

In the second step, which can be viewed as a correction step, we limit the vector $U^\star$ to obtain an approximate solution U^{n+1}.

1 - Predictor step : Finite element calculation

$$\left\{ \begin{array}{l} \text{Compute } U^\star \in W^4 \text{ such that} \\ \\ \displaystyle\int_{K_i} \frac{U^\star - U^n}{\Delta t^n} \varphi_{j,i}\, dS = \int_{K_i} \left(F(U^n) \frac{\partial \varphi_{j,i}}{\partial x} + G(U^n) \frac{\partial \varphi_{j,i}}{\partial y} \right) dS \\ \\ \qquad\qquad\qquad\qquad - \displaystyle\int_{\partial K_i} \left(\vec{\mathbb{F}}(U^n) \cdot \boldsymbol{n} \right) \varphi_{j,i}\, d\sigma \qquad (3.5) \\ \\ \text{for each } K_i \in \mathcal{T}_h \text{ and } \varphi_{j,i} (j = 1, 2, 3) \end{array} \right.$$

2 - Limitation step

This step limits the variation range of components of the vector of physical variables $V^\star = \begin{pmatrix} \rho^\star \\ u^\star \\ v^\star \\ p^\star \end{pmatrix}$ obtained in the predictor step. It leads to V^{n+1} and then U^{n+1}, the final approximation of the vector of conservative variables at time t^{n+1}.

We observe that the two steps are independent from each other, and the limitation process is distinct from the flux calculation, contributing to the originality of the method.

3.4. Riemann solver.

In this subsection we present a short description of Roe's approximate Riemann solver, which will be used here in conjunction with an entropic correction recently proposed by Dubois and Mehlman ([24]) described below.

Let $\mathcal{A}$ be an edge in the grid, and $\vec{\boldsymbol{n}}_{\mathcal{A}}$ a unit vector normal to $\mathcal{A}$; let $F_n(U)$ be the flux in the direction of $\vec{\boldsymbol{n}}_{\mathcal{A}}$, with $U \in W^4$. Let U_l and U_r be the limiting values of U obtained when approaching the edge $\mathcal{A}$ along $\vec{\boldsymbol{n}}_{\mathcal{A}}$ from the upwind and downwind side, respectively, with respect to the direction of $\vec{\boldsymbol{n}}_{\mathcal{A}}$; for the description of Roe's numerical flux, we will provisionally assume that U_l and U_r are two constant states along each side of $\mathcal{A}$ (while we will later use limits of piecewise linear functions, see below).

Let $\widetilde{F}_{\mathcal{A}}(U_l, U_r)$ denote the numerical flux of a specific Riemann solver. It satisfies the consistency relation $\widetilde{F}_{\mathcal{A}}(U, U) = F_n(U)$, for all $U \in W^4$, where

$$F_n(U) = n_1 F(U) + n_2 G(U) = \vec{\mathbb{F}}(U).\vec{\boldsymbol{n}}\,. \tag{3.6}$$

is the physical flux across $\mathcal{A}$, and

$$\vec{\boldsymbol{n}} \;=\; \vec{\boldsymbol{n}}_{\mathcal{A}} = \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \in \mathbb{R}^2_* \;\; (\vec{\boldsymbol{n}} \neq \vec{\mathbf{0}})$$

If $\vec{\boldsymbol{V}} = \begin{pmatrix} u \\ v \end{pmatrix}$ denotes the velocity, we can write it in the edge-normal to edge local basis formed by $\vec{\boldsymbol{n}}$ and a unit vector $\vec{\boldsymbol{a}}$ along edge $\mathcal{A}$ (fig.5.56) as

$$\vec{\boldsymbol{V}}_{\boldsymbol{n}} = \begin{pmatrix} u_n = \;\; u\cos\theta + v\sin\theta \\ v_n = -u\sin\theta + v\cos\theta \end{pmatrix}. \tag{3.7}$$

We can then define an invertible linear transformation T in $\mathbb{R}^4$:

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad , \quad T^{\perp 1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{3.8}$$

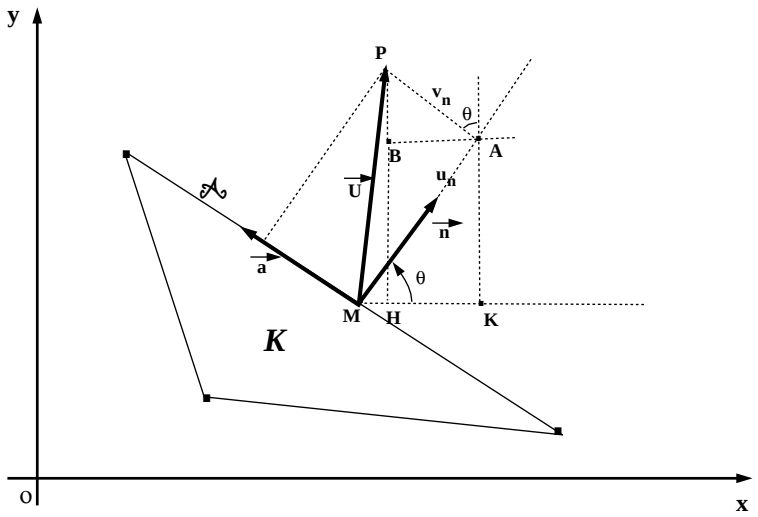

FIGURE 5.56. Local basis for the calculation of the flux

An arbitrary state vector $U = \begin{pmatrix} \rho \\ \rho\vec{\boldsymbol{V}} \\ \rho E \end{pmatrix}$ is then mapped into $\widehat{U}$:

$$U \longrightarrow T \cdot U = \begin{pmatrix} \rho \\ \rho\vec{\boldsymbol{V}}_{\boldsymbol{n}} \\ \rho E \end{pmatrix} = \widehat{U}\,. \tag{3.9}$$

The flux in the direction $\vec{\boldsymbol{n}}_{\mathcal{A}}$

$$F_n(U) = n_1 F(U) + n_2 G(U)\,.$$

can be written in the edge-normal to edge local basis as

$$F_n(U) = cos\theta \begin{pmatrix} \rho u \\ p + \rho u^2 \\ \rho uv \\ (\rho E + p)u \end{pmatrix} + sin\theta \begin{pmatrix} \rho v \\ \rho uv \\ p + \rho v^2 \\ (\rho E + p)v \end{pmatrix} = \begin{pmatrix} \rho u_n \\ (\rho u)u_n + p\cos\theta \\ (\rho v)u_n + p\sin\theta \\ (\rho H)u_n \end{pmatrix}, \tag{3.10}$$

where $H = E + \frac{p}{\rho}$ is the specific total enthalpy (per unit mass).
It can be easily verified that if $\hat{U} = TU = \left(\rho, \rho\vec{\boldsymbol{V}}_{\boldsymbol{n}}, \rho E\right)^T$, then

$$F_n(U) = T^{\perp 1} \circ F(\hat{U}) \tag{3.11}$$

where

$$F(\hat{U}) = (\rho u_n, p + \rho u_n^2, \rho u_n v_n, (\rho E + p) u_n)^T \tag{3.12}$$

This relation allows for the computation of $F_n(U)$ with the help of the first (vector) component F of the flux $\vec{\mathbb{F}}$ only, without resorting to G.

This leads to a reduction of the computing time, since the calculation of the matrices T and $T^{\perp 1}$ and the flux component F requires fewer operations than the calculation of the global flux F_n from its original definition.

In the sequel, the flux calculations will therefore be performed for $n_1 = 1$, $n_2 = 0$ (corresponding to $\theta = 0$).

3.5. Numerical flux of Roe's original scheme.

In Roe's scheme for the one-dimensional Euler equations $U_t + F(U)_x = 0$, one computes the exact solution of the linearized Riemann problem defined at the cell interface $x_{i+\frac{1}{2}}$ between cells C_i and C_{i+1} with corresponding constant states U_i, U_{i+1}, by replacing the Jacobian matrix $\frac{\partial F(U)}{\partial U} = A(U)$, computed at $x_{i+\frac{1}{2}}$, by a special constant matrix $\bar{A}_{i+\frac{1}{2}} = \bar{A}(U_i, U_{i+1})$ called Roe' average matrix, with the properties

(i) $F(U_{i+1}) - F(U_i) = \bar{A}_{i+\frac{1}{2}}(U_{i+1} - U_i)$

(ii) For $U_i = U_{i+1} = U$, we have $\bar{A}(U, U) = A(U) \equiv \frac{\partial F(U)}{\partial U}$

(iii) The eigenvalues λ_k of $\bar{A}$ are real and its eigenvectors r_k are linearly independant ($k = 1, \ldots, 3$).

It can be shown that $\bar{A}_{i+\frac{1}{2}} = \bar{A}(U_i, U_{i+1})$ is in fact equal to $A(\bar{U})_{i+\frac{1}{2}} \equiv \frac{\partial F(U)}{\partial U}|_{U=\bar{U}_{i+\frac{1}{2}}}$ (see [27]) where $\bar{U}_{i+\frac{1}{2}} = (\bar{\rho}_{i+\frac{1}{2}}, \bar{u}_{i+\frac{1}{2}}, \bar{H}_{i+\frac{1}{2}})$ is a specific average of the vectors U_i, U_{i+1} called Roe's average, defined by

$$\bar{\rho}_{i+\frac{1}{2}} = \sqrt{\rho_i \rho_{i+1}}, \quad \bar{u}_{i+\frac{1}{2}} = \frac{(u\sqrt{\rho})_i + (u\sqrt{\rho})_{i+1}}{\sqrt{\rho_i} + \sqrt{\rho_{i+1}}}, \quad \bar{H}_{i+\frac{1}{2}} = \frac{(H\sqrt{\rho})_i + (H\sqrt{\rho})_{i+1}}{\sqrt{\rho_i} + \sqrt{\rho_{i+1}}}$$

In the present context of the two-dimensional Euler equations, and owing to the previous remarks at the end of section 3.4, we will consider two adjacents states U_l, U_r separated by a triangle edge $\mathcal{A}$. Roe's average matrix will be given by (see [27] for a description of this matrix)

$$\bar{A}(U_l, U_r) = \frac{\partial F(U)}{\partial U}|_{U=\bar{U}_{lr}} \tag{3.13}$$

where Roe's average vector

$$\bar{U}_{lr} = (\bar{\rho}, \bar{u}, \bar{v}, \bar{H}) \tag{3.14}$$

is defined by

$$\begin{cases} \overline{\rho} = \sqrt{\rho_l \rho_r} \\ \text{and for } \ a = u,\ v \ or\ H \\ \bar{a} = \frac{a_l\sqrt{\rho_l} + a_r\sqrt{\rho_r}}{\sqrt{\rho_l}+\sqrt{\rho_r}} \end{cases} \tag{3.15}$$

The Roe averages of the remaining dependent variables E, p and c are then computed from $\overline{\rho}, \overline{u}, \overline{v}, \overline{H}$:

$$\begin{cases} \bar{E} = \frac{1}{\gamma}\bar{H} + \frac{\gamma\perp 1}{2\gamma}(\bar{u}^2 + \bar{v}^2) \\ \bar{c} = \{(\gamma-1)(\bar{H} - \frac{1}{2}(\bar{u}^2+\bar{v}^2))\}^{\frac{1}{2}} \\ \bar{p} = \frac{\bar{\rho}\bar{c}^2}{\gamma} \end{cases}$$

Properties (i),(ii) above now read

(i) $\quad F(U_l) - F(U_r) = \bar{A}(U_l, U_r)(U_r - U_l)$

(ii) $\quad \bar{A}(U,U) = dF(U) \equiv \frac{\partial F(U)}{\partial U}$

We write $\lambda_k(U_l, U_r)(k = 1, \ldots, 4)$ and $r_k(U_l, U_r)$ for the eigenvalues and eigenvectors of $\bar{A}(U_l, U_r)$ (see [27]) and let $\triangle U \equiv U_r - U_l$.

Since the $r_k(U_l, U_r)$ are linearly independent, there exists $(\alpha_k)_{k=1,\ldots,4} \in R^4$ such that

$$\triangle U \equiv U_r - U_l = \sum_{k=1}^{4} \alpha_k r_k(U_l, U_r) \tag{3.16}$$

the numerical flux of Roe's scheme can then be written ([28],[29])

$$\tilde{F}_{\mathcal{A}}(U_l, U_r) = \frac{1}{2}(F(U_l) + F(U_r)) - \frac{1}{2}d(U_l, U_r) \tag{3.17}$$

where $d(U_l, U_r)$ is the viscous term, given by

$$d(U_l, U_r) = |\bar{A}(U_l, U_r)|(U_r - U_l) = \sum_{k=1}^{4} \alpha_k |\lambda_k(U_l, U_r)| r_k(U_r, U_l) \tag{3.18}$$

Unfortunately, Roe's scheme does not satisfy an entropy inegality ([27],[29],[30]) and allows non-physical expansion shocks in the vicinity of sonic points ([21]), violating the Lax-Oleinik entropy condition; it is not an E-scheme ([43],[30],[27]), so that the

numerical solution does not necessarily converge to the (unique) entropy solution of the original nonlinear initial value problem.

To modify Roe's scheme so as to satisfy an entropy inequality, several ideas have been proposed (e.g. [29],[32],[33],[44],[45]), which are based on the representation of sonic rarefaction waves, and are equivalent to introducing a certain amount of artificial viscosity, the exact amount depending on a parameter which requires a case-dependent adjustment.

Dubois and Mehlman ([24]) have recently introduced a parameter-free modification of Roe's scheme based on nonlinear Hermite interpolation of an approximate flux function, leading to an entropic scheme.

3.6. Dubois-Mehlman's entropic correction of Roe's scheme.

This approach is based on a nonlinear modification of the flux function defined by (3.17) , in the neighbourhood of sonic points only.

For a given state U we write $U - U_l$ and $U_r - U_l$ as linear combinations of the eigenvectors.

Denoting by $(w_k)_{k=1,\dots,4}$ and $(\alpha_k)_{k=1,\dots,4}$ the characteristic variables associated[1] with $U - U_l$ and $U_r - U_l$, respectively, we have

$$U - U_l = \sum_{k=1}^{4} w_k \cdot \boldsymbol{r}_k(U_l, U_r) \tag{3.19}$$

$$U_r - U_l = \sum_{k=1}^{4} \alpha_k \cdot \boldsymbol{r}_k(U_l, U_r) \tag{3.20}$$

Following [29], we define the intermediate states

$$\left\{ \begin{array}{rcl} U_0 &=& U_l \\ &\vdots& \\ U_k &=& U_{k\perp 1} + \alpha_k \boldsymbol{r}_k(U_l, U_r) \\ &\vdots& \\ U_4 &=& U_r \end{array} \right. \tag{3.21}$$

Let $\mathbb{S}$ denote the set of sonic indices

$$\mathbb{S} = \{i,\ \lambda_i(U_{i\perp 1}) < 0 < \lambda_i(U_i)\} \tag{3.22}$$

[1] In this quasi one-dimensional context (due to the remarks following (3.12)) we can assume the existence of characteristic variables as in the purely one-dimensional case (cf. [27] Vol II p.155).

We now introduce the Dubois-Mehlman modified flux function, parameterized by the states U_l and U_r

$$F^{\mathrm{dm}}(U_l, U_r; U) = F(U_l) + \sum_{k=1}^{4} g_k(w_k) \cdot \boldsymbol{r}_k(U_l, U_r) \tag{3.23}$$

where the $w_k's$ are the characteristic variables introduced in (3.19), and the $g_k's$ are parameterized by the state $(U_j)_{j=1,\ldots,4}$, and defined as follows:

$$\begin{aligned} &\text{if } k \notin \mathbb{S} \qquad \text{then} \qquad g_k(w) = \lambda_k(U_l, U_r) \cdot w \quad \forall w \\ &\text{if } k \in \mathbb{S} \quad g_k(w) = \begin{cases} p_k(w) & \text{for } 0 < w < \alpha_k \\ \lambda_k(U_l, U_r) \cdot w & \text{for } w \notin (0, \alpha_k) \end{cases} \end{aligned} \tag{3.24}$$

where $p_k(w)$ is the unique Hermite polynomial of degree 3 determined by the conditions

$$p_k(0) = 0,\ p_k(\alpha_k) = \lambda_k(U_l, U_r) \cdot \alpha_k,\ p_k^{'}(0) = \lambda_k(U_{k\perp 1}),\ p_k^{'}(\alpha_k) = \lambda_k(U_k)$$

We recall that $\lambda_k(U_j)$ denotes the k-th eigenvalue of the jacobian matrix $A(U) \equiv \frac{\partial F}{\partial U}$ calculated at the intermediate state U_j, while $\lambda_k(U_l, U_r)$ is the k-th eigenvalue of the Roe matrix $\hat{A}(U_l, U_r)$. It can be verified that away from sonic points, the modified flux F^{dm} coincides with Roe's linearized flux:

$$F^R(U) = F(U_l) + \hat{A}(U_l, U_r)(U - U_l) \tag{3.25}$$

If the original physical flux $F(U)$ is at least continuously differentiable and if $\hat{A}(U_l, U_r)$ is a continuous function of U_l and U_r, F^{dm} is a continuous function of U_l, U_r and U.

LEMMA 3.1.

The Riemann problem

$$\begin{cases} \dfrac{\partial U}{\partial t} + \dfrac{\partial F^{dm}(U_l, U_r; U)}{\partial x} = 0 \\ U(x;0) = \begin{cases} U_l & x < 0 \\ U_r & x > 0 \end{cases} \end{cases}$$

has a unique entropy solution.

See [24] for the proofs of lemmas 3.1 and 3.2.

LEMMA 3.2. *For $k \in \mathbb{S}$, the Hermite polynomial $p_k(w)$ is defined by*

$$p_k(w) = a_k w^3 + b_k w^2 + c_k w \, ,$$

with

$$\begin{aligned} a_k &= \frac{\lambda_k(U_k) + \lambda_k(U_{k\perp 1}) - 2\lambda_k(U_l, U_r)}{\alpha_k^2} \\ b_k &= \frac{3\lambda_k(U_l, U_r) - 2\lambda_k(U_{k\perp 1}) - \lambda_k(U_k)}{\alpha_k} \\ c_k &= \lambda_k(U_{k\perp 1}) \end{aligned}$$

The modified numerical flux can be written as

$$\widetilde{F}^{dm}_{\mathcal{A}}(U_l, U_r) = F(U_l) + \sum_{k \notin \mathbb{S}, \lambda_k(U_l,U_r)<0} \alpha_k \cdot \lambda_k(U_l, U_r) \cdot \boldsymbol{r}_k + \sum_{k \in \mathbb{S}} g_k(w_k^\star) \cdot \boldsymbol{r}_k \quad (3.26)$$

where

$$w_k^\star = \frac{-\lambda_k(U_{k\perp 1}) \cdot \alpha_k}{3\lambda_k - 2\lambda_k(U_{k\perp 1}) - \lambda_k(U_k) + \sqrt{(3\lambda_k - \lambda_k(U_{k\perp 1}) - \lambda_k(U_k))^2 - \lambda_k(U_{k\perp 1}) \cdot \lambda_k(U_k)}} \quad (3.27)$$

is the argument of the unique extremum of $g_k(w)$ in the interval $0 < w < \alpha_k$, and λ_k in (3.27) denotes $\lambda_k(U_l, U_r)$.

Description of the eigenvectors $r_k(U_l, U_r)$ and characteristic variables α_k

To complete our description of the Dubois-Mehlman entropy correction to Roe's scheme, we must define the eigenvectors $r_k(U_l, U_r)$ and coefficients α_k appearing in (3.19),(3.20), (3.27).

From [28], we have

$$\boldsymbol{r}_1 = \begin{pmatrix} 1 \\ \widehat{u} - \widehat{c} \\ \widehat{v} \\ \widehat{H} - \widehat{u}\widehat{c} \end{pmatrix} , \; \boldsymbol{r}_2 = \begin{pmatrix} 0 \\ 0 \\ \widehat{v} \\ \widehat{v}^2 \end{pmatrix} , \; \boldsymbol{r}_3 = \begin{pmatrix} 1 \\ \widehat{u} \\ \widehat{v} \\ \frac{1}{2}(\widehat{u}^2 + \widehat{v}^2) \end{pmatrix} , \; \boldsymbol{r}_4 = \begin{pmatrix} 1 \\ \widehat{u} + \widehat{c} \\ \widehat{v} \\ \widehat{H} + \widehat{u}\widehat{c} \end{pmatrix}$$

For $w = \rho, \rho u, \rho v$ et ρE, we set $\Delta w = w_r - w_l$. Equation (3.20) defines a linear system in the 4 unknowns $\alpha_1, \ldots, \alpha_4$. For $\widehat{v} \neq 0$, the unique solution is given by

$$\left\{ \begin{array}{rcl} \alpha_1 & = & \dfrac{\gamma-1}{2\,\widehat{c}^2}\left\{\Delta(\rho E) - \widehat{v}\Delta(\rho v) - \left(\widehat{u} + \dfrac{\widehat{c}}{\gamma-1}\right)\Delta(\rho u) + \left(\dfrac{\widehat{c}\cdot\widehat{u}}{\gamma-1} + \dfrac{1}{2}(\widehat{u}^2+\widehat{v}^2)\right)\Delta\rho\right\} \\ \alpha_2 & = & \dfrac{1}{\widehat{v}}\Delta(\rho v) - \Delta\rho \\ \alpha_3 & = & \left(1 + \dfrac{\widehat{u}}{\widehat{c}}\right)\Delta\rho - \dfrac{1}{\widehat{c}}\Delta(\rho u) - 2\,\alpha_1 \\ \alpha_4 & = & \alpha_1 + \dfrac{1}{\widehat{c}}\left(\Delta(\rho u) - \widehat{u}\Delta\rho\right) \end{array} \right.$$

If $\widehat{v} = 0$, α_2 is indeterminate. One can set it equal to 0, in which case there exists a unique $(\alpha_k)_{k=1,\ldots,4} \in \mathbb{R}^4$ satisfying (3.20), given by

$$\left\{ \begin{array}{rcl} \alpha_1 & = & \dfrac{\gamma-1}{2\,\widehat{c}^2}\left\{\Delta(\rho E) - \left(\widehat{u} + \dfrac{\widehat{c}}{\gamma-1}\right)\Delta(\rho u) + \left(\dfrac{\widehat{c}\cdot\widehat{u}}{\gamma-1} + \dfrac{\widehat{u}^2}{2}\right)\Delta\rho\right\} \\ \alpha_2 & = & 0 \\ \alpha_3 & = & \left(1 + \dfrac{\widehat{u}}{\widehat{c}}\right)\Delta\rho - \dfrac{1}{\widehat{c}}\Delta(\rho u) - 2\,\alpha_1 \\ \alpha_4 & = & \alpha_1 + \dfrac{1}{\widehat{c}}\left(\Delta(\rho u) - \widehat{u}\Delta\rho\right) \end{array} \right.$$

The Dubois-Mehlman correction has proved to be useful for explicit schemes, but difficult to implement for implicit schemes, owing to the difficulty of the linearization process.

3.7. Note on the implementation of the time discretization.

In this paper, we limit our applications to the computation of stationary solutions, and will therefore use a local time stepping process described below.
Introducing the diagonal matrix $\Sigma = \mathrm{diag}\left[\dfrac{\mathrm{Area}(K_i)}{\Delta t_i^n}\right]_{K_i \in \mathcal{T}_h}$, we can write (3.5) as follows:

$$\left\{ \begin{array}{l} \text{Find } U^\star \in W^4 \quad \text{such that} \\ \\ U^\star - U^n \;=\; -\Sigma^{\perp 1}\,\mathcal{R}(U^n) \end{array} \right. \tag{3.28}$$

where $\mathcal{R}(U^n)$ is the residual defined from the right-hand side of (3.5)

$$\mathcal{R}(U^n) = \int_{\partial K_i} \left(\vec{\mathbb{F}}(U^n)\cdot \boldsymbol{n}\right) \phi_{j,i}\, d\sigma \;-\; \int_{K_i} \left(F(U^n)\frac{\partial \phi_{j,i}}{\partial x} + G(U^n)\frac{\partial \phi_{j,i}}{\partial y}\right) dS\,. \quad (3.29)$$

The scheme (3.28) is stable under an appropriate $CFL-condition$.

Let μ denote the $CFL-number$ (assumed to be uniform on the whole grid). For each element $K_i \in \mathcal{T}_h$ we note

- ν_i : mean value, in element K_i, of the characteristic speed corresponding to the largest eigenvalue

$$\nu_i = \sqrt{\overline{u}_i^2 + \overline{v}_i^2} + \overline{c}_i\ ,$$

- h_i : ratio of the area of K_i by its perimeter

$$h_i = \frac{\mathrm{Area}(K_i)}{\mathrm{L}(K_i)}\ .$$

The local time step is then chosen so that

$$\Delta t_i^n \leq \mu \frac{h_i}{\nu_i}\,. \quad (3.30)$$

In most cases we have used a $CFL-number$ $\mu = 0.5$.

In the numerical experiments presented in this paper, we study the distribution on the body B (fig.1) and the isolines of the Mach number M, the pressure p, or the pressure coefficient C_p, defined by

$$Cp \;=\; \frac{p_\infty - p}{\dfrac{1}{2}\rho_\infty \parallel \boldsymbol{V}_\infty \parallel^2} \quad (3.31)$$

To study the convergence of the method, we will present graphs of the $\mathbb{L}_2$-norm of the residual $\mathcal{R}(U^n)$ as a function of the number of iterations.

4. Finite volume methods on unstructured triangular grids

4.1. A two-dimensional finite volume method inspired by the Lax-Friedrichs scheme.

We consider the solution $U(x,y,t)$ of the two-dimensional Euler equations (2.1)-(2.2)

$$U_t + F(U)_x + G(U)_y = 0 \quad (4.1)$$

in some region Ω of the $x-y$ plane. In one space dimension [5], both the staggered form of the Lax-Friedrichs scheme and the Nessyahu-Tadmor scheme use two alternate grids $\{x_j\}$ and $\{x_{j+1/2}\}$ at odd and even time steps, respectively. In two dimensions, we proceed in a similar way, starting from an arbitrary FEM triangular grid $\mathcal{T}_h$ such that

$$\Omega = \bigcup_{T\in\mathcal{T}_h} T \text{ and } T\cap T' = \begin{cases} \phi \\ \text{one vertex} \\ \text{one side} \end{cases} \qquad \text{for any } T, T' \in \mathcal{T}_h \tag{4.2}$$

The nodes of the FEM triangulation are the vertices a_i of the triangles, and in this subsection the degrees of freedom are the vector values of U at the nodes, which can also be considered as cell average values for the cell C_i centered at each individual node a_i (defined below).

For the first grid associated with our finite volume extension of the Lax-Friedrichs scheme, the nodes are the vertices a_i of $\mathcal{T}_h$ while the finite volume cells are the barycentric cells C_i associated with these nodes, obtained by joining the midpoints M_{ij} of the sides originating in a_i to the centroids G_{ij} of the triangles of $\mathcal{T}_h$ which meet at a_i (fig.5.57).

For the second grid the nodes are the midpoints M_{ij} of the sides of the original triangulation, while the cells are the quadrilaterals of the form $L_{ij} = a_i G_{ij} a_j G_{i,j+1}$ having M_{ij} as midpoint of one diagonal, obtained by joining two adjacent nodes a_i, a_j to the centroids of the two triangles of $\mathcal{T}_h$ of which $a_i a_j$ is a side.

Let $U_i^n \cong U(a_i, t^n)$ and $U_{ij}^{n+1} \cong U(M_{ij}, t^{n+1})$ denote the nodal (or cell average) vector values in the first and second grid at time $t = t^n$ and $t = t^{n+1}$, respectively (n even).

For the barycentric cell C_i, let $\vec{\nu}_{ij}^{\,1}$ and $\vec{\nu}_{ij}^{\,2}$ denote the unit outer normal vectors to $G_{ij}M_{ij}$ and $M_{ij}G_{i,j+1}$ respectively, pointing out of cell C_i (fig.5.58) , and for the quadrilateral cell L_{ij}, let $\vec{n}^{\,1}{}_{ij}, \ldots, \vec{n}_{ij}^{\,4}$ be the normal vectors to the cell edges a_iG_{ij}, $G_{ij}a_j$, $a_jG_{i,j+1}$ and $G_{i,j+1}a_i$, respectively, pointing out of cell L_{ij} (fig.5.59).

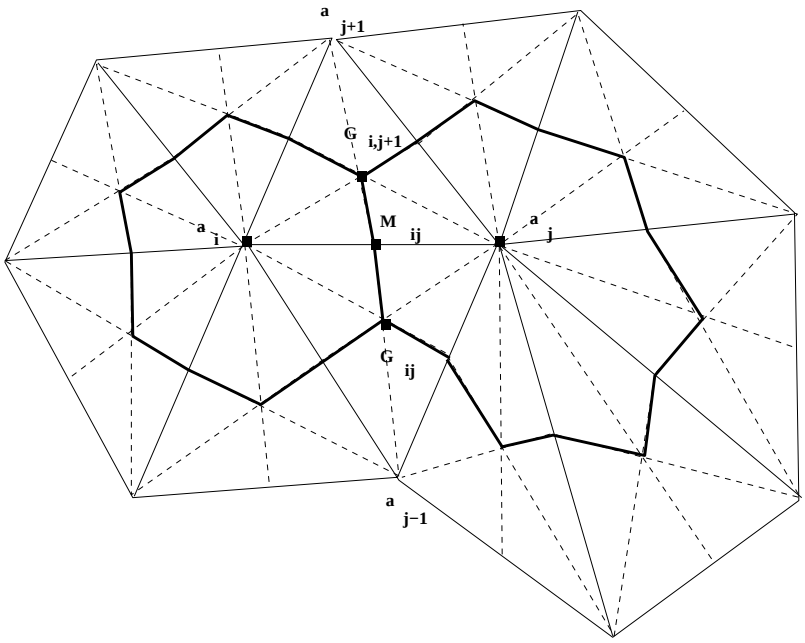

FIGURE 5.57. Barycentric cells around nodes $\mathbf{a_i}$, $\mathbf{a_j}$; quadrilateral cell $\mathbf{a_i G_{ij} a_j G_{i,j+1}}$.

We must also define the elementary flux vectors

$$\vec{\eta}_{ij} = \int_{\Gamma_{ij}=G_{ij}M_{ij}G_{i,j+1}} \vec{\nu} d\sigma = |\overrightarrow{G_{ij}M_{ij}}|\vec{\nu}_{ij}^{\,1} + |\overrightarrow{M_{ij}G_{i,j+1}}|\vec{\nu}_{ij}^{\,2} \tag{4.3}$$

and

$$\begin{cases} \vec{\theta}_{ij} = |\overrightarrow{a_iG_{ij}}|\vec{n}_{ij}^{\,1} + |\overrightarrow{a_iG_{i,j+1}}|\vec{n}_{ij}^{\,4} \\ \vec{\theta}_{ji} = |\overrightarrow{a_jG_{ij}}|\vec{n}_{ij}^{\,2} + |\overrightarrow{a_jG_{i,j+1}}|\vec{n}_{ij}^{\,3}. \end{cases} \tag{4.4}$$

We write furthermore

$$\vec{\nu}_{ij}^{\,k} = \begin{pmatrix} \nu_{ijx}^k \\ \nu_{ijy}^k \end{pmatrix} \text{ for } k = 1, 2, \quad \vec{n}_{ij}^{\,k} = \begin{pmatrix} n_{ijx}^k \\ n_{ijy}^k \end{pmatrix} \text{ for } k = 1, \ldots, 4$$

$$\text{and } \vec{\eta}_{ij} = \begin{pmatrix} \eta_{ijx} \\ \eta_{ijy} \end{pmatrix}, \quad \vec{\theta}_{ij} = \begin{pmatrix} \theta_{ijx} \\ \theta_{ijy} \end{pmatrix}. \tag{4.5}$$

An advantage of using a finite volume formulation where the degrees of freedom are values of the unknown function at the triangulation vertices lies in the possibility to couple (4.1) with an elliptic equation. This can be very convenient in the case of a

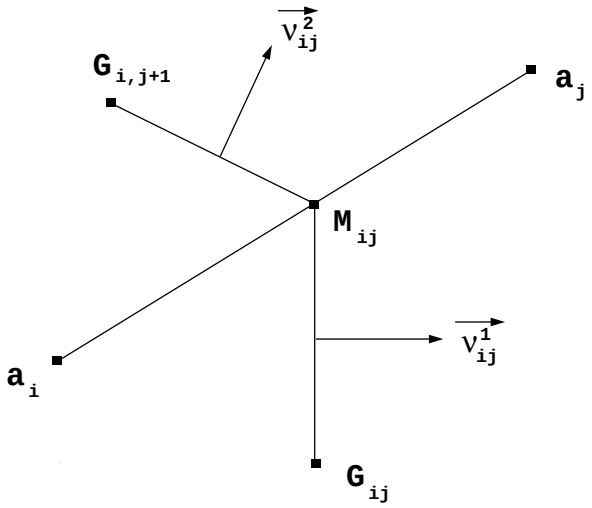

FIGURE 5.58. Barycentric cell boundary element $\Gamma_{ij} = \mathbf{G_{ij}M_{ij}} \cup \mathbf{M_{ij}G_{i,j+1}}$.

scalar conservation law where

$$\begin{cases} f(u) = w_1\, h(u) \\ g(u) = w_2\, h(u) \end{cases} \tag{4.6}$$

i.e. when (4.1) can be written as $u_t + \operatorname{div}(\vec{W}\, h(u)) = 0$ with $\vec{W} = (w_1, w_2)^T$ and $\vec{W}$ stems from an elliptic problem. This situation arises in the study of polyphase flows in porous media, where fluid mechanical and thermodynamical considerations are combined, leading to coupled hyperbolic and elliptic equations [42].

The first step of the two-dimensional finite volume extension of the Lax-Friedrichs scheme is defined by integrating (4.1) on the 3-dimensional cell $L_{ij} \times [t^n, t^{n+1}]$, assuming that the (barycentric) cell values U_i^n at the vertices a_i of the original triangulation are known:

$$\int_{t^n}^{t^{n+1}} \int_{L_{ij}} (U_t + F(U)_x + G(U)_y) dx dy\, dt = 0. \tag{4.7}$$

Applying the divergence theorem and observing that $L_{ij} = (L_{ij} \cap C_i) \cup (L_{ij} \cap C_j)$ we get

$$\int_{L_{ij}} U(x,y,t^{n+1}) dA - \int_{L_{ij}\cap C_i} U(x,y,t^n) dA - \int_{L_{ij}\cap C_j} U(x,y,t^n) dA + \int_{t^n}^{t^{n+1}} \int_{\partial L_{ij}} (F(U)n_x + G(U)n_y) d\sigma\, dt = 0. \tag{4.8}$$

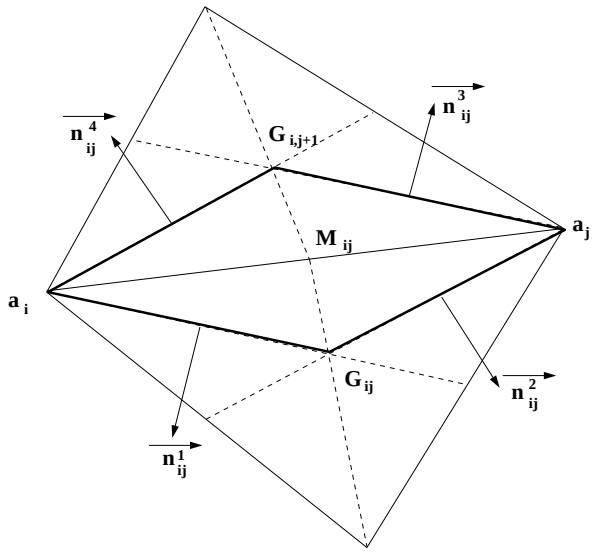

FIGURE 5.59. Quadrilateral cells $\mathbf{L_{ij}}$.

Since $U(x,t)$ is approximated by U_i^n in C_i, U_j^n in C_j, we can choose the approximation U_i^n on $\partial L_{ij} \cap C_i$ and U_j^n on $\partial L_{ij} \cap C_j$, whence the first step of our finite volume Lax-Friedrichs scheme:

$$A(L_{ij})U_{ij}^{n+1} - A(L_{ij} \cap C_i) \cdot U_i^n - A(L_{ij} \cap C_j) \cdot U_j^n + \Delta t\big(F(U_i^n)\theta_{ijx} + G(U_i^n)\theta_{ijy}\big) + \Delta t\big(F(U_j^n)\theta_{jix} + G(U_j^n)\theta_{jiy}\big) = 0. \quad (4.9)$$

We note that this approximation corresponds to choosing the approximate time derivative

$$U_t \cong \left(U_{ij}^{n+1} - \frac{A(L_{ij} \cap C_i) \cdot U_i^n + A(L_{ij} \cap C_j) \cdot U_j^n}{A(L_{ij})}\right) \Big/ \Delta t. \quad (4.10)$$

For the second step we proceed similarly with the help of the barycentric cells C_i:

$$A(C_i)U_i^{n+2} - \sum_{j \text{ neighbour of } i} A(L_{ij} \cap C_i)U_{ij}^{n+1} + \Delta t \sum_{j \text{ neighbour of } i} \big(F(U_{ij}^{n+1})\eta_{ijx} + G(U_{ij}^{n+1})\eta_{ijy}\big) = 0 \quad (4.11)$$

where the value of $U(x,y,t^{n+1})$ on the boundary $\partial C_i = \bigcup_{j \text{ neighbour of } i} \Gamma_{ij}$ is approximated locally, on $\partial C_i \cap L_{ij} = \Gamma_{ij}$, by U_{ij}^{n+1} (i.e. the approximate value of U^{n+1} on the quadrilateral cell L_{ij}. We thus alternately define an approximate solution U_{ij}^{n+1} which is piecewise constant on the quadrilateral cells L_{ij}, at odd time steps ($n = 0, 2, \dots$), and a solution U_i^{n+2} constant on the barycentric cells C_i, at even time steps.

4.2. A two-dimensional finite volume inspired by the Nessyahu-Tadmor scheme.

The construction of a finite volume method inspired by [5] implies the computation of gradients of piecewise linear functions, and the limitation of the gradients. To simplify the presentation, we shall describe our finite volume method in the case of scalar conservation laws in this subsection. The extension to the vector case is obtained from a field by field decomposition (see, e.g. [47] ,[46],[48]).

We consider a scalar nonlinear conservation equation

(4.1′) $$u_t + f(u)_x + g(u)_y = 0$$

At the beginning of the $(n+1)^{\text{st}}$ time step (n even), we have obtained approximate barycentric cell values u_i^n (a_i : a vertex of $\mathcal{T}_h$). We must now, in order to follow the van Leer MUSCL approach used by Nessyahu and Tadmor, construct a piecewise linear profile on the barycentric cells C_i; this can be achieved as follows.

We first construct a piecewise linear approximant on each triangle T of the original triangulation, continuous on the whole computational domain Ω_h, with the help of the barycentric cell/nodal values u_i^n: if $T \in \mathcal{T}_h$ is a triangle with vertices a_{i_j}, $(j = 1, 2, 3)$, we construct $p_T \in P^1$ such that

$$p_T(a_{i_j}) = u_{i_j}^n \ (j = 1, 2, 3)$$

$p_T(x, y)$ is easily obtained from the barycentric coordinates of (x, y) with respect to the vertices of T:

$$p_T(\vec{x}) = \sum_{j=1}^{3} \lambda_j(\vec{x}) u_j^n$$

where the vertices of T have been relabelled $1, 2, 3$ or a_1, a_2, a_3, and $\vec{x} = (x, y) \in T$.

The gradient of the (barycentric) cellwise *piecewise linear interpolant* $L(x, y, t^n)$ to be defined will now be chosen (as e.g. in [39] p.28), for cell C_i, as the arithmetic average of the gradients of the polynomials p_T for all triangles T such that $a_i \in T$: on C_i we take

$$L = L_i(x, y, t^n) = u_i^n + (x - x_i)P_i^n + (y - y_i)Q_i^n \qquad (x, y) \in C_i \tag{4.12}$$

where

$$\begin{pmatrix} P_i^n \\ Q_i^n \end{pmatrix} = \underset{a_i \in T}{\text{Average}} \{\text{grad } p_T\}. \tag{4.13}$$

Contrary to what prevailed in the one-dimensional case, where the average value $\bar{u}_j^n$ of the piecewise linear interpolant $L_j(x,t^n)$ was also its value at the node x_j, we can no longer identify the average value of the piecewise linear interpolant (4.12) , on cell C_i, with its nodal value u_i^n at the "center" a_i of C_i, since a_i need not be the centroid of C_i, and $\frac{1}{A(C_i)}\int_{C_i} L_i(x,y,t^n)dA \neq u_i^n$ in general.

The new cell values at t^{n+1} and t^{n+2} will nevertheless again be *defined* by formulas similar to (4.9) (first step), and (4.11) (second step), obtained by integrating (4.1') on $L_{ij}\times[t^n,t^{n+1}]$ for the first step, and on $C_i\times[t^n,t^{n+1}]$ for the second step:

$$A(L_{ij})u_{ij}^{n+1} \equiv \text{numerical approximation of } \int_{L_{ij}} u(x,y,t^{n+1})dA \tag{4.14}$$

$$A(C_i)u_i^{n+2} \equiv \text{numerical approximation of } \int_{C_i} u(x,y,t^{n+2})dA. \tag{4.15}$$

For the first step of our scheme we write

$$\int_{t^n}^{t^{n+1}}\int_{L_{ij}} (u_t + f(u)_x + g(u)_y)dAdt = 0 \tag{4.16}$$

which leads to

$$\int_{L_{ij}} u(x,y,t^{n+1})dA = \int_{L_{ij}\cap C_i} L(x,y,t^n)dA + \int_{L_{ij}\cap C_j} L(x,y,t^n)dA - \int_{t^n}^{t^{n+1}}\int_{\partial L_{ij}} (f(u)n_x + g(u)n_y)d\sigma dt. \tag{4.17}$$

The numerical approximation of the right-hand side, and (4.14), will thus lead to u_{ij}^{n+1}, which will be our cell value for the quadrilateral cell L_{ij}.

4.3. Approximation of $\int_{L_{ij}\cap C_i} L(x,y,t^n)dA$.

$L(x,y,t^n)$ is the piecewise linear function defined, on cell C_i, by (4.12) . Let $A_iA_{ij}B_{ij}\ A_{i,j+1}$ be the points of the plane defined by the linear function L_i on C_i which correspond to the four vertices of $L_{ij}\cap C_i = [a_iG_{ij}M_{ij}G_{i,j+1}]$ where [...] denotes the quadrilateral generated by the corresponding vertices (fig. 5.60).

The integral of L on $L_{ij}\cap C_i$ is equal to the total volume of the two prisms with triangular base $a_iG_{ij}M_{ij}A_iA_{ij}B_{ij}$ and $a_iM_{ij}G_{i,j+1}A_iB_{ij}A_{i,j+1}$, constructed on the triangular bases $L_{ij}^r\cap C_i$ and $L_{ij}^\ell\cap C_i$ where L_{ij}^r = triangle $(a_iG_{ij}a_j)$ and L_{ij}^ℓ = triangle $(a_ia_jG_{i,j+1})$; r,ℓ stand for right, left (for an observer at a_i).

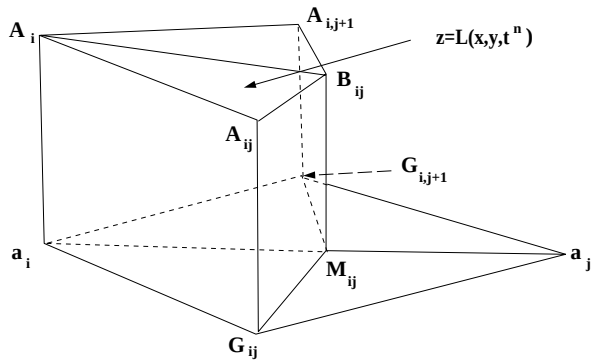

FIGURE 5.60. Prismatic regions for the computation of $\int_{L_{ij}\cap C_i} \boldsymbol{L(x,y,t^n)dA}$.

The volume of the first prism, for instance, is given by

$$\begin{aligned}
\text{Vol}\{a_iG_{ij}M_{ij}A_iA_{ij}B_{ij}\} &= \frac{1}{3}\ \text{Area}\,(a_iG_{ij}M_{ij})\cdot[a_iA_i + G_{ij}A_{ij} + M_{ij}B_{ij}] \\
&= \frac{1}{3}\ \text{Area}\,(L_{ij}^r\cap C_i)\{u_i^n + u_i^n + (x_{G_{ij}} - x_i)P_i^n \\
&\qquad + (y_{G_{ij}} - y_i)Q_i^n + u_i^n + (x_{M_{ij}} - x_i)P_i^n \\
&\qquad + (y_{M_{ij}} - y_i)Q_i^n\} \\
&= \text{Area}\,(L_{ij}^r\cap C_i)\{u_i^n + \frac{1}{3}(x_{G_{ij}} + x_{M_{ij}} - 2x_i)P_i^n \\
&\qquad + \frac{1}{3}(y_{G_{ij}} + y_{M_{ij}} - 2y_i)Q_i^n\} \qquad (4.18)
\end{aligned}$$

where

$$\begin{aligned}
\text{Area}\,(L_{ij}^r\cap C_i) &\equiv \text{Area triangle}\,(a_iG_{ij}M_{ij}) \\
&= \frac{1}{2}\{(x_{G_{ij}} - x_i)(y_{M_{ij}} - y_i) - (x_{M_{ij}} - x_i)(y_{G_{ij}} - y_i)\}. \qquad (4.19)
\end{aligned}$$

Similarly, the volume of the second prism is

$$\begin{aligned}
\text{Vol}\{a_iM_{ij}G_{i,j+1}\ A_iB_{ij}A_{i,j+1}\} &= \text{Area}(L_{ij}^\ell\cap C_i)\cdot \\
&\{u_i^n + \frac{1}{3}(x_{M_{ij}} + x_{G_{i,j+1}} - 2x_i)P_i^n + \frac{1}{3}(y_{M_{ij}} + y_{G_{i,j+1}} - 2y_i)\} \qquad (4.20)
\end{aligned}$$

Summing (4.18) and (4.20), we get

$$\begin{aligned}
\int_{L_{ij}\cap C_i} L(x,y,t^n)dA &= \text{Area}(L_{ij}\cap C_i)\{u_i^n + \frac{1}{3}(x_{M_{ij}} - x_i)P_i^n + \frac{1}{3}(y_{M_{ij}} - y_i)Q_i^n\} \\
&+ \frac{1}{3}\text{Area}(L_{ij}^r\cap C_i)\{(x_{G_{ij}} - x_i)P_i^n + (y_{G_{ij}} - y_i)Q_i^n\} \qquad (4.21) \\
&+ \frac{1}{3}\text{Area}(L_{ij}^\ell\cap C_i)\{(x_{G_{i,j+1}} - x_i)P_i^n + (y_{G_{i,j+1}} - y_i)Q_i^n\}
\end{aligned}$$

with

$$\text{Area}(L_{ij}^{\ell} \cap C_i) \;=\; \frac{1}{2}\{(x_{M_{ij}} - x_i)(y_{G_{i,j+1}} - y_i) - (x_{G_{i,j+1}} - x_i)(y_{M_{ij}} - y_i)\}. \quad (4.22)$$

4.4. Approximation of $\int_{L_{ij}\cap C_j} L(x,y,t^n)dA$.

Proceeding in the same way we find

$$\begin{aligned}\int_{L_{ij}\cap C_j} L(x,y,t^n)dA = \text{Area}(L_{ij}\cap C_j)\{u_j^n + \frac{1}{3}(x_{M_{ij}} - x_j)P_j^n + \frac{1}{3}(y_{M_{ij}} - y_j)Q_j^n\} \\ + \frac{1}{3}\text{Area}(L_{ij}^r \cap C_j)\{(x_{G_{ij}} - x_j)P_j^n + (y_{G_{ij}} - y_j)Q_j^n\} \\ + \frac{1}{3}\text{Area}(L_{ij}^{\ell} \cap C_j)\{(x_{G_{i,j+1}} - x_j)P_j^n + (y_{G_{i,j+1}} - y_j)Q_j^n\}\end{aligned} \quad (4.23)$$

where

$$\text{Area}(L_{ij}^r \cap C_j) = \frac{1}{2}\{(x_{M_{ij}} - x_j)(y_{G_{ij}} - y_j) - (x_{G_{ij}} - x_j)(y_{M_{ij}} - y_j)\}. \quad (4.24)$$

4.5. Approximation of $\int_{t^n}^{t^{n+1}} \int_{\partial L_{ij}} \{f(u)n_x + g(u)n_y\}d\sigma dt$.

This is achieved with the midpoint rule for integration with respect to time:

$$\begin{aligned}\int_{t^n}^{t^{n+1}} \int_{\partial L_{ij}} (f(u)n_x + g(u)n_y)d\sigma dt \\ \cong \Delta t \int_{\partial L_{ij}} \Big\{f\big(u(x,y,t^{n+1/2})\big)n_x + g\big(u(x,y,t^{n+1/2})\big)n_y\Big\}d\sigma\end{aligned} \quad (4.25)$$

where a first order Taylor expansion is used for $u(x,y,t^{n+1/2})$; using (4.1') we write

$$\begin{aligned}u(x,y,t^{n+1/2}) \cong u(x,y,t^n) - \frac{\Delta t}{2}\{f'\big(u(x,y,t^n)\big)u_x(x,y,t^n) \\ + g'\big(u(x,y,t^n)\big)u_y(x,y,t^n)\}\end{aligned} \quad (4.26)$$

On $L_{ij}\cap C_i$ we have chosen $u_x \equiv P_i^n$ and $u_y \equiv Q_i^n$, but we must find an approximate value of $u(x,y,t^n)$ on the line segments a_iG_{ij} and $a_iG_{i,j+1}$ (and similarly on a_jG_{ij} and $a_jG_{i,j+1}$). One possible choice consists in choosing the value of $L(x,y,t^n)$, our linear interpolant, at the midpoints of these segments; we then take, for any (x,y) on a_iG_{ij}:

$$u(x,y,t^n) \cong u_i^n + \frac{1}{2}(x_{G_{ij}} - x_i)P_i^n + \frac{1}{2}(y_{G_{ij}} - y_i)Q_i^n \equiv u_{a_i,G_{ij}}^n \quad (4.27)$$

thus defining our value $u_{a_i,G_{ij}}^n$ for the side a_iG_{ij} of L_{ij}.

In view of (4.26), we can now define an approximate average value of $u(x,y,t^{n+1/2})$ along the side a_iG_{ij} to be used in (4.25):

$$u_{a_iG_{ij}}^{n+1/2} = u_{a_i,G_{ij}}^n - \frac{\Delta t}{2}\{f'(u_{a_i,G_{ij}}^n)P_i^n + g'(u_{a_i,G_{ij}}^n)Q_i^n\}. \quad (4.28)$$

Introducing these values in (4.25), we finally get

$$\begin{aligned}
&\frac{1}{\Delta t}\int_{t^n}^{t^{n+1}}\int_{\partial L_{ij}}\{f(u)n_x+g(u)n_y\}d\sigma dt \qquad (4.29)\\
&\quad\cong f(u^{n+1/2}_{a_i,G_{ij}})n^1_{ijx}\cdot|a_iG_{ij}|+f(u^{n+1/2}_{a_i,G_{i,j+1}})n^4_{ijx}\cdot|a_iG_{i,j+1}|\\
&\qquad+f(u^{n+1/2}_{a_j,G_{ij}})n^2_{ijx}\cdot|a_jG_{ij}|+f(u^{n+1/2}_{a_j,G_{i,j+1}})n^3_{ijx}\cdot|a_jG_{i,j+1}|\\
&\qquad+g(u^{n+1/2}_{a_i,G_{ij}})n^1_{ijy}\cdot|a_iG_{ij}|+g(u^{n+1/2}_{a_i,G_{i,j+1}})n^4_{ijy}\cdot|a_iG_{i,j+1}|\\
&\qquad+g(u^{n+1/2}_{a_j,G_{ij}})n^2_{ijy}\cdot|a_jG_{ij}|+g(u^{n+1/2}_{a_j,G_{i,j+1}})n^3_{ijy}\cdot|a_jG_{i,j+1}|.
\end{aligned}$$

4.6. First step of the finite volume extension of the Nessyahu-Tadmor scheme.

Collecting our approximations (4.21),(4.23),(4.29) of the three terms appearing in the R.H.S. of (4.17) and taking (4.14) into account, we obtain the following approximation u^{n+1}_{ij} for the first (odd) time step of our scheme:

$$\text{Area}(L_{ij})u^{n+1}_{ij}=\text{R.H.S.}(4.21)+\text{R.H.S.}(4.23)-\Delta t\cdot\{\text{R.H.S.}(4.29)\} \qquad (4.30)$$

where u^{n+1}_{ij} can be considered as a cell value for cell L_{ij} at time t^{n+1}, or as a nodal value at the midpoint M_{ij}, at time t^{n+1}.

In preparation for the second (even) time step, we now construct a piecewise linear approximation of u on the quadrilaterals L_{ij}:

$$u(x,y,t^{n+1})\cong L^{(o)}(x,y,t^{n+1})\equiv u^{n+1}_{ij}+(x-x_{M_{ij}})P^{n+1}_{ij}+(y-y_{M_{ij}})Q^{n+1}_{ij} \qquad (4.31)$$

where the slopes P^{n+1}_{ij}, Q^{n+1}_{ij} can be computed as follows.

First we construct a piecewise linear approximate function defined on the triangles $T\in\mathcal{T}_h$ of the original triangulation. On triangle $T=a_ia_ja_k$, we can use for that purpose the newly obtained values u^{n+1}_{ij} at the midpoints of the sides of T. We then compute the average of the slopes of the linear interpolants in the two triangles $T,T'\in\mathcal{T}_h$ sharing a_ia_j as a common side (fig. 5.59), and use these averages in (4.31).

4.7. Second step of the finite volume Nessyahu-Tadmor scheme.

The second step is obtained by integrating (4.1') on the cylindric region $C_i\times[t^{n+1},t^{n+2}]$, using the same finite volume approach as for the first step, to define a cell average value u^{n+2}_i on cell C_i:

$$\text{Area}(C_i)u^{n+2}_i-\sum_{j\ \text{neighbour of}\ i}\int_{L_{ij}\cap C_i}u(x,y,t^{n+1})dxdy$$

$$= -\Delta t \sum_{j \text{ neighbour of } i} \int_{\Gamma_{ij}} \{f(u(x,y,t^{n+3/2}))\nu_x + g(u(x,y,t^{n+3/2}))\nu_y\} d\sigma \quad (4.32)$$

where $u(x,y,t^{n+1})$ is approximated, on $L_{ij} \cap C_i$, by the piecewise linear interpolant (4.31). Its integral on $L_{ij} \cap C_i$ is computed as described in section 4.3.

To obtain an approximate value of $u(x,y,t^{n+3/2})$ we use a Taylor expansion with respect to time combined with (4.1'), and we subdivide the cell-boundary element Γ_{ij} into $G_{ij}M_{ij} \cup M_{ij}G_{i,j+1}$. On $G_{ij}M_{ij}$ (resp. $M_{ij}G_{i,j+1}$), $u(x,y,t)$ is then approximated by its value of the midpoint of the line segment $G_{ij}M_{ij}$ (resp. $M_{ij}G_{i,j+1}$).

4.8. Approximation of the slopes.

In order to compute the gradient (P_i^n, Q_i^n) of the piecewise linear interpolant $L(x,y,t^n)$ for the cell C_i, we must first compute the gradient of the first degree polynomials P_T for all triangles $T \in \mathcal{T}_h$ such that $a_i \in T$. Although we could then in principle directly take the average of the gradients of the polynomials P_T, obtained as described at the beginning of section 4.2, we shall consider here a least-squares technique (cf. [40]). For simplicity, we shall describe it for the case of triangular (finite volume) cells.

Let T be a triangle with centroid G, and let T_j, $j = 1,2,3$ be the neighbouring triangles, with centroids $G_j (j = 1,\ldots,3)$ (fig.5.61); assume the values of the numerical approximation of the solution u at the four points $\{G, G_j\}_{j=1}^3$ are known at time t^n, equal to u_T^n, $u_{T_j}^n$ $(j = 1,\ldots,3)$ (these values can be considered as cell values playing for the triangular cells T, T_j the same role as u_i^n, u_{ij}^n for the cells C_i, L_{ij}).

The least-squares gradient $(\widetilde{grad\ u})_T^n = (\tilde{a}_T^n, \tilde{b}_T^n)$ for triangle T will then be chosen such as to minimize the functional

$$I = \sum_{j=1}^{3} \{u_T^n + \overset{\perp\rightarrow}{GG_j} \cdot (grad\ u)_T^n - u_{T_j}^n\}^2 \quad (4.33)$$

where

$$(grad\ u)_T^n = (a_T^n, b_T^n)$$

The minimum is obtained when

$$\frac{\partial I}{\partial a_T^n} = \frac{\partial I}{\partial b_T^n} = 0 \quad (4.34)$$

and is shown in [40] to lead to the following least-squares gradient:

$$\tilde{a}_T^n = \frac{1}{D} \sum_{j=1}^{3} (y_{G_j} - y_G)^2 \sum_{j=1}^{3} (u_{T_j}^n - u_T^n)(x_{G_j} - x_G)$$

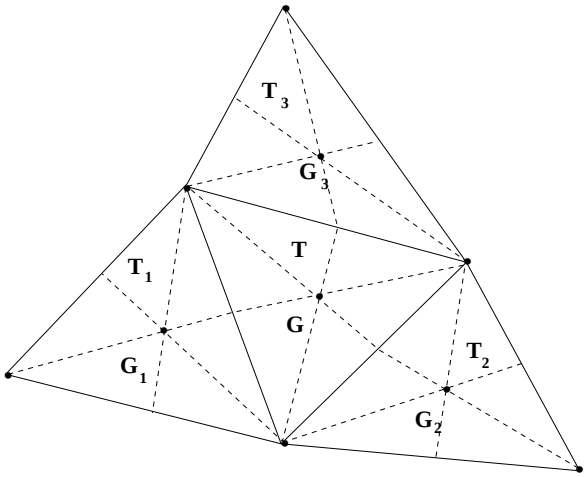

FIGURE 5.61. Computation of the least-squares gradient for a triangular cell $\mathbf{T}$.

$$- \frac{1}{D} \sum_{j=1}^{3} (x_{G_j} - x_G)(y_{G_j} - y_G) \sum_{j=1}^{3} (u^n_{T_j} - u^n_T)(y_{G_j} - y_G) \quad (4.35)$$

$$\tilde{b}^n_T = \frac{1}{D} \sum_{j=1}^{3} (x_{G_j} - x_G)^2 \sum_{j=1}^{3} (u^n_{T_j} - u^n_T)(y_{G_j} - y_G)$$
$$- \frac{1}{D} \sum_{j=1}^{3} (x_{G_j} - x_G)(y_{G_j} - y_G) \sum_{j=1}^{3} (u^n_{T_j} - u^n_T)(x_{G_j} - x_G) \quad (4.36)$$

where the denominator

$$D = \sum_{j=1}^{3} (x_{G_j} - x_G)^2 \sum_{j=1}^{3} (y_{G_j} - y_G)^2 - \left[\sum_{j=1}^{3} (x_{G_j} - x_G)(y_{G_j} - y_G)\right]^2 \quad (4.37)$$

is strictly positive for any non-degenerated triangle.

For the barycentric cells C_i or the quadrilateral cells L_{ij}, the procedure is quite similar to the one described above for triangular cells. Alternately, for a barycentric cell C_i with center a_i, we could first compute the least squares gradient grad $u|_{T_j} = (\tilde{a}^n_{T_j}, \tilde{b}^n_{T_j})$ of each neighbouring triangle T_j (such that $a_i \in T_j$), and then take the cell gradient grad $u|_{C_i} = \text{average}\{\text{grad } u|_{T_j}\}$, with a similar procedure for a quadrilateral cell L_{ij}.

Unfortunately, this procedure does not preserve monotonicity of the data in the usual van Leer sense described below, and allows the creation of local extremas between the nodes; this phenomenon may lead to (or amplify already existing) spurious oscillations, with the associated loss of stability and convergence difficulties in the case of

steady flows. We have therefore introduced some slope limitation in the computation of the gradients.

4.8.1. Slope limitation.

To ensure the stability of the scheme and prevent the generation of oscillations in regions of strong gradients, we must perform a slope correction. Following van Leer's approach ([26], [25]), in which the value at some interface point $x_{i+1/2}$ (in the one dimensional case) must fall within the range of values spanned by the adjacent grid averages, $u_{i\perp 1}$ and u_{i+1}, we limit the slopes of the linear interpolant L defined by (4.12) (resp. $L^{(o)}$ defined by (4.31) to ensure that its value at the boundary points $G_{ij}, M_{ij}, G_{i,j+1}$ of ∂C_i (resp. at the vertices $a_i, a_j, G_{ij}, G_{i,j+1}$ of ∂L_{ij}) are bounded by the values at the cell center u_i^n (resp. u_{ij}^{n+1}) and the value u_j^n at the corresponding neighbouring node a_j (resp $u_{i,j\perp 1}^{n+1}$ and $u_{i,j+1}^{n+1}$ at the adjacent quadrilateral cell"midpoints" $M_{i,j\perp 1}$ and $M_{i,j+1}$). The limitation procedure is implemented on each cell as follows. Let

$$(grad\ u)_i = (P_i, Q_i)^T$$

denote the gradient for cell i, where $P_i \approx u_x$, $Q_i \approx u_y$ at node i. If u satisfies the van Leer requirement we choose

$$P_i^{\lim} = \underset{\mathrm{j}\ \in\ \mathcal{N}(i)}{\operatorname{minmod}} \{P_j\}$$

$$= \begin{cases} \min_{\mathrm{j}\ \in\ \mathcal{N}(i)} |P_j| .(\text{common sign of all values } P_j) \\ \qquad\qquad \text{if all the values } P_j\ (\mathrm{j}\ \in\ \mathcal{N}(i)) \text{ have the same sign} \\ 0 \qquad\qquad \text{otherwise} \end{cases}$$

where $\mathcal{N}(i)$ is the set of nodes j adjacent to node i. If u does not satisfy the van Leer requirement, we set $P_{\mathrm{i}}^{\lim} = 0$. The computation of $\mathrm{Q}_{\mathrm{i}}^{\lim}$ is done in the same manner. For quadrilateral cells L_{ij} we proceed in a similar way.

5. Numerical Experiments

In order to assess the relative advantages of both methods , we applied them to several typical test problems; we present results for

- Euler flow around a Double-ellipse at supersonic regime.

- Supersonic flow past a blunt body at 0^o of angle of attack and $M_\infty = 2$

Both methods, which are second-order accurate and non-oscillatory thanks to the use of limiters, have been applied on the same grids, to compute the steady flows by stationarization, with the help of grid adaptation to improve the resolution. For the grid adaptation, we have used a procedure developed by M.J.Castro Diaz and F.Hecht at INRIA (France) ([37],[38]).For the finite volume method, the extension of section 4.2 to two-dimensional systems of conservation laws is achieved by the procedure described in ([46],[47],[48]).

On the whole, the finite volume version of the Nessyahu-Tadmor scheme has proved to be less time-consuming and significantly more accurate, for given grids. . The discontinuous finite element method required more computing time to reach convergence. Nevertheless, both methods seem to provide a reliable alternative to other well established schemes, and lend themselves to an extension to three-dimensional problems for unstructured grids.

Example 1 Supersonic flow past a double-ellipse at 20^o of angle of attack and $M_\infty = 2$

For this problem inspired by ([35]), but with Mach number $M_\infty = 2$ instead of the range of hypersonic Mach numbers considered there, and 20^o of angle of attack, the geometry is a double ellipse ([35]), defined by

$$\begin{cases} x \leq 0 & \begin{cases} z \leq 0 & \left(\frac{x}{0.06}\right)^2 + \left(\frac{z}{0.015}\right)^2 = 1 \\ z \geq 0 & \left(\frac{x}{0.035}\right)^2 + \left(\frac{z}{0.025}\right)^2 = 1 \end{cases} \\ 0 \leq x \leq 0.016 & \begin{cases} z \geq 0 & z = 0.025 \\ z \leq 0 & z = -0.015 \end{cases} \end{cases}$$

For this steady flow problem we used the same three grids with both methods.We present the results obtained with the initial and final grid only.

For the initial mesh (1558 vertices), both methods give fairly comparable results; notice that the C_p curves can be nearly superposed, which is an indication that both methods are indeed doing some reasonable calculation. The same is true for the pressure contours of both methods, with perhaps a very small advantage for our finite volume method (FV) which gives slightly sharper shocks.

The final grid (5055 vertices) shows a clear advantage for the FV method, which gives a nearly perfect shock resolution with very smooth contours, while the DFE method shows a serious breach of monotonicity in the lower part of the bow shock.

As was the case with the initial grid, the C_p curves can again be nearly exactly superposed.

The major difference between the two methods appears to lie in the convergence history and computing times. Fig. 5.64 shows a clear advantage for our finite volume method, for the initial grid (1558 vertices). Computing times for the initial grid (CPU : 3564 for FV and 48288 for DFE) confirm the advantage of the proposed Finite Volume Method.

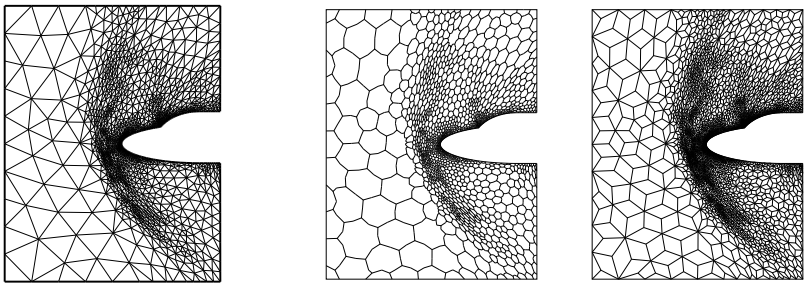

FIGURE 5.62. Euler flow around a double ellipse. Original grid, barycentric cells C_i and quadrilateral cells L_{ij}

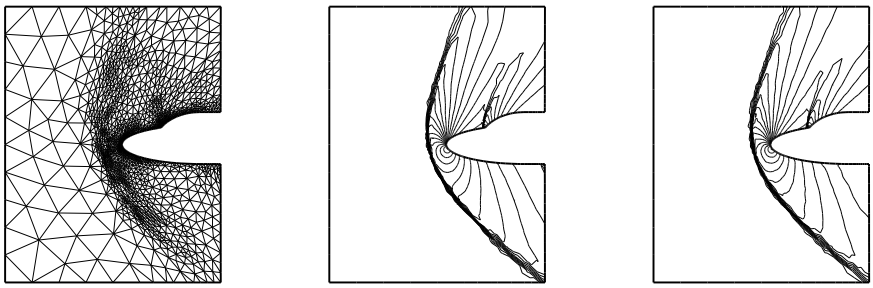

FIGURE 5.63. Double ellipse : Initial grid (1558 vertices) and solution (pressure contours) (FV, middle) and (DFE, right)

Example 2 Supersonic flow past a blunt body at 0^o of angle of attack. see ([4]).

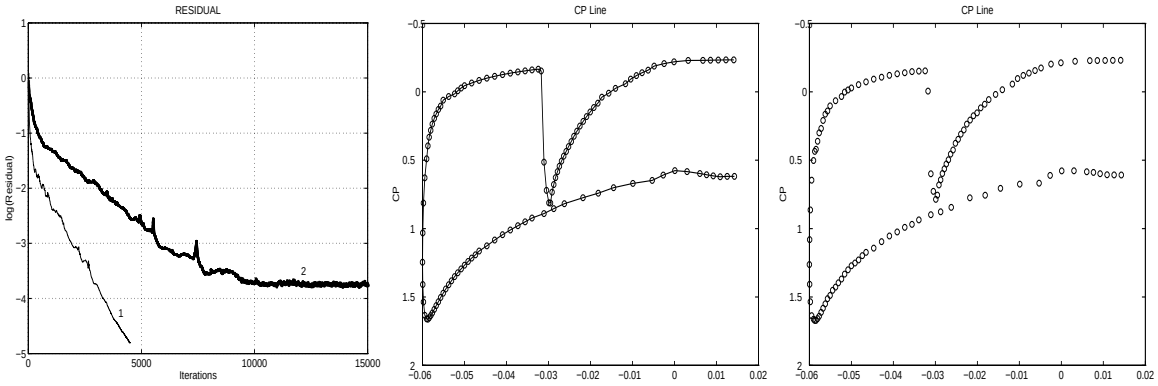

FIGURE 5.64. Residual for initial grid (1558 vertices)(1=FV,2=DFE) and C_p body cuts (FV, middle) and (DFE, right) (Double ellipse)

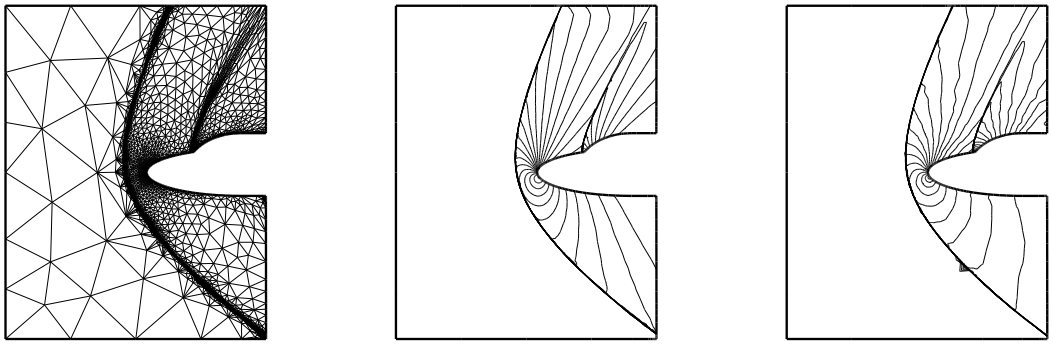

FIGURE 5.65. Double ellipse : Final grid (5055 vertices) and solution (pressure contours) (FV, middle) and (DFE, right)

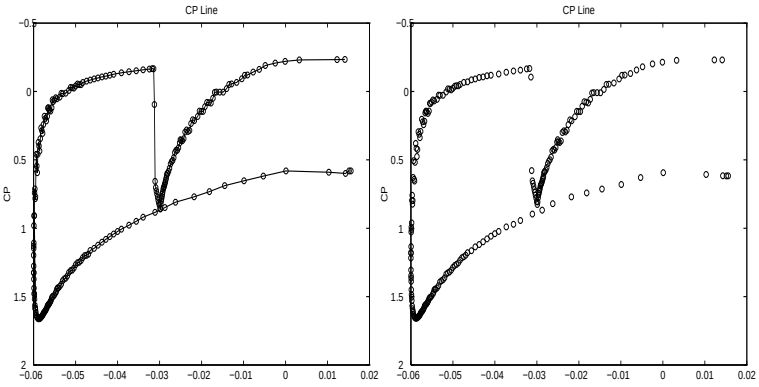

FIGURE 5.66. C_p body cuts for the final grid (5055 vertices) (FV, left) and (DFE) (Double ellipse)

This is a standard test problem which has been considered by many authors. The problem definition is borrowed from [4].The Mach number is $M_\infty = 2.0$ with an incidence $\alpha = 0^o$. The initial and final grid had 2737 and 7039 nodes, respectively. Fig.5.68 and 5.69 show the corresponding pressure contours. Our results, obtained with grid adaptation ([37],[38]), sustain comparison with other results for which grid adaptation has also been used.

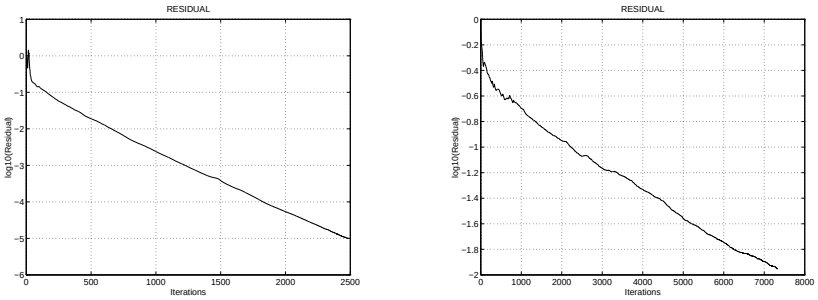

FIGURE 5.67. Convergence histories for Finite Volume scheme (left) and Discontinuous finite element method (right) (Blunt body)

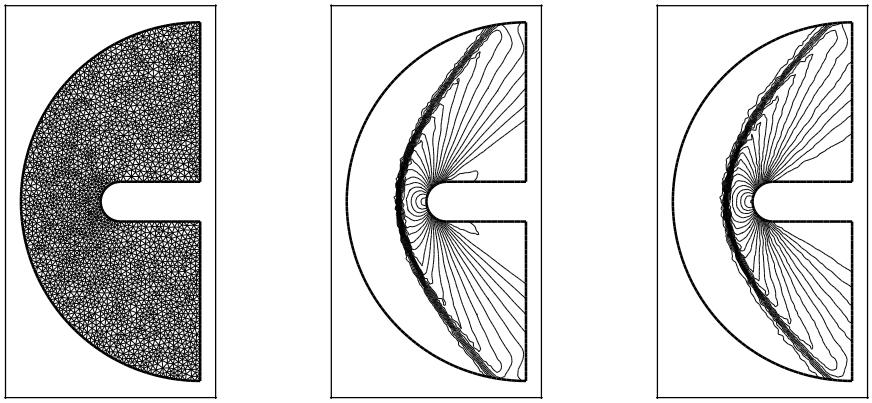

FIGURE 5.68. Initial grid and solution (pressure contours) FV (middle) and DFE (right)

Fig. 5.67 , which presents the residuals for both methods, shows a strong advantage in favour of our finite volume method, considering the results on fig.5.68 and 5.69, a fact which is confirmed by table 5.5: both the CPU-times and the speed of convergence are substantially better.

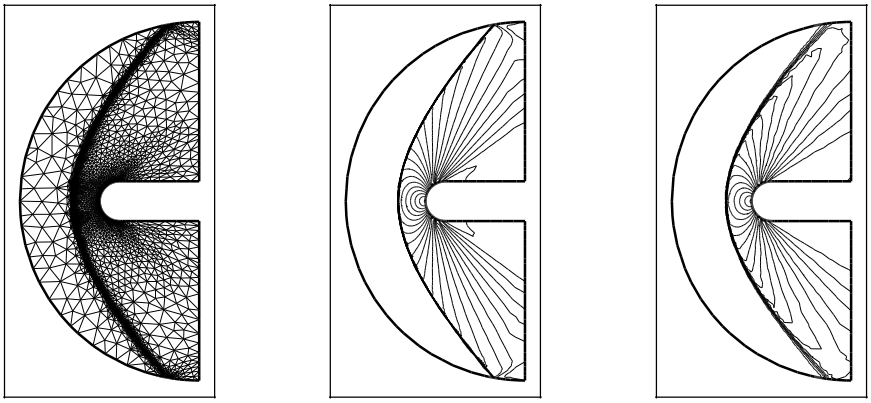

FIGURE 5.69. Final grid and solution (pressure contours) FV (middle) and DFE (right)

Method	CPU	nodes	elements	Solution
F.V	2701	2737	5244	without adaptation
D.F.E.M	11700	2737	5244	without adaptation
F.V	16905	7039	13801	with adaptation
D.F.E.M	46886	7039	13801	with adaptation

TABLE 5.5. Characteristics of F.V and D.F.E.M (Initial and Final Grids, Blunt Body)

This test case displays the robustness and the accuracy of the finite volume method, given the fact that the discontinuous finite element method is already an excellent and well established method.

Concluding remarks

The examples presented here suggest that the finite volume method proposed by the authors is capable of a very high resolution, with a clear advantage, as compared with the discontinuous finite element method studied here, at least at the level of the convergence histories, computing times, smoothness of the pressure contours, and monotonicity enforcement.

We are presently adapting the finite volume method to a mixed finite volume/Galerkin finite element method for the Navier-Stokes equations ([49]) where the convective part of the Navier-Stokes equations are treated with the finite volume method presented here while the diffusive term will be handled by a finite element method.

In another paper ([50],[51]) we consider the particular case of the linear equation $u_t + div(u\vec{V}) = 0$ where $\vec{V} = (V_1(x,y), V_2(x,y))^T$ with $div\,\vec{V} = 0$; we obtain an L^∞ bound for the numerical solution, from which we deduct the existence of a subsequence $\{u_{\mathcal{T}_k,\Delta t_k}\}$ which converges weakly to some function u in $L^\infty(\mathbb{R}^2 \times \mathbb{R}^+)$-weak. Using another set of inequalities, we obtain a total variation-type bound, slightly weaker than a classical bound on the total variation of the numerical solution, called a bound on the "weighted total variation", from which we are then able to show that the above weak limit function u is indeed a weak solution of $u_t + div(u\vec{V}) = 0$.

Références

[1] P. Arminjon et M.C. Viallon , Généralisation du schéma de Nessyahu-Tadmor pour une équation hyperbolique à deux dimensions d'espace, Comptes Rendus de l'Acad. des Sciences, Paris, t.320, série I , 85-88, January 1995.

[2] P. Arminjon, M.C. Viallon, A. Madrane , From Lax-Friedrichs to a Multidimensional Finite Volume Extension of the Nessyahu-Tadmor Scheme for Compressible Flows, Proc. Int. Conf. on Numerical Meth. for the Euler and Navier-Stokes Equations. Centre de Rech.Math. Univ de Montréal, P.Arminjon and A.Dervieux, editors, september 1995, to appear in the Am.Math.Soc-CRM series.

[3] P. Arminjon, M.C. Viallon and A. Madrane (1997), A Finite Volume Extension of the Lax-Friedrichs and Nessyahu-Tadmor Schemes for Conservation Laws on Unstructured Grids, Int.J. Comp.Fluid Dynamics, in press.

[4] P.Arminjon, D.Stanescu and M.C. Viallon , A two-dimensional finite volume extension of the Lax-Friedrichs and Nessyahu-Tadmor schemes for compressible flows, Proc. of the 6 th.Int.Symp. on Comp. Fluid Dynamics, Lake Tahoe (Nevada), September 4-8,1995, M. Hafez and K.Oshima, editors,vol.4.

[5] H. Nessyahu and E. Tadmor , Non-oscillatory Central Differencing for Hyperbolic Conservation Laws, J. Comp. Physics 87, No. 2, 408–463, 1990.

[6] G.D.V Gowda and J.Jaffré , A discontinuous finite element method for non linear scalar conservation laws, INRIA Res. Rep. No. 1848, Rocquencourt,78153 Le Chesnay, France, 1993.

[7] W. Reed and T. Hill, Triangular mesh methods for the neutron transport equation, Tech.Rep. LA-UR-73-479, Los Alamos Scientific Laboratory, 1973.

[8] P. Lesaint and P. Raviart, On a finite element method for solving the neutron transport equations, in Mathematical Aspects of Finite Elements in Partial Differential Equations, Academic Press, New York, 1974.

[9] P.Jamet and R.Bonnerot, Numerical solution of the Eulerian equations of compressible flow by a finite element method which follows the free boundary and the interfaces, J. Comput. Phys., 18, 21-45, 1975.

[10] J.T.Oden , Mathematical Aspect of Finite Element Approximations in Theoretical Mechanics, in Nemat-Nasser, Ed., Mechanics Today, Vol II, Pergamon Press, Oxford, pp. 159-250, 1975.

[11] T.J.R. Hughes et al., A new finite element formulation for computational fluid dynamics 1,2,3,4, Computer Methods in Applied Mechanics and Engineering, 54 (1986), pp. 223-234 and 341-355 and 58,(1986), pp.305-336.

[12] P.Arminjon and A.Rousseau, Discontinuous finite elements and Godunov-type methods for the Eulerian equations of gas dynamics. Comput.Meths.Appl.Mech.Engrg., 49, pp.17-36, 1985.

[13] P.Arminjon, Finite element analysis of moving loads on a floating ice sheet, in Topics in Numerical Analysis III, Proc.Royal Irish Acad. Conf. on Num. Anal., 1976, J.J.H.Miller, editor, Academic Press, 1977.

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

[15] 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 , pp. 93-118, 1984.

[16] 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, J.Periaux, A.Chaudouet, J.Wu, C.Marino and C.A. Bredia, Eds, Berlin, 1986.

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

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

[19] C. Johnson and J.Pitkaranta, An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation, Math. Comp., 46 , pp. 1-26, 1986.

[20] G. Richter, An optimal order error estimate for the discontinuous Galerkin method, Math. Comp., 50 , pp. 75-88, 1988.

[21] J.J.Quirk, A Contribution to the Great Riemann Solver Debate, NASA-ICASE Rep. No. 92-64, Nov. 1992, and Int.J.Num.Meth.Fluids,Vol.18, pp.555-574, 1994.

[22] P.Woodward and P.Colella, The Numerical Simulation of Two-Dimensional Fluid Flow with Strong Shocks, J.Comp.Phys.54,p.115-173, 1984.

[23] J. Jaffré and L. Kaddouri, Discontinuous finite elements for the Euler equations, Proc. Third Int. Conf. Hyperbolic Problems, Uppsala 1990, B.Engquist and B.Gustafsson, Eds. Studentlitteratur (Lund,Sweden)- Chartwell-Bratt (Bromley, England), Vol.III, pp.602-610.

[24] F. Dubois and G. Mehlman, A non-parameterized Entropy Correction for Roe's Approximate Riemann Solver, rapport CMAP-Ecole Polytechnique, n^o 248, 1991; see also Proc. 10th AIAA CFD Conf., Honolulu, June 1991, and AIAA Journal.

[25] G.D.V. Gowda, Discontinuous finite elements for nonlinear scalar conservation laws, Thèse de Doctorat, Université Paris IX-Dauphine, 1988.

[26] B.V. Leer, Towards the ultimate conservative difference scheme: IV. A new approach to numerical convection,J.Comput.Physics,23,276–299,1977.

[27] C. Hirsch, Numerical Computation of Internal and External Flows, Vol.2 : Computational Methods for Inviscid and Viscous Flows, Wiley Series in Numerical Methods in Engineering, J. Wiley & Sons, 1990.

[28] P.L. Roe , Approximate Riemann solvers, parameter vectors, and difference schemes, J. Comp. Phys. 43, 357–372, 1981.

[29] A. Harten,P.D. Lax and B. V. Leer, On Upstream Differencing and Godunov Type Schemes for Hyperbolic Conservation Laws, SIAM Review, 25(1), 35–61, 1983.

[30] E.Godlewski and P.A.Raviart , Hyperbolic Systems of Conservation Laws. Soc.Math. Appl.Ind., Ellipses, Paris, 1990.

[31] S. Osher and F. Solomon, Upwind difference schemes for hyperbolic systems of conservation laws, Math. Comp, 38, 339–374, 1982.

[32] A. Harten and J.M. Hyman, Self-Adjusting Grid Methods for One-Dimensional Hyperbolic Conservation Laws, J. Comput. Physics, 50, 235–269, 1983.

[33] Harten A., High resolution schemes for hyperbolic conservation laws, J. Comp Phys., 49, pp. 357-393, 1983.

[34] D.W.Levy, K.G.Powell and B.Van Leer, An implementation of a grid independent upwind scheme for the Euler equations, AIAA paper, 89-1931, 1989.

[35] INRIA and GAMNI-SMAI , Workshop on hypersonic flows for reentry problems, Problem 6 : Flow over a double ellipse, test case 6.1 : Non-Reactive Flows. Antibes, France, January 22-25, 1990.

[36] J.Peraire, M.Vahdati, K.Morgan and O.C.Zienkiewicz , Adaptative Remeshing for Compressible Flow Computations, J.Comp.Phys.72,449-466,1987.

[37] M.J.Castro Diaz and F.Hecht, Anisotropic Surface Mesh Generation, INRIA Res. Rep. No. 2672, October 1995, INRIA, Rocquencourt, 78153 Le Chesnay, France.

[38] M.J.Castro Diaz, Mesh Refinement over Triangulated Surfaces, INRIA Res. Rep. No. 2462, June 1994, INRIA, Rocquencourt, 78153 Le Chesnay, France.

[39] F. Fezoui , Résolution des équations d'Euler par un schéma de van Leer en élément finis. Res.Rep. INRIA No. 358, Rocquencourt, 78153 Le Chesnay, France, 1985.

[40] S. Champier , Convergence de schémas numériques type Volumes Finis pour la résolution d'équations hyperboliques. Thèse, Univ. de St-Etienne, 1992.

[41] B. van Leer , Towards the Ultimate Conservative Difference Scheme.II. Monotonicity and conservation combined in a second order scheme,J. Comp. Phys., Vol.14, 361–370, 1974.

[42] R.Eymard and T.Gallouët, Convergence d'un schéma de type éléments finis-volumes finis pour un système d'une équation elliptique et d'une équation hyperbolique, M^2AN , Vol.27 , No.7, 843-861, 1993.

[43] S. Osher, Riemann solvers, the entropy condition and difference approximations, SIAM J.Num.Analysis, 21, 217-235, 1984

[44] P.L.Roe, Some contributions to the modelling of discontinuous flows, in Lectures in Applied Mathematics, vol.22, pp.163-193, (B.Engquist, S.Osher and R.C.J.Somerville, Editors), SIAM, Philadelphia, 1985.

[45] P.L.Roe and J.Pike, Efficient construction and utilisation of approximate Riemann solutions, in Computing Meth. in Appl.Sciences and Engineering, VI (R.Glowinski and J.L.Lions, Editors) North-Holland (1984), pp. 499-518.

[46] F.Angrand, A. Dervieux, L. Loth and G. Vijayasundaram , Simulations of Euler transonic flows by means of explicit finite-element type schemes, INRIA Res. Rep. No. 250, Rocquencourt, 78153 Le Chesnay, France, 1983.

[47] F. Angrand and A. Dervieux , Some explicit triangular finite element schemes for the Euler equations, Int. J. Num. Meth. in Fluids, Vol. 4, 749–764, 1984.

[48] P. Arminjon and A. Dervieux , Construction of TVD-like Artificial Viscosities on Two-Dimensional Arbitrary FEM Grids, J. Comp. Phys., 106, No.1, 176–198, 1993.

[49] F.Beux, S.Lanteri, A.Dervieux and B.Larrouturou, Upwind stabilization of Navier-Stokes solvers, INRIA Res.Rep.No. 1885, Rocquencourt, 78153 Le Chesnay, France, 1993.

[50] M.C. Viallon et P. Arminjon , Convergence du schéma de Nessyahu-Tadmor sur des maillages non structurés pour une équation hyperbolique linéaire bidimensionnelle, Rapport de recherche No. C.N.R.S. U.R.A. 740, Équipe d'Analyse Numérique, Universités de Lyon et Saint-Étienne, septembre 1994.

[51] M.C. VIALLON AND P. ARMINJON, Convergence of a finite volume extension of the Nessyahu-Tadmor scheme on unstructured grids for a two-dimensional linear hyperbolic equation, Rapport de recherche, No. 2239, Centre de Recherches mathématiques, Université de Montréal, january 1995, to appear in SIAM J.Num.Anal.

6. APPENDIX

Solution of the projection problem for the slope limitation

The slope limitation process consists in solving a series of local minimization problems of dimension $nv(K) = 3$, with two constraints.

In this appendix we will show how these minimization problems can be solved; we follow the description given in [6] for the scalar case.

For $w = \rho,\ u,\ v,\ p$: we denote by

- P_K, the plane with equation

$$\sum_{i=1}^{3} X_i = 3\overline{w}_K^\star = w_{K,A_1} + w_{K,A_2} + w_{K,A_3}$$

- Q_K, the cube

$$\prod_{i=1}^{3} [wmin(A_i), wmax(A_i)] \quad \text{with} \quad \left\{ \begin{array}{rcl} wmin(A_i) & = & (1-\alpha)\overline{w}_K^\star + \alpha w_{min}(A_i) \\ wmax(A_i) & = & (1-\alpha)\overline{w}_K^\star + \alpha w_{max}(A_i) \end{array} \right.$$

where $w_{min}(A_i), w_{max}(A_i)$ are defined by (3.4).

For $0 \leq \alpha \leq 1$, $\overline{w}_K^\star$ satisfies the inequalities

$$wmin(A_i) \leq \overline{w}_K^\star \leq wmax(A_i)\ ,\ i = 1, \ldots, 3 \tag{6.1}$$

- $J(X) = \dfrac{1}{2}\|X - w_K^\star\|^2$ where $w_K^\star = (w_{K,A_i})_{i=1,\ldots,3}$
- $V = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ is a vector normal to P_K

With these notations, the computation of w_K^{n+1} amounts to the following minimization problem:

$$\left\{ \begin{array}{l} \text{Find } w_K^{n+1} \in P_K \cap Q_K \quad \text{such that} \\ \\ J(w_K^{n+1}) = \min\limits_{X \in P_K \cap Q_K} J(X) \end{array} \right. \tag{6.2}$$

This is a convex minimization problem which has a unique solution since $P_K \cap Q_K \neq \emptyset$ (the point $X = (X_i)_{i=1,\dots,3}$ with $X_i = \bar{w}_K^\star, i = 1, \dots, 3$, belongs to $P_K \cap Q_K$ by (6.1)) .

In order to solve the minimization problem (6.2), we dualize the constraint $X \in P_K$.

We introduce the Lagrangian L defined in $\mathbb{R}^3 \times \mathbb{R}$ by

$$L(X, \mu) = J(X) + \mu \left\{ \left(\sum_{i=1}^{3} X_i \right) - 3\overline{w}_K^\star \right\} \tag{6.3}$$

Problem (6.2) is then equivalent to the following associated saddle point problem:

$$\begin{cases} \text{Find the saddle point } (w_K^{n+1}, \lambda) \in Q_K \times \mathbb{R} \text{ such that} \\ \\ L(w_K^{n+1}, \lambda) = \min\limits_{X \in P_K} \max\limits_{\mu \in \mathbb{R}} L(X, \mu) = \max\limits_{\mu \in \mathbb{R}} \min\limits_{X \in P_K} L(X, \mu) \end{cases} \tag{6.4}$$

To solve problem (6.4), we first solve, for any given $\mu \in \mathbb{R}$, the minimization problem

$$\begin{cases} \text{Find } \widehat{X}(\mu) \in Q_K \quad \text{such that} \\ \\ L(\widehat{X}(\mu), \mu) = \min\limits_{X \in Q_K} L(X, \mu) \end{cases} \tag{6.5}$$

We observe that $w_K^\star = (w_{K,A_i})_{i=1,\dots,3} \in P_K$, so that $< V, w_K^\star > = 3\overline{w}_K^\star$.

The Lagrangian can therefore be written as

$$L(X, \mu) = \frac{1}{2} \{ \|X - w_K^\star\|^2 + 2\mu < V, X - w_K^\star > \}$$

or

$$L(X, \mu) = \frac{1}{2} \|X - (w_K^\star - \mu V)\|^2 - \frac{\mu^2}{2} \|V\|^2 \tag{6.6}$$

The minimization problem (6.5) is thus equivalent to minimizing the distance between $w_K^\star - \mu V$ and the cube Q_K; the solution $\widehat{X}(\mu)$ is therefore given by the projection of $w_K^\star - \mu V$ on Q_K : $\quad \widehat{X}(\mu) = \text{Proj}_{(Q_K)} (w_K^\star - \mu V)$.

$\widehat{X}(\mu)$ can be obtained by truncation of the components of $w_K^\star - \mu V$:

$$\begin{cases} \text{if} \quad w_{K,A_i}^\star - \mu \in [wmin(A_i), wmax(A_i)] \quad \text{then} \quad \left(\widehat{X}(\mu) \right)_i = w_{K,A_i}^\star - \mu \\ \\ \text{if} \quad w_{K,A_i}^\star - \mu > wmax(A_i) \quad \text{then} \quad \left(\widehat{X}(\mu) \right)_i = wmax(A_i) \\ \\ \text{if} \quad w_{K,A_i}^\star - \mu < wmin(A_i) \quad \text{then} \quad \left(\widehat{X}(\mu) \right)_i = wmin(A_i) \end{cases} \tag{6.7}$$

To complete the solution of problem (6.4), we must now find the real number λ which maximizes the real function

$$\mu \longmapsto F(\mu) = L\left(\widehat{X}(\mu), \mu\right)$$

The desired limited vector $w_K^{n+1} = (w_{K,A_i}^{n+1})_{i=1,\dots,3}$ will then be given by

$$w_K^{n+1} = \widehat{X}(\lambda) \quad \text{where} \quad \lambda \quad \text{yields} \quad \max_{\mu \in \mathbb{R}} F(\mu)$$

To find the maximum of F, let us write its first and second derivatives:

$$F'(\mu) =< V, \widehat{X}(\mu) - w_K^\star >= \sum_{i=1}^{3} \left(\widehat{X_i}(\mu) - w_{K,A_i}^\star\right)$$

$$F''(\mu) = \sum_{i=1}^{3} \widehat{X_i}'(\mu)$$

where $\widehat{X}(\mu)$ is given by (6.7), so that

$$\begin{cases} \text{if} \quad w_{K,A_i}^\star - \mu \in [wmin(A_i), wmax(A_i)] \,, \quad \text{then} \quad \widehat{X}'(\mu) = -1 \\ \text{otherwise,} \quad \widehat{X}'(\mu) = 0 \,. \end{cases}$$

We therefore get

$$F''(\mu) = -\ \text{card}\ \left\{ i \in \{0, 1, \dots, 3\} / w_{K,A_i}^\star - \mu \in [wmin(A_i), wmax(A_i)] \right\} \leq 0 \,.$$

F is thus a continuous and differentiable, piecewise quadratic concave function.

F'' has $2 \times 3 = 6$ points of discontinuity which arise when $\widehat{X}(\mu)$ crosses a face of the cube Q_K.

The function F is therefore unambiguously defined and readily available, so that one can easily find the value λ for which is attains its maximum.

Fig.5.70 shows how the duality method works in the case of the slope limitation of a variable w in the one-dimensional case ($nv(K) = 2$).

Conclusion

The limitation process, for the slopes of the physical variables, reduces to the maximization, for each element K of the grid, of a piecewise quadratic function of one real variable, which is differentiable and easy to calculate; as the limitation is only active in the neighborhood of strong gradients or fronts of the vector of physical variables, it turns out to be a simple, relatively inexpensive part of the Jaffré-Kaddouri-Gowda discontinuous finite element method.

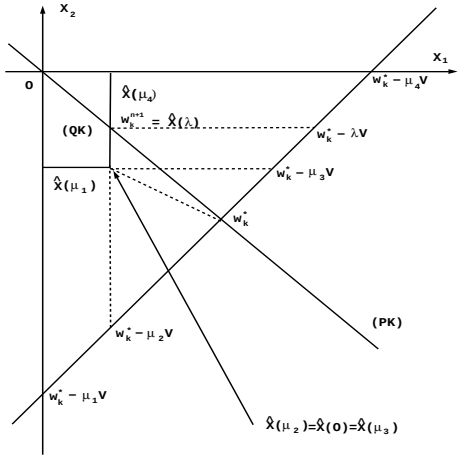

FIGURE 5.70. Determination of w_K^{n+1} by the duality method in the one-dimensional case $(nv(K) = 2)$