Construction of second order accurate monotone and stable residual distributive schemes : the unsteady case.

Rémi Abgrall [1] and Mohamed Mezine
Mathématiques Appliquées de Bordeaux,
Université Bordeaux I,
351 cours de la Libération, 33 405 Talence Cedex
abgrall,mezine@math.u-bordeaux.fr

Keywords : Compressible flow solvers, Residual schemes, fluctuation splitting schemes, unstructured meshes, multidimensional up-winding.
AMS Subject Classification : 65C99, 65M60, 76N10 65C99, 65M60, 76N10

[1] institut Universitaire de France

Proposed running head : A stable, monotone, accurate residual distributive scheme for unsteady problems.

Rémi Abgrall
Mathématiques Appliquées de Bordeaux
351 Cours de la Libération
33 405 Talence Cedex
France
email : abgrall@math.u-bordeaux.fr
fax : +33 5 56 84 26 26

Send proof to Rémi Abgrall

Construction of second order accurate monotone and stable residual distributive schemes : the unsteady case.

R. Abgrall[†]and M. Mezine
Mathématiques Appliquées de Bordeaux
Université Bordeaux 1, 351 Cours de la libération,
33405 Talence

May 3, 2002

Abstract

The aim of this paper is to construct upwind residual distribution schemes for the time accurate solution of hyperbolic conservation laws. To do so, we evaluate a space-time fluctuation based on a space-time approximation of the solution, and develop new residual distribution schemes which are extensions of classical steady upwind residual distribution schemes. This method has been applied to the solution of scalar advection equation and to the solution of the compressible Euler equations both in two space dimensions. The scheme is unconditionally stable and oscillation free. Several numerical examples are shown to demonstrate the stability and accuracy of the method.

Key Words - residual schemes; fluctuation splitting schemes; Euler equations; unsteady flows; unstructured meshes.

AMS Subject Classification - 65C99, 65M60, 76N10

[†]Institut Universitaire de France

Contents

List of Figures

1 Introduction

This paper is devoted to the construction of a class of stable, accurate and *compact* methods for the solution of unsteady compressible flows on triangular unstructured meshes.

In the past decades, many numerical schemes have been devoted to the simulation of unsteady compressible flows with possibly discontinuities within the flow. In most cases, this is done with methods that are extension of the so–called high order upwind schemes (see [18, 19] for a review), even in unstructured meshes [16]. There is at least two important exceptions to this, the finite element type schemes constructed via the stream–line diffusion method, [26, 22], and the Discontinuous Galerkin schemes [8].

The schemes constructed following the upwind high order philosophy are, for unstructured meshes, most of the time quite disappointing because they suffer from excessive numerical diffusion. To overcome this problem, one has try to adapt ideas forming from the ENO/WENO [20, 21] philosophy. Example of such attempts are [1, 17, 21]. If these extensions had some success, the price to pay is very high : the coding is rather complex, and more disturbing, these schemes do not have a compact stencil, so parallel implementation becomes difficult. Last the analysis of the accuracy of the solution is rather unclear, to the least difficult to carry out.

In order tackle these two problems – compacity of the stencil, effective accuracy of the solution–, at least two classes of methods have emerged in the recent years, the Discontinuous Galerkin schemes (DG), and the Residual Distributive schemes (RD) or Fluctuation Splitting schemes. Though different in spirit, they have a common core at least in their "unstabilised" versions, the residual property. This property that we discuss below in the framework of RD schemes, allows to show the formal accuracy of the scheme on a very general mesh. The DG schemes use a discontinuous polynomial representation of the unknowns that is a generalization of what is done in finite volume schemes. The solution is updated via evaluation of fluxes, and the stabilization mechanism is obtained by very similar techniques as in classical finite volume. The net effect of this is to lose the residual property.

On the contrary, the RD schemes use a pointwise representation of the solution, like in finite difference schemes. The unknowns are updated by evaluation of the amount of residual sent to the vertices, and the stabilization mechanism can be similar to artificial viscosity, as in the SUPG finite element method [23] or the Stream line diffusion method [26], or inspired by the nonlinear techniques of the so–called high resolution schemes [24, 19]. One can also take into account the genuinely multidimensional structure of the problem [34].

In this paper, we consider schemes of the RD class, borrowing ideas coming from the upwind high order philosophy, as well as ideas coming from the stream line diffusion one. Our goal is to propose a systematic construction of robust and high order schemes, on general meshes. This problem has received recently a lot of interest, one may quote the pioneering work of Roe, Sidilkover, Deconinck and their coworkers [13, 33, 34], and also more recently in [35, 2, 3, 12], but only for *steady* problems. Here we consider unsteady problems.

The paper is organized as follows. First, we recall some basic fact about upwind residual distributive schemes. Then we construct, for scalar problems, a class of upwind, monotonocity preserving and second order schemes. Several numerical examples are considered to show the effectiveness of our approach. Then, we extend this method, following the ideas of [5], first to symetrizable systems, then to the Euler equations. The main drawback of this method is that we are constraint to a CFL like condition to preserve the monotonicity property of the scheme (not for stability purposes). Following an idea of Deconinck et al. [10], we then construct an unconditionally, monotonicity preserving, second order accurate scheme for unsteady flow problems, several numerical example illustrate our technique. Some conclusions end the paper.

2 Review of residual distribution schemes

2.1 Scalar conservation law

Let us consider the two-dimensional conservation law

$$\frac{\partial u}{\partial t} + \text{ div } f(u) = 0 \text{ in } \Omega \times (0,T),\ \Omega \subset \mathbb{R}^2 \tag{2.1}$$

or in quasi linear form

$$\frac{\partial u}{\partial t} + \langle \lambda(u), \nabla u \rangle = 0, \tag{2.2}$$

where $\lambda(u) = \partial f / \partial u$ is the advection speed. Let τ_h a triangulation of the domain Ω. The set of triangles is $\{T_j\}_{j=1,\dots,nt}$, the mesh points are $\{M_i\}_{i=1,\dots,ns}$. The vertices of a triangle T are M_{i_1}, M_{i_2}, M_{i_3}, or i_1, i_2, i_3. When there is no ambiguity i_1, i_2, i_3 are replaced by 1, 2, 3. The discretization techniques use $\mathbb{P}_1$ finite element meshes, i.e. triangular meshes with a piece-wise linear solution representation $u^h = \sum_i u_i \Lambda_i$, where Λ_i is the linear shape function associated to node i.

We define the fluctuation or cell residual over a triangle T as the integral over this element of the spatial differential operator :

$$\Phi^T = \int_T \text{div } f^h \, dx = |T| \, \overrightarrow{\lambda}.\nabla u^h \tag{2.3}$$

since ∇u^h is constant, $\overrightarrow{\lambda}$ is a cell-averaged advection speed. If we take

$$\overrightarrow{\lambda} = \frac{1}{|T|} \int_T \lambda(u^h) \, dx, \tag{2.4}$$

it makes an exact equivalence between the conservative form and the quasi linear form of the cell residual, and allows us to work with the quasi-linear expression while maintaining local conservation. Expression (2.4) is called a conservative Roe-linearisation. One obtains fluctuation

$$\Phi^T = \sum_{M_i \in T} k_i u_i, \tag{2.5}$$

where $k_i = \frac{1}{2} \overrightarrow{\lambda}.n_i^T$ is called the inflow parameter, n_i^T are the scaled inward normals of the edge oppposite to node i (see Figure 1) .

The Residual Distribution method consists in distributing fractions of the cell residual Φ^T to the vertices of the element. One notes by Φ_i^T the fraction of the cell residual sent to node i. For consistency and conservativity one has :

$$\sum_{M_i \in T} \Phi_i^T = \Phi^T. \tag{2.6}$$

A residual distribution scheme reads :

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{|C_i|} \sum_{T,\, M_i \in T} \phi_i^T, \tag{2.7}$$

where $|C_i|$ is the area of the median dual cell C_i (see Figure 2).

Since u^h is continuous across edges, it can be shown [4] that, under the assumption of the standard Lax Wendroff theorem, the limit solution is a solution of (2.1) in the linear and non linear case. This lets a lot of flexibility, because we are not any more constraint by the geometry of the mesh, hence a few design principles are needed. The are (i) Upwinding; (ii) Linearity Preservation; (iii) Monotonicity.

(i) **The upwinding property (U)**

No fraction of the element residual is sent to upstream nodes

$$\phi_i^T = 0 \text{ when } k_i \leq 0. \tag{2.8}$$

(ii) **Second order accuracy at steady state : the Linear Preserving condition (LP)**

A way to get second order accuracy at steady state (see [2]) is to require that the residual Φ_i^T satisfies the property

$$\phi_i^T(u^h) = \mathcal{O}(h^3), \tag{2.9}$$

for any smooth solution of (2.1), u^h the piecewise linear interpolation of this solution. If Φ_i^T/Φ^T remains bounded when $\Phi^T \to 0$ we get second order accuracy. It follows from the accuracy properties of the piecewise linear interpolation. It is easy to see that the cell residual satisfies $\Phi^T = \mathcal{O}(h^3)$, so it implies (2.9).

(iii) **The monotonicity condition (P)**

The scheme does not create local extrema or, if the residual sent at node i can be written as

$$\Phi_i^T = \sum_j c_{ij}(u_i - u_j), \tag{2.10}$$

we impose $c_{ij} \geq 0$. The scheme is L^∞ stable under a CFL condition.

2.2 Examples of schemes for scalar problems

Several schemes satisfying some or all of these design principles have been developed both in 2D and 3D. For a complete review see [30]. Some of the most popular schemes are given in Table 1: the N-scheme is a linear monotone (P) first order upwind scheme ; the LDA scheme is a second order (LP) linear upwind scheme ; the PSI scheme is a nonlinear second order (LP) and monotone (P) upwind scheme.

Notes that a Godunov's theorem states that linear schemes (i.e. schemes for which the distribution does not depend on the solution itself) cannot be (LP) and (P) at the same time. One can get both properties only if some non linearity is introduced into the scheme. An example is the PSI–blending below.

One notes $k_i^+ = \max(0, k_i)$, $k_i^- = \min(0, k_i)$.

Scheme	residual at node i	type	U	P	LP
N	$\Phi_i^N = k_i^+(u_i - \widetilde{u})$	linear	√	√	
LDA	$\Phi_i^{LDA} = -k_i^+ N \Phi^T$	linear	√		√
PSI	$\Phi_i^{PSI} = l\Phi_i^N + (1-l)\Phi_i^{LDA}$	nonlinear	√	√	√

Table 1: Residual distribution schemes.

The residual at node i for the N-scheme must satisfy the conservation relation (2.6), so we must have

$$\sum_{M_i \in T} k_i^- \widetilde{u} = \sum_{M_i \in T} k_i^- u_i, \tag{2.11}$$

one gets for $\widetilde{u}$

$$\widetilde{u} = N \sum_{M_i \in T} k_i^- u_i, \tag{2.12}$$

with $N = (\sum_j k_j)^{-1}$.

The PSI of Struijs [34] scheme can be obtained as a combination of the first order monotone N- scheme and the second order non-monotone LDA scheme, where $l \in \mathbb{R}$ is chosen to make the scheme monotone (P) and second order accurate at steady state (LP). It is shown in [2] that if we take $l = \max(\varphi(r_1), \varphi(r_2), \varphi(r_3))$ where

$$r_i = \frac{\Phi_i^{LDA}}{\Phi_i^N}, \qquad \varphi(x) = \begin{cases} x/(1-x) \text{ if } x < 0 \\ 0 \text{ else} \end{cases}, \tag{2.13}$$

the scheme satisfy the above properties.

A recent development is the proposition of modified schemes resulting from first order scheme like the N-scheme, see [5] and then [6]. The problem is formulated as follows on a general N–vertex simplex : from a low order monotone scheme $\{\Phi_i\}$, construct a high order scheme $\{\Phi_i^*\}$ such that

$$\begin{array}{ll} \sum_j \Phi_j^* = \sum_j \Phi_j = \Phi & \text{Conservation} \\ \Phi_j^*\Phi_j \geq 0 & \text{Monotonicity} \end{array}$$

We get second order accuracy under the following conditions

$$\begin{array}{ll} \Phi_j^* = \mathcal{O}(\Phi) & \\ \text{and} & \\ \Phi = \mathcal{O}(h^3) & \text{at steady state for smooth solutions.} \end{array}$$

The two conditions are important ; in particular they impose second order accuracy constraints on the flux approximation $f^h(u^h)$.

After introducing

$$\beta_j = \frac{\Phi_j}{\Phi}, \qquad \hat{\beta}_j = \frac{\Phi_j^*}{\Phi}$$

these conditions can be reformulated as constraints on the weights with which the residuals are distributed to the nodes

$$\begin{array}{ll} \sum_{j=1}^N \beta_j = \sum_{j=1}^N \hat{\beta}_j = 1 & \text{Conservation} \\ \beta_j\hat{\beta}_j \geq 0 & \text{Monotonicity} \\ \hat{\beta}_j \text{ is bounded} & \text{High-order accuracy if } \Phi^T \text{ is small enough} \end{array}$$

Our procedure is to impose these constraints on a mapping, which takes a set of weights $\{\beta_j\}$ corresponding to a monotone scheme (which may only be first-order accurate) to a set of weights $\{\hat{\beta}_j\}$ corresponding to a scheme that is both monotone and of maximum accuracy. In reformulating, we allow for an arbitrary number N of degrees of freedom per simplex. The mapping from the N-vector $\{\beta_j\}$ to the N-vector$\{\hat{\beta}_j\}$ cannot be linear, because of the Godunov Theorem, but there are many such nonlinear mappings, They are the truly multidimensional analogues of limiter functions [19] and we expect their thorough investigation to take considerable time.

Here, we simply offer two examples for the case $N = 3$, presenting the vectors $\beta = (\beta_1, \beta_2, \beta_3)$ and $\hat{\beta} = (\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)$ as the barycentric coordinates of a point in space with respect to an equilateral triangle. To ensure boundedness, we insist that, for all j, $0 \leq \hat{\beta}_j \leq 1$, so that the point β lies within the triangle, or on its boundary. A weaker condition, constraining the point β to a finite neighborhood of the triangle, seems possible but has not yet been explored.

If the monotone weights are all positive, then β already lies within the triangle, and it is natural to take simply $\hat{\beta} = \beta$. If β lies outside the triangle, one possibility is simply to project β onto the boundary of the triangle. For example, we may take

$$\hat{\beta}_j = \frac{\beta_j^+}{\sum_j \beta_j^+} \tag{2.14}$$

In Figure 3 this is shown geometrically on the left. The mapping from β to $\hat{\beta}$ is always a translation toward one of the vertices.

An alternative is to take $\hat{\beta}$ as the point on the boundary of the triangle that is closest to β. The following logic accomplishes that

```
IF (B1.GE.0.) AND (B2.GE.0.) AND (B3.GE.0.) GOTO END
BS1 = B1+B2/2, BS2 = B2+B3/2, BS3 = B3+B1/2
```

```
B1 = MAX(0.,MIN(BS1,1.-BS2,1.))
B2 = MAX(0.,MIN(BS2,1.-BS3,1.))
B3 = MAX(0.,MIN(BS3,1.-BS1,1.))
END
```

Since $\hat{\beta}_j$ is homogeneous of degree 1 in x_j, this can be coded without division. In Figure 3 this map is shown geometrically on the right. Extension to $N \geq 4$ is straightforward for both maps.

2.3 Compressible Euler Equations

Let us consider the compressible Euler equations in the domain $\Omega \subset \mathbb{R}^2$,

$$\frac{\partial W}{\partial t} + \operatorname{div} \mathcal{F}(W) = 0 \text{ in } \Omega \times (0,T), \tag{2.15}$$

or in quasi-linear form

$$\frac{\partial W}{\partial t} + A(W)\frac{\partial W}{\partial x} + B(W)\frac{\partial W}{\partial y} = 0 \text{ in } \Omega \times (0,T). \tag{2.16}$$

where $A(W)$, $B(W)$ are the Jacobians of the fluxes. The flux $\mathcal{F} = (F, G)$ and the conserved variables are given by

$$\begin{aligned} &W = (\rho, \rho u, \rho v, E)^T, \; F(W) = (\rho u, \rho u^2 + p, \rho uv, u(E+p))^T, \\ &G(W) = (\rho v, \rho uv, \rho v^2 + p, v(E+p))^T \end{aligned} \tag{2.17}$$

where ρ is the density, u and v the components of the velocity, ϵ is the internal energy, and $E = \rho\epsilon + \frac{1}{2}\rho(u^2+v^2)$ is the total energy. The equation of state is the perfect gas law

$$p = (\gamma - 1)\rho\epsilon = (\gamma - 1)\left(E - \frac{1}{2}\rho(u^2+v^2)\right),$$

where γ is the ratio of specific heats ($\gamma = 1.4$).

The construction of the residual distribution schemes is based on the quasi-linear form of the conservation law. For reasons of conservation, a particular conservative linearisation is needed to generalize (2.3) and (2.4). Then the non linear system (2.15) is locally approximated by the linear system

$$\frac{\partial W}{\partial t} + \bar{A}\frac{\partial W}{\partial x} + \bar{B}\frac{\partial W}{\partial y} = 0, \tag{2.18}$$

where $\bar{A}$ and $\bar{B}$ are the Jacobians of the fluxes evaluated at some average state $\bar{W}$. The conservative linearisation is obtained by a multidimensional extension of Roe's linearisation [14, 11]. Introducing Roe's parameter vector $Z = \sqrt{\rho}(1, u, v, H)$, where H is the total enthalpy, one notes that $W = W(Z)$ and $\mathcal{F} = \mathcal{F}(Z)$ are quadratic in Z,

$$W(Z) = \frac{1}{2}\, D(Z)Z \qquad \mathcal{F}(Z) = \frac{1}{2}\mathcal{R}(Z)Z, \tag{2.19}$$

where $D(Z)$ and $\mathcal{R}(Z)$ are matrices that depend linearly on Z. Assuming that Z has a linear variation on T, we get for the cell residual

$$\Phi^T = \sum_{M_i \in T} K_i \widetilde{W}_i, \tag{2.20}$$

where $K_i = \frac{1}{2}(\bar{A}n_i^x + \bar{B}n_i^y)$ with

$$\bar{A} = \frac{\partial F}{\partial W}(\bar{Z}), \quad \bar{B} = \frac{\partial G}{\partial W}(\bar{Z}), \quad \bar{Z} = \frac{Z_1 + Z_2 + Z_3}{3},$$

the vector $\widetilde{W}_i$ is defined as $\widetilde{W}_i = D(\bar{Z})Z_i$.

Remark 1. $(\bar{A},\bar{B}) = \mathcal{R}(\bar{Z})D^{-1}(\bar{Z})$, *the Jacobians of the fluxes evaluated at some average state* $\bar{W}$. *It can be shown that the matrices* K_i *are diagonalizable with real eigenvalues. We set* $K_i = L_i\Lambda_i R_i$ *where* L_i *contains the left eigenvectors and* R_i *the right eigenvectors, and* Λ_i *is the diagonal matrix containing the eigenvalues.*

The positive and negative parts of K_i are denoted K_i^+ and K_i^-, they are defined by

$$K_i^+ = L_i\Lambda_i^+ R_i \qquad K_i^- = L_i\Lambda_i^- R_i \tag{2.21}$$

where Λ_i^+ and Λ_i^- contains the positive and the negative eigenvalues of K_i.

The design principles for systems residual distribution schemes are similar to those defined in the previous section for scalar conservation laws, i.e.

(i) **The upwinding property (U)**

The scheme is upwind if the following condition is true : if all the eigenvalues of K_i are negative, then $\Phi_i^T = 0$.

(ii) **Second order accuracy at steady state : the Linear Preserving condition (LP)**

This condition is the same as for scalar conservation law (see [2]). A condition is to require that the residual Φ_i^T satisfies the property

$$\phi_i^T(W^h) = \mathcal{O}(h^3) \tag{2.22}$$

for any smooth solution of (2.15), where W^h is the piecewise linear interpolation of this solution.

(iii) **The monotonicity condition**

One wants to avoid nonphysical oscillations. The formulation of this intuitive condition is difficult, however see [2] and sections 2.4 and 4.2.4 below for details.

2.4 Examples of schemes for systems

- ***The system N-scheme***

We set

$$\Phi_i^N = \sum_j K_i^+ N K_j^-(\widetilde{W}_i - \widetilde{W}_j), \tag{2.23}$$

where $N = (\sum\limits_j K_j^-)^{-1}$. It is shown in [2] that the matrix product K_i^+N has always a meaning, so the scheme is well defined. In [3], it is shown that for a symetrizable system the N scheme is linearly dissipative. In the present case, if the linearization is carried out in the entropy variables $V = \nabla_W(\rho s)$, one has

$$\sum_{M_i\in T} \langle V_i, \Phi_i^N\rangle = \frac{1}{2}\sum_{M_i\in T}\langle V_i, K_iV_i\rangle + \mathcal{Q}^N(V_1,V_2,V_3), \tag{2.24}$$

with $\mathcal{Q}^N(V_1,V_2,V_3)$ a positive quadratic form.

- ***The system LDA scheme***

The system LDA scheme is defined by

$$\Phi_i^{LDA} = -K_i^+ N\Phi^T. \tag{2.25}$$

Like the N-scheme, the system LDA scheme is always well defined. For symetrizable system no dissipation property is known.

• ***The system N-modified scheme***

The generalization of the scalar PSI scheme is not obvious. We recall the extension constructed in [2] because we use the same type of idea here.

Consider a direction $m = (m_x, m_y)$, and the matrix $K_m = m_x\bar{A}+m_y\bar{B}$. Since this matrix is diagonalisable in $\mathbb{R}$, we may consider $\ell_j, j = 1,4$ an $\mathbf{r}_j, j = 1,4$ its left and right eigenvectors. Next we consider the fluctuations $\varphi_i^j = \ell_j \cdot \Phi_i$, i.e. we decompose Φ_j on the basis $\{\mathbf{r}_j\}$,

$$\Phi_j = \sum_{i=1}^{4} \varphi_i^j \mathbf{r}_i.$$

It is clear that

$$\ell_j \cdot \Phi = \sum_{i=1}^{4} \varphi_i^j,$$

so we may apply the procedure defined in equations (2.14), or the limiter defined by orthogonal projection. This produces a set of fluctuations $\{(\varphi_i^j)^*\}$, and we define the limited scheme by

$$\Phi_i^* = \sum_{j=1}^{4} \left(\varphi_i^j\right)^{\mathbf{r}}_j.$$

In [5], it is shown that this scheme is LP. It is also shown that if one starts from the system N scheme, then this scheme has a L^∞ type stability property. This property has been demonstrated on several test case, ranging from subsonic to hypersonic cases.

3 Petrov-Galerkin formulation of Residual Distribution approach

The Residual Distribution schemes can be viewed as a particular class of Petrov-Galerkin finite elements schemes. Denoting by w_i the Petrov-Galerkin weighting function associated to node i : $w_i = \lambda_i + \alpha_i^T$ on T, one can choose α_i^T such that the spatial discretisation of the Petrov-Galerkin formulation is the same of the Residual Distribution approach. Starting from the finite element spatial discretization of (2.2) with constant advection speed one has

$$\sum_{T\in\tau_h} \int_T w_i \frac{\partial u^h}{\partial t}\, dx + \sum_{T\in\tau_h} \int_T w_i \langle \lambda, \nabla u^h\rangle \, dx = 0. \tag{3.1}$$

Since $\langle \lambda, \nabla u^h\rangle$ is constant on T and equal to $\Phi^T/|T|$ one obtains

$$\sum_{T\in\tau_h} \int_T w_i \frac{\partial u^h}{\partial t}\, dx + \sum_{T, M_i\in T} \int_T w_i \, dx \, \frac{\Phi^T}{|T|} = 0. \tag{3.2}$$

Using the definition of w_i in the second integral in equation (3.2) leads to

$$\sum_{T\in\tau_h} \int_T w_i \frac{\partial u^h}{\partial t}\, dx + \sum_{T, M_i\in T} (\frac{1}{3} + \alpha_i^T)\Phi^T = 0. \tag{3.3}$$

Giving $\alpha_i^T = \Phi_i^T/\Phi^T - 1/3$, the spatial discretisation of the Petrov-Galerkin formulation is the same of Residual Distribution approach. The previous equation becomes

$$\sum_{T\in\tau_h} \sum_{M_j\in T} m_{ij}^T \frac{d\, u_j}{dt} + \sum_{T, M_i\in T} \Phi_i^T = 0 \tag{3.4}$$

where (m_{ij}^T) is the consistent mass matrix defined by $m_{ij}^T = \displaystyle\int_T w_i \Lambda_j \, dx$. This expression is the consistent formulation of the Residual Distribution scheme.

$$(m_{ij}^T) = \frac{|T|}{3} \begin{bmatrix} \Phi_1^T/\Phi^T + 1/6 & \Phi_1^T/\Phi^T - 1/12 & \Phi_1^T/\Phi^T - 1/12 \\ \Phi_2^T/\Phi^T - 1/12 & \Phi_2^T/\Phi^T + 1/6 & \Phi_2^T/\Phi^T - 1/12 \\ \Phi_3^T/\Phi^T - 1/12 & \Phi_3^T/\Phi^T - 1/12 & \Phi_3^T/\Phi^T + 1/6 \end{bmatrix}. \tag{3.5}$$

Remark 2. *The first order N-scheme cannot be obtained from a Petrov-Galerkin approach since Φ_i^T/Φ^T are unbounded for vanishing cell residual.*

Remark 3. *Starting from (3.4) we can derive a certain number of schemes, extension of standard Residual Distribution schemes that have the property Φ_i^T/Φ^T bounded (see [15]).*

If we want to recover (2.7) we must replace the consistent mass matrix by an inconsistent mass matrix. This matrix is obtained by lumping its element of each row to the diagonal, and replace Φ_i^T/Φ^T by 1/3 (equidistribution of the fluctuation) giving the inconsistent Galerkin lumped mass matrix (1/3)Id. This leads to the following inconsistent formulation

$$|C_i|\frac{du_i}{dt} + \sum_{T, M_i \in T} \Phi_i^T = 0. \tag{3.6}$$

Explicit Euler scheme for the time derivative leads to (2.7).

For steady computations an accurate discretization of the time derivative is not required, so that (3.6) is a good alternative in place of (3.4) which needs to solve a system of non linear algebraic equations. With this approximation only first order accuracy can be achieved for unsteady problems even if we use second or third accurate schemes for time derivative. So the consistent formulation (3.4) is needed to reach high order accuracy. One of the aim of this paper is to show how to construct consistent mass matrices without loosing stability properties of the schemes for steady state. The development of monotone high accurate schemes using (3.4) has been considered by Ferrante [15] with a flux corrected transport formulation. He showed that positivity is lost with this formulation (the mass matrix is not guaranteed a M-matrix), and to recover monotonicity he used the Flux Corrected Transport to damp the spurious oscillations. This approach works reasonably well, but it is not a natural extension of (2.7). We loose the aspects of the Residual Distribution methods like (i) upwinding; (ii) linearity preservation (iii) compactness of the stencil (due to the FCT limiting).

The aim of this paper is to show how one can recover properties (i), (ii) and (iii).

4 Residual Distribution Schemes for unsteady scalar advection equations

We showed, in the previous section, that the Residual Distribution scheme (2.7) is not suitable to compute unsteady problems due to the inconsistent treatment of the mass matrix. Schemes with consistent mass matrix has been considered by Ferrante. Here we present and illustrate another solution to reach high order accuracy for unsteady scalar advection, and in the last section, we extend it to compressible Euler equations.

4.1 Principles

Let us consider equation (2.1) with constant advection speed

$$\frac{\partial u}{\partial t} + \langle \lambda, \nabla u \rangle = 0 \text{ in } \Omega \times (0, T).$$

The idea is to see in (2.1) a steady problem in $\mathbb{R}_\xi \times \mathbb{R}_x \times \mathbb{R}_t$ where ξ is an iteration parameter. We first need to define the cell residual. For steady problems the cell residual over a triangle T is defined as the integral over this element of the operator div with the piecewise linear approximation of u on T. Here we solve the unsteady problem like a steady problem, so we define the cell residual as the integral of the differential operator $\partial_t + div$ with a space-time approximation of u on $T \times [t_n, t_{n+1}]$. To do this we need a second order approximation of u to get second order accuracy, which implies $\Phi^T = \mathcal{O}(h^3)$. The most natural choice is

$$u^h(x,t) = \frac{t - t_n}{\Delta t} u^{n+1}(x) + \frac{t_{n+1} - t}{\Delta t} u^n(x) \tag{4.1}$$

where u^n and u^{n+1} are respectively the piecewise linear approximation of u at time t_n and t_{n+1}. Hence we define the residual Φ^T by

$$\Phi^T = \int_{t_n}^{t_{n+1}} \int_T \frac{\partial u^h}{\partial t} + \langle \lambda, \nabla u^h \rangle \, dx dt. \tag{4.2}$$

After calculations, we get

$$\Phi^T = \frac{|T|}{3} \sum_{M_i \in T} (u_i^{n+1} - u_i^n) + \frac{\Delta t}{2} \sum_{M_i \in T} k_i (u_i^{n+1} + u_i^n), \tag{4.3}$$

with the same definition of k_i.

We denote $\Phi_{i,n}$ and $\Phi_{i,n+1}$ the residual sent to the nodes (M_i, t_n) and (M_i, t_{n+1}).

The residual (4.2) is the fluctuation computed over the prism $K = T \times [t_n, t_{n+1}]$ (see Figure 4). Our approach is a space-time interpretation of the classical Residual Distribution methods with a $\mathbb{Q}_1$ approximation of the solution. We need to distribute this fluctuation on the vertices of the prism K, upwinding in time, ie a causality principle, leads to distribute it to the nodes located at time t_{n+1}, so one has $\Phi_{i,n} = 0$.

The Residual Distribution scheme reads

$$\forall M_i \in \tau_h, \qquad \sum_{T, M_i \in T} \Phi_{i,n+1} = 0. \tag{4.4}$$

where $\Phi_{i,n+1}$ is the residual sent to node M_i at time t_{n+1}. After-wards we replace $\Phi_{i,n+1}$ by Φ_i since there is no possible confusion. The residuals are assumed to fulfill the conservation relation

$$\sum_{M_i \in T} \Phi_i = \Phi. \tag{4.5}$$

where Φ is given by (4.3).

Design principles

(i) **The upwinding property** No fraction of the element residual is sent to nodes for time $t = t_n$. Equation (4.4) means that all the residual of K is sent to time t_{n+1}.

(ii) **The Linear Preserving condition (LP)** The (LP) property is the same as for the steady case : if the solution is smooth then

$$\Phi_i = \mathcal{O}(h^3, \Delta t^3).^3).$$

(iii) **No spurious oscillations in the solution**

4.2 Extension of classical schemes

We discuss a space-time N-scheme and LDA scheme that reduces to the classical N-scheme and LDA scheme for steady state application. This space-time schemes allow us to construct space-time blending and modified schemes.

4.2.1 The N-scheme

We set

$$\Phi_i^N = \frac{|T|}{3}(u_i^{n+1} - u_i^n) + \frac{\Delta t}{2} k_i^+ (u_i^{n+1} - \widetilde{u}^{n+1}) + \frac{\Delta t}{2} k_i^+ (u_i^n - \widetilde{u}^n). \tag{4.6}$$

where $\widetilde{u}^n$ and $\widetilde{u}^{n+1}$ are computed to have the conservation relation (4.5). One gets

$$\widetilde{u}^n = N \sum_{M_j \in T} k_j^- u_j^n, \qquad \widetilde{u}^{n+1} = N \sum_{M_j \in T} k_j^- u_j^{n+1} \tag{4.7}$$

with the same definition of N (see section 2.4). The residual at node i becomes

$$\Phi_i^N = \frac{|T|}{3}(u_i^{n+1} - u_i^n) + \frac{\Delta t}{2} \sum_{M_j \in T} \left[k_i^+ N k_j^- (u_i^{n+1} - u_j^{n+1}) + k_i^+ N k_j^- (u_i^n - u_j^n) \right]. \tag{4.8}$$

The schemes reads, for all $M_i \in \tau_h$

$$\begin{aligned} &\sum_{T, M_i \in T} \Phi_i^N = 0. \\ &\sum_{T, M_i \in T} \left[\frac{|T|}{3}(u_i^{n+1} - u_i^n) + \frac{\Delta t}{2} \sum_{M_j \in T} \left[k_i^+ N k_j^- (u_i^{n+1} - u_j^{n+1}) + k_i^+ N k_j^- (u_i^n - u_j^n) \right] \right] = 0. \end{aligned} \tag{4.9}$$

Remark 4. *For steady computations, we recover the standard N-scheme.*

The scheme (4.9) is an implicit scheme, it can be written as $AU^{n+1} = BU^n$ where U^{n+1} and U^n are respectively the vectors of unknows at time t_{n+1} and t_n. The entries of the matrices A and B are given by

$$A_{ii} = \sum_{T, M_i \in T} \left(\frac{|T|}{3} + \frac{\Delta t}{2} k_i^+\right) \quad , \quad A_{ij} = \sum_{T, (M_i, M_j) \in T} -\frac{\Delta t}{2} k_i^+ N k_j^- \tag{4.10}$$

$$B_{ii} = \sum_{T, M_i \in T} \left(\frac{|T|}{3} - \frac{\Delta t}{2} k_i^+\right) \quad , \quad B_{ij} = \sum_{T, (M_i, M_j) \in T} \frac{\Delta t}{2} k_i^+ N k_j^- \tag{4.11}$$

and we have the following lemma :

Lemma 1. *The matrix A is non singular and it is a L-matrix i.e. $A_{ii} > 0$ $\forall i$ and $A_{ij} \le 0$ for $j \ne i$.*

Proof. One has $A_{ii} > 0$, and $A_{ij} \le 0$ since $N < 0 \Rightarrow k_i^+ N k_j^- \ge 0$, hence A is a L-matrix.

To prove that A is non-singular we show that A is a diagonally strictly dominant matrix i.e.

$$\forall i \quad |A_{ii}| > \sum_{j \ne i} |A_{ij}|$$

We have

$$\begin{aligned} |A_{ii}| - \sum_{j \ne i} |A_{ij}| &= \sum_{T M_i \in T} \left(\frac{|T|}{3} + \frac{\Delta t}{2} k_i^+\right) - \sum_{M_j \ne M_i} \sum_{T, (M_i, M_j) \in T} \frac{\Delta t}{2} k_i^+ N k_j^- \\ &= \sum_{T, M_i \in T} \left(\frac{|T|}{3} + \frac{\Delta t}{2} k_i^+\right) - \sum_{T, M_i \in T} \sum_{M_j \ne M_i} \frac{\Delta t}{2} k_i^+ N k_j^- \\ &= \sum_{T, M_i \in T} \left(\frac{|T|}{3} + \frac{\Delta t}{2} k_i^+\right) - \sum_{T, M_i \in T} \frac{\Delta t}{2} k_i^+ \\ &= \sum_{T, M_i \in T} \frac{|T|}{3} \\ &= |C_i| > 0 \end{aligned}$$

This shows that A is diagonally strictly dominant. This achieves the proof. ◻

One notes that $B_{ij} \geq 0$ for $j \neq i$, and under the condition $\Delta t \leq \frac{2}{3}\frac{|T|}{k_i^+}$ one has $B_{ii} \geq 0$. The monotonicity property of the scheme is described in the following

Proposition 4.1. *The extension of the N-scheme defined by (4.9) is monotone under the CFL like condition*

$$\Delta t \leq \min_{M_i, T} \frac{2}{3}\frac{|T|}{k_i^+} \tag{4.12}$$

Proof. The scheme can be written as $A\,U^{n+1} = B\,U^n$ with A a non singular L-matrix, and B a matrix with positive elements. It is shown in [27] that if A is a regular L-matrix the following assertions are equivalent

(i) A^{-1} is positive,

(ii) there exists a diagonal positive matrix D such that $D^{-1}AD$ is diagonally strictly dominant.

We apply this results with $D = Id$ so it implies that A^{-1} is positive. B has positive elements, then the matrix $A^{-1}B$ is positive.

• ***Step 1 :*** Suppose there exists $a \in \mathbb{R}$ such that $u_i^n \geq a$ for all $M_i \in \tau_h$. One wants to show $u_i^{n+1} \geq a$ for all $M_i \in \tau_h$.

One has

$$\sum_{T \ni M_i} \left(\frac{|T|}{3}(u_i^{n+1} - u_i^n) + \frac{\Delta t}{2} \sum_{M_j \neq M_i} k_i^+ N k_j^- (u_i^{n+1} - u_j^{n+1} + u_i^n - u_j^n) \right) = 0$$

$$\begin{aligned}
\sum_{T \ni i} \left(\frac{|T|}{3} u_i^{n+1} + \frac{\Delta t}{2} \sum_{j \neq i} k_i^+ N k_j^- (u_i^{n+1} - u_j^{n+1}) \right) &= \sum_{T \ni i} \left(\frac{|T|}{3} u_i^n - \frac{\Delta t}{2} \sum_{j \neq i} k_i^+ N k_j^- (u_i^n - u_j^n) \right) \\
&= \sum_{T \ni i} \left((\frac{|T|}{3} - \frac{\Delta t}{2} k_i^+) u_i^n + \frac{\Delta t}{2} \sum_{j \neq i} k_i^+ N k_j^- u_j^n \right) \\
&\geq \sum_{T \ni i} \left((\frac{|T|}{3} - \frac{\Delta t}{2} k_i^+) a + \frac{\Delta t}{2} \sum_{j \neq i} k_i^+ N k_j^- a \right) \\
&\geq \sum_{T \ni i} \frac{|T|}{3} a
\end{aligned} \tag{4.13}$$

because the matrix B is positive. The inequalities (4.13) imply that

$$\sum_{T \ni i} \left(\frac{|T|}{3}(u_i^{n+1} - a) + \frac{\Delta t}{2} \sum_{j \neq i} k_i^+ N k_j^- ((u_i^{n+1} - a) - (u_j^{n+1} - a)) \right) \geq 0,$$

which can be rewritten as

$$A(u^{n+1} - a) \geq 0.$$

The matrix A^{-1} is positive so one has $u^{n+1} - a \geq 0$, *ie* $u_i^{n+1} \geq a$ for all $M_i \in \tau_h$.

• ***Step 2 :*** Suppose that $\exists a \in \mathbb{R}$ such that $u_i^n \leq a$ for all $M_i \in \tau_h$. One wants to show $u_i^{n+1} \leq a$ for all $M_i \in \tau_h$. We do the same thing as before. This achieves the proof. □

Remark 5. *The condition (4.12) seems to be a severe limit on the allowed time step but computations with CFL great than one showed that monotonocity is preserved. It follows from the fact that this condition is only a sufficient condition. An deeper study would need to consider the matrix $A^{-1}B$ for which we have any expression. In section (6), using a idea of Deconinck and al [10], we show how to avoid this severe constraint.*

4.2.2 The LDA scheme

We set

$$
\begin{aligned}
\Phi_i^{LDA} =& \frac{|T|}{3}(-k_i^+ N + \frac{1}{6})(u_i^{n+1} - u_i^n) + \\
& \frac{|T|}{3}(-k_i^+ N - \frac{1}{12}) \sum_{M_j \neq M_i} (u_j^{n+1} - u_j^n) - \frac{\Delta t}{2} k_i^+ N \sum_{M_j \in T} k_j (u_j^{n+1} + u_j^n)
\end{aligned}
\tag{4.14}
$$

This scheme directly follows from the consistent formulation (3.4) with Crank-Nicolson discretization for the time derivative, using the standard LDA-scheme. Consider expression (3.4) with the standard LDA-scheme

$$
\sum_{T, M_i \in T} \sum_{M_j \in T} m_{ij}^T \frac{d u_j}{dt} + \sum_{T, M_i \in T} (-k_i^+ N) \Phi^T(u^h) = 0
$$

where $u^h = \sum_j u_j(t) \Lambda_j$. If we consider Crank-Nicolson discretization of time derivative one obtains

$$
\begin{aligned}
&\sum_{T, M_i \in T} \sum_{M_j \in T} m_{ij}^T (u_j^{n+1} - u_j^n) + \frac{1}{2} \sum_{T, M_i \in T} (-k_i^+ N) \left(\Phi^T(u^{n+1}) + \Phi^T(u^n) \right) = 0 \\
&\sum_{T, M_i \in T} \sum_{M_j \in T} m_{ij}^T (u_j^{n+1} - u_j^n) + \frac{1}{2} \sum_{T, M_i \in T} (-k_i^+ N) \left(\sum_j k_j (u_j^{n+1} + u_j^n) \right) = 0
\end{aligned}
\tag{4.15}
$$

After lenghtily calculations, expression (4.15) can be written as

$$
\sum_{T, M_i \in T} \Phi_i^{LDA} = 0,
$$

with Φ_i^{LDA} defined by (4.14).

As for the extension of N-scheme, we write LDA-scheme as $A U^{n+1} = B U^n$. The matrices A and B have no apparent properties.

Remark 6. *We recover the LDA scheme of Ferrante [15].*

4.2.3 The PSI scheme

The extension of PSI scheme is straightforward. It results in a combination of the first order monotone N-scheme (4.9) and the second order non-monotone LDA-scheme (4.14)

$$
\Phi_i^{PSI} = l \Phi_i^N + (1 - l) \Phi_i^{LDA}. \tag{4.16}
$$

where $l = \max(\varphi(r_1), \varphi(r_2), \varphi(r_3))$ with

$$
r_i = \frac{\Phi_i^{LDA}}{\Phi_i^N}, \qquad \varphi(x) = \begin{cases} \dfrac{x}{1-x} & \text{if } x < 0 \\ 0 & \text{else} \end{cases}
$$

We have experimentally noticed that the blending parameter, proposed for steady state by Deconinck and van der Weide, defined by

$$
l = \frac{|\Phi^T|}{\sum_j |\Phi_j^N|} \tag{4.17}
$$

works very well, even though it does not satisfy the positivity requirements (we have found numerical counter-examples). We need to solve a system of non linear equations. This is done by Newton's method. One can write the PSI scheme as $F(U^{n+1}) = 0$,

$$\begin{aligned} F: \quad & \mathbb{R}^{ns} \rightarrow \mathbb{R}^{ns} \\ & U \mapsto (F_i(U))_{i=1,\dots,ns} \end{aligned}$$

where $F_i(U) = \sum_{T, M_i \in T} \Phi_i^{PSI}$. The kth step of Newton's method is

$$U^{k+1} = U^k - \widetilde{J}(U^k) F(U^k),$$

where $\widetilde{J}(U^k)$ is an approximation of $\partial F / \partial U$ at point U^k.

4.2.4 The N-Modified scheme

The N-Modified scheme is described in [5]. We set

$$\phi_i^{NM} = \Phi_i^N + \Psi_i, \tag{4.18}$$

Ψ_i is computed to get second order accuracy and positivity *exactly* as for the modified scheme of subsection 2.2. The difference is that the evaluation of the limiter is implicit. This is solved as in the previous section, thanks to a Newton algorithm.

4.3 Numerical Results

4.3.1 The rotating cosine hill

The rotating cosine hill is a classical test-case for numerical schemes of the two-dimensional linear unsteady advection equation. The test consists in the transport of a cosine shape by a circular solid body advection field centered at the origin

$$\frac{\partial u}{\partial t} + \langle \lambda, \nabla u \rangle = 0 \quad \text{in } [-1,1] \times [-1,1], \tag{4.19}$$

where $\lambda = (y, -x)^T$. The initial solution is $(1 + \cos(4\pi \sqrt{(x+0.5)^2 + y^2})/2$. The solution is set to zero at the inflow boundaries, at each time step.

The computation was made on a unstructured grid of 8079 nodes and 15836 elements. The time step was taken to satisfy the condition (4.12) :

$$\Delta t = CFL \times min \frac{2}{3} \frac{|T|}{k_i^+}, \tag{4.20}$$

with $CFL = 0.9$. The results, using the schemes described in the previous section and the MUSCL scheme (with *minmod* limiter and Runge-Kutta integration in time) after one revolution are compared in Figure 5. The N-scheme is clearly the most diffusive (see Table 2), streamwise and crosswise diffusion are considerable. The LDA scheme keeps the height of the peak much better but the monotonicity is not preserved. The PSI and N-modified schemes gives similar results but the N-modified scheme is the less diffusive with a peak value of 0.802. This results are much better than those obtained by the MUSCL scheme.

4.3.2 The rotating cylinder

This test case differs from the previous only for the initial profile

$$u(x,0) = \begin{cases} 1 & \text{for } r < 0.25, \\ 0 & \text{else} \end{cases} \tag{4.21}$$

where $r = \sqrt{(x+0.5)^2 + y^2}$, which is not continuous, contrary to the previous case. The computation was made on the same grid with a CFL of 0.9, the results after one revolution are displayed in figure 6. The

Scheme	min	max
N	0	0.217
MUSCL	0	0.313
LDA	-0.03	0.983
PSI	0	0.756
N-modified	0	0.802

Table 2: Min and Max solution values for the rotating cosine hill test case.

solutions exhibit the same properties as the rotating cosine hill. The PSI scheme and N-modified scheme gives similar results. This results are better than MUSCL schemes. The LDA scheme exhibits spurious oscillations, as expected.

In both cases, the limited N scheme and the blended–PSI behave the same. However, the modified N scheme CPU cost is about half that of the blended scheme, because one has not to evaluate the LDA scheme.

4.4 Conclusions for the scalar problems

We have considered two different techniques for increasing the accuracy of unsteady problems, the PSI–blending and the N–modified schemes. The results are roughly speaking similar, with a slight advantage for the N–modified scheme, in term of accuracy and efficiency. It is possible to extend the PSI–blending technique to systems, following [2]. This has been done. This technique works quite well, but in some cases, for example the interaction of a vortex and a shock (§5.3.4) or the ramp problem (§5.3.3), the results are disappointing : there are some slight oscillations across the shocks.

For this reasons, we have abandon this technique, and we do not consider it anymore.

5 Residual Distribution Schemes for the unsteady Euler equations

Consider the system of Euler equations

$$\frac{\partial W}{\partial t} + \text{ div } \mathcal{F}(W) = 0,$$

where W and $\mathcal{F}$ are given by (2.17). The system is closed by the perfect gas law.

The construction of the Residual Distribution schemes, described in the previous section, is based on the quasi-linear form of the convection equations (like for steady schemes) and a space-time approximation of the unknowns . So extension to the unsteady Euler equations need to use : (i) second order approximation of W and $\mathcal{F}$ to get second order accuracy; (ii) a conservative linearisation.

5.1 The cell residual

The cell residual over a triangle T is given by

$$\Phi^T = \int_{t_n}^{t_{n+1}} \int_T \left(\frac{\partial W^h}{\partial t} + \text{ div } \mathcal{F}^h \right) dx dt,$$

where W^h and $\mathcal{F}^h$ are respectively space-time approximation of solution and fluxes. This approximation is obtained using Roe's parameter vector Z . We set

$$\begin{aligned} W^h &= \frac{t - t_n}{\Delta t} W^{n+1} + \frac{t_{n+1} - t}{\Delta t} W^n, \\ \mathcal{F}^h &= \frac{t - t_n}{\Delta t} \mathcal{F}^{n+1} + \frac{t_{n+1} - t}{\Delta t} \mathcal{F}^n, \end{aligned}$$

where $W^s = W(Z^s)$ and $\mathcal{F}^s = \mathcal{F}(Z^s)$, Z^s the piecewise linear approximation of Z at time s (t_n or t_{n+1}). One has

$$\Phi^T = I_1 + I_2$$
$$\text{with} I_1 = \int_{t_n}^{t_{n+1}} \int_T \frac{\partial W^h}{\partial t}\, dxdt, \qquad I_2 = \int_{t_n}^{t_{n+1}} \int_T \text{ div } \mathcal{F}^h dxdt.$$

For I_1 we have

$$I_1 = \int_{t_n}^{t_{n+1}} \frac{\partial W}{\partial t}\, dxdt = \int_T \frac{D(Z^{n+1})Z^{n+1} - D(Z^n)Z^n}{2}\, dx.$$

Moreover,

$$\begin{aligned}
\int_T D(Z)Z\, dx &= \frac{|T|}{3}\left(D(\frac{Z_1+Z_2}{2})\frac{Z_1+Z_2}{2} + D(\frac{Z_1+Z_3}{2})\frac{Z_1+Z_3}{2} + D(\frac{Z_2+Z_3}{2})\frac{Z_2+Z_3}{2}\right)\\
&= \frac{|T|}{12}\sum_{M_i\in T} D(Z_i)Z_i + \frac{|T|}{4}\sum_{M_i\in T} D(\overline{Z})Z_i\\
&= \frac{|T|}{6}\sum_{M_i\in T} W_i + \frac{|T|}{4}\sum_{M_i\in T} \widetilde{W}_i,
\end{aligned}$$

because $D(Z)Z$ is quadratic in Z. I_1 becomes

$$I_1 = \frac{|T|}{4}\sum_{M_j\in T}\left[(\frac{1}{3}W_j^{n+1} + \frac{1}{2}\widetilde{W}_j^{n+1}) - (\frac{1}{3}W_j^n + \frac{1}{2}\widetilde{W}_j^n)\right].$$

Now we compute I_2

$$\begin{aligned}
I_2 &= \int_{t_n}^{t_{n+1}} \int_T \text{ div } \mathcal{F}\ dxdt\\
&= \frac{\Delta t}{2}\int_T \text{ div } \mathcal{F}^{n+1} + \text{ div } \mathcal{F}^n\ dx\\
&= \frac{\Delta t}{2}\int_T \mathcal{R}(Z^{n+1})\, dx.\nabla Z^{n+1} + \frac{\Delta t}{2}\int_T \mathcal{R}(Z^n)\, dx.\nabla Z^n\\
&= \frac{\Delta t}{2}\sum_{M_j\in T} \frac{\mathcal{R}(\bar{Z}^{n+1}).n_j^T}{2} Z_j^{n+1} + \frac{\mathcal{R}(\bar{Z}^n).n_j^T}{2} Z_j^n.
\end{aligned}$$

Setting $K_j^{n+1} = \frac{1}{2}\langle \mathcal{R}(\bar{Z}^{n+1})D^{-1}(\bar{Z}^{n+1}), n_j^T\rangle$ and $K_j^n = \frac{1}{2}\langle \mathcal{R}(\bar{Z}^n)D^{-1}(\bar{Z}^n), n_j^T\rangle$, the cell residual is given by

$$\begin{aligned}
\Phi^T = &\frac{|T|}{4}\sum_{M_j\in T}\left[(\frac{1}{3}W_j^{n+1} + \frac{1}{2}\widetilde{W}_j^{n+1}) - (\frac{1}{3}W_j^n + \frac{1}{2}\widetilde{W}_j^n)\right]\\
&+ \frac{\Delta t}{2}\sum_{M_j\in T} K_j^{n+1}\,\widetilde{W}_j^{n+1} + K_j^n\,\widetilde{W}_j^n,
\end{aligned} \tag{5.1}$$

where $\widetilde{W}_j^n = D(\bar{Z}^n)Z_j^n$ and $\widetilde{W}_j^{n+1} = D(\bar{Z}^{n+1})Z_j^{n+1}$.

This expression is similar to (4.3), but a difficulty appears with the presence of terms K_j^{n+1} which are computed at time t_{n+1} (due to the non-linearity of the Euler equations). Therefore system Residual Distribution schemes for unsteady computations are intrinsically implicit in time. This is solved thanks to a Newton method.

5.2 System residual distribution schemes

In this section we give generalization of the unsteady scalar schemes described in section 4.2.

5.2.1 The system N-scheme

For a scalar advection equation one has

$$\phi_i^N = \frac{|T|}{3}(u_i^{n+1} - u_i^n) + \frac{\Delta t}{2} k_i^+ (u_i^{n+1} - \widetilde{u}^{n+1}) + \frac{\Delta t}{2} k_i^+ (u_i^n - \widetilde{u}^n), \tag{5.2}$$

where $\widetilde{u}^n$ and $\widetilde{u}^{n+1}$ are computed to have the conservation relation (4.5). We set

$$\begin{aligned}\phi_i^N = & \frac{|T|}{4}\Big[(\frac{1}{3}W_i^{n+1} + \frac{1}{2}\widetilde{W}_i^{n+1}) - (\frac{1}{3}W_i^n + \frac{1}{2}\widetilde{W}_i^n)\Big] \\ & + \frac{\Delta t}{2} K_i^{n+1^+}(\widetilde{W}_i^{n+1} - W_{n+1}^*) + \frac{\Delta t}{2} K_i^{n+}(\widetilde{W}_i^n - W_n^*).\end{aligned}$$

$\widetilde{W}^{n+1}$ and $\widetilde{W}^n$ are computed to make the scheme conservative, one gets

$$W_{n+1}^* = N^{n+1}\sum_j K_j^{n+1^-}\widetilde{W}_j^{n+1}, \quad W_n^* = N^n \sum_j K_j^{n-}\widetilde{W}_j^n \tag{5.3}$$

with $N^n = (\sum_j K_j^{n-})^{-1}$ and $N^{n+1} = (\sum_j K_j^{n+1^-})^{-1}$.

Remark 7. *The matrices $N^n K_j^n$ and $N^{n+1}K_j^{n+1}$ always exist (see [2]), so the system N-scheme is well defined.*

5.2.2 The system LDA scheme

We set

$$\begin{aligned}\Phi_i^{LDA} = & (-(K_i^+ N)^{n+1} + \frac{1}{6}\mathrm{Id})\Phi_i^t + \\ & (-(K_i^+N)^{n+1} - \frac{1}{12}\mathrm{Id})\sum_{M_j \neq M_i}\Phi_j^t - (K_i^+N)^{n+1}\Phi^s\end{aligned}$$

where $\Phi_j^t = \frac{|T|}{4}\Big[(\frac{1}{3}W_j^{n+1} + \frac{1}{2}\widetilde{W}_j^{n+1}) - (\frac{1}{3}W_j^n + \frac{1}{2}\widetilde{W}_j^n)\Big]$ and $\Phi^s = \frac{\Delta t}{2}\sum_{j\in T} K_j^{n+1}\widetilde{W}_j^{n+1} + K_j^n \widetilde{W}_j^n$.

The remark 7 also applies to the LDA scheme.

5.2.3 The system N-modified scheme

The idea is similar than the one developed in [5]. We present here a summary of these results, and we refer to the above reference for a detailed presentation.

For the sake of simplicity, let us explain our procedure on the linearized $d \times d$ system

$$\frac{\partial W}{\partial t} + A\frac{\partial W}{\partial x} + B\frac{\partial W}{\partial y} = 0$$

and we assume A and B symmetric.

Since for any α β the matrix $\alpha A + \beta B$ is symetric, there exists a complete set of eigenvectors $\{r_j\}_{j=1,d}$.

Following [5], we introduce the simple waves

$$U_\sigma(x,t) = \big(t - t_0\big)\ \big(\alpha x + \beta y\big) r_\sigma$$

where r_σ is any eigenvector in $\{r_j\}_{j=1,d}$ and λ_σ is the associated eigenvalue. We first notice that any function

$$W(x,t) = \frac{t - t_n}{\Delta t}W^{n+1}(x) + \frac{t_{n+1} - t}{\Delta t}W^(x)$$

where W^n and W^{n+1} are linear in x is a sum of $3 \times d \times 2$ simple waves plus possibly a constant. This comes from the fact that W^n (resp. W^{n+1}) can be written as

$$\begin{aligned} W^n(x) &= \sum_{j=1}^{3} W_j^n \Lambda_j(x) \\ &= \text{Cte} + \frac{1}{2|T|} \sum_{j=1}^{3} W_j^n \langle n_j, x\rangle \end{aligned}$$

where n_j is the inward normal opposite to the jth vertex of T. Then, taking $n_j = (\alpha_j, \beta_j)^T$, we decompose W_j^n on an eigenbasis of $\alpha_j A + \beta_j B$ $\{r_k^j\}_{k=1,d}$ and we recognize that

$$(t - t_n)\langle W_j^n, r_k^j\rangle\langle n_k, x\rangle r_k^j$$

is precisely a simple wave.

In summary, any function piecewise linear in time and space can be decomposed into the sum of $3 \times d$ simple waves

$$\begin{aligned} U_\sigma(x,t) &= \left(\frac{t-t_n}{\Delta t}(\alpha^{n+1}x + \beta y^{n+1}) + \frac{t_{n+1}-t}{\Delta t}(\alpha^n x + \beta y^n)\right) r_\sigma \\ &:= \left(\frac{t-t_n}{\Delta t}\varphi^{n+1}(x) + \frac{t_{n+1}-t}{\Delta t}\varphi^n(x)\right) r_\sigma. \end{aligned}$$

Then, following once more [5], it is possible to show that the N scheme (5.2) is monotone on any simple wave, that is the residual Φ_i^N sent to the node i can be written, for a simple wave, as a linear combination of terms like

$$\begin{aligned} &\sum_{j=1,3} \left(c_{ij}^k\right)^{n+1} \left(\varphi_j^{n+1} - \varphi_i^{n+1}\right) r_j^k \\ &\text{and} \\ &\sum_{j=1,3} \left(c_{ij}^k\right)^{n} \left(\varphi_j^{n} - \varphi_i^{n}\right) r_j^k \end{aligned} \tag{5.4}$$

with $c_{ij}^k \geq 0$ and $k = 1, d$ and φ_j is the value of φ at the jth vertex of T.

Using this remark, we introduce the modified N scheme as follow. We take a direction (α, β), consider the eigenvectors of $\alpha A + \beta B$, write the residual as

$$\Phi_i^N = \sum_{k=1}^{d} \varphi_i^k r_k, \qquad \varphi_i^k = \langle \Phi_i^N, r_k\rangle$$

and modify the scalar residual $\{\varphi_i^k\}_{i=1,3}$ as in section 4.2.4 [1]

Following [5], and thanks to the relations (5.4), it is possible to show that the modified N scheme is stable.

In the case of a symetrisable system, we proceed along the same lines. The only modification is that the scalar products between a state variable W or a residual Φ_i^N with an eigenvector r_k are replaced by the inner product between W (resp. Φ_i^N) and the *left* eigenvector associated to the right eigenvector r_k.

In the previous description, the choice of the direction (α, β) is arbitrary. Different choice will produce different results, and different schemes. However, all our numerical experiments for steady and unsteady problems lead to the conclusion that the quality of the results is independent of the direction (α, β). By quality, we mean the non oscillatory properties of the schemes, and the mesh resolution of the different features of the solution. In all the numerical experiment we present in this paper for fluid problems, we have chosen the vector (α, β) to be the local flow velocity vector for symmetry reasons mainly.

[1] We notice that $\sum_{i=1}^{3} \varphi_i^k = \langle \Phi, r_k\rangle$.

5.3 Numerical Results

In this section we present results on some classical test cases.

5.3.1 A two-dimensional Riemann Problem

The initial data are chosen in order to represent a "2D Sod tube" in the domain $[-1,1]\times[-1,1]$:

$$\begin{cases} \rho &= 0.1 \text{ if } xy<0, \text{ 1 otherwise} \\ p &= 0.1 \text{ if } xy<0, \text{ 1 otherwise} \end{cases} \tag{5.5}$$

and the velocity (u,v) was set to zero. The solution is computed at time t=0.2 on a structured triangulation where $\Delta x = \Delta y = 0.01$ and the CFL condition has been set to 0.9.

The isolines of the density and pressure are shown in Figures 8 and 9 for the N-scheme and the N-modified scheme.

5.3.2 A Mach 3 Wind Tunnel with a Forward Facing Step

This test case has been extensively studied by Woodward and Collela [9], and is widely present in the literature. The setup of the problem is the following : a right-going Mach 3 uniform flow enters a wind tunnel of 1 unit width and 3 units long. The step is 0.2 units high and is located 0.6 units from the left-hand end of the tunnel. The problem is initialized by a uniform, right going Mach 3 flow. Reflective boundary conditions are applied along the walls of the tunnel, and inflow and outflow boundary conditions are applied at the entrance and the exit of the tunnel. The results at time $t=4$ with the N-scheme and the N-modified scheme are shown. The simulation was done at CFL=0.9.

The corner of the step is a singularity. It is well known that if no special treatment is done, an entropy production is observed in the vicinity of the step corner, and it alters the quality of the second reflected shock. This is not physical because we have a strong expansion wave, so no entropy should be created. However, unlike in [9], we do not modify our scheme near the corner, because we are only interested in its stability properties.

An unstructured mesh have been considered, it contains 10,868 nodes and 21,281 triangles, it is refined near the corner. Portion of the mesh is shown in Figure 15.

The quality of the slip line coming out of the triple point is noticable, as well as the resolution of the shocks, in particular at the exit section of the tunnel. The maximum shock width is no larger than two cells. Between the first order and the second order results, the quality of the fan (at the corner) has dramaticaly been improved : the reflected shock is now correctly set, the weak compression shock after the fan appears, and interact with the first reflected shock, see the slip line coming out of the interaction between the reflected shock and the weak compression shock.

5.3.3 Reflection of a Shock on a Wedge

This problem was studied by Quirk [32]. A planar shock initially enters from the left in a quiescent fluid and is reflected from a 45 degrees ramp. Its Mach number is $M_s = 5.5$ and is defined toward the flow values in the quiescent fluid where the density is set to 1.4 and the pressure to 1. Reflective boundary conditions are applied along the ramp and the bottom and the upper of the problem domain. For such an incident shock wave Mach number and such a reflecting wedge angle, a double Mach reflection is expected. The interest of this test case is that , according to [7], the angle $\theta = 45^o$ and $M_s = 5.5$ is nearly at the transition between a double Mach reflection and a regular reflection. If the scheme were too diffusive, we would get a regular reflection instead of a double Mach reflection. Hence this is a good test of accuracy.

The density is displayed on Figure 11. The resolution of the different structure is quite clean, despites the poor resolution of the mesh.

5.3.4 Shock-Vortex Interaction Problem

This test case describes the interaction between a stationary shock and a vortex. It was first presented by Pao and Salas [31], and was studied by Meadows *et al* [29] with a TVD scheme and by Jiang and Shu [25].

The computational domain is taken to be $[0,2] \times [0,1]$. A stationary Mach 1.1 shock is set at $x = 0.5$ and normal to the x-axis. Its left state is $(\rho, u, v, p) = (1, \sqrt{\gamma}, 0, 1)$. A small vortex is superposed to the flow left to the shock and centers at $(x_c, y_c) = (0.5, 0.25)$. The vortex is described as a perturbation to the velocity (u, v), temperature $T = \frac{p}{\rho}$ and entropy $S = \ln \frac{p}{\rho^\gamma}$ of the mean flow and denote it by tilde values

$$\begin{aligned} \tilde{u} &= \epsilon \tau e^{\alpha(1-\tau^2)} \sin\theta & (5.6)\\ \tilde{v} &= -\epsilon \tau e^{\alpha(1-\tau^2)} \cos\theta & (5.7)\\ \tilde{T} &= -\frac{(\gamma-1)\epsilon^2 e^{2\alpha(1-\tau^2)}}{4\alpha\gamma} & (5.8)\\ \tilde{S} &= 0 & (5.9) \end{aligned}$$

where $\tau = \frac{r}{r_c}$ and $r = \sqrt{(x-x_c)^2 + (y-y_c)^2}$. Here ϵ indicates the strength of the vortex, α controls the decay rate of the vortex, and r_c is the critical radius for which the vortex has the maximum strength. We choose the same value as in [25], ie $\epsilon = 0.3$, $r_c = 0.05$ and $\alpha = 0.204$. The above defined vortex is a steady solution to the 2D Euler equation. The upper and lower boundary are set to be reflective. We use a uniform grid of 251×100. The pressure isolines at three different times are displayed on Figure 12.

6 Toward an unconditionally monotone and LP scheme

In this section we elaborate an unconditionally monotone and LP scheme, starting from the idea developed in [35], for scalar advection equation and give numerical result for the Euler equations. First we construct an unconditionally monotone scheme, and by limiting contributions we obtain an unconditionally monotone and LP scheme.

Let us consider the scalar advection equation

$$\frac{\partial u}{\partial t} + \langle \lambda, \nabla u \rangle = 0 \quad \text{in } \Omega \times [0, T]. \tag{6.1}$$

The N-scheme reads

$$\Phi_i^N = \frac{|T|}{3}(u_i^{n+1} - u_i^n) + \frac{\Delta t}{2} \sum_{M_j \in T} k_i^+ N k_j^- (u_i^{n+1} - u_j^{n+1} + u_i^n - u_j^n). \tag{6.2}$$

It is an upwind scheme in the prismatic element K : no contribution is sent to the past node located at time t_n (upwinding in time).

We have shown that the N-scheme is a monotone scheme under a CFL-like condition. If we write the N-scheme as $Au^{n+1} = Bu^n$, we have shown that A^{-1} has positive elements and B has the same propertie under a certain condition. In [35], the authors construct space-time residual distribution schemes using a space-time mesh containing three level of nodes and two layers of elements in the temporal direction. because two layers of elements appear to be the minimum necessary to allow for a scheme to be unconditionally monotone. The space time elements are tetrahedrons, this induces some unecessary complexity in the scheme.

Here we start from the same remark, and use prismatic elements for simplicity, and the techniques developped above. We consider three levels of nodes with temporal coordinates t_n, $t_{n+\alpha}$ and t_{n+1}, $\alpha \in [0,1]$. This delimits two layers of space-time elements. Considers a triangle T, one notes K_1 and K_2 the prisms delimited by T contained in the first and the second layer, with thickness $\Delta t_1 = t_{n+\alpha} - t_n$ and $\Delta t_2 = t_{n+1} - t_{n+\alpha}$ (see Figure 14).

6.1 Description of an unconditionally monotone scheme

Starting from the space-time N-scheme (4.8), which was made on a single layer of elements, we reformulate it on two layers of elements to allow us an unconditionally stable time marching procedure while maintaining monotonicity both in space and time.

We denote by $\Phi_{i,n}$, $\Phi_{i,n+\alpha}$ and $\Phi_{i,n+1}$ the residual sent to nodes (M_i, t_n), $(M_i, t_{n+\alpha})$ and (M_i, t_{n+1}). In the first layer, the past does not depend from the future, that is to say $\Phi^{K_1}_{i,n} = 0$. We set

$$\Phi^{K_1}_{i,n+\alpha} = \frac{|T|}{3}(u_i^{n+\alpha} - u_i^n) + \frac{\Delta t_1}{2} \sum_{M_j \in T} k_i^+ N k_j^- (u_i^{n+\alpha} - u_j^{n+\alpha} + u_i^n - u_j^n)$$

In the second layer we do not impose the upwind property, ie $\Phi^{K_2}_{i,n+\alpha} \neq 0$. We set

$$\Phi^{K_2}_{i,n+\alpha} = \frac{\Delta t_2}{2} \sum_{M_j \in T} k_i^+ N k_j^- (u_i^{n+\alpha} - u_j^{n+\alpha}) \tag{6.3}$$

$$\Phi^{K_2}_{i,n+1} = \frac{|T|}{3}(u_i^{n+1} - u_i^{n+\alpha}) + \frac{\Delta t_2}{2} \sum_{M_j \in T} k_i^+ N k_j^- (u_i^{n+1} - u_j^{n+1}) \tag{6.4}$$

Finally the scheme reads

$$\sum_{K,(i,n+\alpha)\in K} \Phi^K_{i,n+\alpha} = 0 \quad \text{and} \quad \sum_{K,(i,n+1)\in K} \Phi^K_{i,n+1} = 0. \tag{6.5}$$

We detail now the contribution for each nodes.

• For the node$(M_i, t_{n+\alpha})$ one has

$$\begin{aligned}\Phi^{K_1}_{i,n+\alpha} + \Phi^{K_2}_{i,n+\alpha} =& \frac{|T|}{3}(u_i^{n+\alpha} - u_i^n) + \frac{\Delta t_1}{2} \sum_{M_j \in T} k_i^+ N k_j^- (u_i^{n+\alpha} - u_j^{n+\alpha} + u_i^n - u_j^n) \\ &+ \frac{\Delta t_2}{2} \sum_{M_j \in T} k_i^+ N k_j^- (u_i^{n+\alpha} - u_j^{n+\alpha})\end{aligned} \tag{6.6}$$

$$\begin{aligned}\Phi^{K_1}_{i,n+\alpha} + \Phi^{K_2}_{i,n+\alpha} =& \frac{|T|}{3}(u_i^{n+\alpha} - u_i^n) + \frac{\Delta t_1}{2}(1 + \frac{\Delta t_2}{\Delta t_1}) \sum_{M_j \in T} k_i^+ N k_j^- (u_i^{n+\alpha} - u_j^{n+\alpha}) \\ &+ \frac{\Delta t_1}{2} \sum_{M_j \in T} k_i^+ N k_j^- (u_i^n - u_j^n)\end{aligned} \tag{6.7}$$

Assembling contributions of all elements one obtains $Au^{n+\alpha} = Bu^n$ with

$$A_{ii} = \sum_{T, M_i \in T} (\frac{|T|}{3} + \frac{\Delta t_1}{2}(1 + \frac{\Delta t_2}{\Delta t_1})k_i^+) \quad , \quad A_{ij} = \sum_{T,(M_i,M_j)\in T} -\frac{\Delta t_1}{2}(1 + \frac{\Delta t_2}{\Delta t_1})k_i^+ N k_j^- . \tag{6.8}$$

$$B_{ii} = \sum_{T, M_i \in T} (\frac{|T|}{3} - \frac{\Delta t_1}{2}k_i^+) \quad , \quad B_{ij} = \sum_{T,(M_i,M_j)\in T} \frac{\Delta t_1}{2}k_i^+ N k_j^- . \tag{6.9}$$

The matrix B have positive entries if Δt_1 satisfies condition (4.12). The inverse of A always exists and have positive entries for all time step Δt_1, Δt_2.

• For the node(M_i, t_{n+1}), assembling contributions of all elements gives $Au^{n+1} = Bu^{n+\alpha}$ with

$$A_{ii} = \sum_{T, M_i \in T} (\frac{|T|}{3} + \frac{\Delta t_2}{2}k_i^+) \quad , \quad A_{ij} = \sum_{T,(M_i,M_j)\in T} -\frac{\Delta t_2}{2}k_i^+ N k_j^- . \tag{6.10}$$

$$B_{ii} = \sum_{T, M_i \in T} \frac{|T|}{3}k_i^+ \quad , \quad B_{ij} = 0. \tag{6.11}$$

The matrix B and A^{-1} have always positive elements for all choice of time step Δt_2.

The timestep for the first layer (Δt_1) is limited by a CFL-like condition as we have seen in the previous section for the space-time N-scheme. However, since we do not impose an upwinding condition in time for the second layer, an arbitrary timestep Δt_2 can be choosen. Then the global timestep $\Delta t = \Delta t_1 + \Delta t_2$ is not contraint by any CFL like condition. We have constructed a scheme with unconditionally stable implicit time stepping which maintains full monotonocity.

6.2 An unconditionally monotone and LP scheme

Limiting contributions in the prism K_1 and K_2 gives a monotone and LP scheme. From residuals $\Phi^{K_1}_{i,n+\alpha}$, $\Phi^{K_2}_{i,n+\alpha}$ and $\Phi^{K_2}_{i,n+1}$ we construct modified residuals $\Phi^{K_1*}_{i,n+\alpha}$, $\Phi^{K_2*}_{i,n+\alpha}$ and $\Phi^{K_2*}_{i,n+1}$ in the same way as before (see section 4.2.4).

- In the prism K_1, we construct $\Phi^{K_1*}_{i,n+\alpha}$ such that

$$\sum_{M_j \in T} \Phi^{K_1*}_{j,n+\alpha} = \Phi^{K_1} \tag{6.12}$$

$$\Phi^{K_1*}_{i,n+\alpha} = \mathcal{O}(h^3) + \mathcal{O}(\Delta t^3) \tag{6.13}$$

- In the prism K_2, we construct $\Phi^{K_2*}_{i,n+\alpha}$ and $\Phi^{K_2*}_{i,n+1}$ such that

$$\sum_{M_j \in T} \left[\Phi^{K_2*}_{j,n+\alpha} + \Phi^{K_2*}_{j,n+1}\right] = \Phi^{K_2} \tag{6.14}$$

$$\Phi^{K_2*}_{i,n+\alpha} = \mathcal{O}(h^3) + \mathcal{O}(\Delta t^3) \tag{6.15}$$

$$\Phi^{K_2*}_{i,n+1} = \mathcal{O}(h^3) + \mathcal{O}(\Delta t^3) \tag{6.16}$$

With this construction the resulting scheme is second order accurate both in time and space, and unconditionally monotone.

6.3 A numerical example : the Mach 3 Wind Tunnel with a Forward Facing Step

We use the same mesh as before (see section 5.3.2), which is refined near the corner, with a fixed ratio $\Delta t_2/\Delta t_1$ equal to 10. Isolines of density , Mach number and entropy at time $t = 4$ are shown. The density, mach number and entropy deviation are shown on Figure 20. The results look very similar to those shown on Figure 16,17 and 18.

7 Concluding remarks

We have developed a family of schemes for unsteady flow problems. These scheme have several characteristics. First they are adapted to unstructured triangular type meshes. The degrees of freedom are located at the vertices of the mesh. They use the most compact stencil for second order accuracy, namely the closest neighbors of a vertex. They are second order accurate in the sense that if the flow is smooth enough, the truncation error is $O(h^2)$, so that the error is formally second order even on irregular meshes. They are upwind, non oscilatory and parameter free. Last, it is possible to make them unconditionally stable without loosing their basic properties (monotonicity, upwind and second order accurate).

Our future work will follow two paths. First, how is it possible to introduce viscous terms without losing the accuracy ? A simple splitting is not sufficient, we refer to the work of Lerat et al. [28]. Second, how is it possible to increase the accuracy of theses scheme ? A possibility may be to follow the ideas of [6] : a fourth order scheme for steady scalar problems is introduced, it has all the desired properties.

A Energy-Stability of the Narrow scheme

This appendix contains the energy stability proof of the Narrow scheme for the equation

$$\frac{\partial W}{\partial t} + \sum_{i=1}^{2} \frac{\partial}{\partial x_i}(\mathbf{f}^{\,i}(W)) = 0 \quad (x,t) \in \mathbb{R}^2 \times \mathbb{R}^+, \tag{1.1}$$

where $W \in \mathbb{R}^m$ denotes the vector of conserved variables and $\mathbf{f} = (\mathbf{f}^{\,1}, \ldots, \mathbf{f}^{\,d}) : \mathbb{R}^m \to \mathbb{R}^{m\times 2}$ the flux vector, with $\mathbf{f}^{\,i} = A_i W$. The matrices A_i are assumed to be constant. In quasilinear form one has

$$\frac{\partial W}{\partial t} + \sum_{i=1}^{2} A_i \frac{\partial W}{\partial x_i} = 0 \tag{1.2}$$

Equation (1.1) is assumed to be equiped with a convex entropy extension so that the additional scalar inequality holds

$$\frac{\partial}{\partial t} H(W) + \sum_{i=1}^{2} \frac{\partial}{\partial x_i}(G^i(\mathbf{f}\,)) \le 0. \tag{1.3}$$

The existence of a convex entropy insures that the quasilinear form (1.2) is symmetrizable via the change of variables $W \mapsto V$ where $V = H_W^T$ the entropy variables. Performing this change of variables one obtains

$$\widetilde{A}_0 \frac{\partial V}{\partial t} + \sum_{i=1}^{2} \widetilde{A}_i \frac{\partial V}{\partial x_i} = 0, \tag{1.4}$$

where $\widetilde{A}_0 = \partial W/\partial V = (H_{WW})^{-1}$ is symmetric positive definite and the matrices $\widetilde{A}_i = \partial\, \mathbf{f}^{\,i}/\partial V = A_i A_0$ are symmetric. The Narrow scheme reads

$$\sum_{T, M_i \in T} \left(\frac{|T|}{3}(W_i^{n+1} - W_i^n) + \frac{\Delta t}{2} \sum_{M_j \in T} K_i^+ N K_j^- (W_j^{n+1} - W_i^{n+1} + W_j^n - W_i^n) \right) = 0. \tag{1.5}$$

In symmetrization variables one has

$$\sum_{T, M_i \in T} \left(\frac{|T|}{3}\widetilde{A}_o(V_i^{n+1} - V_i^n) + \frac{\Delta t}{2} \sum_{M_j \in T} \widetilde{K}_i^+ N \widetilde{K}_j^- (V_j^{n+1} - V_i^{n+1} + V_j^n - V_i^n) \right) = 0. \tag{1.6}$$

We first multiply this equation by $V_i^{n+1} + V_i^n$ to obtain

$$\begin{aligned} &\sum_{M_i \in T} \frac{|T|}{3} \langle \widetilde{A}_o(V_i^{n+1} - V_i^n), V_i^{n+1} + V_i^n \rangle \\ &+ \frac{\Delta t}{2} \sum_{T, M_i \in T} \sum_{M_j \in T} \langle \widetilde{K}_i^+ N \widetilde{K}_j^- (V_j^{n+1} - V_i^{n+1} + V_j^n - V_i^n), V_i^{n+1} + V_i^n \rangle = 0 \end{aligned} \tag{1.7}$$

Summing on the nodes of the mesh we get

$$\begin{aligned} &\sum_{T \in \tau_h} \Bigg(\sum_{M_i \in T} \frac{|T|}{3} < \widetilde{A}_o(V_i^{n+1} - V_i^n), V_i^{n+1} + V_i^n > \\ &+ \frac{\Delta t}{2} \sum_{M_i \in T} \sum_{M_j \in T} < \widetilde{K}_i^+ N \widetilde{K}_j^- (V_j^{n+1} - V_i^{n+1} + V_j^n - V_i^n), V_i^{n+1} + V_i^n > \Bigg) = 0 \end{aligned} \tag{1.8}$$

Using the fact that $\widetilde{A}_0$ is symmetric, one has

$$\begin{aligned} < \widetilde{A}_o(V_i^{n+1} - V_i^n), V_i^{n+1} + V_i^n > &= < \widetilde{A}_o V_i^{n+1}, V_i^{n+1} > + < \widetilde{A}_o V_i^{n+1}, V_i^n > - < \widetilde{A}_o V_i^n, V_i^n > - < \widetilde{A}_o V_i^n, V_i^{n+1} > \\ &= < \widetilde{A}_o V_i^{n+1}, V_i^{n+1} > - < \widetilde{A}_o V_i^n, V_i^n > \end{aligned}$$

The second term of (1.8) can be written as

$$\sum_{M_i\in T}\left(\sum_{M_j\in T} < \widetilde{K}_i^+ N\widetilde{K}_j^-(V_j^{n+1}-V_i^{n+1}+V_j^n-V_i^n), V_i^{n+1}+V_i^n >\right) =$$

$$\sum_{M_i\in T}\left(\sum_{M_j\in T} < \widetilde{K}_i^+ N\widetilde{K}_j^-(V_j^{n+1}+V_j^n), V_i^{n+1}+V_i^n > - \sum_{M_j\in T} < \widetilde{K}_i^+ N\widetilde{K}_j^-(V_i^{n+1}+V_i^n), V_i^{n+1}+V_i^n >\right)$$

$$+\sum_{M_i\in T}\left(\sum_{M_j\neq M_i} < \widetilde{K}_i^+ N\widetilde{K}_j^-(V_j^{n+1}+V_j^n), V_i^{n+1}+V_i^n > + < \widetilde{K}_i^+(V_i^{n+1}+V_i^n), V_i^{n+1}+V_i^n >\right)$$

In equation (1.8) we can add the term $-\sum_{T\in\tau_h}\sum_{i\in T} < \widetilde{K}_i(V_i^{n+1}+V_i^n), V_i^{n+1}+V_i^n >$ because this term is equal to zero since the geometry surrounding a vertex is closed.

$$\begin{aligned}
\sum_{T\in\tau_h}\sum_{M_i\in T} < \widetilde{K}_i(V_i^{n+1}+V_i^n), V_i^{n+1}+V_i^n > &= \sum_{M_i\in\tau_h}\sum_{T,M_i\in T} < \widetilde{K}_i(V_i^{n+1}+V_i^n), V_i^{n+1}+V_i^n > \\
&= \sum_{M_i\in\tau_h} < \left(\sum_{T,M_i\in T}\widetilde{K}_i\right)(V_i^{n+1}+V_i^n), V_i^{n+1}+V_i^n > \\
&= 0
\end{aligned}$$

Equation (1.8) becomes

$$\sum_{T\in\tau_h}\sum_{M_i\in T}\frac{|T|}{3}(< \widetilde{A}_o V_i^{n+1}, V_i^{n+1} > - < \widetilde{A}_o V_i^n, V_i^n >)$$

$$+\frac{\Delta t}{2}\sum_{T\in\tau_h}\sum_{M_i\in T}\left(\frac{1}{2} < |\widetilde{K}_i|(V_i^{n+1}+V_i^n), V_i^{n+1}+V_i^n >\right. \tag{1.9}$$

$$\left. + \sum_{M_j\in T} < \widetilde{K}_i^+ N\widetilde{K}_j^-(V_j^{n+1}+V_j^n), V_i^{n+1}+V_i^n >\right) = 0$$

Moreover one has

$$\begin{aligned}
\sum_{M_i\in T}\sum_{M_j\in T} < \widetilde{K}_i^+ N\widetilde{K}_j^-(V_j^{n+1}+V_j^n), V_i^{n+1}+V_i^n > &= \sum_{M_i\in T}\sum_{M_j\in T} < V_j^{n+1}+V_j^n, \widetilde{K}_j^- N\widetilde{K}_i^+(V_i^{n+1}+V_i^n) > \\
&= \sum_{M_j\in T}\sum_{M_i\in T} < \widetilde{K}_j^- N\widetilde{K}_i^+(V_i^{n+1}+V_i^n), V_j^{n+1}+V_j^n >
\end{aligned}$$

Equation (1.9) becomes

$$\sum_{T\in\tau_h}\sum_{M_i\in T}\frac{|T|}{3}(< \widetilde{A}_o V_i^{n+1}, V_i^{n+1} > - < \widetilde{A}_o V_i^n, V_i^n >)$$

$$+\frac{\Delta t}{2}\sum_{T\in\tau_h}\sum_{M_i\in T}\left(\frac{1}{2} < |\widetilde{K}_i|(V_i^{n+1}+V_i^n), V_i^{n+1}+V_i^n >\right.$$

$$\left. + \sum_{M_j\in T}\frac{1}{2} < \widetilde{K}_i^+ N\widetilde{K}_j^-(V_j^{n+1}+V_j^n), V_i^{n+1}+V_i^n > + \frac{1}{2} < \widetilde{K}_i^- N\widetilde{K}_j^+(V_j^{n+1}+V_j^n), V_i^{n+1}+V_i^n >\right) = 0 \tag{1.10}$$

Next, rewrite

$$\widetilde{K}_i^+ N \widetilde{K}_j^- + \widetilde{K}_i^- N \widetilde{K}_j^+$$

in the following form

$$K_i^+ N \widetilde{K}_j^- + \widetilde{K}_i^- N \widetilde{K}_j^+ = \widetilde{K}_i N \widetilde{K}_j - \widetilde{K}_i^+ N \widetilde{K}_j^+ + \widetilde{K}_i^- N \widetilde{K}_j^-,$$

leads to

$$\begin{aligned}
&\sum_{T \in \tau_h} \sum_{M_i \in T} \frac{|T|}{3} (< \widetilde{A}_o V_i^{n+1}, V_i^{n+1} > - < \widetilde{A}_o V_i^n, V_i^n >) \\
&\frac{\Delta t}{2} \sum_{T \in \tau_h} \sum_{M_i \in T} \Bigg(\frac{1}{2} < \widetilde{K}_i N \widetilde{K}_j (V_j^{n+1} + V_j^n), V_i^{n+1} + V_i^n > \\
&+ \frac{1}{2} < \widetilde{K}_i^+ (V_i^{n+1} + V_i^n), V_i^{n+1} + V_i^n > - \frac{1}{2} \sum_{M_j \in T} < \widetilde{K}_i^+ N \widetilde{K}_j^+, (V_j^{n+1} + V_j^n), V_i^{n+1} + V_i^n > \\
&+ \frac{1}{2} < -\widetilde{K}_i^- (V_i^{n+1} + V_i^n), V_i^{n+1} + V_i^n > + \frac{1}{2} \sum_{M_j \in T} < \widetilde{K}_i^- N \widetilde{K}_j^-, (V_j^{n+1} + V_j^n), V_i^{n+1} + V_i^n > \Bigg) = 0
\end{aligned} \tag{1.11}$$

Finally one obtains

$$\begin{aligned}
&\sum_{T \in \tau_h} \sum_{M_i \in T} \frac{|T|}{3} (< \widetilde{A}_o V_i^{n+1}, V_i^{n+1} > - < \widetilde{A}_o V_i^n, V_i^n >) \\
&+ \sum_{T \in \tau_h} \big(Q_a^T(V_1, V_2, V_3) + Q_b^T(V_1, V_2, V_3) + Q_c^T(V_1, V_2, V_3) \big) = 0
\end{aligned} \tag{1.12}$$

In [4], it is shown that Q_a^T, Q_b^T and Q_c^T are positive quadratic forms : the system N scheme is energy stable.

References

[1] R. Abgrall. On Essentially Non-Oscillatory Schemes on Unstructured Meshes : Analysis and Implementation. *J. Comp. Phys.*, 114(1):45–54, September 1994.

[2] R. Abgrall. Toward the ultimate conservative scheme : Following the quest. *J. Comput. Phys.*, accepted for publication.

[3] R. Abgrall and T. Barth. Weighted residual distribution schemes for conservation laws via adaptive quadrature. *Siam J. Sci. Comput.*, 2002. accepted for publication.

[4] R. Abgrall, K. Mer, and B. Nkonga. A lax–wendroff type theorem for residual schemes. In M. Haffez and J.J Chattot, editors, *Innovative methods for numerical solutions of partial differential equations*, pages 243–266. World Scientific, 2002.

[5] R. Abgrall and M. Mezine. Construction of second order accurate monotone and stable residual schemes : the steady case. Technical report, MAB, 2001. submitted to J. Comput. Phys.

[6] R. Abgrall and P.L. Roe. Construction of very high order fluctuation schemes. Technical report, MAB, 2001. submitted in J. Sci. Comp.

[7] G. Ben-Dor. *Shock Wave Reflection Phenomena*. Springer–Verlag, Berlin/New York, 1991.

[8] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35(6):2440–2463 (electronic), 1998.

[9] P. Collela and P. Woodward. The numerical simulation of two-dimensionnal fluid with strong shocks. *J. Comput. Phys.*, 54:115–173, 1984.

[10] A. Csík, M. Ricchiuto, H. Deconinck, and S. Poedts. Space-time residual distribution schemes for hyperbolic conservation laws, 2001. manque des choses.

[11] H. Deconinck, P.L. Roe, and R. Struijs. A multidimensional generalisation of Roe's difference splitter for the Euler equations. *Computer and Fluids*, 22(2/3):215–222, 1993.

[12] H. Deconinck, K. Sermeus, and R. Abgrall. Status of multidimensional upwind residual distribution schemes and applications in aeronautics. AIAA paper 2000-2328, June 2000. AIAA CFD Conference, Denver (USA).

[13] H. Deconinck, R. Struijs, G. Bourgeois, and P.L. Roe. Compact advection schemes on unstructured meshes. VKI Lecture Series 1993–04, Computational Fluid Dynamics, 1993.

[14] H. Deconinck, R. Struijs, G. Bourgeois, and P.L. Roe. Compact advection schemes on unstructured meshes. VKI Lecture Series 1993–04, Computational Fluid Dynamics, 1993.

[15] A. Ferrante. Solution of the unsteady euler equations using residual distribution and flux correcte transport. Technical report, von Kàrman Institute, Chaussée de Waterloo, 72, B-1640 Rhode Saint Genèse, Belgium, 1997. Project report.

[16] L. Fezoui and B. Stoufflet. A class of implicit upwind schemes for Euler simulations with unstructured meshes. *J. Comput. Phys.*, 84(1):174–206, 1989.

[17] O. Friedrich. Weighted essentially non-oscillatory schemes for the interpolation of mean values on unstructured grids. *J. Comput. Phys.*, 144(1):194–212, 1998.

[18] E. Godlewski and P.A. Raviart. *Hyperbolic systems of conservations laws, I.* Ellipse, 1991.

[19] E. Godlewski and P.A. Raviart. *Hyperbolic systems of conservations laws.* Applied Mathematical Sciences. Springer Verlag, 1995.

[20] A. Harten, S. Osher, B. Engquist, and S.R. Chakravarthy. Uniformly high–order accurate non–oscillatory schemes iii. *J. Comput. Phys.*, 71:231–303, 1987.

[21] C. Hu and C.W. Shu. Weighted essentially non-oscillatory schemes on triangular meshes. *J. Comput. Phys.*, 150(1):97–127, 1999.

[22] Th.J.R. Hughes and M. Mallet. A new finite element formulation for Computational Fluid Dynamics : III the generalized streamline operator for multidimensional advective–diffusive systems. *Computer Methods in Applied Mechanics and Engineering*, 58:305–328, 1986.

[23] Th.J.R. Hughes and M. Mallet. A new finite element formulation for Computational Fluid Dynamics : IV a discontinuity–capturing operator for multidimensional advective–diffusive systems. *Computer Methods in Applied Mechanics and Engineering*, 58:329–336, 1986.

[24] M.Y. Hussaini, B. van Leer, and J. Van Rosendal, editors. *Upwind and High Resolution Schemes.* Springer Verlag, 1997.

[25] G.-S. Jiang and C.-W. Shu. Efficient implementation of weighted eno schemes. *J. Comput. Phys.*, 126:202–228, 1996.

[26] C. Johnson, U. Nävert, and J. Pitkäranta. Finite element methods for linear hyperbolic problems. *Computer methods in applied mechanics and engineering*, pages 285–312, 1984.

[27] P. Lascaux and R. Théodor. *Analyse numérique matricielle appliquée à l'art de l'ingénieur.* Masson, Paris, France, 1986.

[28] A. Lerat and Ch. Corre. A residual-based compact scheme for the compressible Navier-Stokes equations. *J. Comput. Phys.*, 170(2):642–675, 2001.

[29] K. R. Meadows, A. Kumar, , and M. Y. Hussaini. A computational study of the interaction between a vortex and a shock wave. AIAA Paper 89-1043, 1989.

[30] H. Paillère. *Multidimensional Upwind residual Discretisation Schemes for the Euler and Navier Stokes Equations on Unstructured Meshes.* PhD thesis, Université Libre de Bruxelles, 1995.

[31] S. P. Pao and M. D. Salas. AIAA Paper 81-1205, 1981.

[32] J.J. Quirk. A contribution to the great riemann solver debate. *Int. J. Numer. Methods Fluids*, 18(555-574), 1994.

[33] P.L. Roe and D. Sidilkover. Optimum positive linear schemes for advection in two and three dimensions. *Siam J. Numer. Anal.*, 29(6):1542–1588, 1992.

[34] R. Struijs, H. Deconinck, and P. L. Roe. Fluctuation Splitting Schemes for the 2D Euler equations. *VKI LS 1991-01, Computational Fluid Dynamics*, 1991.

[35] E. van der Weide and H. Deconinck. Positive matrix distribution schemes for hyperbolic systems. In *Computational Fluid Dynamics '96*, pages 747–753. Wiley, 1996.

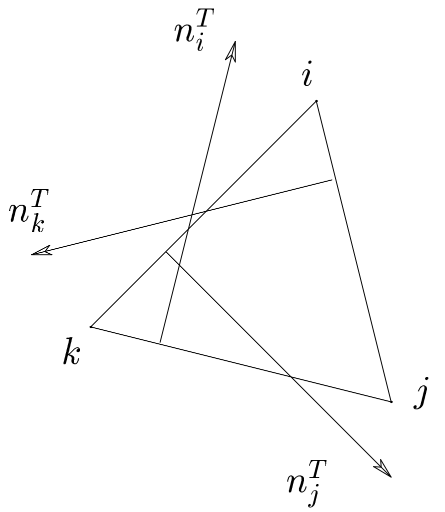

Figure 1: Triangle with scaled inward normals.

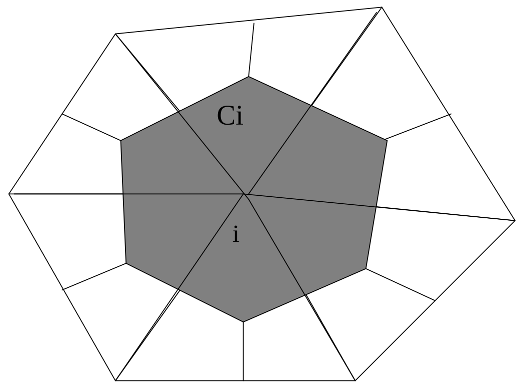

Figure 2: The dual cell is obtained by joining the midpoints of the edges starting from M_i and the centroids of the triangles containing M_i as a vertex.

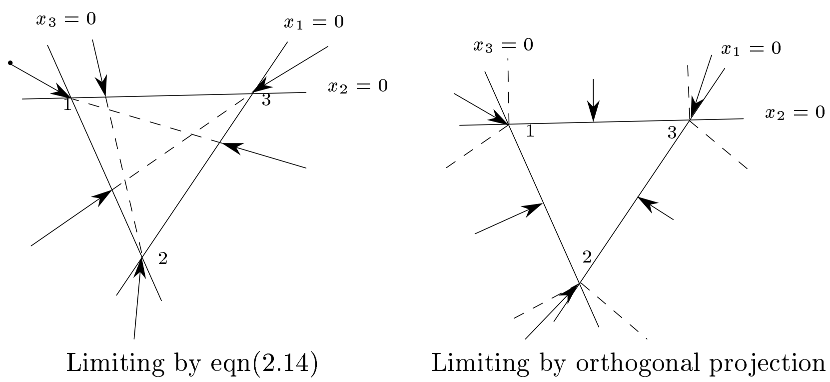

Figure 3: Geometrical illustration of the two limiters

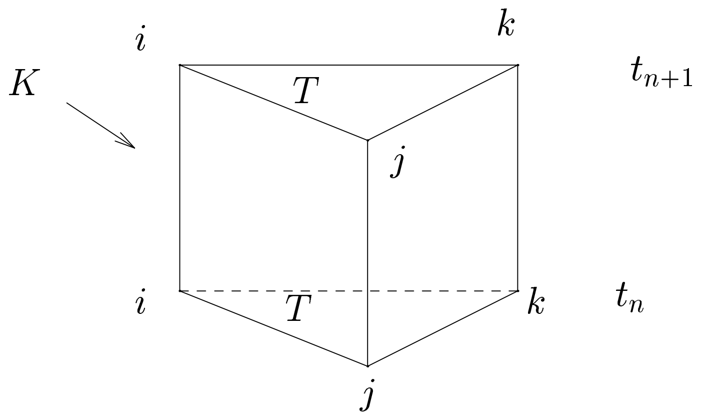

Figure 4: the prism K defined by the triangle T

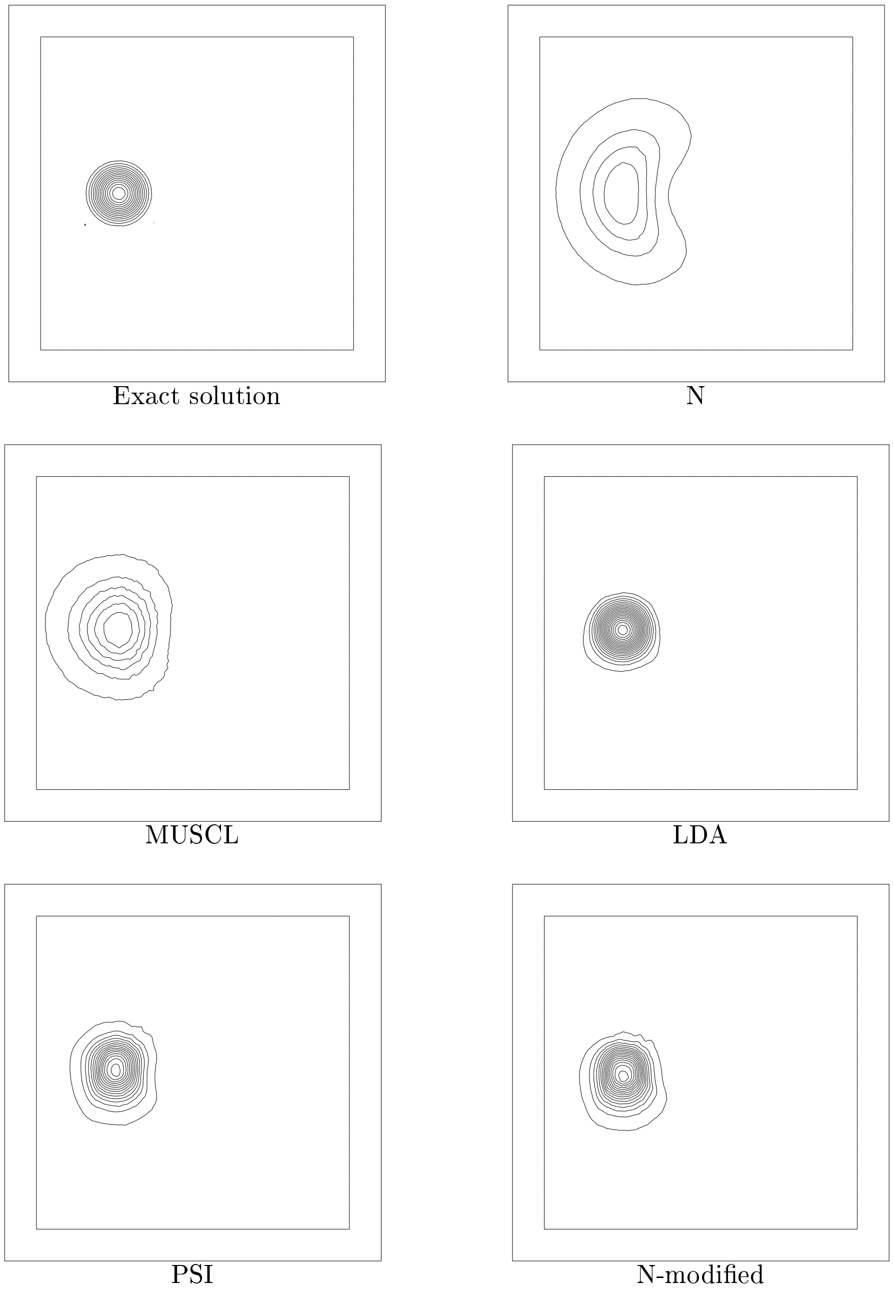

Figure 5: Solutions for the rotating cosine hill after one revolution

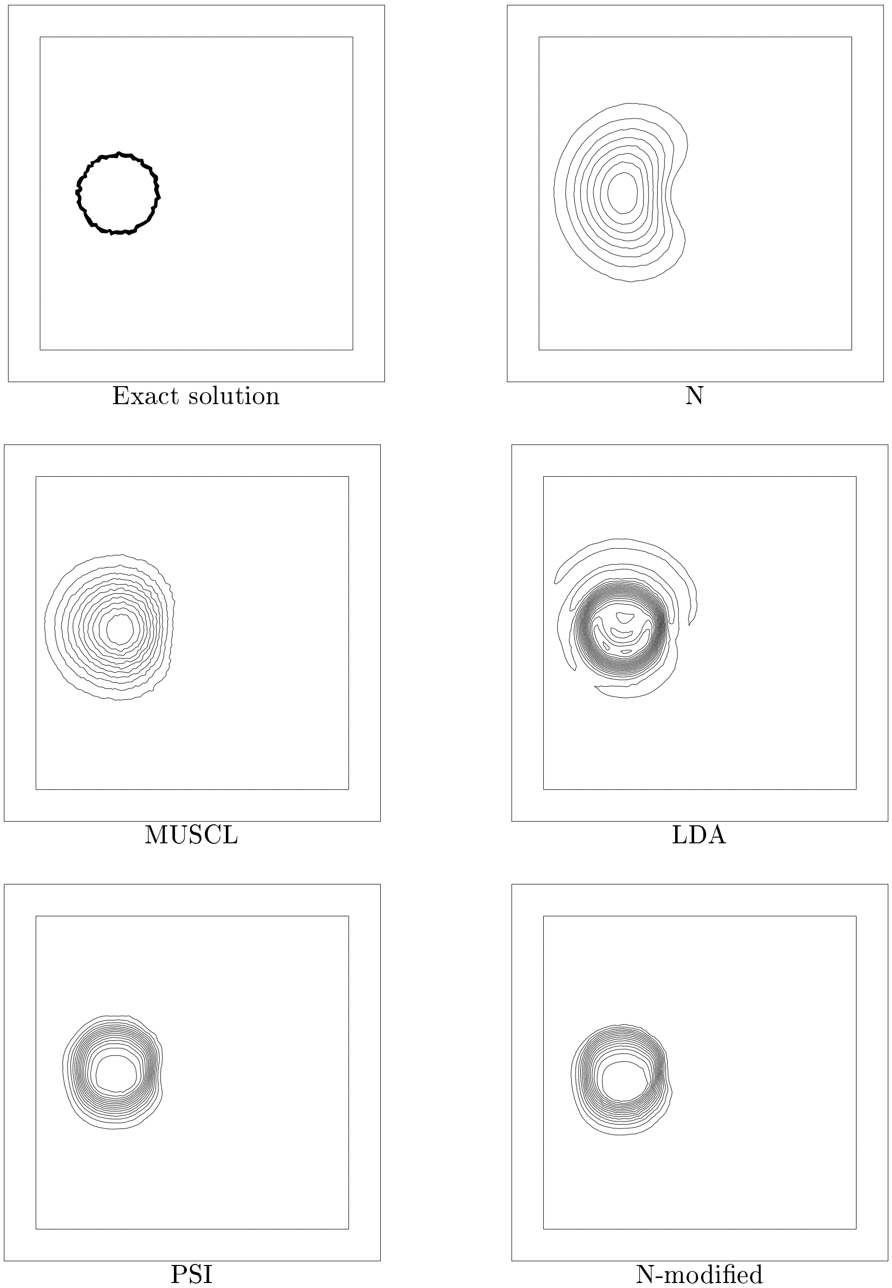

Figure 6: Solutions for the rotating cylinder after one revolution.

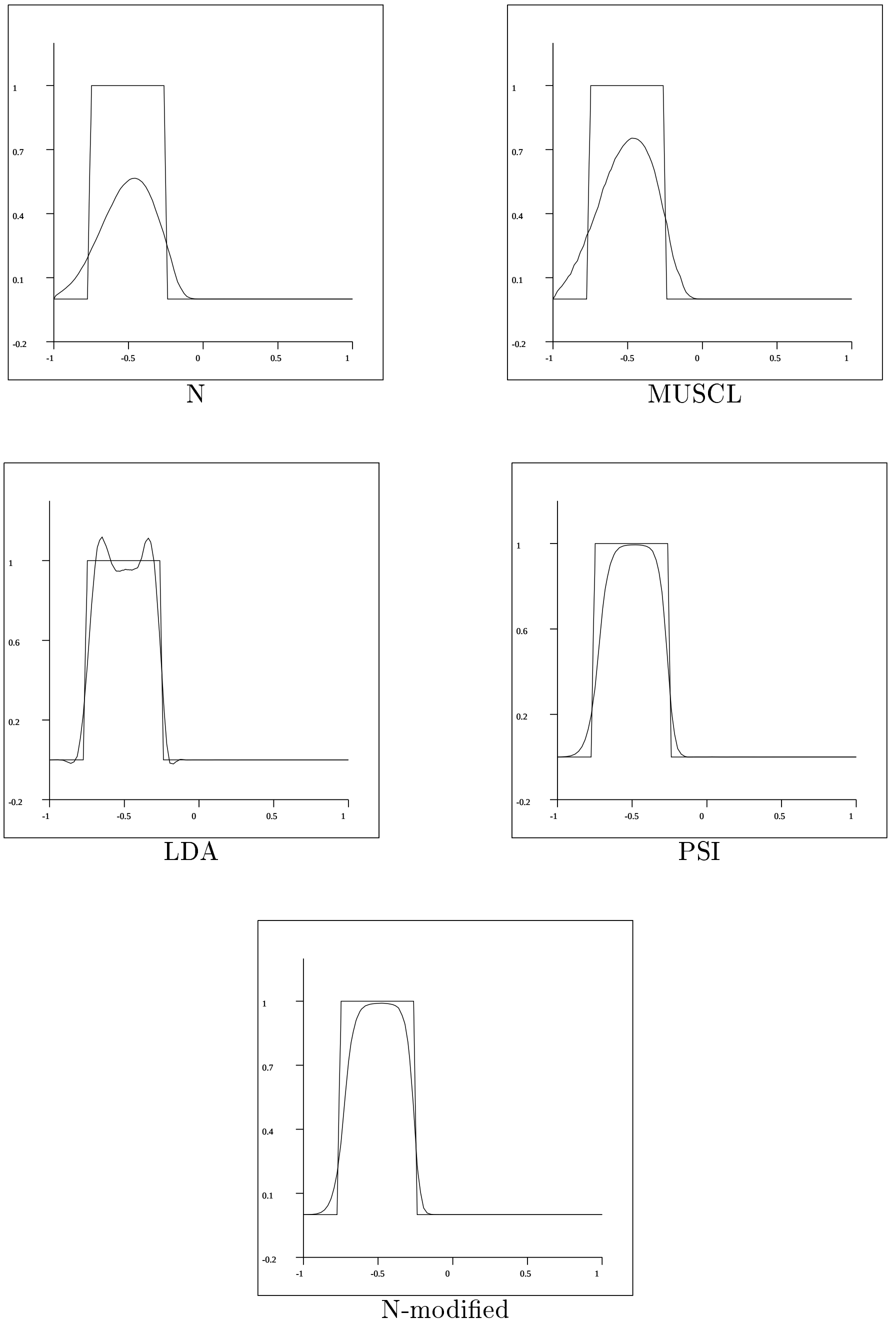

Figure 7: Sections at y=0 of the solutions for the rotating cylinder

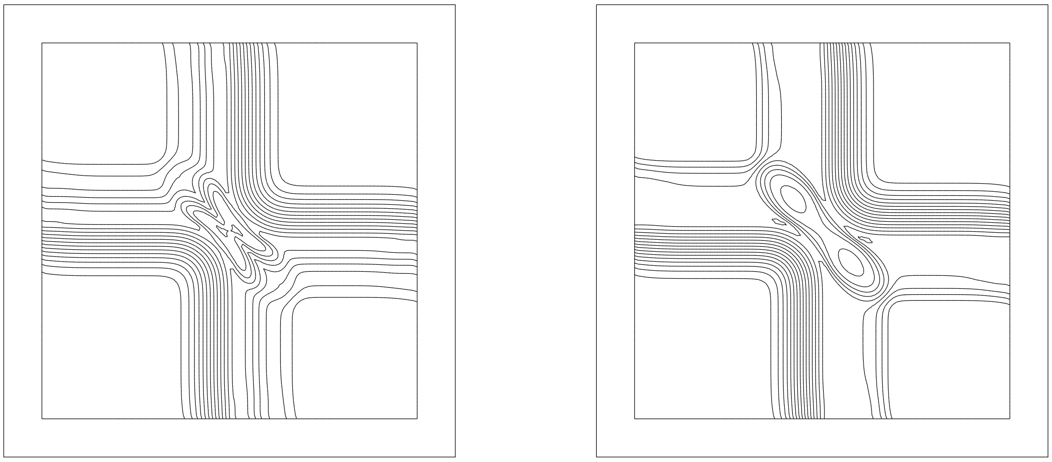

Figure 8: 2D Riemann problem computed by the N-scheme at time t=0.2, density (left) and pressure (right).

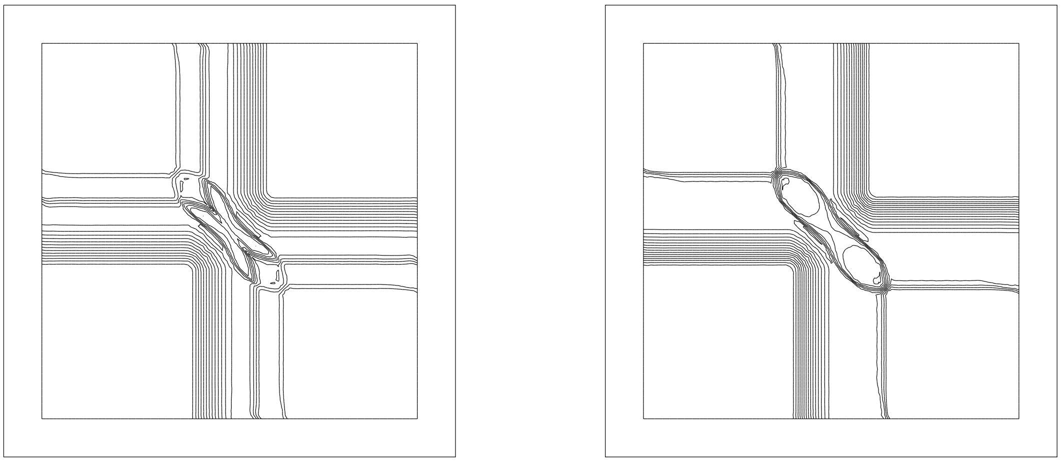

Figure 9: 2D Riemann problem computed by the N-modified scheme at time t=0.2, density (left) and pressure (right).

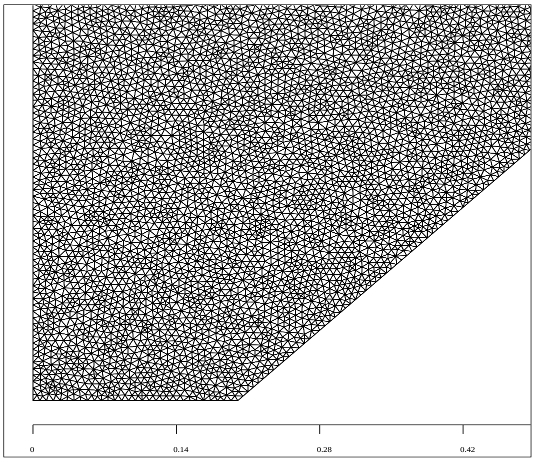

Figure 10: Part of the unstructured grids

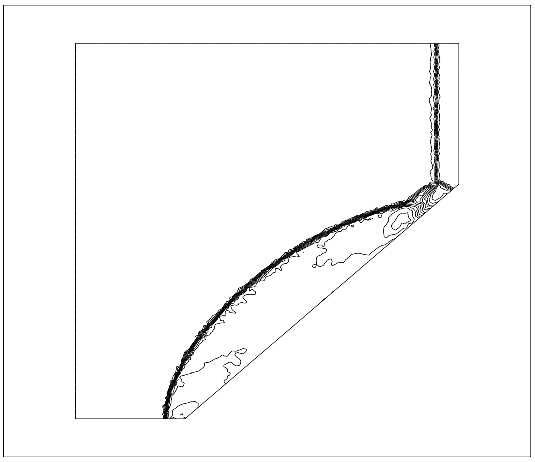

Figure 11: Reflection of a planar shock from a ramp. Density 20 contour lines from 1.18 to 20.12.

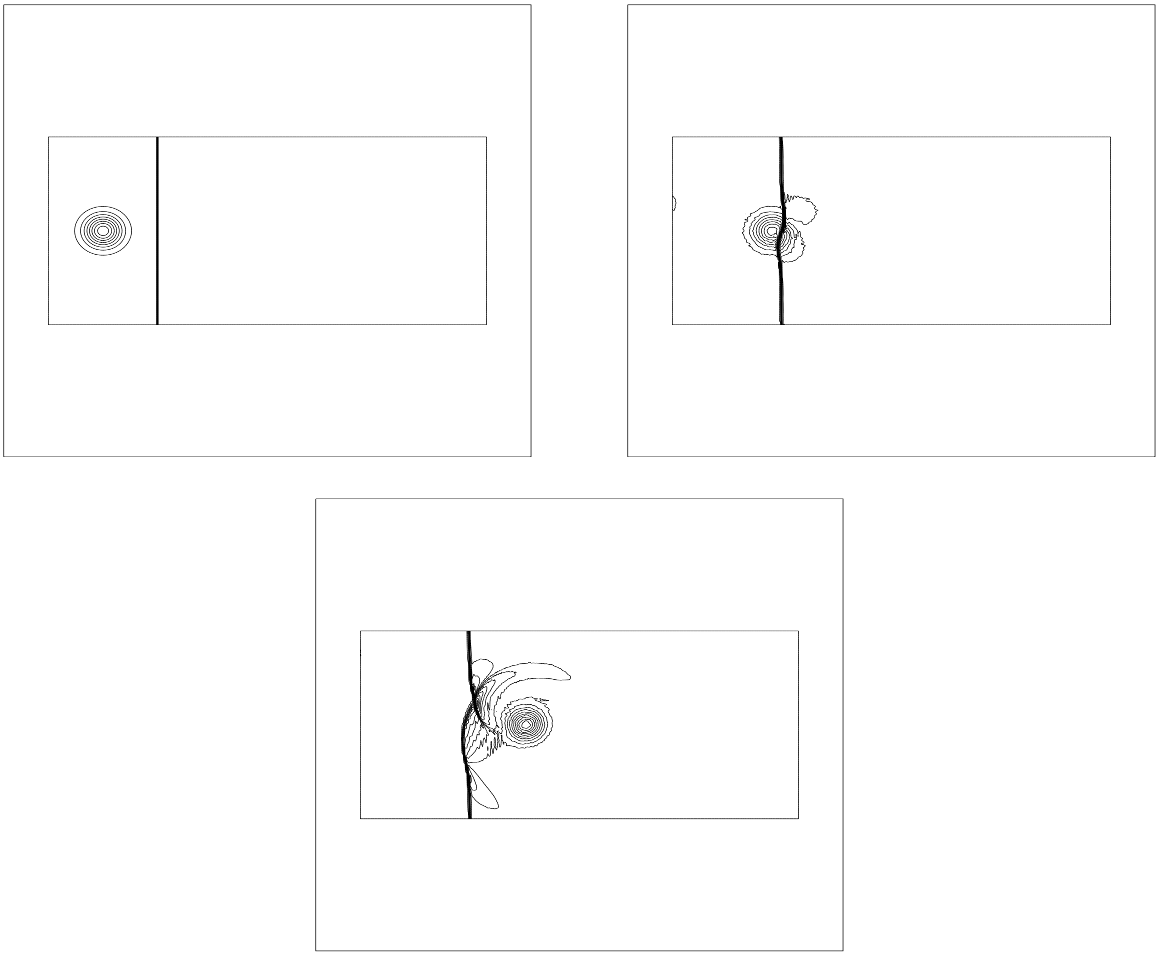

Figure 12: Shock vortex interaction. Pressure. N-modified scheme. 30 contours lines from 0.84 to 1.4. Top left: t=0; Top right: t=0.2, Bottom: t=0.4.

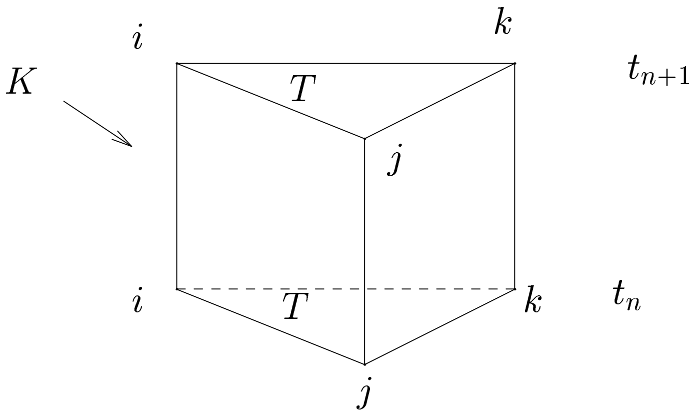

Figure 13: Prismatic element K

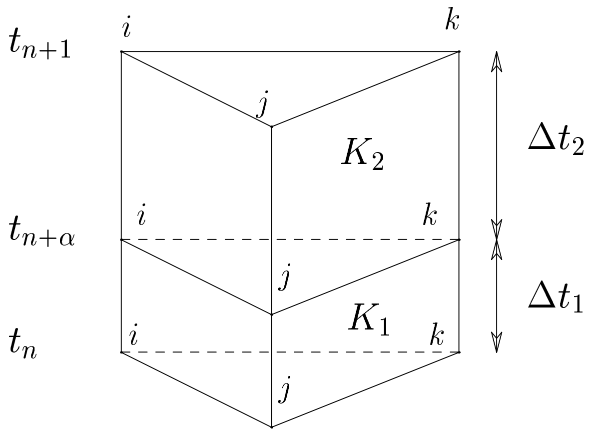

Figure 14: Space-time grid

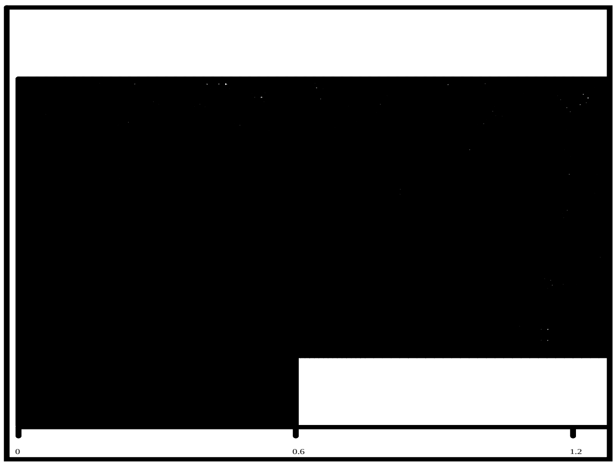

Figure 15: Part of the unstructured grid.

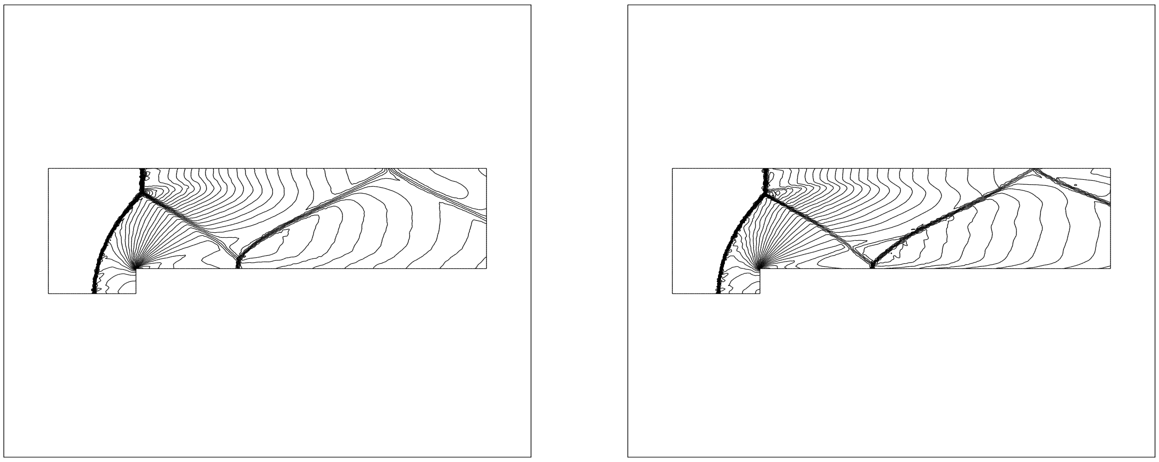

Figure 16: Forward-facing step problem. Density iso-lines: 30 equally spaced contour lines from 0.090338 to 6.2365. Left : N scheme. Right : N-modified scheme.

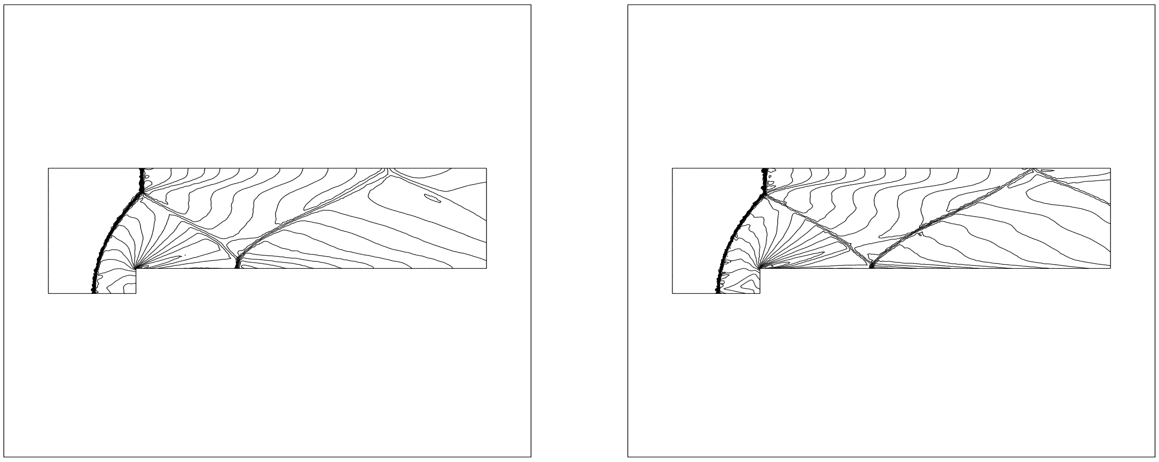

Figure 17: Forward-facing step problem. Mach number iso-lines: 25 equally spaced contour lines from 0.02 to 3.82. Left : N scheme. Right : N-modified scheme.

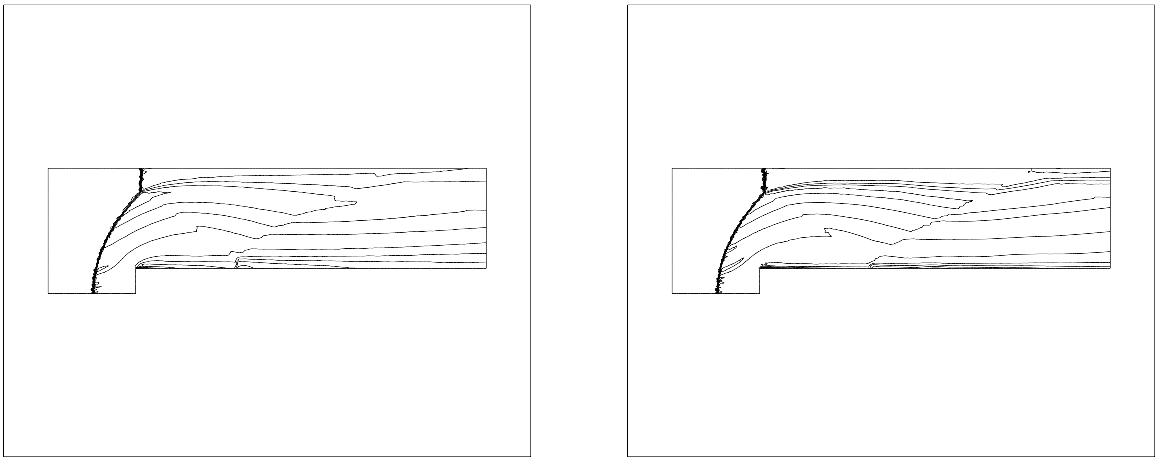

Figure 18: Forward-facing step problem. Entropy production near the step corner : 17 equally spaced contour lines from 0.63 to 1.5. Left : N scheme. Right : N-modified scheme.

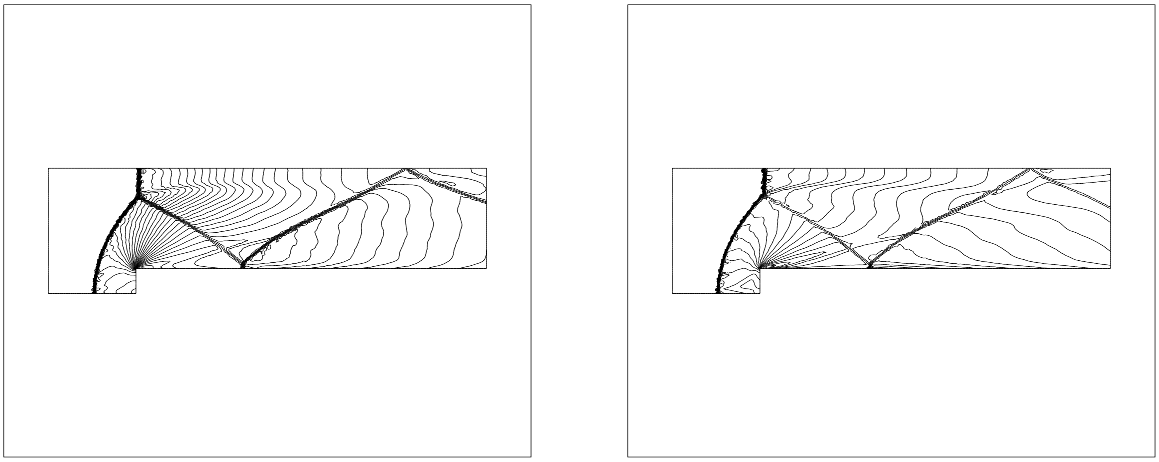

Figure 19: Forward-facing step problem. Left: Density iso-lines: 30 equally spaced contour lines from 0.090338 to 6.2365. Right: Mach number iso-lines: 25 equally spaced contour lines from 0.02 to 3.82.

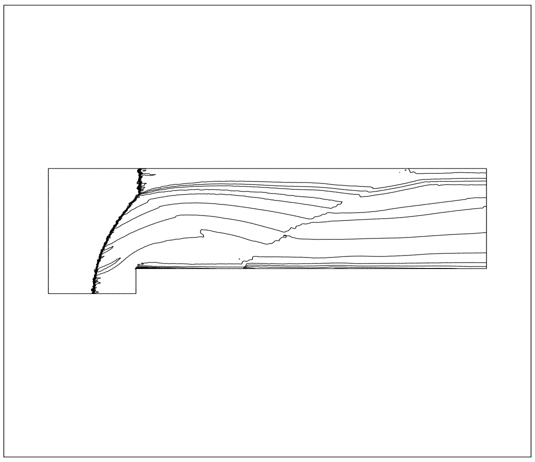

Figure 20: Forward-facing step problem. Entropy production near the step corner : 17 equally spaced contour lines from 0.63 to 1.5. Left : N scheme. Right : N-modified scheme.