ERROR ESTIMATE AND THE GEOMETRIC CORRECTOR FOR THE UPWIND FINITE VOLUME METHOD APPLIED TO THE LINEAR ADVECTION EQUATION

DANIEL BOUCHE*, JEAN-MICHEL GHIDAGLIA†, AND FRÉDÉRIC PASCAL‡

Abstract. This paper deals with the upwind finite volume method applied to the linear advection equation on a bounded domain and with natural boundary conditions. We introduce what we call the geometric corrector which is a sequence associated with every finite volume mesh in $\mathbf{R}^{nd}$ and every non vanishing vector $\mathbf{a}$ of $\mathbf{R}^{nd}$. First we show that if the continuous solution is regular enough and if the norm of this corrector is bounded by the mesh size then an order one error estimate for the finite volume scheme occurs. Then we prove that this norm is indeed bounded by the mesh size in various cases including the one where an arbitrary coarse conformal triangular mesh is uniformly refined in two dimension. Computing numerically exactly this corrector allows us to state that this result might be extended under conditions to more general cases like the one with independent refined meshes.

Key words. Finite Volume Method ; Consistency and Accuracy ; Geometric Corrector ; Unstructured Meshes

AMS subject classifications. 65M06, 65M12, 65M15, 65M50

1. Introduction. Finite Volume Methods (FVM) have been first used in the context of computational mechanics in situations in which solutions present discontinuities (see the monograph of Kröner [16], Godlewski and Raviart [13], Eymard *et al* [10]). A reason for this is that they fundamentally rely on an integral version of the equations. This contrasts with Finite Difference Methods (FDM) for which smoothness of solutions is used in order to approximate derivatives by differential quotients. For Finite Element Methods (FEM), the situation is somewhat similar since traces theorems on hyper-surfaces in H^1-type spaces exclude also the approximation of discontinuous solutions. It is the main reason why FVM are widely used for the approximation of hyperbolic systems of conservation laws. Indeed, even when initial and boundary data are smooth, solutions of such equations produce in finite time discontinuous solutions. Let us also mention that other advantages of FVM, like compact numerical stencil, effectiveness on unstructured grids, simplicity in coding or data structures, are reasons for which these methods are more and more used for Elliptic and Parabolic equations, albeit smoothness of the solution is guarantee.
Concerning FVM for hyperbolic equations, there is a lot of methods and especially in Computational Fluid Dynamics ; historically these methods were natural extensions of the famous Godunov method for the Euler equations of gas dynamics to multidimensional problems. The Godunov method was known to be too diffusive and this was attributed to the fact that it is a first order method. In the beginning of the eighties, Bram Van Leer proposed MUSCL methods which can be seen as corrections that lead to second order schemes.

*Commissariat à l'Énergie Atomique, DPTA, BP 12, F-91680 Bruyères-le-Châtel (daniel.bouche@cea.fr) and CMLA, ENS Cachan et CNRS UMR 8536, 61 avenue du Président Wilson, F-94235 Cachan Cedex.

†Centre de Mathématiques et de Leurs Applications, ENS Cachan et CNRS UMR 8536, 61 avenue du Président Wilson, F-94235 Cachan Cedex (jmg@cmla.ens-cachan.fr).

‡Laboratoire de Mathématiques, UMR 8628, CNRS et Université Paris-Sud, Bâtiment 425, Université Paris-Sud, F-91405 Orsay (Frederic.Pascal@math.u-psud.fr) and CMLA, ENS Cachan et CNRS UMR 8536, 61 avenue du Président Wilson, F-94235 Cachan Cedex.

However, it is an open problem to determine the optimal rate of convergence of FVM. This contrasts strongly with FDM and FEM where such rates are known. Hence there is an apparent paradox between the fact that some methods are termed as "first order" or even "second order" and the fact that corresponding error estimates are not proved. Indeed, there is some confusion in the literature : what is usually called "first order" method for example corresponds to the order of the truncation error on *uniform cartesian* grids. In practice, and this is the main interest of FVM, these methods are used on unstructured meshes and therefore it is not known whether "first order" method leads to first order error estimates. In fact the situation is even more puzzling since "first order" methods on unstructured grids *do not lead in general to a truncation error that goes to zero* as the mesh size goes to zero (see the discussion at the end of Section 3.3).
Even for the scalar linear advection equation, to obtain *a priori* optimal error estimates is still a challenging task. One of the main difficulties holds, as said already, in the fact that the non uniformity of the mesh brings up an apparent loss of consistency as it has been observed by Hoffman [14], Turkel [27] and Pike [24]. In fact this loss of consistency is an artifact of standard convergence proof : for instance the Lax theorem states that stability and consistency are sufficient conditions for a scheme to be convergent at the same rate than the truncation error converges to zero. Actually consistency is not necessary, the scheme maintains the accuracy and the global error behaves better than the local error would indicate. This property of enhancement of the truncation error is called supra-convergence and for second and higher order boundary value problems, this phenomenon discovered by Tikhonov and Samarskij [26] was widely analyzed in various cases by Manteuffel and his co-authors in [19], [20], [15], [21] and in Garcia-Archilla and Sanz-Serna [11].

In these papers, the analysis relies on the fact that the truncation error (defined by applying the discrete operator to the exact solution) can be rewritten in the special form $L_h\tau_1 + \tau_2$ where τ_1 and τ_2 are of the optimal order $\mathcal{O}(h^p)$, L_h is the discrete operator and $L_h\tau_1$ is only $\mathcal{O}(h^{p-1})$. Then an optimal discrete energy estimate for the global error can be derived. This idea is extended to finite difference schemes for hyperbolic systems with non uniform one dimensional grid and with irregular cartesian multidimensional grid in Levermore *et al* [18] and in Monk and Süli [22].

This way to rewrite the truncation error can be seen as a correction of the error in order to cancel the leading part of the local error due to the non uniformity of the mesh. Actually in Wendroff and White [31], [30] and Wendroff [29], a grid function is introduced for suitably smooth solutions of hyperbolic system in order to prove the optimal order of convergence of upwind finite difference and Lax-Wendroff schemes in one dimension space and in two dimension when an Alternate Direction Method is used. Berger *et al* [4], [3] successfully used the idea to get the optimal rate of convergence of the "h-box" scheme defined on a non-uniform cartesian grid with embedded irregular small cells.
Recently Després [9], [8], by using energy estimates, has given an elegant proof of an order $h^{1/2}$ error estimates with respect to the L^2 norm for the linear advection equation on regular finite element meshes in the particular case of periodic boundary conditions.
Finally, let us mention that when no regularity conditions are imposed on the exact solution, by suitably application of Kuznetsov [17] approximation theory for instance, it is possible to establish rate of convergence of $h^{1/2}$ in the L^1 norm for non regular cartesian grid : see Eymard *et al* [10], Cockburn and Gremaud [5], [6], [7], Vila and

Villedieu [28], Teng and Zhang [25]. All these results deal with data (and solutions) which belongs to the set of functions with bounded variations. As shown in the aforementioned references, the $h^{1/2}$ is optimal for these solutions. However these papers raise the question whether such a rate is due to the irregularity of the mesh. Our present work answers to this open question by showing that it is not the case : we show that on irregular meshes, provided the solution is smooth, the error estimate behaves like h. Hence the poor convergence behavior is due to the lake of smoothness of solutions.

In this paper we study the initial and boundary value problem for the linear advection equation posed on a polygonal domain of $\mathbf{R}^{nd}$ (nd is the space dimension of the problem under consideration). We construct first what we call "geometric correctors". They form a family of vectors in $\mathbf{R}^{nd}$, $(\mathbf{\Gamma}_j)_{j=1,\dots,N}$, where N is the number of control volumes. This set of geometric correctors depends only on the mesh and on the advection vector but not on the solutions to the advection equation.
Our first result in Section 3.5 shows that if the solution is regular and if this family of vectors is uniformly bounded by the mesh size *i.e.* $||\mathbf{\Gamma}_j|| \leq Ch$ then under a Courant-Friedrichs-Lewy (CFL) condition, the classical explicit first order upwind scheme for the advection equation is indeed first order : $||u_h - u|| \leq C'h$ where u is the exact solution and u_h the FVM approximation.
Since we are able to compute exactly the $(\mathbf{\Gamma}_j)_{j=1,\dots,N}$, numerical simulations allow us to study in which cases the hypothesis $||\mathbf{\Gamma}_j|| \leq Ch$ is satisfied. We prove in Section 4.2 this estimation in various cases including the one where an arbitrary coarse conformal triangular mesh is refined. Finally, we present numerical experiments that lead us to conjecture that this result holds true in the case of independent refined meshes if the advection vector is not parallel to a side of the polygonal domain. On the other hand, on the basis of numerical evidence, we conjecture that if the advection vector is parallel to a side of the polygonal domain then the best estimate should be in $h^{1/2}$.

2. The continuous problem. We consider Ω a polygonal domain in $\mathbf{R}^{nd}$ with $nd \geq 1$ and we denote by $\mathbf{n}$ the unitary external normal vector on $\partial\Omega$. Let $\mathbf{a}$ be a non zero vector and let denote $\partial\Omega^- = \{x \in \partial\Omega, \mathbf{a}\cdot\mathbf{n}(x) < 0\}$. Given a function φ defined on Ω and a function ψ defined on $\partial\Omega^- \times [0,+\infty[$, the initial and boundary value problem for the advection equation on $\bar{\Omega} \times [0,+\infty[$ reads :

$$\frac{\partial u}{\partial t} + (\mathbf{a}\cdot\nabla)u = 0\,, \quad (x,t) \in \Omega\times]0,+\infty[\,, \tag{2.1}$$

$$u(x,0) = \varphi(x)\,, \quad x \in \Omega\,, \tag{2.2}$$

$$u(x,t) = \psi(x,t)\,, \quad (x,t) \in \partial\Omega^- \times [0,+\infty[\,. \tag{2.3}$$

As it well known and understood, this problem has a unique smooth solution provided the data φ and ψ are smooth and satisfy the so called compatibility conditions (for example the first compatibility condition is given below in (2.6)).
In fact there are various methods for obtaining this solution. The first one, which uses functional analysis, is due to Bardos [1]. It consists in considering smooth solutions to the parabolic equation ($\nu > 0$) :

$$\frac{\partial u^\nu}{\partial t} + (\mathbf{a}\cdot\nabla)u^\nu = \nu\Delta u^\nu\,, (x,t) \in \Omega\times]0,+\infty[\,, \tag{2.4}$$

with the complete Dirichlet boundary conditions :

$$u^\nu(x,t) = \psi^\nu(x,t)\,, (x,t) \in \partial\Omega \times [0,+\infty[\,, \tag{2.5}$$

where ψ^ν is a smooth extension of ψ to $\Omega \times [0, +\infty[$. Then the solution to (2.1)-(2.3) is obtained at the limit $\nu = 0$. This produces a weak solution, whose regularity is obtained using the compatibility conditions. This method has the great advantage to apply to general (*i.e.* with variable non necessarily smooth coefficients) first order linear hyperbolic equations.
However this method is not constructive in contrast with the so called method of characteristics. In the case of the linear advection equation (2.1), this last method consists in considering the backward characteristics (a straight line here) : $(x - s\mathbf{a}, t - s)_{s\geq 0}$ starting from an arbitrary point $(x, t) \in \Omega \times [0, +\infty[$ and then look for its intersection with the boundary of this cylinder. One sees easily that this point belongs either to the set $\Omega \times \{0\}$ or to the set $\partial\Omega^- \times [0, +\infty[$ and since (2.1) simply means that u is constant along this characteristics, one finds $u(x, t)$ by using either (2.2) or (2.3). Let us be more precise. Given $x \in \Omega$, we denote by $s(x)$ the first positive real number τ such that $x - \tau\mathbf{a}$ meets the boundary of Ω (one sees easily that necessarily $x - s(x)\mathbf{a} \in \partial\Omega^-$). Then there are three cases

(i) if $0 < s(x) < t$ then $u(x, t) = \psi(x - s(x)\mathbf{a}, t - s(x))$,
(ii) if $s(x) > t$ then $u(x, t) = \varphi(x - t\mathbf{a})$,
(iii) if $s(x) = t$ then $u(x, t) = \psi(x - t\mathbf{a}, 0) = \varphi(x - t\mathbf{a})$,

the last case shading light on the first compatibility condition between ψ and φ :

$$\psi(x, 0) = \varphi(x)\,, \quad \forall x \in \partial\Omega^-\,. \tag{2.6}$$

3. A cell centered finite volume discretization and associated error estimates.

3.1. Notations and geometric properties of meshes. Let $\mathcal{T} = \{K_j : j = 1, \ldots, N\}$ be a partition of the domain Ω in polyhedral volumes K_j (the control volumes) that forms a structured or unstructured triangulation of Ω and such that the hyper-face between two adjacent volumes is included in an hyper-plane. For a given j between 1 and N, there are two cases.

- In the first one, the volume K_j has no hyper-face on the boundary $\partial\Omega$ and then we denote by $\mathcal{N}(j)$ the set of indices $k \neq j$ between 1 and N such that $K_k \cap K_j$ has $(nd - 1)$ positive measure.
- In the second one, the boundary of the volume K_j meets the boundary $\partial\Omega$ in a set of $(nd - 1)$ positive measure and then we denote by $\mathcal{N}_0(j)$ the set of indices $k \neq j$ between 1 and N such that $K_k \cap K_j$ has $(nd - 1)$ positive measure and we complete this set into $\mathcal{N}(j)$ by negative integers numbering the hyper-faces of K_j which are on the boundary $\partial\Omega$. We denote the set of these negative integers by $\mathcal{N}_b(j)$.

In both cases we have $\mathcal{N}(j) = \mathcal{N}_0(j) \cup \mathcal{N}_b(j)$, since $\mathcal{N}_b(j)$ is empty when K_j has no hyper-face on the boundary $\partial\Omega$.
Let $k \in \mathcal{N}(j)$. If $k \in \mathcal{N}_0(j)$, we denote by $\mathbf{n}_{j,k}$ the unit normal on $K_j \cap K_k$ which points out from K_j and by $\mathbf{N}_{j,k}$ the product $\mathbf{N}_{j,k} = |K_j \cap K_k|\mathbf{n}_{j,k}$ where $|K_j \cap K_k|$ denotes the $(nd-1)$ positive measure of $K_j \cap K_k$. If $k \in \mathcal{N}_b(j)$, we denote by K_k the symmetric of K_j with respect to the hyper-face $K_j \cap \partial\Omega$ and keep the same notations as above. We shall use in what follows the partition of $\mathcal{N}(j)$:

$$\mathcal{N}(j) = \mathcal{N}^+(j) \cup \mathcal{N}^-(j) \cup \mathcal{N}^0(j)\,, \tag{3.1}$$

where $\mathcal{N}^+(j) = \{k \in \mathcal{N}(j), \mathbf{a} \cdot \mathbf{n}_{j,k} > 0\}$, $\mathcal{N}^-(j) = \{k \in \mathcal{N}(j), \mathbf{a} \cdot \mathbf{n}_{j,k} < 0\}$ and $\mathcal{N}^0(j) = \{k \in \mathcal{N}(j), \mathbf{a} \cdot \mathbf{n}_{j,k} = 0\}$. Similar definitions are extended to $\mathcal{N}_0^\varepsilon(j)$ and

$\mathcal{N}_b^\varepsilon(j)$, $\varepsilon \in \{0,+,-\}$. Lastly, the centroid of K_j will be denoted by g_j while the one of $K_j \cap K_k$ will be denoted by $g_{j,k}$.
Since we are interested in convergence results, we are going to consider families of triangulation $\mathcal{T}^h$ indexed by the real number $h = \max_{K_j \in \mathcal{T}^h} h_j$ where h_j is the diameter of the volume K_j. By definition of the parameter h, we have $|K_j| \leq h^{nd}$ and $|K_j \cap K_k| \leq h^{nd-1}$ for all $K_j, K_k \in \mathcal{T}^h$ and we assume that there exist $h_0 > 0$ and positive constants κ_1 and κ_2 such that for every $h < h_0$ we have

$$\frac{h_j^{nd}}{|K_j|} \leq \kappa_1 \quad \text{and} \quad \sharp\mathcal{N}(j) \leq \kappa_2\,, \quad \forall K_j \in \mathcal{T}^h\,. \tag{3.2}$$

Remark 1. *The first assumption is equivalent to the shape-regularity assumption and the second one means that the number of neighbors of each volume remains bounded when h tends to zero.*
We shall also use the following properties.
Proposition 3.1. *With the previous notations, we have the two vector identities :*

$$\sum_{k\in\mathcal{N}(j)} \mathbf{N}_{j,k} = 0 \quad \textit{and} \quad \sum_{k\in\mathcal{N}(j)} \mathbf{X}\cdot\mathbf{N}_{j,k}\; g_{j,k} = \mathbf{X}\,|K_j|\,, \quad \forall \mathbf{X} \in \mathbf{R}^{nd}\,. \tag{3.3}$$

Proof. We shall simply rely on the divergence theorem (here the superscript ℓ denotes the ℓ-th component of a vector) :

$$\int_{K_j} \frac{\partial f}{\partial x_\ell}\,dx = \int_{\partial K_j} f\mathbf{n}^\ell\,d\sigma = \sum_{k\in\mathcal{N}(j)} \int_{K_j\cap K_k} f\mathbf{n}_{j,k}^\ell\,d\sigma\,, \tag{3.4}$$

valid for smooth scalar functions f on K_j. Taking $f = 1$ and observing that

$$\int_{K_j\cap K_k} \mathbf{n}_{j,k}^\ell\,d\sigma = \mathbf{N}_{j,k}^\ell\,,$$

we readily obtain the first identity in (3.3). Concerning the second one, in (3.4) we take for f the m-th coordinate of the point x i.e. $f(x) = x_m$ and use that

$$\int_{K_j\cap K_k} x_m\,d\sigma = |K_j \cap K_k|(g_{j,k})_m$$

so that (3.4) leads this time to $|K_j|\delta_{m,\ell} = \sum_{k\in\mathcal{N}(j)} (g_{j,k})_m \mathbf{N}_{j,k}^\ell$ for all ℓ and m in $[1,\ldots,nd]$ which indeed shows the second identity in (3.3). ∎

Proposition 3.2. *For every non vanishing vector $\mathbf{a} \neq 0$ we have :*

$$\mathcal{N}^+(j) \neq \emptyset \quad \textit{and} \quad \sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} \neq 0\,. \tag{3.5}$$

Proof. According to Proposition 3.1 and (3.1), we have

$$\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} + \sum_{k\in\mathcal{N}^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} = 0\,.$$

If $\mathcal{N}^+(j) = \emptyset$ then the first term vanishes so does the second. But since the terms involved in the second sum have the same sign, we deduce that $\mathcal{N}^-(j) = \emptyset$ and $\mathcal{N}(j) = \mathcal{N}^0(j)$ i.e. $\mathbf{a}\cdot\mathbf{N}_{j,k} = 0$ for each $k \in \mathcal{N}(j)$. Using the second identity in (3.3) we conclude that $\mathbf{a} = 0$. ∎

Concerning the time discretization of (2.1), we are going to use a *finite difference* approach and therefore we consider an increasing sequence $0 = t_0 < t_1 < t_2 < \ldots < t_n < t_{n+1} \longrightarrow \infty$ and set $\Delta t_n = t_{n+1} - t_n$.

We shall consider sequences $\xi = (\xi_j)_{j=1,\ldots,N}$ of scalars (respectively of vectors in $\mathbf{R}^{nd}$) and we will estimate the ℓ^p norm for $p \in [1, +\infty]$ with ($|\xi_j|$ being respectively the ℓ^p norm in $\mathbf{R}^{nd}$) :

$$||\xi||_\infty = \max_{1\leq j\leq N} |\xi_j| \quad \text{and} \quad ||\xi||_p = \left(\sum_{j=1}^{N} |K_j||\xi_j|^p \right)^{1/p} \quad \text{for } p \geq 1\,.$$

3.2. The first order explicit upwind finite volume scheme. This scheme reads as follows :

$$\frac{u_j^{n+1} - u_j^n}{\Delta t_n} + \frac{1}{|K_j|} \left(\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} u_j^n + \sum_{k\in\mathcal{N}^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} u_k^n \right) = 0\,. \tag{3.6}$$

Indeed the underlying philosophy of the finite volume scheme is to approximate on each control volume K_j for $j \in 1, \ldots, N$ the mean value of the exact solution to the continuous equation (2.1)

$$U_j^n = \frac{1}{|K_j|} \int_{K_j} u(x, t_n) dx\,, \tag{3.7}$$

by taking into account the direction where the information comes from. System (3.6) allows to compute an approximation of U_j^n once the initial values u_j^0 for $j \geq 1$ and the boundary ones u_k^n for $k \in \cup_j \mathcal{N}_b^-(j)$ and for $n \geq 0$ have been provided.
The classical way to check if the goal is achieved is to evaluate the truncation error which consists in replacing u_j^n by U_j^n in (3.6). Here we are going to use a truncation error based on the value $u(g_j, t_n)$ of the exact solution at the centroid of the control volumes. More precisely one computes

$$\begin{aligned} E_j^n = \frac{u(g_j, t_{n+1}) - u(g_j, t_n)}{\Delta t_n} + \frac{1}{|K_j|} \Bigg(& \sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} u(g_j, t_n) \\ & + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} u(g_k, t_n) + \sum_{k\in\mathcal{N}_b^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} u_k^n \Bigg)\,. \end{aligned} \tag{3.8}$$

Then from equations (3.6) and (3.8), the error denoted by $\alpha_j^n = u_j^n - u(g_j, t_n)$ satisfies the following formula

$$\frac{\alpha_j^{n+1} - \alpha_j^n}{\Delta t_n} + \frac{1}{|K_j|} \left(\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} \alpha_j^n + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} \alpha_k^n \right) = -E_j^n\,, \tag{3.9}$$

that can also be written as

$$\alpha_j^{n+1} = (\mathcal{L}^n \alpha^n)_j - \Delta t_n E_j^n \tag{3.10}$$

where we denote by $\mathcal{L}^n$ the following operator which acts on sequences $\xi = (\xi_j)_{j=1,\dots,N}$:

$$(\mathcal{L}^n\xi)_j = \xi_j - \frac{\Delta t_n}{|K_j|}\left(\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\xi_j + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\xi_k\right). \tag{3.11}$$

In classical finite difference theory, one transfers information on the smallness of the truncation error E_j^n to the error α_j^n via a stability property of the scheme which amounts here to show that the norm of the operator $\mathcal{L}^n$ is not greater than 1. Courant, Friedrichs and Lewy (C.F.L.), in their early study of the discretization of the one dimensional advection equation by finite differences, have introduced the C.F.L. number as a limitation on the time step Δt_n in order to achieve stability. In the case of the scheme (3.6), their construction can be mimicked as follows. First we take on each volume the local time given by

$$\tau_j = \frac{|K_j|}{\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}}, \tag{3.12}$$

(observe that thanks to Proposition 3.2, $\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} \neq 0$) then we set :

$$\Delta t_* = \min_j \tau_j . \tag{3.13}$$

The C.F.L. number is then the ratio :

$$\lambda = \frac{\Delta t_n}{\Delta t_*} . \tag{3.14}$$

When $\lambda \leq 1$, the operator $\mathcal{L}^n$ has a norm not greater than 1. More precisely, we have the following result whose proof is given in Appendix 6.1.

THEOREM 3.3. *Under the C.F.L. condition $\lambda \leq 1$, for every $p \in [1, +\infty]$ the operator $\mathcal{L}^n$ satisfies :*

$$||\mathcal{L}^n\xi||_p \leq ||\xi||_p . \tag{3.15}$$

This result, when it is combined with (3.10), has the following straightforward corollary.

COROLLARY 3.4. *Under the C.F.L. condition $\lambda \leq 1$ and for every $p \in [1, +\infty]$ we have the estimate :*

$$||\alpha^n||_p \leq ||\alpha^0||_p + \sum_{i=0}^{n-1} \Delta t_i ||E^i||_p . \tag{3.16}$$

This inequality shows that estimations on E^i transfers to ones on the error α^n.

3.3. On the truncation error. Let us consider the volume of control K_j. We see that the error E_j^n can be divided into three sums

$$E_j^n = G_j^n + I_j^n + \frac{1}{|K_j|}\sum_{k\in\mathcal{N}_b^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}(u_k^n - u(g_{j,k}, t_n)) \tag{3.17}$$

where

$$G_j^n = \frac{u(g_j,t_{n+1}) - u(g_j,t_n)}{\Delta t_n} + \frac{1}{|K_j|}\sum_{k\in\mathcal{N}(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}u(g_{j,k},t_n) \tag{3.18}$$

and

$$\begin{aligned} I_j^n = \frac{1}{|K_j|}\Bigg(&\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}(u(g_j,t_n) - u(g_{j,k},t_n)) \\ &+ \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}(u(g_k,t_n) - u(g_{j,k},t_n))\Bigg)\,. \end{aligned} \tag{3.19}$$

Recalling the estimates (3.2) on the volumes, the third term in (3.17) that concerns the boundary conditions treatment, behaves like $\mathcal{O}(h)$ if we assume for instance that

$$u_k^n = \frac{1}{\Delta t_n |K_j\cap K_k|}\int_{t_n}^{t_{n+1}}\int_{K_j\cap K_k} \psi(\sigma,t)d\sigma dt\,, \quad k\in\mathcal{N}_b^-(j)\,, \quad n\geq 0 \tag{3.20}$$

i.e. that the discrete numerical boundary treatment satisfies

$$|u_k^n - u(g_{j,k},t_n)| = \mathcal{O}(h^2)\,, \quad k\in\mathcal{N}_b^-(j)\,, \quad n\geq 0\,. \tag{3.21}$$

Using the relationship (3.3) on the normals, the assumptions (3.2) on the mesh and an intensive use of Taylor's expansions we see that the second term in G_j^n can be written as

$$\begin{aligned} \sum_{k\in\mathcal{N}(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}u(g_{j,k},t_n) &= \sum_{k\in\mathcal{N}(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}(u(g_{j,k},t_n) - u(g_j,t_n))\,, \\ &= \sum_{k\in\mathcal{N}(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}(g_{j,k}-g_j)\cdot\nabla u(g_j,t_n) + \mathcal{O}(h^{nd+1})\,, \\ &= \left(\sum_{k\in\mathcal{N}(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}g_{j,k}\right)\cdot\nabla u(g_j,t_n) + \mathcal{O}(h^{nd+1})\,, \\ &= |K_j|\mathbf{a}\cdot\nabla u(g_j,t_n) + \mathcal{O}(h^{nd+1})\,. \end{aligned} \tag{3.22}$$

Finally, since

$$\frac{u(g_j,t_{n+1}) - u(g_j,t_n)}{\Delta t_n} = \frac{\partial u}{\partial t}(g_j,t_n) + \mathcal{O}(\Delta t_n) \tag{3.23}$$

and since u is the solution of (2.1), we deduce that if $\Delta t_n \leq ch$

$$G_j^n = \mathcal{O}(h)\,. \tag{3.24}$$

But concerning I_j^n, a similar computation leads this time to $I_j^n = \mathcal{O}(1)$ and we find that $E_j^n = \mathcal{O}(1)$. Hence E_j^n does not even converge to zero as h goes to zero, in a certain sense the scheme is inconsistent. One might think that the estimate above is not optimal, but in fact it is not the case since in the one dimensional case the previous computations are much simpler and show that indeed E_j^n does not converge

to zero as h goes to zero. This fact (the scheme is not pointwise-consistent) was already observed in 2D by Kröner [16], Lemma 3.2.8. On the other hand, Lemma 4.1 of Chapter IV in Godlewski and Raviart [13] cannot be applied to the upwind scheme on a uniform triangulation by equilateral triangles as they wrongly claim. In their proof, they use the identity $\frac{\partial \Phi}{\partial u}(u,u,\mathbf{n}) + \frac{\partial \Phi}{\partial u}(u,u,-\mathbf{n}) = 0$ where $\Phi(u,v,\mathbf{n})$ denotes the numerical flux, that is $\Phi(u,v,\mathbf{n}) = (\mathbf{a}\cdot\mathbf{n})u$ for $\mathbf{a}\cdot\mathbf{n} > 0$, $\Phi(u,v,\mathbf{n}) = (\mathbf{a}\cdot\mathbf{n})v$ for $\mathbf{a}\cdot\mathbf{n} < 0$ and $\Phi(u,v,\mathbf{n}) = 0$ for $\mathbf{a}\cdot\mathbf{n} = 0$. And one checks readily that this identity is wrong. However the result (that on such a triangulation the upwind scheme is first order accurate) is true as it will follow from Theorem 4.2 proved in this paper.

3.4. A geometric corrector. Our goal is to construct a sequence of scalars $(\gamma_j^n)_{j=1,\dots,N}$ that satisfies the N equations :

$$\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\gamma_j^n + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\gamma_k^n = |K_j| I_j^n \tag{3.25}$$

in order to correct the errors $(\alpha_j^n)_{j=1,\dots,N}$ introduced in Section 3.2 and to prove that under smoothness assumptions the finite volume scheme is first order accurate in spite of the fact that the truncation error is inconsistent. We first slightly modify the linear system (3.25) to be able to estimate in terms of h the norm of the sequence. Thus, we will consider the construction of a sequence of vectors in $\mathbf{R}^{nd}$, $(\mathbf{\Gamma}_j)_{j=1,\dots,N}$ that satisfies the N following vector equations (let us recall that g_j and $g_{j,k}$ belong to $\mathbf{R}^{nd}$) :

$$\begin{aligned}\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\mathbf{\Gamma}_j + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\mathbf{\Gamma}_k = \\ \sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}(g_{j,k}-g_j) + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}(g_{j,k}-g_k)\,.\end{aligned} \tag{3.26}$$

We are going to prove thanks to Theorem 3.8 that this system has one and only one solution but let us give the following definitions :

DEFINITION 3.5. *Given a non zero vector $\mathbf{a}\in\mathbf{R}^{nd}$ and a triangulation $(K_j)_{j=1,\dots,N}$ of a polygonal domain Ω in $\mathbf{R}^{nd}$ defined as in Section 3.1, the geometric corrector for the advection equation (2.1) is the sequence of vectors in $\mathbf{R}^{nd}$, $(\mathbf{\Gamma}_j)_{j=1,\dots,N}$, defined by the $N\times N$ system of equations (3.26).*

DEFINITION 3.6. *For every control volume $J\in\mathcal{T}$, we denote by $\mathcal{C}(J)$ the cone of dependence of J :*

(3.27)

$$\mathcal{C}(J) = \{K\in\mathcal{T}\,/\,\exists\, J_1,\dots,J_p\in\mathcal{T}\,,\ \mathbf{a}\cdot\mathbf{N}_{J_1,K}<0,\ \mathbf{a}\cdot\mathbf{N}_{J_2,J_1}<0,\ \dots,\ \mathbf{a}\cdot\mathbf{N}_{J,J_p}<0\}\,.$$

Let us remark that

$$\forall K\in\mathcal{C}(J)\,,\ \big(j\in\mathcal{N}_0^-(K)\Rightarrow K_j\in\mathcal{C}(J)\big)\,. \tag{3.28}$$

PROPOSITION 3.7. *For every control volume $J\in\mathcal{T}$, there is at least one volume $K_k\in\mathcal{C}(J)$ that shares a face with $\partial\Omega^-$ i.e. such that*

$$\mathcal{N}_b^-(k)\neq\emptyset\,. \tag{3.29}$$

Proof. Indeed let us assume that there is no such volume in $\mathcal{C}(J)$. Since $\mathcal{C}(J)$ is composed of polyhedron, it is also a polyhedron and there is at least one face (on the boundary of $\mathcal{C}(J)$) for which the external normal forms an obtuse angle with $\mathbf{a}$. Then for each volume $K \in \mathcal{C}(J)$ that meets this face, since from the assumption that K does not intersect $\partial\Omega^-$, there is an adjacent volume $L \in \mathcal{T} \setminus \mathcal{C}(J)$ such that $\mathbf{a} \cdot \mathbf{N}_{K,L} < 0$. From the definition of $\mathcal{C}(J)$, this last relation yields that L is in $\mathcal{C}(J)$, that is contradictory. ∎

According to Proposition 3.2, $\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}$ does not vanish, we then define the $N \times N$ matrix B such that for an arbitrary sequence $X = (X_j)_{j=1,\dots,N}$ in $\mathbf{C}^N$ we have

$$(BX)_j = \frac{\sum_{k\in\mathcal{N}_0^-(j)}(-\mathbf{a}\cdot\mathbf{N}_{j,k})X_k}{\sum_{k\in\mathcal{N}^+(j)}\mathbf{a}\cdot\mathbf{N}_{j,k}}\,. \tag{3.30}$$

We denote by $(\mathbf{\Delta}_j)_{j=1,\dots,N}$ the sequence of vectors in $\mathbf{R}^{nd}$ defined by

$$\mathbf{\Delta}_j = \frac{\sum_{k\in\mathcal{N}^+(j)}\mathbf{a}\cdot\mathbf{N}_{j,k}(g_{j,k}-g_j) + \sum_{k\in\mathcal{N}_0^-(j)}\mathbf{a}\cdot\mathbf{N}_{j,k}(g_{j,k}-g_k)}{\sum_{k\in\mathcal{N}^+(j)}\mathbf{a}\cdot\mathbf{N}_{j,k}}\,, \tag{3.31}$$

it is clear that (3.26) can be written as

$$(Id - B)\mathbf{\Gamma} = \mathbf{\Delta}\,. \tag{3.32}$$

We are now able to prove that $Id - B$ is a non-singular M-matrix :
THEOREM 3.8. *The spectrum $\sigma(B)$ of B satisfies*

$$\sigma(B) \subset \{z \in \mathbf{C}\,,\ |z| < 1\}\,. \tag{3.33}$$

Thus

$$(Id - B)^{-1} = \sum_{l=0}^{\infty} B^l\,. \tag{3.34}$$

Proof.
i) Let $X \neq 0$, $X \in \mathbf{C}^N$ such that $BX = \lambda X$. From definition (3.30), by observing that $-\mathbf{a}\cdot\mathbf{N}_{j,k} \geq 0$ for $k \in \mathcal{N}_0^-(j)$ and by using the relation $\sum_{k\in\mathcal{N}(j)}\mathbf{a}\cdot\mathbf{N}_{j,k} = 0$ rewritten as $\sum_{k\in\mathcal{N}_0^-(j)}\mathbf{a}\cdot\mathbf{N}_{j,k} + \sum_{k\in\mathcal{N}_b^-(j)}\mathbf{a}\cdot\mathbf{N}_{j,k} + \sum_{k\in\mathcal{N}^+(j)}\mathbf{a}\cdot\mathbf{N}_{j,k} = 0$, we have

$$|(BX)_j| \leq \frac{\sum_{k\in\mathcal{N}_0^-(j)} -\mathbf{a}\cdot\mathbf{N}_{j,k}}{\sum_{k\in\mathcal{N}^+(j)}\mathbf{a}\cdot\mathbf{N}_{j,k}}||X||_\infty = \left(1 + \frac{\sum_{k\in\mathcal{N}_b^-(j)}\mathbf{a}\cdot\mathbf{N}_{j,k}}{\sum_{k\in\mathcal{N}^+(j)}\mathbf{a}\cdot\mathbf{N}_{j,k}}\right)||X||_\infty \tag{3.35}$$

then $||BX||_\infty \leq ||X||_\infty$ and $|\lambda| \leq 1$.

ii) Now let us assume that $\lambda \in \mathbf{C}$ with $|\lambda| = 1$ is an eigenvalue of B i.e. that there is $X \neq 0$ such that $BX = \lambda X$ and $X \in \mathbf{C}^N$. Let us denote by $K_j \in \mathcal{T}$ a volume such that

$$|X_j| = \max_{1\leq k\leq N}|X_k|\,. \tag{3.36}$$

From the j-th component of $BX = \lambda X$ and since $\sum_{k\in\mathcal{N}(j)} \mathbf{N}_{j,k} = 0$, we get

$$\sum_{k\in\mathcal{N}_0^-(j)} (-\mathbf{a}\cdot\mathbf{N}_{j,k}) \left(|X_j| - \frac{X_k}{\lambda} \right) + \sum_{k\in\mathcal{N}_b^-(j)} (-\mathbf{a}\cdot\mathbf{N}_{j,k}) |X_j| = 0. \tag{3.37}$$

Thus

$$\sum_{k\in\mathcal{N}_0^-(j)} (-\mathbf{a}\cdot\mathbf{N}_{j,k})\, \mathcal{R}e \left(|X_j| - \frac{X_k}{\lambda} \right) + \sum_{k\in\mathcal{N}_b^-(j)} (-\mathbf{a}\cdot\mathbf{N}_{j,k}) |X_j| = 0\,, \tag{3.38}$$

$$\sum_{k\in\mathcal{N}_0^-(j)} (-\mathbf{a}\cdot\mathbf{N}_{j,k})\, \mathcal{I}m \left(|X_j| - \frac{X_k}{\lambda} \right) = 0\,. \tag{3.39}$$

Since $\mathcal{R}e(|X_j| - X_k/\lambda) \geq 0$ and $\mathcal{I}m(|X_j| - X_k/\lambda) \geq 0$, all the terms above are positive and thus equal to zero so that $\mathcal{N}_b^-(j) = \emptyset$ and for all $k \in \mathcal{N}_0^-(j)$, $|X_j| = X_k/\lambda$ i.e. $|X_j| = |X_k|$. By induction, we find that for all k such that K_k belongs to $\mathcal{C}(K_j)$, $|X_k| = |X_j|$ and $\mathcal{N}_b^-(k) = \emptyset$ i.e. K_k does not meet $\partial\Omega^-$, which is in contradiction with Proposition 3.7. ∎

REMARK 2. *Let us denote :*

$$\delta_j = \sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} = \sum_{k\in\mathcal{N}^-(j)} |\mathbf{a}\cdot\mathbf{N}_{j,k}| > 0\,, \tag{3.40}$$

so that (3.30) reads also

$$(BX)_j = \frac{\sum_{k\in\mathcal{N}_0^-(j)} |\mathbf{a}\cdot\mathbf{N}_{j,k}| X_k}{\delta_j}\,. \tag{3.41}$$

We then observe that $(BX)_j$ is a barycenter of the X_k for $k \in \mathcal{N}^-(j)$ when K_j does not meet the boundary $\partial\Omega^-$ and of the X_k for $k \in \mathcal{N}_0^-(j)$ and 0 for $k \in \mathcal{N}_b^-(j)$ when K_j meets the boundary $\partial\Omega^-$.
Now, given $m \geq 2$, we have :

$$(B^m X)_j = \frac{\sum_{j_1\in\mathcal{N}_0^-(j)} \cdots \sum_{j_m\in\mathcal{N}_0^-(j_{m-1})} |\mathbf{a}\cdot\mathbf{N}_{j,j_1}| \ldots |\mathbf{a}\cdot\mathbf{N}_{j_{m-1},j_m}| X_{j_m}}{\delta_j \delta_{j_1} \ldots \delta_{j_{m-1}}}\,. \tag{3.42}$$

REMARK 3. *The great advantage of the geometric corrector $\mathbf{\Gamma}_k$ defined by (3.26) on the corrector γ_j^n defined by (3.25) is that the former depends only on the advection vector $\mathbf{a}$ and on the geometry of the problem Ω and the elements K_j of the mesh and not on the initial and boundary data. On the other hand it might be that the introduction of this geometric corrector leads to a weaker estimate. We are going to show that it is not the case by considering the case where the data are affine functions. More precisely, we take for initial conditions* (2.2) *and boundary conditions* (2.3)

$$\varphi(x) = \boldsymbol{\delta_1}\cdot x + \delta_2 \quad \text{and} \quad \psi(x,t) = \boldsymbol{\delta_1}\cdot x + \delta_2 - \boldsymbol{\delta_1}\cdot\mathbf{a}t\,, \quad t \geq 0 \tag{3.43}$$

where $\boldsymbol{\delta_1}$ is an arbitrary vector in $\mathbf{R}^{nd}$ and δ_2 is an arbitrary real number. Then the solution u to (2.1) is

$$u(x,t) = \boldsymbol{\delta_1}\cdot x + \delta_2 - \boldsymbol{\delta_1}\cdot\mathbf{a}t \tag{3.44}$$

with $\nabla u = \boldsymbol{\delta_1}$. In this case a straightforward computation yields that the exact solution of (3.25) *has the form*

$$\gamma_j^n = -\mathbf{\Gamma}_j\cdot\boldsymbol{\delta_1} \equiv -\mathbf{\Gamma}_j\cdot\nabla u \tag{3.45}$$

and indicates the optimality of the geometric corrector approach.

3.5. The error estimate. The previous remark suggests to use the vector geometrical corrector $(\mathbf{\Gamma}_j)_{j=1,\dots,N}$ defined by (3.26) to construct the scalar corrector γ_j^n by taking the formula

$$\gamma_j^n = -\mathbf{\Gamma}_j \cdot \nabla u(g_j, t_n)\,. \tag{3.46}$$

Main result : we are now in a position to prove the following general result.

THEOREM 3.9. *Let u be the smooth solution to (2.1)-(2.3) where φ and ψ are arbitrary smooth functions satisfying the compatibility conditions (see Bardos [1]).*
Let assume the local quasi-uniformity of the family of meshes i.e. that there is a positive constant κ_3 such that

$$\frac{1}{\kappa_3}|K_k| \le |K_j| \le \kappa_3 |K_k|\,, \quad \forall h < h_0, \quad \forall K_j \in \mathcal{T}^h, \quad \forall k \in \mathcal{N}(j). \tag{3.47}$$

Let make the following hypotheses on the discretization of the initial and boundary data : we assume that there exist two constants κ_4 and κ_5 such that for all $h < h_0$

$$||(u_j^0 - \varphi(g_j))_{j=1,\dots,N}||_p \le \kappa_4 h\,, \tag{3.48}$$

$$|u_k^n - \psi(g_{j,k}, t_n)| \le \kappa_5 h^2\,, \quad \forall j\,, \quad \forall k \in \mathcal{N}_b^-(j) \ \text{ and } t_n \le T\,. \tag{3.49}$$

Under the C.F.L. condition $\lambda \le 1$, for every $p \in [1, +\infty]$, if there exists C_p such that the geometric corrector $\mathbf{\Gamma}$ satisfies the estimate

$$||\mathbf{\Gamma}||_p \le C_p h\,, \tag{3.50}$$

then the error for the explicit upwind finite volume scheme satisfies the first order estimate :

$$||(u_j^n - U_j^n)_{j=1,\dots,N}||_p \le C_p' h, \quad t_n \le T\,. \tag{3.51}$$

REMARK 4. *If we have instead of (3.50) the weaker estimate*

$$||\mathbf{\Gamma}||_p \le C_p h^\alpha\,, \tag{3.52}$$

for some $\alpha \in]0, 1]$ and provided we keep the same hypotheses, the proof of Theorem 3.9 leads to the estimate

$$||(u_j^n - U_j^n)_{j=1,\dots,N}||_p \le C_p' h^\alpha, \quad t_n \le T\,. \tag{3.53}$$

Proof. Let us recall that we have (3.10)

$$\alpha_j^{n+1} = (\mathcal{L}^n \alpha^n)_j - \Delta t_n E_j^n\,. \tag{3.54}$$

We correct the error α with the corrector γ given by (3.46)

$$\underline{\alpha}_j^n = \alpha_j^n + \gamma_j^n \tag{3.55}$$

so that

$$\underline{\alpha}_j^{n+1} = (\mathcal{L}^n \underline{\alpha}^n)_j - \Delta t_n \underline{E}_j^n \tag{3.56}$$

where the corrected truncation error has the form

$$\underline{E}_j^n = E_j^n - \frac{1}{|K_j|}\left(\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\gamma_j^n + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\gamma_k^n\right) - \frac{\gamma_j^{n+1}-\gamma_j^n}{\Delta t_n}$$

$$= G_j^n + \underline{I}_j^n + \frac{1}{|K_j|}\sum_{k\in\mathcal{N}_b^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}(u_k^n - u(g_{j,k},t_n)) - \frac{\gamma_j^{n+1}-\gamma_j^n}{\Delta t_n}\,. \tag{3.57}$$

Now the previous bad behavior part in the error is changed into

$$\underline{I}_j^n = \frac{1}{|K_j|}\left(\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\left(u(g_j,t_n) - u(g_{j,k},t_n) - \gamma_j^n\right)\right.$$

$$\left. + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\left(u(g_k,t_n) - u(g_{j,k},t_n) - \gamma_k^n\right)\right)\,. \tag{3.58}$$

Using the assumptions (3.2) and (3.47) on the mesh and the following Taylor expansions

$$u(g_j,t_n) - u(g_{j,k},t_n) = (g_j - g_{j,k})\cdot\nabla u(g_j,t_n) + \mathcal{O}(h^2)\,, \tag{3.59}$$

$$u(g_k,t_n) - u(g_{j,k},t_n) = (g_k - g_{j,k})\cdot\nabla u(g_j,t_n) + \mathcal{O}(h^2)\,, \tag{3.60}$$

we have

$$\underline{I}_j^n = \frac{1}{|K_j|}\left(\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\left((g_j - g_{j,k})\cdot\nabla u(g_j,t_n) - \gamma_j^n\right)\right.$$

$$\left. + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\left((g_k - g_{j,k})\cdot\nabla u(g_j,t_n) - \gamma_k^n\right)\right) + \mathcal{O}(h)\,. \tag{3.61}$$

Definitions (3.26) and (3.46) of the geometric corrector and of the corrector γ_j^n yields that

$$\underline{I}_j^n = \frac{1}{|K_j|}\sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\,\mathbf{\Gamma}_k\cdot(\nabla u(g_k,t_n) - \nabla u(g_j,t_n)) + \mathcal{O}(h)\,. \tag{3.62}$$

Since $\mathbf{\Gamma}$ satisfies the estimate (3.50), we obtain

$$\underline{I}_j^n = \mathcal{O}(h). \tag{3.63}$$

On the other hand

$$\frac{\gamma_j^{n+1}-\gamma_j^n}{\Delta t_n} = \mathbf{\Gamma}_j\cdot\left(\nabla\frac{\partial u}{\partial t}(g_j,t_n) + \mathcal{O}(\Delta t_n)\right)\,, \tag{3.64}$$

then again from the estimate (3.50) on $\mathbf{\Gamma}$, we have

$$\frac{\gamma_j^{n+1}-\gamma_j^n}{\Delta t_n} = \mathcal{O}(h). \tag{3.65}$$

Thus using the assumption on the boundary conditions (3.49) and gathering results (3.24), (3.63), (3.65), we conclude that

$$\underline{E}_j^n = \mathcal{O}(h)\,. \tag{3.66}$$

According to the estimate (3.16) applied to (3.56) we deduce that

$$||(u_j^n - u(g_j,t_n) + \gamma_j^n)_{j=1,\dots,N}||_p \leq ||(u_j^0 - u(g_j,0) + \gamma_j^0)_{j=1,\dots,N}||_p + \mathcal{O}(h)\,. \tag{3.67}$$

Now, the assumption on the initial values (3.48) shows that for $t_n \leq T$,

$$||(u_j^n - u(g_j,t_n))_{j=1,\dots,N}||_p = \mathcal{O}(h)\,. \tag{3.68}$$

Finally the required results (3.51) follows from the Taylor expansion :

$$U_j^n = u(g_j,t_n) + \mathcal{O}(h^2)\,. \tag{3.69}$$

■

REMARK 5. *All the results in this section extend to the implicit first order upwind scheme and to more general meshes where the intersection of two adjacent volumes are not anymore included in an hyper-plan but are composed of several $nd-1$ dimension polygons. All these polygons have to be numbered and they are distributed according to the angle which they form with vector* **a** *as in* (3.1). *These extensions will be the subject of a forthcoming article.*

REMARK 6. *In the definition* (3.26) *of the geometric corrector, the centroid g_j of the volume K_j can be replaced by any point $\tilde{g}_j$ such that the distance between g_j and $\tilde{g}_j$ such that $||g_j - \tilde{g}_j|| \leq C\,h$. Indeed, this will simply change $\mathbf{\Gamma}_j$ into $\mathbf{\Gamma}_j + g_j - \tilde{g}_j$.*

4. On the geometric corrector. We are now going to study from a theoretical point of view some various case where we have the following estimate on the geometric corrector : there exists C_p such that the sequence $\mathbf{\Gamma} \equiv (\mathbf{\Gamma}_j)_{j=1,\dots,N}$ satisfies

$$||\mathbf{\Gamma}||_p \leq C_p h\,. \tag{4.1}$$

4.1. The one dimensional case. We consider the case $nd = 1$. Here the advection vector is a scalar and we assume that $a > 0$. The finite volume scheme reads here as :

$$\frac{u_j^{n+1} - u_j^n}{\Delta t_n} + a\frac{u_j^n - u_{j-1}^n}{\Delta x_j} = 0\,. \tag{4.2}$$

We have $\Omega =]A,B[$ and the "triangulation" of Ω is done by the control volumes $K_j =]x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}[$ for $j = 1,\dots,N$. The centroid of K_j is given by $g_j = x_j = (x_{j+\frac{1}{2}} + x_{j-\frac{1}{2}})/2$ and $|K_j| = \Delta x_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$. Equations (3.26) read in this case as :

$$\left\{ \begin{array}{lcll} a\,\Gamma_1 & = & a\,(x_{\frac{3}{2}} - x_1)\,, & j = 1\,, \\ a\,(\Gamma_j - \Gamma_{j-1}) & = & a\,(x_{j+\frac{1}{2}} - x_j) - a\,(x_{j-\frac{1}{2}} - x_{j-1})\,, & j \geq 2\,, \end{array} \right. \tag{4.3}$$

whose solution is simply :

$$\Gamma_j = x_{j+\frac{1}{2}} - x_j = \frac{\Delta x_j}{2}\,, \quad j \geq 1\,. \tag{4.4}$$

In this case, the estimation (4.1) is straightforward for all p and therefore Theorem 3.9 applies : the upwind finite volume method for the advection equation in one dimension leads to a first order approximation.

THEOREM 4.1. *We denote by* $h = \max_{j=1,\dots,N} \Delta x_j$ *and we assume that there is a constant* κ_3 *such that*

$$\frac{1}{\kappa_3} \leq \frac{\Delta x_j}{\Delta x_{j-1}} \leq \kappa_3 \,, \quad \forall j \geq 2 \,, \quad \forall h < h_0 \,. \tag{4.5}$$

If there exist two constants κ_4 *and* κ_5 *such that discretization of the initial and boundary data satisfies*

$$|| \left(u_j^0 - \varphi(x_j)\right)_{j=1,\dots,N} ||_p \leq \kappa_4 h \quad \text{and} \quad |u_0^n - \psi(x_{\frac{1}{2}}, t_n)| \leq \kappa_5 h^2 \,, \quad t_n \leq T \,. \tag{4.6}$$

then under the C.F.L. condition $\lambda \equiv a\Delta t_n / \min_{j=1,\dots,N} \Delta x_j \leq 1$, *for every* $p \in [1, +\infty]$, *the error* $u_j^n - U_j^n$ *satisfies the first order estimate :*

$$||u^n - U^n||_p \leq C'_p h, \quad t_n \leq T \,. \tag{4.7}$$

Let us observe that the corrected error has the form

$$\underline{\alpha}_j^n = u_j^n - \Big(u(x_j, t_n) + (x_{j+\frac{1}{2}} - x_j) \cdot \nabla u(x_j, t_n)\Big) \tag{4.8}$$

and this has a clear interpretation : u_j^n provided by the scheme should be compared to $u(x_{j+\frac{1}{2}}, t_n)$ rather than to $u(x_j, t_n)$. The upwinding introduces a bias which leads to compare u_j^n with $u(x_{j+\frac{1}{2}}, t_n)$. In the case where $a < 0$ we would have found that u_j^n should be compared with $u(x_{j-\frac{1}{2}}, t_n)$. This has already been observed in Eymard *et al.* [10] in their study of the one dimensional linear advection equation.

Let us now make explicit the operator B in the one dimensional case. For $a > 0$ one sees easily that for all $X \in \mathbf{C}^N$:

$$(BX)_1 = 0 \text{ and } (BX)_j = X_{j-1} \,, \text{ for } j \geq 2 \,, \tag{4.9}$$

so that $B^N = 0$ while $B^{N-1} \neq 0$.

4.2. Back to the *nd* **dimensional case.** In this section, we show by straightforward computations that the geometric corrector is of order h for some two-dimensional structured meshes. We first consider the case of meshes composed of parallelograms.

4.2.1. Non-uniform non-orthogonal quadrilateral grid. We consider a quadrilateral domain bounded from below by two half lines, that we use as coordinate axis and that are uniquely defined such that the advection vector $\mathbf{a}$ is oriented from the exterior to the interior on these lines. Now we assume that the mesh is generated by half lines parallel to these axis as it is shown in Figure 4.1.

An element of the mesh is described by two indices (m, n), $m \geq 1$ along x axis, $n \geq 1$ along y axis and $h(m)$ (respectively $k(n)$) designs the length of the element (m, n) along x (respectively y). Here for instance sufficient conditions on the mesh to satisfy (3.2) and (3.47) read as

$$c_1 h \leq h(m) \leq c_2 h \quad \forall m \geq 1 \,, \quad c_3 h \leq k(n) \leq c_4 h \quad \forall n \geq 1 \,. \tag{4.10}$$

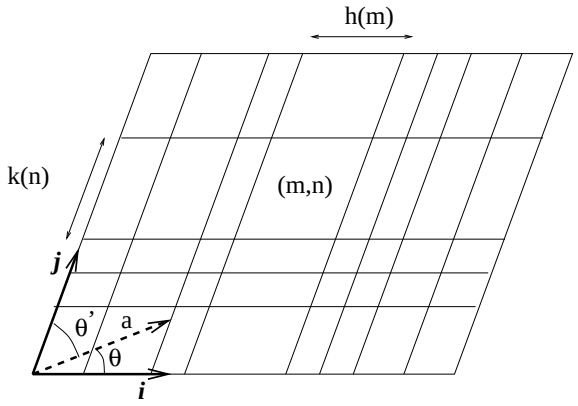


FIG. 4.1. *Non-uniform non-orthogonal quadrilateral grid*

The corrector for the element (m, n) is therefore denoted by $\mathbf{\Gamma}_{m,n}$ and the formula (3.26) for the geometric corrector suggests using

$$G_{m,n} = \mathbf{\Gamma}_{m,n} - \frac{h(m)}{2}\vec{i} - \frac{k(n)}{2}\vec{j} \tag{4.11}$$

where $\vec{i}$ (respectively $\vec{j}$) designs the unit vector along x (respectively y) axis. If we denote θ (respectively θ') the angle of the advection vector $\mathbf{a}$ with the x axis (respectively the y axis), it is straightforward to obtain that $G_{m,n}$ satisfy the following recursive formula

$$G_{m,n} = \lambda_{m,n} G_{m,n-1} + (1 - \lambda_{m,n}) G_{m-1,n} \quad \forall m \geq 1\,, \quad \forall n \geq 1 \tag{4.12}$$

where

$$\lambda_{m,n} = \frac{h(m)\sin(\theta)}{h(m)sin(\theta) + k(n)\sin(\theta')} \quad \forall m \geq 1, \quad \forall n \geq 1 \tag{4.13}$$

and where by convention we set

$$G_{m,0} = -\frac{h(m)}{2}\vec{i} \quad \text{and} \quad G_{0,n} = -\frac{h(n)}{2}\vec{j}\,. \tag{4.14}$$

Thus, $G_{m,n}$ is the barycenter of $G_{m,n-1}$ and $G_{m-1,n}$ with respective weights $\lambda_{m,n}$ and $1 - \lambda_{m,n}$. Then a recursive computation leads $G_{m,n}$ to be a barycenter of $G_{p,0}$ with $1 \leq p \leq m$ and of $G_{0,q}$ with $1 \leq q \leq n$. From this property, we deduce that the absolute value of both components of $\mathbf{\Gamma}_{m,n}$ is smaller than $h/2$ up to a multiplicative constant depending on c_i, $i = 1, \ldots, 4$ and that the estimate (4.1) is true for a non-uniform non-orthogonal quadrilateral mesh.

REMARK 7. *The above result can also be proved for quadrilateral meshes that only satisfy analogue conditions to* (4.5) *in each direction.*

REMARK 8. *The above result is easily extended to a mesh composed of parallelepipeds in dimension 3.*

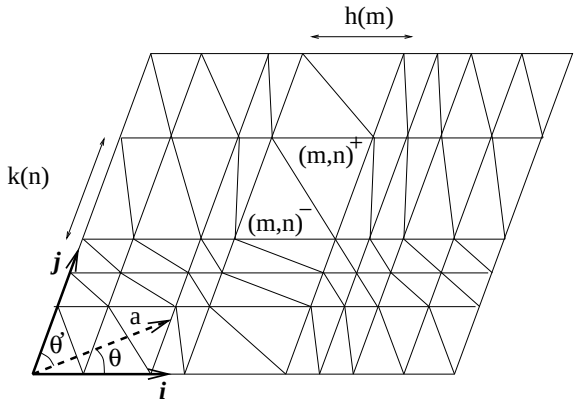


FIG. 4.2. *Non-uniform non-orthogonal triangular grid*

4.2.2. Non-uniform non-orthogonal triangular grid. Let us divide each parallelogram of the previous mesh in two triangles according to the Figure 4.2. The corrector for the downwind triangle of parallelogram (m,n) is denoted by $\mathbf{\Gamma}^+_{m,n}$ and the corrector for the upwind one is denoted by $\mathbf{\Gamma}^-_{m,n}$. Writing the equations that satisfy the geometric corrector in both triangles and eliminating $\mathbf{\Gamma}^-_{m,n}$ we get an equation on $\mathbf{\Gamma}^+_{m,n}$ of the same form as for the mesh composed of parallelogram. We now introduce

$$G^+_{m,n} = \mathbf{\Gamma}^+_{m,n} - \frac{h(m)}{3}\vec{i} - \frac{k(n)}{3}\vec{j} \tag{4.15}$$

and obtain that $G^+_{m,n}$ satisfy the recursive formula (4.12). As for the previous study, we conclude that $\mathbf{\Gamma}^+_{m,n}$ is of order h. Using the expression of $\mathbf{\Gamma}^-_{m,n}$ which is the sum of terms of order h, we prove estimate (4.1) for this type of meshes.

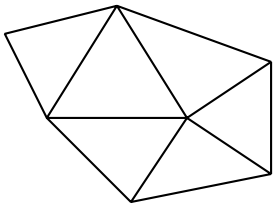


FIG. 4.3. $\mathcal{T}_0$ *: a coarse triangular conformal mesh*

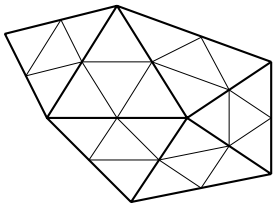


FIG. 4.4. $\mathcal{T}_1$ *: 1-refinement of* $\mathcal{T}_0$

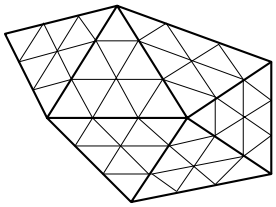


FIG. 4.5. $\mathcal{T}_2$ *: 2-refinement of* $\mathcal{T}_0$

4.3. Asymptotically structured triangular meshes in dimension two. The purpose of this Section is to prove the estimate (4.1) in the case where a global refinement technique is applied to a given arbitrary and unstructured triangular conformal (in the finite element sense) mesh $\mathcal{T}_0$ as the one plotted in Figure 4.3. Given $\ell \in \mathbf{N}$ a positive integer, we denote by $\mathcal{T}_\ell$ the mesh obtained from $\mathcal{T}_0$ by dividing each triangle into $(\ell+1)^2$ congruent triangles i.e by introducing ℓ regularly spaced points on each edge as it can be seen on Figures 4.4 and 4.5.

THEOREM 4.2. *For every triangular conformal mesh $\mathcal{T}_0$ and for every non vanishing vector $\mathbf{a}$ in $\mathbf{R}^2$ there exists a constant $C(\mathbf{a}, \mathcal{T}_0)$ such that for every $\ell \in \mathbf{N}$ we have the estimate*

$$||\mathbf{\Gamma}^{(\ell)}||_\infty \leq \frac{C(\mathbf{a}, \mathcal{T}_0)}{\ell + 1}\,, \tag{4.16}$$

where $\mathbf{\Gamma}^{(\ell)}$ denotes the geometric corrector sequence on $\mathcal{T}_\ell$.
Denoting by $h^{(\ell)} = \max_{K_j \in \mathcal{T}_\ell} h_j > 0$, we have by construction $h^{(\ell)} = \frac{h^{(0)}}{\ell+1}$ and since $h^{(0)}$ depends only on $\mathcal{T}_0$, we deduce from (4.16) the proof of the estimate (4.1) in this case.
COROLLARY 4.3. *For every triangular conformal mesh $\mathcal{T}_0$ of a polygonal domain Ω and for every non vanishing vector $\mathbf{a}$ in $\mathbf{R}^2$ there exists a constant $C(\mathbf{a}, \mathcal{T}_0)$ such that for every $\ell \in \mathbf{N}$ and for every real number $p \in [1, \infty]$ we have the estimate*

$$||\mathbf{\Gamma}^{(\ell)}||_p \leq C(\mathbf{a}, \mathcal{T}_0)|\Omega|^{1/p} \max_{K_j \in \mathcal{T}_\ell} h_j\,. \tag{4.17}$$

Actually we are going to show a stronger result than Theorem 4.2. To state this result, we introduce the notation $\mathcal{N}^-(\Omega) = \{j \text{ s.t. } K_j \text{ meets } \partial\Omega^- \text{ according a set of } nd-1 \text{ positive measure}\}$ and consider instead of (3.26)

$$\begin{aligned} \sum_{k \in \mathcal{N}^+(j)} \mathbf{a} \cdot \mathbf{N}_{j,k} \mathbf{\Gamma}_j + \sum_{k \in \mathcal{N}_0^-(j)} \mathbf{a} \cdot \mathbf{N}_{j,k} \mathbf{\Gamma}_k = \\ \sum_{k \in \mathcal{N}^+(j)} \mathbf{a} \cdot \mathbf{N}_{j,k}(g_{j,k} - g_j) + \sum_{k \in \mathcal{N}_0^-(j)} \mathbf{a} \cdot \mathbf{N}_{j,k}(g_{j,k} - g_k) + \epsilon_j\,, \end{aligned} \tag{4.18}$$

where ϵ_j vanishes for j not in $\mathcal{N}^-(\Omega)$.
THEOREM 4.4. *For every triangular conformal mesh $\mathcal{T}_0$ and for every non vanishing vector $\mathbf{a}$ in $\mathbf{R}^2$, we consider on $\mathcal{T}_\ell$ a sequence $\epsilon_j = \frac{\delta_j}{(\ell+1)^2}$ where the sequence δ_j of vectors in $\mathbf{R}^2$ is uniformly bounded by a constant $C_0(\mathbf{a}, \mathcal{T}_0)$ and vanishes for j not in $\mathcal{N}^-(\Omega)$. Then there exists a constant $C(a, \mathcal{T}_0)$ such that for every $\ell \in \mathbf{N}$ we have the estimate (4.16) where $\mathbf{\Gamma}^{(\ell)}$ denotes the solution to (4.18) on $\mathcal{T}_\ell$.*
The proof of Theorem 4.4 proceeds by induction on M the number of elements of $\mathcal{T}_0$.

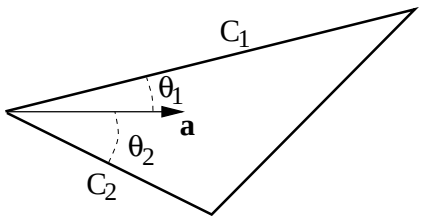


FIG. 4.6. *The case of $\sharp\mathcal{N}^+(T) = 1$*

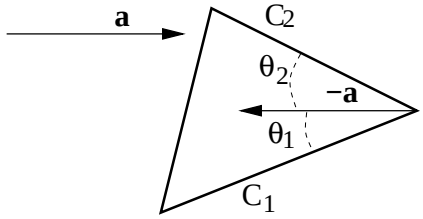


FIG. 4.7. *The case of $\sharp\mathcal{N}^+(T) = 2$*

4.3.1. The case where $\mathcal{T}_0$ has one element. In this case Ω is composed of a single triangle T and $\mathcal{T}_0 = T$. Besides the simpler case where the vector $\mathbf{a}$ is tangent to one of the edges of T, there are two cases that are presented in Figures 4.6 and 4.7 where $\theta_1 \neq 0$ and $\theta_2 \neq 0$: either $\sharp\mathcal{N}^+(T) = 1$ or $\sharp\mathcal{N}^+(T) = 2$. We index as it

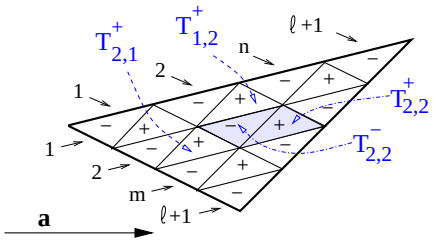


FIG. 4.8. $\sharp\mathcal{N}^+(T) = 1$: *a ℓ-refinement*

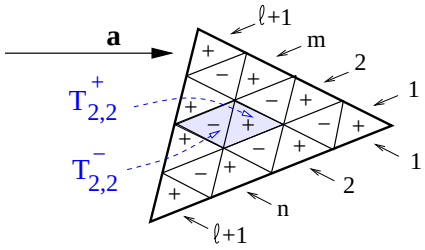


FIG. 4.9. $\sharp\mathcal{N}^+(T) = 2$: *a ℓ-refinement*

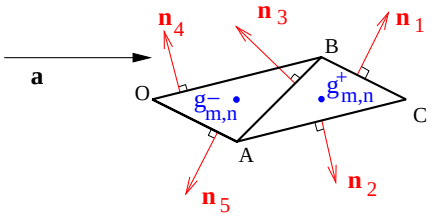


FIG. 4.10. $\sharp\mathcal{N}^+(T) = 1$: *the two sub-triangles $T^+_{m,n}$ and $T^-_{m,n}$, ($\mathbf{n}_i \equiv \mathbf{N}_i/||\mathbf{N}_i||$)*

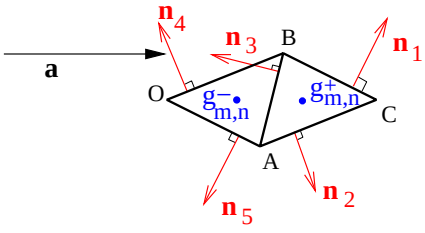


FIG. 4.11. $\sharp\mathcal{N}^+(T) = 2$: *the two sub-triangles $T^+_{m,n}$ and $T^-_{m,n}$, ($\mathbf{n}_i \equiv \mathbf{N}_i/||\mathbf{N}_i||$)*

is shown in Figure 4.8 and 4.9 the $(\ell+1)^2$ triangles in $\mathcal{T}_\ell$, for $\ell \geq 1$, by a triplet (n, m, ϵ) where n and m are natural numbers satisfying $1 \leq n \leq \ell+1$, $1 \leq m \leq \ell+1$, $n+m \leq \ell+2$ and $\epsilon = \pm$. When $\epsilon = -$, we have $\sharp\mathcal{N}^+(T^\epsilon_{m,n}) = 1$ while when $\epsilon = +$, we have $\sharp\mathcal{N}^+(T^\epsilon_{m,n}) = 2$. We say that the triangle $T^\epsilon_{m,n}$ is interior when it has no edge on the boundary of T.

The case where $\sharp\mathcal{N}^+(T) = 1$.

- All the sub-triangles $T^+_{m,n}$ (with $1 \leq m \leq \ell$, $1 \leq n \leq \ell$ and $m+n \leq \ell+1$) are interior ones. Hence if we adopt the notations explained in Figure 4.10 then we can write equation (4.18) in $T^+_{m,n}$ as follows :

(4.19)
$$\mathbf{a}\cdot\mathbf{N}_1\mathbf{\Gamma}^+_{m,n} + \mathbf{a}\cdot\mathbf{N}_2\mathbf{\Gamma}^+_{m,n} + \mathbf{a}\cdot\mathbf{N}_3\mathbf{\Gamma}^-_{m,n} = \mathbf{a}\cdot\mathbf{N}_1\left(\frac{B+C}{2} - g^+_{m,n}\right) + \mathbf{a}\cdot\mathbf{N}_2\left(\frac{C+A}{2} - g^+_{m,n}\right) + \mathbf{a}\cdot\mathbf{N}_3\left(\frac{A+B}{2} - g^-_{m,n}\right)$$

If c_1, c_2, θ_1 and θ_2 are defined as in Figure 4.6, since $\mathbf{a}\cdot\mathbf{N}_1 + \mathbf{a}\cdot\mathbf{N}_2 = -\mathbf{a}\cdot\mathbf{N}_3 = \frac{c_1\sin\theta_1 + c_2\sin\theta_2}{l+1}$ and since the right hand side of (4.19) is equal to $\frac{\zeta_1}{(l+1)^2}$ where $\boldsymbol{\zeta}_1$ is a vector which depends only on T and $\mathbf{a}$, we deduce that

(4.20)
$$\mathbf{\Gamma}^+_{m,n} - \mathbf{\Gamma}^-_{m,n} = \frac{1}{\ell+1}\frac{\boldsymbol{\zeta}_1}{c_1\sin\theta_1 + c_2\sin\theta_2} \quad \text{for } 1 \leq m, n \leq \ell\,,\ m+n \leq \ell+1\,.$$

- Now let us consider the interior triangles $T^+_{m,n}$ for which we have $2 \leq m \leq \ell$, $2 \leq n \leq \ell$. Then the four triangles $T^+_{m,n}$, $T^-_{m,n}$, $T^+_{m-1,n}$, $T^+_{m,n-1}$ are also

interior triangles and we can write equation (4.18) in the triangle $T^-_{m,n}$ as

$$\mathbf{a}\cdot\mathbf{N}_4\mathbf{\Gamma}^+_{m-1,n}+\mathbf{a}\cdot\mathbf{N}_5\mathbf{\Gamma}^+_{m,n-1}-\mathbf{a}\cdot\mathbf{N}_3\mathbf{\Gamma}^-_{m,n}= \tag{4.21}$$
$$\mathbf{a}\cdot\mathbf{N}_4\left(\frac{O+B}{2}-g^+_{m-1,n}\right)+\mathbf{a}\cdot\mathbf{N}_5\left(\frac{A+O}{2}-g^+_{m,n-1}\right)-\mathbf{a}\cdot\mathbf{N}_3\left(\frac{A+B}{2}-g^-_{m,n}\right).$$

By summing up equations (4.19) and (4.21), since their right hand sides are opposite, we obtain

$$\mathbf{\Gamma}^+_{m,n}=\frac{c_1\sin\theta_1\mathbf{\Gamma}^+_{m-1,n}+c_2\sin\theta_2\mathbf{\Gamma}^+_{m,n-1}}{c_1\sin\theta_1+c_2\sin\theta_2}\,. \tag{4.22}$$

- Let us now discuss the case where $T^-_{m,n}$ has at least one edge on the boundary of T. In this case the only changes in (4.19) and (4.21) is that corresponding terms in the right and left hand sides are not present when a triangle $T^\epsilon_{p,q}$ does not exist. Let us give these cases in full details.
  - The case $T^-_{1,1}$. The analogue of (4.21) reads as

$$-\mathbf{a}\cdot\mathbf{N}_3\mathbf{\Gamma}^-_{1,1}=-\mathbf{a}\cdot\mathbf{N}_3\left(\frac{A+B}{2}-g^-_{1,1}\right)+\frac{\delta^-_{1,1}}{(\ell+1)^2}\,. \tag{4.23}$$

    Hence from (4.23) and (4.20)

$$\mathbf{\Gamma}^+_{1,1}=\frac{\boldsymbol{\zeta}_2}{\ell+1} \tag{4.24}$$

    where $\boldsymbol{\zeta}_2$ is a vector which depends only on T and $\mathbf{a}$.
  - The case $T^-_{m,1}\,, 2\le m\le\ell$. Equations (4.19), (4.21) and (4.22) are still valid in this case if we set $\mathbf{\Gamma}^+_{m,0}\equiv\frac{B+C}{2}-g^+_{m,1}+\frac{\delta^-_{m,1}}{(\ell+1)c_2\sin\theta_2}$. Let us observe that this terms behaves like

$$\mathbf{\Gamma}^+_{m,0}=\frac{\boldsymbol{\zeta}_3}{\ell+1} \tag{4.25}$$

    where $\boldsymbol{\zeta}_3$ is a vector which depends only on T and $\mathbf{a}$.
  - The case $T^-_{1,n}\,, 2\le n\le\ell$. Here again equations (4.19), (4.21) and (4.22) are still valid in this case if we set $\mathbf{\Gamma}^+_{0,n}\equiv\frac{A+C}{2}-g^+_{1,n}+\frac{\delta^-_{1,n}}{(\ell+1)c_1\sin\theta_1}$ and observe that

$$\mathbf{\Gamma}^+_{0,n}=\frac{\boldsymbol{\zeta}_4}{\ell+1} \tag{4.26}$$

    where $\boldsymbol{\zeta}_4$ is a vector which depends only on T and $\mathbf{a}$.
  - The case $T^-_{m,n}$ with $m+n=\ell+2$. Here we can only write (4.21) and we obtain

$$\mathbf{\Gamma}^-_{m,n}=\frac{c_1\sin\theta_1\mathbf{\Gamma}^+_{m-1,n}+c_2\sin\theta_2\mathbf{\Gamma}^+_{m,n-1}}{c_1\sin\theta_1+c_2\sin\theta_2}+\frac{\boldsymbol{\zeta}_5}{\ell+1}\,. \tag{4.27}$$

    where $\boldsymbol{\zeta}_5$ is a vector which depends only on T and $\mathbf{a}$ and where we set $\mathbf{\Gamma}^+_{\ell+1,0}\equiv\frac{\delta^-_{\ell+1,1}}{(\ell+1)c_2\sin\theta_2}$ and $\mathbf{\Gamma}^+_{0,\ell+1}\equiv\frac{\delta^-_{1,\ell+1}}{(\ell+1)c_1\sin\theta_1}$.

The barycenter structure of formula (4.22) is the key point and permits to conclude : we find by induction on (m,n) the estimation on $\mathbf{\Gamma}^+_{m,n}$ and then by relation (4.20) we find the estimation on $\mathbf{\Gamma}^-_{m,n}$:

$$|\mathbf{\Gamma}^\epsilon_{m,n}| \leq \frac{C(\mathbf{a},T)}{\ell+1}\,, \tag{4.28}$$

where the constant $C(\mathbf{a},T)$ depends only on T and $\mathbf{a}$.

The case where $\sharp\mathcal{N}^+(T)=2$. This case is slightly simpler since all the sub-triangles $T^-_{m,n}$ are interior ones. Hence if we adopt the notations explained in Figure 4.11, for every (m,n) with $1\leq n\leq \ell$, $1\leq m\leq \ell$, $n+m\leq \ell+1$, equation (4.18) in $T^-_{m,n}$ reads as :

$$\mathbf{a}\cdot\mathbf{N}_4\mathbf{\Gamma}^+_{m+1,n}+\mathbf{a}\cdot\mathbf{N}_5\mathbf{\Gamma}^+_{m,n+1}-\mathbf{a}\cdot\mathbf{N}_3\mathbf{\Gamma}^-_{m,n}= \tag{4.29}$$
$$\mathbf{a}\cdot\mathbf{N}_4\left(\frac{O+B}{2}-g^+_{m+1,n}\right)+\mathbf{a}\cdot\mathbf{N}_5\left(\frac{A+O}{2}-g^+_{m,n+1}\right)-\mathbf{a}\cdot\mathbf{N}_3\left(\frac{A+B}{2}-g^-_{m,n}\right).$$

Concerning the sub-triangles $T^+_{m,n}$, we have two cases.

- In the first one, $1\leq n\leq \ell$, $1\leq m\leq \ell$, $m+n\leq \ell+1$ then equation (4.18) in $T^+_{m,n}$ reads again as (4.19) and (4.20) is still valid. Combining (4.19) with (4.29), formula (4.22) is now replaced by

$$\mathbf{\Gamma}^+_{m,n}=\frac{c_1\sin\theta_1\mathbf{\Gamma}^+_{m+1,n}+c_2\sin\theta_2\mathbf{\Gamma}^+_{m,n+1}}{c_1\sin\theta_1+c_2\sin\theta_2}\,. \tag{4.30}$$

- In the second one, $m+n=\ell+2$ and then equation (4.18) in $T^+_{m,n}$ reads as :

$$\mathbf{a}\cdot\mathbf{N}_1\mathbf{\Gamma}^+_{m,n}+\mathbf{a}\cdot\mathbf{N}_2\mathbf{\Gamma}^+_{m,n}= \tag{4.31}$$
$$\mathbf{a}\cdot\mathbf{N}_1\left(\frac{B+C}{2}-g^+_{m,n}\right)+\mathbf{a}\cdot\mathbf{N}_2\left(\frac{C+A}{2}-g^+_{m,n}\right)+\frac{\delta^+_{m,n}}{(\ell+1)^2}\,.$$

  Hence we can say that :

$$\mathbf{\Gamma}^+_{m,n}=\frac{\boldsymbol{\zeta}}{\ell+1}\,,\ \text{for } m+n=\ell+2\,, \tag{4.32}$$

  where $\boldsymbol{\zeta}$ is a vector which depends only on T and $\mathbf{a}$.

In conclusion we are able to show first by induction on $p=\ell+3-m-n$ that

$$|\mathbf{\Gamma}^+_{m,n}|\leq\frac{|\boldsymbol{\zeta}|}{\ell+1}\,. \tag{4.33}$$

Indeed, for $p=1$ this assertion follows from (4.32). Assuming that (4.33) holds true for p, we take (m,n) with $m+n=\ell+2-p=\ell+3-(p+1)$ and write (4.30). Since $m+n+1=\ell+3-p$ we can apply (4.33) :

$$|\mathbf{\Gamma}^+_{m+1,n}|\leq\frac{|\boldsymbol{\zeta}|}{\ell+1}\,,|\mathbf{\Gamma}^+_{m,n+1}|\leq\frac{|\boldsymbol{\zeta}|}{\ell+1}\,, \tag{4.34}$$

so that thanks to (4.30), (4.33) holds true for $p+1$. Now that (4.33) is shown, using (4.20) we conclude that (4.28) holds true in the case $\sharp\mathcal{N}^+(T)=2$.
This ends the proof of Theorem 4.4 in the case where $\mathcal{T}_0$ has only one element.

4.3.2. The case where $\mathcal{T}_0$ has two elements. In this case, the domain Ω is the union of two triangles which share an edge. If $\mathbf{a}$ is tangent to this edge, then sequences of geometric corrector on one triangle are independent from sequences defined on the other one. Hence Theorem 4.4 follows in this case. Now when $\mathbf{a}$ is not tangent to this edge we denote these two triangles by T_1 and T_2 in order that $\mathbf{a}$ points from T_1 into T_2. We denote by $(\mathbf{\Gamma}^{(\ell)})$ the solution to (4.18) on $\mathcal{T}_\ell$ the ℓ-refinement of $\mathcal{T}_0 = T_1 \cup T_2$. We also denote $\mathcal{T}_{\ell,1}$ and $\mathcal{T}_{\ell,2}$ the ℓ-refinement of $\mathcal{T}_{0,1} = T_1$ and of $\mathcal{T}_{0,2} = T_2$. We are going to analyze sub-sequences of $(\mathbf{\Gamma}^{(\ell)})$ associated to the triangle T_1 and the triangle T_2. The first key observation is that since on $T_1 \cap T_2$, all the outer normals in sub-triangles which are in T_1 forms an acute angle with $\mathbf{a}$, the sub-sequence of $(\mathbf{\Gamma}^{(\ell)})$ whose indices corresponds to triangles which are in T_1 satisfies exactly (4.18) when it is written only for the refinement $\mathcal{T}_{\ell,1}$. But according to Theorem 3.8 we know that this solution is unique and therefore this sub-sequence is equal to the sequence of geometric correctors defined on $\mathcal{T}_{\ell,1}$. We already proved that Theorem 4.4 applies to T_1 and deduce that

$$||\mathbf{\Gamma}_j^{(\ell)}|| \leq \frac{C_1(\mathbf{a}, \mathcal{T}_0)}{\ell + 1}\,, \quad \text{for all } j \text{ corresponding to a sub-triangle in } T_1\,. \tag{4.35}$$

The next step is to estimate the remaining geometric correctors that correspond to sub-triangles in T_2. By inspection of the equation (4.18) when it is written for $\mathcal{T}_{\ell,2}$, one notices two facts.

1. If one considers a sub-triangle which is in T_2 but which has no edge on $T_1 \cap T_2$, then the corrector of this sub-triangle satisfies the same equation in both $\mathcal{T}_{\ell,2}$ and $\mathcal{T}_\ell$.
2. While for a sub-triangle j which is in T_2 and which has an edge on $T_1 \cap T_2$ we do not have the same equation since one term on each side of the equation is missing. But the difference is equal to

$$\mathbf{a} \cdot \mathbf{N}_{j,j_1} \left(\mathbf{\Gamma}_{j_1}^{(\ell),1} + g_{j,j_1} - g_{j_1} \right) \tag{4.36}$$

where we denote by j_1 the sub-triangle in T_1 which shares an edge with the sub-triangle j.

From (4.35), we observe that these differences slightly modify the r.h.s. of equations associated to sub-triangles of T_2 that share a boundary with ∂T_2^- by adding a term of the form $\frac{\delta(\mathbf{a},\mathcal{T}_0)}{(\ell+1)^2}$. Applying Theorem 4.4 to T_2, we deduce that

$$||\mathbf{\Gamma}_j^{(\ell)}|| \leq \frac{C_2(\mathbf{a}, \mathcal{T}_0)}{\ell + 1}\,, \quad \text{for all } j \text{ corresponding to a sub-triangle in } T_2\,. \tag{4.37}$$

Then (4.35) and (4.37) allow us to conclude the proof of Theorem 4.4 in this case.

4.3.3. Proof of Theorem 4.4. Let $M \geq 1$ be given. We assume that this result is true for all triangulation $\mathcal{T}_0$ with a number of elements less or equal to M. Let $\mathcal{T}_0$ be a triangulation with $M + 1$ elements. First we are going to decompose this triangulation in two parts separated by a broken line made with consecutive edges having the same orientation with respect to $\mathbf{a}$ (see Figure 4.12). Then we will finish the proof by simply observing that it follows by extending the argument given in the case $M = 2$ in the previous Section.

We prove the decomposition in two parts for a more general triangulation :

LEMMA 4.5. *Let $\mathcal{T}_0$ be a conformal mesh made with strictly convex polygons i.e. the interior angle at vertices are strictly less than π. We assume that there is at least*

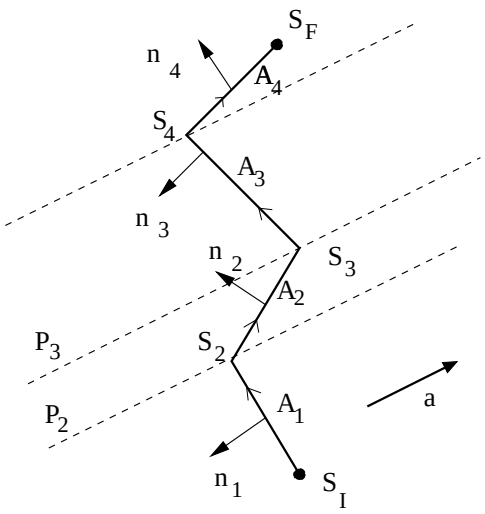


FIG. 4.12. *Broken line made with consecutive oriented edges, (*$\mathbf{n}_i \equiv \mathbf{N}_i/||\mathbf{N}_i||$*)*

one interior edge not parallel to the vector $\mathbf{a}$*. There exists a broken line made with consecutive oriented edges* $A_1 = (S_I, S_2)$*, ...,* $A_i = (S_i, S_{i+1})$*, ...,* $A_K = (S_K, S_F)$ *which links two distinct vertices on the boundary* S_I *and* S_F *and such that* $\mathbf{N}_i \cdot \mathbf{a} < 0\,, \forall i = 1, \ldots, K$*, where the angle* $(A_i, \mathbf{N}_i)$ *is equal to* $\frac{\pi}{2}$*.*

Proof. First let us choice an arbitrary vertex on the boundary that is a vertex of an interior edge $A_1 = (S_I, S_2)$. We assume that the normal $\mathbf{N}_1$ to A_1 such that the angle $(A_1, \mathbf{N}_1)$ is equal to $\frac{\pi}{2}$, satisfies $\mathbf{N}_1 \cdot \mathbf{a} < 0$. If not, we take $\mathbf{a} = -\mathbf{a}$ and switch the role of S_I and S_F.

If S_2 is a vertex on the boundary then we conclude with $S_F = S_2$. If not, let us denote $\mathcal{P}_2$ the half-plane defined by the straight line $(S_2, \mathbf{a})$ (dashed line in Figure 4.12) and that does not contain the edge A_1. Since S_2 is an interior vertex and since interior angles of volumes sharing this vertex are strictly less than π, there is necessarily an edge $A_2 = (S_2, S_3)$ strictly in the half-plane $\mathcal{P}_2$. From the definition of $\mathcal{P}_2$, we have $\mathbf{N}_2 \cdot \mathbf{a} < 0$ where $\mathbf{N}_2$ is the normal to A_2 such that the angle $(A_2, \mathbf{N}_2)$ is equal to $\frac{\pi}{2}$. Now by induction, we can determine a broken line of oriented edges A_1, ... A_i such that $\mathbf{N}_1 \cdot \mathbf{a} < 0$, ..., $\mathbf{N}_i \cdot \mathbf{a} < 0$ with the same convention for the orientation of $\mathbf{N}_i$. Since the distance of S_I to the line $(S_{i+1}, \mathbf{a})$ is strictly superior to the distance of S_I to the line $(S_i, \mathbf{a})$, there exists K such $S_{K+1} \equiv S_F$ is on the boundary.

■

Thus we can realize Ω as the union of two adjacent polygonal sets Ω_1 and Ω_2 where $\mathbf{a}$ points from Ω_1 into Ω_2. Here again, the sub-sequence of geometric correctors on Ω whose indices correspond to triangles in Ω_1 is identical to the sequence of geometric correctors only defined in Ω_1. Since the coarse triangulation of Ω_1 has less than M elements, we can apply Theorem 4.4.

For each sub-triangle j of Ω_2 that shares an edge with the broken line defined in Lemma 4.5, the equation that satisfies the corrector in Ω and the equation that satisfies the corrector in Ω_2 differs from a combination of terms like (4.36) which are

of the form $\frac{\delta_j}{(\ell+1)^2}$ and the δ_j are uniformly bounded by a constant which only depends on $\mathbf{a}$ and $\mathcal{T}_0$. Then, since Ω_2 has less than M elements, we can again apply Theorem 4.4 and therefore show this result for $M+1$. This concludes the proof of order h convergence of the upwind scheme for a uniform refinement of an arbitrary coarse conformal triangular mesh.

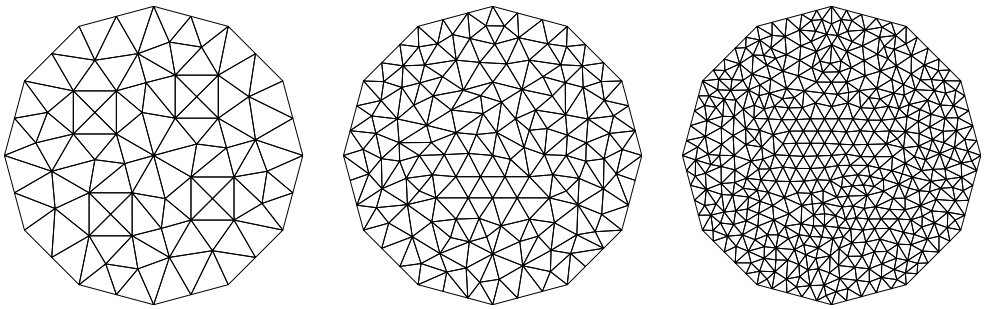


FIG. 5.1. *Some terms of the sequence of independent meshes of the dodecagon domain*

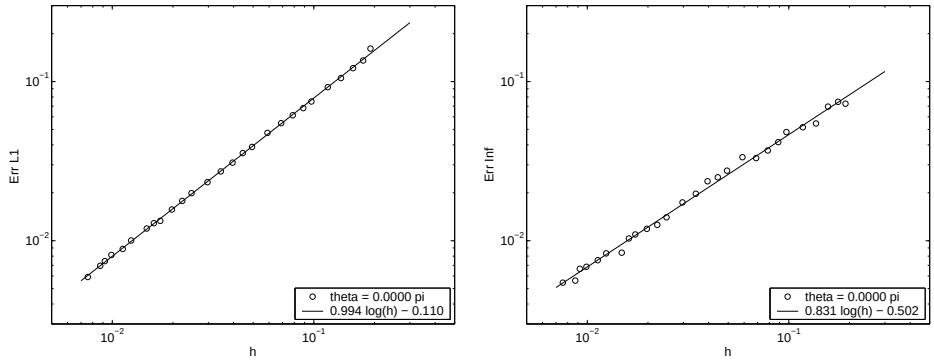


FIG. 5.2. $\theta = 0$: $\|\mathbf{\Gamma}\|_1$ *and* $\|\mathbf{\Gamma}\|_\infty$ *versus* h *for a sequence of independent meshes of the dodecagon domain*

5. Numerical estimates for independent refinements. Since our analysis developed in Section 3 is valid for arbitrary type of meshes, in order to validate from a numerical point of view the estimate (4.1) we perform some tests with a sequence of independent unstructured meshes where the mesh size decreases. In the present simulation, if we take two consecutive grids, one is not the refinement of the other one (by dividing for instance each triangle into four congruent sub-triangles) but the mesh size is reduced. We consider a dodecagon with several meshes composed from 274 triangles to 286514 triangles computed with the software Gmsh [12] developed by Jean-François Remacle and Christophe Geuzaine. In Figure 5.1, the pictures correspond to the three independent meshes of the sequence. We compute the corrector as solution of equation (3.26) for two advection vectors defined by the angle θ with the x-axis : $\theta = 0$ and $\theta = \pi/4$.
In Figure 5.2, 5.3 and 5.4 the L^1 and L^∞ estimates of the geometric corrector and the L^1 norm of $\mathbf{\Delta}$ defined by (3.31) are plotted versus the mesh size. The straight lines are the least-squares fits to the point and the slopes are 0.994 ($\theta = 0$) and 0.988

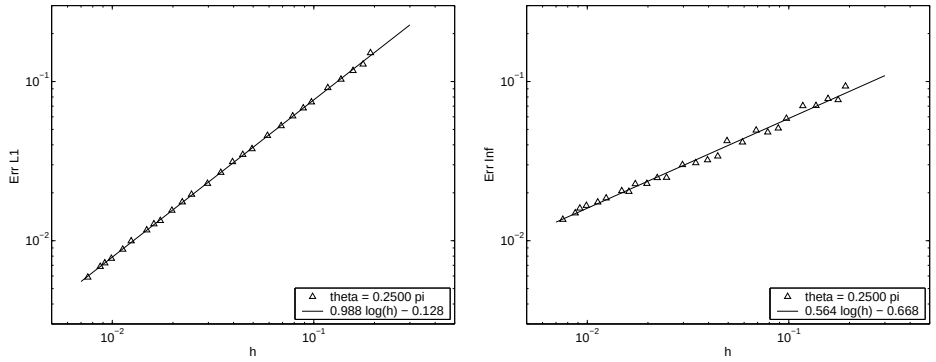


FIG. 5.3. $\theta = \frac{\pi}{4}$: $\|\mathbf{\Gamma}\|_1$ *and* $\|\mathbf{\Gamma}\|_\infty$ *versus* h *for a sequence of independent meshes of the dodecagon domain*

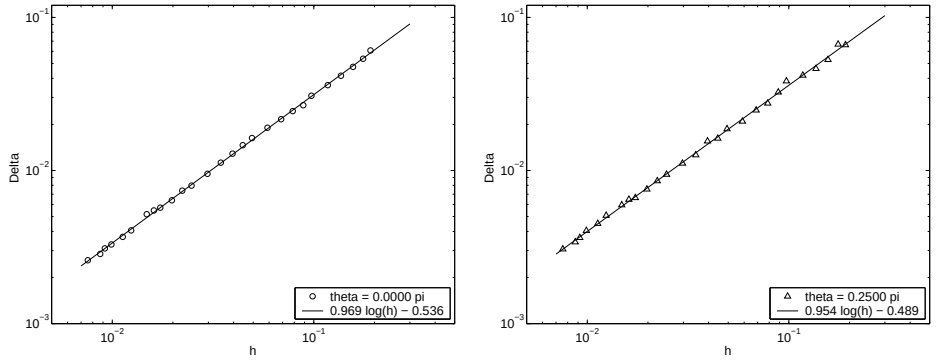


FIG. 5.4. $\theta = 0$ *and* $\theta = \frac{\pi}{4}$: $\|\mathbf{\Delta}\|_1$ *versus* h *for a sequence of independent meshes of the dodecagon domain*

($\theta = \pi/4$) for the L^1 norm of the corrector. We find 0.831 ($\theta = 0$) and 0.564 ($\theta = \pi/4$) for the L^∞ norm of the corrector and 0.969 ($\theta = 0$) and 0.954 ($\theta = \pi/4$) for $\|\mathbf{\Delta}\|_1$.

The L^1 norm of the geometric corrector and the norm of $\mathbf{\Delta}$ behaves like $\mathcal{O}(h)$. Concerning the L^∞ norm of $\mathbf{\Gamma}$, we have an unexpected behavior when the angle θ that define the advection vector $\mathbf{a}$ is equal to $\pi/4$ (as it is shown in figure 5.2) i.e. when the advection vector is parallel to two sides of the domain. Here, asymptotically we have $\|\mathbf{\Gamma}\|_\infty = \mathcal{O}(h^{1/2})$ and although the r.h.s of the equation (3.32) tends to zero with h in all cases, we observe that the estimation of the norm of the solution of this equation depends on the relative position of the advection vector with the boundary.

This behavior is similar to the loss of accuracy proved in Peterson [23]. More precisely, for the mesh and the subtle alignment with the direction of transport proposed by Peterson, we are able to prove that the L^∞ norm of the geometric corrector behaves like $\mathcal{O}(h^{1/2})$ and the L^1 norm is of order one. The technical proof is quite long and out of the scope of this paper and will be published elsewhere.

6. Appendices.

6.1. Proof of Theorem 3.3. We rewrite (3.11) as follows :

$$(\mathcal{L}^n\xi)_j = \left(1 - \frac{\Delta t_n}{|K_j|}\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\right)\xi_j + \frac{\Delta t_n}{|K_j|}\sum_{k\in\mathcal{N}_0^-(j)}(-\mathbf{a}\cdot\mathbf{N}_{j,k})\xi_k\,. \tag{6.1}$$

According to the C.F.L. condition, the coefficient in front of ξ_j is non negative, therefore :

$$|(\mathcal{L}^n\xi)_j| \le \left(1 - \frac{\Delta t_n}{|K_j|}\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\right)||\xi||_\infty + \frac{\Delta t_n}{|K_j|}\sum_{k\in\mathcal{N}_0^-(j)}(-\mathbf{a}\cdot\mathbf{N}_{j,k})||\xi||_\infty\,, \tag{6.2}$$

$$\le \left(1 - \frac{\Delta t_n}{|K_j|}\sum_{k\in\mathcal{N}^+(j)\cup\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\right)||\xi||_\infty \tag{6.3}$$

But since $\sum_{k\in\mathcal{N}(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} = 0$, one gets

$$\sum_{k\in\mathcal{N}^+(j)\cup\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k} = \sum_{k\in\mathcal{N}_b^-(j)}(-\mathbf{a}\cdot\mathbf{N}_{j,k}) \ge 0$$

and therefore the r.h.s. of (6.3) is not greater than $||\xi||_\infty$ and the theorem is proved for $p = \infty$.

Concerning the case $p = 1$, we have to estimate $\sum_j |K_j||(\mathcal{L}^n\xi)_j|$. Using (6.1) and the C.F.L. condition we have :

$$\sum_j |K_j||(\mathcal{L}^n\xi)_j| \le \sum_j \left(\left(|K_j| - \Delta t_n \sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}\right)|\xi_j| + \Delta t_n \sum_{k\in\mathcal{N}_0^-(j)}(-\mathbf{a}\cdot\mathbf{N}_{j,k})|\xi_k|\right)\,,$$

$$\le \sum_j |K_j||\xi_j| - \Delta t_n\Theta\,, \tag{6.4}$$

where

$$\Theta = \sum_j \left(\sum_{k\in\mathcal{N}^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}|\xi_j| + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}|\xi_k|\right)\,,$$

$$= \sum_j \left(\sum_{k\in\mathcal{N}_0^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}|\xi_j| + \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}|\xi_k|\right) + \sum_j \sum_{k\in\mathcal{N}_b^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}|\xi_j|\,.$$

Denoting by :

$$\delta = \sum_j \sum_{k\in\mathcal{N}_0^+(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}|\xi_j| + \sum_j \sum_{k\in\mathcal{N}_0^-(j)} \mathbf{a}\cdot\mathbf{N}_{j,k}|\xi_k|\,,$$

we clearly have $\Theta \ge \delta$. On the other hand observing that in the r.h.s. of δ each interior hyper-face $K_j \cap K_k$ occurs exactly twice and the contributions are opposite by the change of sign of the external normal. Hence $\delta = 0$ and then $\Theta \ge 0$. It follows finally from (6.4) that the theorem is proved for $p = 1$. Since (3.15) has been proved for $p = 1$ and $p = \infty$, it holds true for every p by interpolation ([2]).

REFERENCES

[1] C. Bardos. Problèmes aux limites pour les équations aux dérivées partielles du premier ordre à coefficients réels; théorèmes d'approximation; application à l'équation de transport. *Ann. Sci. École Norm. Sup. (4)*, 3:185–233, 1970.

[2] J. Berg and L. Löfström. *Interpolation spaces. An introduction*. Springer Verlag, Berlin, 1976.

[3] M. J. Berger, C. Helzel, and R. J. LeVeque. H-box methods for the approximation of hyperbolic conservation laws on irregular grids. *SIAM J. Numer. Anal.*, to appear.

[4] M. J. Berger, R. J. LeVeque, and L. G. Stern. Finite volume methods for irregular one-dimensional grids. In *Mathematics of Computation 1943–1993: a half-century of computational mathematics (Vancouver, BC, 1993)*, volume 48 of *Proc. Sympos. Appl. Math.*, pages 255–259. Amer. Math. Soc., Providence, RI, 1994.

[5] B. Cockburn and P.-A. Gremaud. A priori error estimates for numerical methods for scalar conservation laws. i: The general approach. *Math. Comput.*, 65(214):533–573, 1996.

[6] B. Cockburn and P.-A. Gremaud. A priori error estimates for numerical methods for scalar conservation laws. ii: Flux-splitting monotone schemes on irregular cartesian grids. *Math. Comput.*, 66(218):547–572, 1997.

[7] B. Cockburn, P.-A. Gremaud, and J. X. Yang. A priori error estimates for numerical methods for scalar conservation laws. iii: Multidimensional flux-splitting monotone schemes on non-cartesian grids. *SIAM J. Numer. Anal.*, 35(5):1775–1803, 1998.

[8] B. Després. Théorème de lax et volumes finis. In *Proceedings du 32e CANUM, La Grande-Motte*. Université de Montpellier et S.M.A.I., 2003.

[9] B. Després. Lax theorem and finite volume schemes. *Math. Comp.*, 2004.

[10] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In P.-A. Ciarlet and J.-L. Lions, editors, *Handbook of Numerical Analysis*, volume 7, pages 713–1020. North-Holland, 2000.

[11] B. Garcia-Archilla and J.M. Sanz-Serna. A finite difference formula for the discretization of d^3/dx^3 on nonuniform grids. *Math. Comput.*, 57(195):239–257, 1991.

[12] C. Geuzaine and J.-F. Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. http://www.geuz.org/gmsh/, 1997-2003.

[13] E. Godlewski and P.A. Raviart. *Numerical approximation of Hyperbolic Systems of Conservation Laws*. Springer Verlag, New-York, 1996.

[14] J. D. Hoffman. Relationship between the truncation errors of centered finite-difference approximations on uniform and nonuniform meshes. *J. Comput. Phys.*, 46:469–474, 1982.

[15] H.-O. Kreiss, T.A. Manteuffel, B. Swartz, B. Wendroff, and A. B. White, Jr. Supra-convergent schemes on irregular grids. *Math. Comput.*, 47:537–554, 1986.

[16] D. Kröner. *Numerical schemes for conservation laws*. Wiley-Teubner Series Advances in Numerical Mathematics. Chichester: Wiley., 1997.

[17] N.N. Kuznetsov. Accuracy of some approximate methods for computing the weak solutions of a first-order quasi-linear equation. *U.S.S.R. Comput. Math. Math. Phys. 16(1976), No.6, 105-119 (1978)*, 16(6):105–119, 1976.

[18] C. D. Levermore, T. A. Manteuffel, and A. B. White, Jr. Numerical solution of partial differential equations on irregular grids. In *Computational techniques and applications: CTAC-87 (Sydney, 1987)*, pages 417–426. North-Holland, Amsterdam, 1988.

[19] T. A. Manteuffel and A. B. White, Jr. The numerical solution of second order boundary value problems on nonuniform meshes. *Math. Comput.*, 47:511–535, 1986.

[20] T. A. Manteuffel and A. B. White, Jr. On the efficient numerical solution of systems of second order boundary value problems. *SIAM J. Numer. Anal.*, 23:996–1006, 1986.

[21] T. A. Manteuffel and A. B. White, Jr. A calculus of difference schemes for the solution of boundary-value problems on irregular meshes. *SIAM J. Numer. Anal.*, 29(5):1321–1346, 1992.

[22] P. Monk and E. Süli. A convergence analysis of yee's scheme on nonuniform grids. *SIAM J. Numer. Anal.*, 31(2):393–412, 1994.

[23] T. E. Peterson. A note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation. *SIAM J. Numer. Anal.*, 28(1):133–140, 1991.

[24] J. Pike. Grid adaptive algorithms for the solution of the euler equations on irregular grids. *J. Comput. Phys.*, 71:194–223, 1987.

[25] Z.-H. Teng and P. Zhang. Optimal L^1-rate of convergence for the viscosity method and monotone scheme to piecewise constant solutions with shocks. *SIAM J. Numer. Anal.*, 34(3):959–978, 1997.

[26] A.N. Tikhonov and A.A. Samarskij. Homogeneous difference schemes on non-uniform nets. *U.S.S.R. Comput. Math. Math. Phys.*, 1963:927–953, 1964.

[27] E. Turkel. Accuracy of schemes with nonuniform meshes for compressible fluid flows. *Appl. Numer. Math.*, 2:529–550, 1986.

[28] J.-P. Vila and P. Villedieu. Convergence of an explicit finite volume scheme for first order symmetric systems. *Numer. Math.*, 94(3):573–602, 2003.

[29] B. Wendroff. Supraconvergence in two dimensions. Technical Report LA-UR 95-3068, Los Alamos National Laboratory, 1995.

[30] B. Wendroff and A. B. White, Jr. Some supraconvergent schemes for hyperbolic equations on irregular grids. In *Nonlinear hyperbolic equations—theory, computation methods, and applications (Aachen, 1988)*, volume 24 of *Notes Numer. Fluid Mech.*, pages 671–677. Vieweg, Braunschweig, 1989.

[31] B. Wendroff and A. B. White, Jr. A supraconvergent scheme for nonlinear hyperbolic systems. *Comput. Math. Appl.*, 18(8):761–767, 1989.