Journal of Scientific Computing, Vol. 19, Nos. 1–3, December 2003 (© 2003)

High Order Fluctuation Schemes on Triangular Meshes

R. Abgrall[1] **and P. L. Roe**[2]

Received March 29, 2002; accepted (in revised form) November 14, 2002

We develop a new class of schemes for the numerical solution of first-order steady conservation laws. The schemes are of the residual distribution, or fluctuation-splitting type. These schemes have mostly been developed in the context of triangular or tetrahedral elements whose degrees of freedom are their nodal values. We work here with more general elements that allow high-order accuracy. We introduce, for an arbitrary number of degrees of freedom, a simple mapping from a low-order monotone scheme to a monotone scheme that is as accurate as the degrees of freedom will allow. Proofs of consistency, convergence and accuracy are presented, and numerical examples from second, third and fourth-order schemes.

KEY WORDS: Upwind stabilized schemes; hyperbolic problems; residual distributive schemes; finite elements; high order schemes.

1. INTRODUCTION

In this paper, we consider high order discretization of the problem

$$\operatorname{div}(f(u)) = 0 \qquad x \in \Omega \tag{1}$$

with Dirichlet boundary conditions on the inflow part of the domain $\Omega \subset \mathbb{R}^d$

$$u = u_{\Gamma^-} \qquad x \in \Gamma^- \tag{2}$$

Here, f is a C^2 vector valued function, and u is a real valued function. The inflow boundary Γ^- is defined as

$$\Gamma^- = \{x \in \partial\Omega, \vec{n}(x) \cdot \vec{\lambda} > 0\}$$

where $\vec{n}(x)$ is the inward normal vector at $x \in \partial\Gamma$.

[1] Mathématiques Appliquées de Bordeaux, Université Bordeaux I, 33 405 Talence Cedex, France and Institut Universitaire de France. E-mail: Remi.Abgrall@math.u-bordeaux.fr

[2] W. M. Keck Laboratory for Computational Fluid Dynamics, Department of Aerospace Engineering, University of Michigan, Ann Arbor, Michigan 48109.

0885-7474/03/1200-0003/0 © 2003 Plenum Publishing Corporation

We consider entropy weak solutions of (1). Throughout the paper, we assume that u_{Γ^-} in Eq. (2) is regular enough so that the boundary conditions can be set strongly. This assumption is only aimed at simplifying the text.

An approximate solution of (1) is sought on a triangular mesh $\mathcal{T}_h$. The nodes or $\mathcal{T}_h$ are denoted by $\{M_i\}$ and T is a generic triangle.

There are many schemes for approximating (1), but here we consider upwind residual schemes, as in [1–3]. These schemes are distributive schemes; in their simplest forms they can be written as

$$\sum_{T,\, M_i \in T} \Phi_i^T = 0. \tag{3}$$

In practical calculations, the solution of (3) is sought for as the limit, when the iteration count n tends to infinity, of the pseudo unsteady scheme

$$|C_i| \frac{u_i^{n+1} - u_i^n}{\Delta t} + \sum_{T,\, M_i \in T} \Phi_i^T = 0. \tag{4}$$

Here, $|C_i|$ is the area of the dual cell associated with M_i.

In the schemes (3) or (4) the sum is over those triangles T that share the vertex M_i. On each triangle T we define a residual

$$\Phi^T = \int_T \operatorname{div} f^h(u_h)\, dx. \tag{5}$$

We also define Φ_i^T as the amount of Φ^T associated with vertex M_i, such that a conservation property is satisfied

$$\sum_{i,\, M_i \in T} \Phi_i^T = \Phi^T \tag{6}$$

Roe [4] refers to Φ^T as the fluctuation in triangle T and to the quantities Φ_i^T as signals sent to the vertices. It is known, see [1, 5, 6], how to construct *second order* accurate schemes for *steady* problems. The unsteady case has been considered more recently, see for example [7, 6].

One major objective of this paper is to generalize this form in order to get higher accurate schemes, i.e., more than second order accurate schemes, for *steady* problems. The unsteady case will be considered elsewhere. To do so, each triangle is equipped with additional degrees of freedom in addition to the values at the vertices, and signals and fluctuations are defined in order to get a very high order accurate scheme. For example, values can be assigned to the mid-points of the sides.

In this paper, we give precise conditions on f^h and the signals for which, under the conditions of a Lax Wendroff theorem, the limit solution of a scheme of the type (4) is a weak solution of

$$\frac{\partial u}{\partial t} + \operatorname{div} f(u) = 0.$$

We note that under the same conditions, the limit solution of the scheme (3) is a weak solution of (1). Then, we provide conditions for which the solution is formally high order accurate. In particular, we explain in details the difference between steady and unsteady problems. In a second part, we consider from this perspective some known examples having second order accuracy, including some that are monotonicity preserving. Then we show a stability property of these schemes, in the case $f(u) = \vec{\lambda} u$, that yields an error estimate in some cases. Using these principles, we construct third- and fourth-order monotonicity preserving schemes and give some numerical applications. This construction can be formally extended to any order of accuracy, so these schemes can be seen as an alternative to the DG schemes. However, we believe that our schemes will prove less expensive, especially for very high-order and/or three-dimensional cases.

2. NOTATIONS AND BASIC CONCEPTS

2.1. Notation and Degrees of Freedom

We explain our notations, and specify the degrees of freedom. In a mesh $\mathcal{T}$ whose triangles are denoted by T, and whose vertices by M_j, we seek a solution that is piecewise polynomial of degree k in each triangle, and therefore need to provide each T with $(k+1)(k+2)/2$ degrees of freedom. A simple way to do this is to use P^k elements whose degrees of freedom are the values of the solution at the points whose barycentric coordinates within T are $\{i/k, j/k, (k-i-j)/k\}$ for $0 \leqslant i, j \leqslant k$. This creates k^2 subtriangles within T; denote a typical subtriangle by T'_T. See Fig. 1 for the cases $k = 1$ and $k = 2$.

This way of adding points in T provide a natural sub-triangulation of T. The collection of all these sub-triangles furnishes a natural conformal refined triangulation of Ω.

We denote by σ a point that may belong to the set M_j, or which may be one of the additional points introduced by the sub-triangulation. In the case of a P^1 interpolation, the degrees of freedom are the vertices M_j of $\mathcal{T}$ and we will not use the

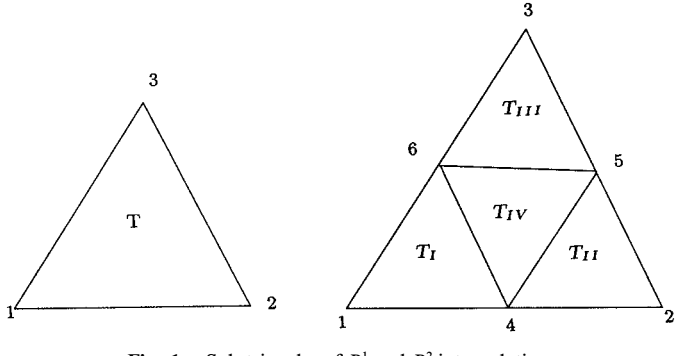


Fig. 1. Sub-triangles of P^1 and P^2 interpolation.

notation σ. For any triangle T, we denote by h_T its diameter, $|T|$ its area and $h = \max_T h_T$.

The schemes that we study here are of the following form: for any degree of freedom σ

$$u_\sigma^{n+1} = u_\sigma^n - \frac{\Delta t}{|C'_\sigma|} \sum_{T, \sigma \in T} \Psi_\sigma^T \qquad \text{with}$$
$$\Psi_\sigma^T = \sum_{T'_T \subset T, \sigma \in T'_T} \Phi_\sigma^{T'_T} \tag{7}$$

Once more, C'_σ is the dual control volume associated with any new or original point σ of the refined mesh. The residuals $\{\Phi_\sigma^{T'_T}\}$ satisfy the conservation relation

$$\sum_{\sigma \in T} \Psi_\sigma^T = \sum_{\sigma \in T} \sum_{T'_T \subset T, \sigma \in T} \Phi_\sigma^{T'_T} = \int_T \operatorname{div} f^h(u^h)\, dx.$$

For $k = 1$, the only subtriangle of T is t itself; for $k = 2$ the subtriangles are $T_I, T_{II}, T_{III}, T_{IV}$, and so on.

2.2. Consistency, Accuracy

Very natural questions are then

- Under which assumptions do we have consistency, and can we exhibit a Lax–Wendroff theorem under reasonable assumptions,
- What accuracy can we reach in principle and how could this be achieved in practice?

In this section, we assume that the problem

$$\frac{\partial u}{\partial t} + \operatorname{div} f(u) = 0 \tag{8}$$

is vector valued: $u \in \mathbb{R}^p \times \mathbb{R}^+$ is defined on $\mathbb{R}^d$. The flux $f = (f_1, \ldots, f_d)$ is also defined on $\mathbb{R}^d$, and $f_i \in \mathbb{R}^p$.

2.2.1. A Lax Wendroff Theorem

Under which reasonable conditions can the scheme (7) converge towards a weak solution of (8)? To answer this question, we give a Lax Wendroff-like theorem. Let us give some conditions on the mesh, the fluctuations $\Phi_\sigma^{T'_T}$, and define some notations and functional spaces.

Assumption 2.1. The mesh $\mathcal{T}_h$ is conformal and regular. By regular we mean that all triangles are roughly the same size, more precisely that there exist constants C_1 and C_2 such that for any triangle

$$T, \qquad C_1 \leqslant \sup_{T \in \mathcal{T}_h} \frac{h^2}{|T|} \leqslant C_2$$

where for this purpose the triangles considered are those at the finest level of subtriangulation. Of course, if the original triangulation is regular in this sense, so is the subtriangulation. We say that two triangles are neighbors if they have a common edge.

Let $\mathscr{T}_h$ be a triangulation satisfying Assumption 2.1, and $\mathscr{C}_h$ be a set of dual volumes associated with the degrees of freedom σ. Then we define the following subspaces,

$$V_h^k = \{v_h \in C^0(\mathbb{R}^2)^p;\ v_{h|T} \text{ polynomial of degree } k,\ \forall\, T \in \mathscr{T}_h\}$$
$$X^h = \{v_h;\ v_{h|C} \text{ constant } \in \mathbb{R}^p,\ \forall\, C \in \mathscr{C}_h\}.$$

Here, $f_{|T}$ denotes the restriction of f to T.

We denote by $\pi_h^1 v$ the piecewise linear interpolation of a continuous function. Let $L_h: V_h^k \to X^h$ be the mass lumping operator, $L_h(v) = \sum_\sigma v(x_\sigma)\ \chi_\sigma$ where χ_σ is the characteristic function of the cell C_σ.

We assume that the residual $\Phi_\sigma^{T'_T}$ and the numerical solution satisfy the following conditions.

Assumption 2.2. Let $\mathscr{T}_h$ be a triangulation satisfying the Assumption 2.1. For any $C \in \mathbb{R}^+$, there exists $C'(C, \mathscr{T}_h) \in \mathbb{R}^+$ which depends only on C and $\mathscr{T}_h$ such that for any $u \in (X^h)^2$, with $\|u\|_{L^\infty(\mathbb{R}^2)} \leqslant C$ we have

$$\forall T,\ \forall \sigma,\ \|\Phi_\sigma^{T'_T}\| \leqslant C'(C, \mathscr{T}_h)\ h \sum_{\sigma' \in T} \|u(\sigma') - u(\sigma)\|. \tag{9}$$

Remark 2.1. One should see this assumption as asserting continuity of the residual components (or signals) $\Phi_\sigma^{T'_T}$ with respect of the nodal values of u; in particular, when u is constant, $\Phi_\sigma^{T'_T} = 0$. The proof of Theorem 2.2 when $\Phi_\sigma^{T'_T}$ satisfy Assumption 2.2 is still valid if the number of arguments in $\Phi_\sigma^{T'_T}$ is bounded independently of h and the triangle T. In practice, this is always true if the triangulation is uniform, since the arguments of $\Phi_\sigma^{T'_T}$ are contained in some neighborhood of σ comprising a finite number of points.

Assumption 2.3. There exists an approximation f^h of the flux f such that

(i) $\forall u^h \in X^h$, $\Phi^T := \int_T \operatorname{div} f^h(u^h)\, dx = \sum_{T'_T \subset T} \sum_{\sigma \in T'_T} \Phi_\sigma^{T'_T}(u^h)$,

(ii) $\forall u^h \in X^h$, $\forall T_1, T_2$ neigbors,

$$f^h(u^h)_{|T_1} \cdot \vec{n} = f^h(u^h)_{|T_2} \cdot \vec{n}\ p.p. \text{ on } T_1 \cap T_2$$

where $\vec{n}$ is a normal of $T_1 \cap T_2$.

(iii) For any $C > 0$, there exists $C'(C)$ such that for any $u^h \in X^h$ with $\|u^h\|_{L^\infty(\mathbb{R}^2)} \leqslant C$, one has for $T \in \mathscr{T}_h$ and $f_T^h = f_{|T}^h$, $\|\operatorname{div} f_T^h(u^h)\| \leqslant \frac{C'}{h} \sum_{\sigma,\sigma'} \|u_\sigma^h - u_{\sigma'}^h\|$ a.e. on T.

(iv) For any sequence $(u^h)_h$ bounded in $L^\infty(\mathbb{R}^2 \times \mathbb{R}^+)^p$ independantly of h and convergent in $L^2_{loc}(\mathbb{R}^2 \times \mathbb{R}^+)^p$ to u, we have

$$\lim_h \|f^h(u^h) - f(u)\|_{L^1_{loc}(\mathbb{R}^2 \times \mathbb{R}^+)^p} = 0.$$

Later, we give examples of schemes that satisfy the previous assumptions. We have the following result,

Theorem 2.2. Let be $u_0 \in L^\infty(\mathbb{R}^d)^p$ and u^h the approximation given by (7). We assume that the scheme satisfies the Assumptions 2.2 and 2.3. We also assume there exists a constant C that depends only on C_1, C_2 and u_0 and a function $u \in (L^2(\mathbb{R}^d \times \mathbb{R}^+))^p$ such that

$$\sup_h \sup_{x,y,t} |u^h(x,y,t)| \leqslant C$$

$$\lim_h \|u - u_h\|_{L^2_{loc}(\mathbb{R}^d \times \mathbb{R}^+)^p} = 0$$

Then u is a weak solution of (8).

The proof of Theorem 2.2 is given in Appendix A.

2.2.2. Accuracy

Here, we consider the steady case

$$\begin{aligned} &\operatorname{div} f(u) = 0 \qquad x \in \Omega \\ &\text{Dirichlet boundary conditions on the inflow boundary} \end{aligned} \tag{10}$$

and derive a formal sufficient condition that ensures the scheme is $r+1$-th order.

In the following, we denote by ψ_σ the Lagrange basis function of degree r associated with the degree of freedom σ, and for any smooth function φ, we denote by φ^h the interpolant

$$\varphi^h = \sum_\sigma \varphi(\sigma)\, \psi_\sigma.$$

We also introduce $\Phi_\sigma^T = \int_T \psi_\sigma \operatorname{div} f^h(u^h)$.

Consider any test function $\varphi \in C^{r+1}(\mathbb{R}^d)^p$ and Φ_σ^T defined in (7). We have

$$\begin{aligned} \sum_\sigma \varphi(\sigma) \sum_{T, \sigma \in T} \Psi_\sigma^T &= \sum_T \sum_{\sigma \in T} \varphi(\sigma)\, \Psi_\sigma^T \\ &= \sum_T \sum_{\sigma \in T} \varphi(\sigma)\, \Phi_\sigma^T + \sum_T \sum_{\sigma \in T} \varphi(\sigma)(\Psi_\sigma^T - \Phi_\sigma^T) \\ &= \int \varphi^h \operatorname{div} f^h(u^h) + \sum_T \sum_{\sigma \in T} \varphi(\sigma)(\Psi_\sigma^T - \Phi_\sigma^T) \end{aligned}$$

We note that $\sum_{\sigma \in T} \Psi_\sigma^T = \sum_{\sigma \in T} \Phi_\sigma^T$, so, for any $\sigma' \in T$, we have

$$\sum_{\sigma \in T} \varphi(\sigma)(\Psi_\sigma^T - \Phi_\sigma^T) = \sum_{\sigma \in T} (\varphi(\sigma) - \varphi(\sigma'))(\Psi_\sigma^T - \Phi_\sigma^T)$$

If N is the number of degrees of freedom in T, and we assume here that it is independant of T, there are $C = \frac{N(N+1)}{2}$ such possibilities. Finaly we get the truncation error

$$\sum_{\sigma} \varphi(\sigma) \sum_{T,\sigma\in T} \Psi_\sigma^T = \int \varphi^h \operatorname{div} f^h(u^h) + \frac{1}{C} \sum_{T} \sum_{\sigma,\sigma'\in T} (\varphi(\sigma)-\varphi(\sigma'))(\Psi_\sigma^T - \Phi_\sigma^T). \tag{11}$$

Then we notice that $\int_T \psi_\sigma \operatorname{div} f(u^h) = \mathcal{O}(h^{r+2})$ for a steady problem if u^h is the interpolant of a regular enough function, say $u \in H^s$ for s large enough. In fact, if u^h represents the interpolant of such a function u, then

$$\begin{aligned}\int_T \psi_\sigma \operatorname{div} f(u^h)\, dx &= \int_{\partial T} \psi_\sigma f^h(u^h)\cdot \vec{n}\, dl - \int_T \nabla\psi_\sigma f^h(u^h) \\ &= \int_{\partial T} (f^h(u^h)-f(u))\cdot \vec{n}\, dl - \int_T \nabla\psi_\sigma (f^h(u^h)-f(u)) \\ &= \mathcal{O}(h)\times\mathcal{O}(h^{r+1})\end{aligned}$$

if $f^h(u^h)$ is a $r+1$th approximation of f. Similarly,

$$\int \varphi \operatorname{div} f(u) = \int \varphi^h \operatorname{div} f^h(u^h) + \mathcal{O}(h^{r+1})$$

under the same assumptions.

If we assume that $\Phi_\sigma^q = \mathcal{O}(h^{r+2})$ for a steady state solution, it is clear that the term

$$\sum_{\sigma,\sigma'\in T} (\varphi(\sigma)-\varphi(\sigma'))(\Psi_\sigma^T - \Phi_\sigma^T)$$

is $\mathcal{O}(h^{r+3})$, and proceeding as in [5], we conclude that

$$\left| \sum_{\sigma} \sum_{T,\sigma\in T} \varphi(\sigma)\, \Phi_\sigma - \int_\Omega \varphi(x) \operatorname{div} f(u)\, dx \right| \leqslant C(\mathscr{T})\, \|\nabla\varphi\|_\infty\, h^{r+1}$$

where $C(\mathscr{T})$ is a constant that depends on the constants C_1 introduced in the Assumption 2.1.

We have been rather vague on how the boundary conditions are implemented in practice. Since they are given on the inflow part of $\partial\Omega$, we set strongly at the inflow nodes. Doing that, we assume they are smooth enough, and the calculation of the truncation error becomes valid.

Hence, we have shown the following result

Proposition 2.3. Let be $r > 0$ an integer, and consider the steady problem (10). Under the following assumptions

1. The mesh satisfy Assumption 2.1,
2. For any H^{r+1} function u, denoting by u^h its interpolant of degree r, $f^h(u^h)$ is an approximation of order $r+1$ in the L^∞ norm,

3. For any triangle T and any sub-triangle T'_T of T, the sub-residuals $\Phi_\sigma^{T'_T}$ evaluated at u^h, the interpolant of the steady solution of $\operatorname{div} f(u) = 0$ satisfy $\Phi_\sigma^{T'_T} = \mathcal{O}(h^{r+2})$.

Then for any test function $\varphi \in C^{r+1}(\Omega)$, we have the following consistency estimate

$$\left| \sum_\sigma \sum_{T, \sigma \in T} \varphi(\sigma)\, \Phi_\sigma - \int_\Omega \varphi(x) \operatorname{div} f(u)\, dx \right| \leqslant C(\mathcal{T})\, \|\nabla \varphi\|_\infty\, h^{r+1}.$$

If so, we say that the scheme is $r+1$th order accurate.

Remark 2.4. It is important to see that the facts that (i) u^h is an interpolant of the *steady* solution and (ii) $f^h(u^h)$ is a good approximation of the true flux play a key role. In particular, the result is not true if one wishes to generalize it to unsteady problems with the same assumptions on the residuals. Here, what plays the role of residuals should include the time "residual" $\frac{\partial u}{\partial t}$ in addition to the space residual $\operatorname{div} f(u)$. This has been done for second order accuracy in [6]. If this is not done properly, the scheme is only first order accurate.

In Section 3, we recall schemes that are second order (for steady state) and monotonicity preserving. In Section 5, we construct similar schemes for $r \geqslant 3$ and provide a general construction method.

3. EXAMPLES OF SCHEMES

For second order accurate schemes, the degrees of freedom are the vertices of $\mathcal{T}_h$: we do not use the notation σ but M_i instead.

We recall the N schemes, LDA and PSI schemes developed by Roe, Deconinck and coworkers in the context of non-refined grids and P_1 elements [2]. They are examples of upwind schemes, meaning that the residuals Φ^T are split into signals Φ_i^T in a manner that reflects the local flow of information by enhancing those signals when M_i is a downwind node of T. These schemes satisfy the technical requirements of the previous sections. The PSI scheme is the only one that is both second-order and monotone. Schemes having both of these properties must be constructed in an essentially nonlinear way to avoid the restrictions of Godunov's Theorem, that linear monotone schemes are only first order. See for example [8].

We begin this section by a systematic construction of high order monotone scheme. Then we provide examples for linear and non linear problems.

3.1. Monotone Schemes

To begin with, assume that Φ_i^T, the part of the fluctuation assigned to the ith vertex, can be written as

$$\Phi_i^T = \sum_{j,\, M_j \in T} c_{ij}^T (u_i - u_j).$$

The coefficients c_{ij}^T may be independent of the data (linear schemes) or may depend on the data (non linear schemes.) The scheme is monotonicity preserving provided $c_{ij} \geqslant 0$ under the CFL like condition

$$\Delta t \left[\frac{\sum_{T, M_i \in T} \sum_{M_j \in T} c_{ij}^T}{|C_i|} \right] \leqslant 1 \qquad \text{for any} \quad M_i.$$

It is possible to construct monotonicity preserving second-order schemes for steady problems, but we have already noted that Godunov's Theorem states that they cannot be linear. A systematic construction is to begin with a monotone, linear, and therefore first-order scheme, and then to modify it so that it achieves higher-order accuracy.

In order to simplify the notations, from now on, unless explicitely mentionned, all the summations will refer to the vertices or degree of freedom within one given element. Similarily, we omit the superscript T in the residuals, since all the calculations refer to a single generic element.

Denote by Φ_i^M the part of the fluctuation assigned by some monotone scheme to the jth vertex, and by Φ_i^H the part assigned by a nonlinear high-order scheme. The requirements are that

$$\sum_j \Phi_j^H = \sum_j \Phi_j^M = \Phi \qquad \text{Conservation}$$

$$\Phi_j^H \Phi_j^M \geqslant 0 \qquad \text{Monotonicity preserving property}$$

We get second order accuracy under the following conditions

$$\Phi_j^H = \mathcal{O}(\Phi) \qquad \text{and}$$

$$\Phi = \mathcal{O}(h^3) \qquad \text{at steady state for smooth solutions.}$$

The two conditions are important from the analysis of Section 2.2.2; in particular they impose second order accuracy constraints on the flux approximation $f^h(u^h)$.

After introducing

$$\beta_j = \frac{\phi_j^M}{\phi}, \qquad \hat{\beta}_j = \frac{\phi_j^H}{\phi}$$

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

$$\sum_{j=1}^N \beta_j = \sum_{j=1}^N \hat{\beta}_j = 1 \qquad \text{Conservation}$$

$$\beta_j \hat{\beta}_j \geqslant 0 \qquad \text{Monotonicity}$$

$$\hat{\beta}_j \text{ is bounded} \qquad \text{High-order accuracy if } \phi^T \text{ is small enough}$$

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 triangle. 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 [9] and we expect their thorough investigation to take considerable time.

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

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

$$\hat{\beta}_j = \frac{\beta_j^+}{\sum_j \beta_j^+} \qquad \text{with} \quad x^+ = \max(0, x) \tag{12}$$

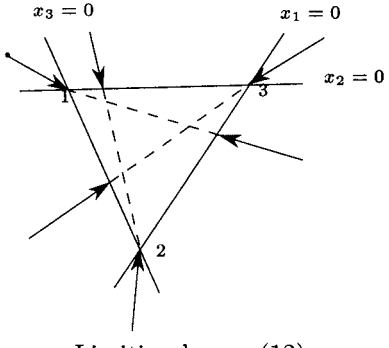
In Fig. 2 this is shown geometrically on the left. The mapping from $\boldsymbol{\beta}$ to $\hat{\boldsymbol{\beta}}$ is always a translation toward one of the vertices.

An alternative is to take $\hat{\boldsymbol{\beta}}$ as the point on the boundary of the triangle that is closest to $\boldsymbol{\beta}$ where the reference triangle (1, 2, 3) of Figure is materialized by an equilateral triangle. The following logic accomplishes that

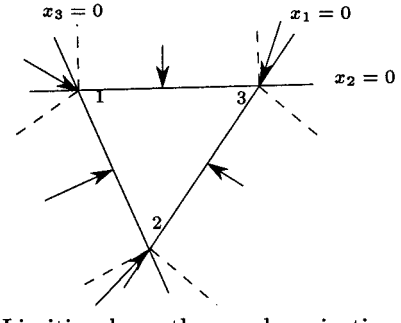
1. If $\beta_1, \beta_2, \beta_3$ are positive, define $\hat{\beta}_j = \beta_j$,
2. Else

(a) If β_1 is negative, define $\beta_2^* = \beta_2 + \frac{\beta_1}{2}$ and $\beta_3^* = \beta_3 + \frac{\beta_1}{2}$

(i) If $\beta_2^* \leqslant 0$, $\hat{\beta}_1 = \hat{\beta}_2 = 0$, $\hat{\beta}_3 = 1$



Limiting by eqn(12)



Limiting by orthogonal projection

Fig. 2. Geometrical illustration of the two limiters.

(ii) If $\beta_3^* \leqslant 0$, $\hat{\beta}_1 = \hat{\beta}_3 = 0$, $\hat{\beta}_2 = 1$

(iii) If $\beta_2^* \geqslant 0$ and $\beta_3^* \geqslant 0$, $\hat{\beta}_1 = 0$, $\hat{\beta}_2 = \beta_2^*$ and $\hat{\beta}_3 = \beta_3^*$

(b) If β_2 is negative, consider (a) with the transformation of indices and change the indices according the rules $1 \to 2$, $2 \to 1$, $3 \to 3$.

(c) If β_2 is negative, consider (a) with the transformation of indices and change the indices according the rules $1 \to 3$, $2 \to 2$, $3 \to 1$

Since $\hat{\boldsymbol{\beta}}_j$ is homogeneous of degree 1 in x_j, this can be coded without division. In Fig. 2 this map is shown geometrically on the right. Extension to $N \geqslant 4$ is straightforward for both maps.

3.2. Case of a Linear Flux

We start by $f(u) = \vec{\lambda} u$ where $\vec{\lambda}$ is some constant vector. If u^h is a piecewise linear function, the residual on T, Φ, can be written

$$\Phi = \sum_{M_i \in T} k_i u_i \tag{13}$$

where

$$k_i = \int_T \nabla \Lambda_i \cdot \vec{\lambda}\, dx = \tfrac{1}{2} \vec{\lambda} \cdot \vec{n}_i.$$

The vector $\vec{n}_i$ is the inward normal vector to T opposite the vectex M_i, see Fig. 3.

3.2.1. The N *Scheme*

The N scheme may be written in several ways, one of which is

$$\Phi_i = k_i^+ (u_i - \tilde{u}). \tag{14}$$

For consistency with the definition (13) we must have

$$\sum_j k_j^+ u_j - \left(\sum_j k_j^+ \right) \tilde{u} = \sum_j k_j u_j$$

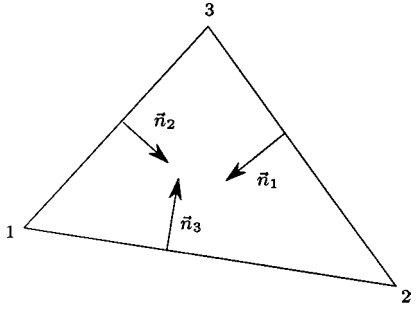


Fig. 3. Inward normal vectors $\vec{n}_i$.

so that

$$-\left(\sum_j k_j^+\right)\tilde{u} = \sum_j k_j^- u_j$$

because $k_j = k_j^+ + k_j^-$. Since $\sum_j k_j = 0$ because $\sum_j n_j = 0$, we have

$$\tilde{u} = \left(\sum_j k_j^-\right)^{-1}\left(\sum_j k_j^- u_j\right). \tag{15}$$

The N scheme is defined by Eqs. (14) and (15). We define

$$n = \left(\sum_j k_j^-\right)^{-1} \leqslant 0$$

and note that yet another definition of the N scheme would be

$$\Phi_i^T = \sum_j k_i^+ n k_j^- (u_i - u_j) = \sum_j c_{ij}(u_i - u_j). \tag{16}$$

where $c_{ij} = k_i^+ n k_j^-$. Since $c_{ij} \geqslant 0$, the N scheme is monotonicity preserving under the constraint

$$\frac{\Delta t}{|T|}\max_i \sum_{T,\, M_i \in T}\ \sum_{M_j \in T} c_{ij} \leqslant 1.$$

There are two types of triangles : those for which only one k_j is positive, and those for which two of the k_js are positive. In the first case, the triangle is called one target ; in the second case, it is a two target triangle. For a one target triangle, assume $k_1 > 0$, $k_2, k_3 \leqslant 0$, so that

$$\Phi_1^T = \Phi^T, \qquad \Phi_2^T = 0, \qquad \Phi_3^T = 0.$$

In this case the point $\boldsymbol{\beta}$ lies at one vertex of the triangle; either of the above mappings will leave it there . In a two target triangle, say k_1, $k_2 > 0$ and $k_3 < 0$. We have

$$\Phi_1^T = k_1(u_1 - u_3), \qquad \Phi_2^T = k_2(u_2 - u_3), \qquad \Phi_3^T = 0.$$

In this case $\boldsymbol{\beta}$ lies at one side of the triangle, but not necessarily between the vertices. In fact, in a two target case, we have, in a smooth steady state, $\Phi^T = \mathcal{O}(h^3)$, but Φ_1^T , $\Phi_2^T = \mathcal{O}(h^2)$. So the point $\boldsymbol{\beta}$ may lie arbitrarily far from the triangle. This example shows that the scheme is only first order accurate at steady state, and this is confirmed by numerical experiments.

3.2.2. The LDA Scheme

A second order upwind scheme is the LDA (Low Diffusion Advection) scheme defined by

$$\Phi_i = -n k_i^+ \Phi. \tag{17}$$

Since $\Phi_i^T = \sum_j -(k_i^+ nk_j)(u_i - u_j) = \sum_j c_{ij}(u_i - u_j)$ where $c_{ij} = -k_i^+ nk_j$, we see that c_{ij} may be either positive or negative, so that the LDA is not a monotonicity preserving scheme in general. However, (17) shows that each Φ_i is of the same order as Φ^T, and the scheme should be second-order in the steady state. Again this is confirmed by numerical experiments. However, we see that in the one target case, the two schemes coincide.

3.2.3. The PSI Scheme

The PSI (Positive Streamwise Invariant) scheme of Struijs [2] is constructed from the N scheme using one of the limiter presented in paragraph 3.1. Remarkably the scheme does not depend on which of the two limiters is chosen! This is because, in the two target case, $\boldsymbol{\beta}$ lies on an edge, and if it is outside the triangle, either limiter will move it to the nearest vertex. This gives precisely the successful PSI scheme of Struijs.

3.3. Case of a Non Linear Flux

Similar schemes can be constructed for the problem

$$\operatorname{div} f(u) = 0$$

provided we can find, for each element T, a *local linearization*, that is to say, a vector $\hat{\vec{\lambda}}$ such that, for a piecewise interpolant u^h,

$$\int_T \operatorname{div} f(u^h)\, dx = \int_T \hat{\vec{\lambda}} \cdot \nabla u^h\, dx.$$

In the case of Burgers' equation $f(u) = (u^2/2, u)$, which we employ later as a test problem in numerical examples, a solution is given by $\hat{\vec{\lambda}} = (\frac{u_1+u_2+u_3}{3}, 1)$, (u_1, u_2, u_3) being the values at the three vertices.

4. VARIATIONAL FORMULATION OF THE PSI AND LDA SCHEMES

In this section, we describe, for the linear PDE, a variational formulation of the PSI and LDA schemes that is useful in the sequel. The case of a non linear flux can be discussed in the same way, at least formaly.

We assume that u_{Γ^-} is a trace, that of a function also denoted by u_{Γ^-}. The solution u of the problem (1–2) can be rewritten as $u = u_{\Gamma^-} + v$, and v satisfies

$$\begin{aligned} &\vec{\lambda} \cdot \nabla v = g \qquad && x \in \Omega \\ &v = 0 && \text{on } \Gamma^-. \end{aligned} \tag{18}$$

with $g = \vec{\lambda} \cdot \nabla u_{\Gamma^-}$. For the sake of simplicity, we assume that Ω is polygonal, and that the mesh is adapted to the boundary ($\partial\Omega = \partial(\bigcup_{T \in \tau_h} T)$). To be consistent with the previous sections, the unknown function v in (18) is still denoted by the letter u in the rest of this section.

Taking φ piecewise linear, we have

$$\begin{aligned} &\sum_i \varphi(x_i) \sum_{T,\, M_i \in T} \Phi_i = 0, \qquad \text{for} \quad M_i \notin \Gamma^-, \\ &u_i = 0 \qquad\qquad\qquad\qquad\quad \text{if} \quad M_i \in \Gamma^- \end{aligned} \tag{19}$$

We assume that the residual Φ_i is

$$\Phi_i = \beta_i \Phi, \qquad \Phi = \int_T (\vec{\lambda} \cdot \nabla u^h - g)\, dx$$

with $\sum_j \beta_j = 1$ and β_j bounded. The relation (19) can be rewritten as

$$\sum_i \varphi(x_i) \sum_{T,\, M_i \in T} \Phi_i = \sum_T \left(\sum_{M_j \in T} \beta_j \varphi(x_j) \right) \Phi$$

Then, we have

$$\sum_T \left(\sum_{M_j \in T} \beta_j \varphi(x_i) \right) \Phi = \int_\Omega \ell(\varphi)(\vec{\lambda} \cdot \nabla u^h - g)\, dx \tag{20}$$

where, for any triangle T,

$$\ell(\varphi)_{|T} = \sum_{M_j \in T} \beta_j \varphi(x_j).$$

Consider the following space $V_h^1 = \{u^h$ piecewise linear and continuous, $u^h_{|\Gamma^-} = 0\}$ equipped with the graph norm

$$\|u^h\|^2_{V_h^1} = \int_\Omega (\vec{\lambda} \cdot \nabla u^h)^2.$$

Compared to Section 2.2.1, we have introduced the homogeneous Dirichlet boundary condition in the definition of V_h^1.

We denote by a the bilinear form on $V_h^1 \times V_h^1$ defined by

$$a(u^h, v^h) = \int_\Omega \ell(v^h)\, \vec{\lambda} \cdot \nabla u^h\, dx.$$

An element of V_h^1 is a solution of (19) in V_h^1 if and only if for any $v^h \in V_h^1$ we have

$$a(u^h, v^h) = \int_\Omega \ell(v^h)\, g$$

This is the variational formulation of (19).

We specialise now to the PSI scheme. The case of the LDA scheme is similar. The first remark is that for any vertex M_i, there is at least one triangle having M_i as vertex that is one target on M_i.

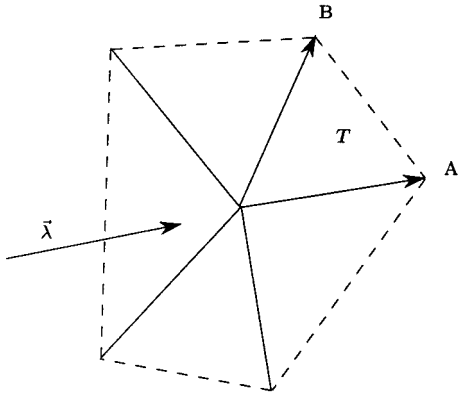


Fig. 4. There is at least one one–target triangle.

To prove that, consider all the triangle T sharing M_i. As in Fig. 4, we number the triangles in the positive orientation. Consider T a triangle, its vertices are $\{M_i, A, B\}$. The triangle T is one target to M_i if $\vec{\lambda}$ belongs to the angular sector defined by $\vec{M}_iA$, $\vec{M}_iB$, that is if $\det(\vec{\lambda}, \vec{M}_iA) > 0$ and $\det(\vec{\lambda}, \vec{M}_iB) < 0$. There is clearly at least one one-target triangle because as a point A rotates completely around M_i, the value of $\det(\vec{\lambda}, \vec{M}_iA)$ must change sign twice (when $\vec{M}_iA$ is parallel or antipallel to $\vec{\lambda}$). One of the sectors must span the transition from positive to negative. The proof extends to 3D.

The consequence of that is, denoting by T_{M_i} the one–target triangle at M_i, that

$$\int_\Omega \ell(v^h)^2\, dx \geqslant \sum_i |T_{M_i}|\, (v_i^h)^2.$$

If the mesh is regular, there exist a positive constant C such that

$$\sum_i |T_{M_i}|\, (v_i^h)^2 \geqslant C \int_\Omega (v^h)^2\, dx$$

(here we use the fact that $v_i = 0$ on Γ^-), so we get

$$\int_\Omega \ell(v^h)^2\, dx \geqslant C \int_\Omega (v^h)^2\, dx. \tag{21}$$

Using this remark, we show in Appendix 2, Proposition 2.1, that there exists $C > 0$ such that

$$\sup_{v_h \in V_h^1} \frac{\int_\Omega \ell(v^h)\, \vec{\lambda} \cdot \nabla u^h}{\sqrt{\int_\Omega (v^h)^2}} \geqslant C \sqrt{\int_\Omega (\vec{\lambda} \cdot \nabla u^h)^2}. \tag{22}$$

The inequality (22) shows that we have a uniform inf–sup condition, that is we have stability of the scheme (see [10]).

Moreover, it is possible to derive an error estimate. First, we notice that

$$a(u^h, v^h) \leqslant \|u^h\|_{V_h^1} \|\ell(v^h)\|_{L^2} \leqslant \|u^h\|_{V_h^1} \|v^h\|_{V_h^1}.$$

The last inequality comes from the fact $0 \leqslant \beta_i \leqslant 1$.

By Céa's lemma, we have the inequality

$$\|u^h - u\|_{V_h^1} \leqslant \left(1 + \frac{1}{C}\right) \inf_{w^h \in V_h^1} \|u - w^h\|_{V_h^1}. \tag{23}$$

In particular, if $u \in H^1(\Omega)$, we have the existence of C' such that

$$\|u^h - u\|_{V_h^1} \leqslant C'h. \tag{24}$$

The important fact are the relation (20) and the inequality (21). This inequality comes from the upwind property of the scheme and enables us to show the coercivity of the scheme. We note that this proof works also for the LDA scheme.

5. CONSTRUCTION OF HIGH ORDER UPWIND MONOTONICITY PRESERVING RESIDUAL DISTRIBUTIVE SCHEME

We show the explicit construction of a third order scheme. The extension to higher order accuracy is straightforward.

Denote V_h^2 the space of continuous functions that are quadratic in each triangle T and are vanishing on the inflow boundary Γ^-. The space V_h^2 is equipped with the graph norm. We want to construct a scheme whose variational formulation reads: find $u^h \in V_h^2$ such that

$$\forall v^h \in V_h^2, \qquad \int_\Omega \ell(v^h)\, \lambda \cdot \nabla u^h = \int_\Omega \ell(v^h)\, g.$$

We need to specify the mapping ℓ. The key is to have an inequality of the type (22). If so, the error estimate (23) is obvious, and (20) becomes

$$\|u^h - u\|_{V_h^2} \leqslant C'h^2. \tag{25}$$

Consider T a triangle, and T_I, T_{II}, T_{III} and T_{IV} as in Fig. 1, and we still denote by σ the degrees of freedom of u^h. We can construct the N scheme on the subtriangles of T. We denote by $\Phi_\sigma^{T'_T}$ for $T'_T = I, II, III, IV$ the corresponding signals sent to the points contained in T. Then we consider

$$\Phi_{T'_T} = \int_{T'_T} \vec{\lambda} \cdot \nabla u^h.$$

Since u^h is now quadratic,

$$\Phi_{T'_T} \neq \sum_{\sigma \in T'_T} \Phi_\sigma^{T'_T},$$

but we still can use the N scheme for a comparison purpose. We insist that the signals sent by the subtriangles to their vertices be limited to the same order of

magnitude as the restriction of the high-order residual $\Phi_{T'_T}$. Thus we define the scheme by a set of modified signals

$$\Psi_\sigma^{T'_T} = \beta_\sigma^{T'_T} \Phi_{T'_T} \tag{26}$$

with

$$\beta_\sigma^{T'_T} = \frac{\left(\dfrac{\Phi_\sigma^{T'_T}}{\Phi_{T'_T}}\right)^+}{\displaystyle\sum_{\sigma' \in T'_T} \left(\frac{\Phi_{\sigma'}^{T'_T}}{\Phi_{T'_T}}\right)^+} \tag{27}$$

and

$$\Psi_\sigma = \sum_{T'_T,\, \sigma \in T'_T} \Psi_\sigma^{T'_T}. \tag{28}$$

The conservation relation is

$$\sum_{\sigma \in T} \Psi_\sigma = \sum_{T'_T \subset T} \int_{T'_T} \vec{\lambda} \cdot \nabla u^h = \int_T \vec{\lambda} \cdot \nabla u^h.$$

Provided the stability conditions of the Lax Wendroff condition are true, the scheme will be convergent: u^h is continuous across the edges of T, thanks to the continuity of u^h, so is the flux $f^h(u^h) = \vec{\lambda} \cdot \nabla u^h$.

Let us come back to the stability property of the scheme. The variational formulation is given by

$$\ell(v^h)_{T'_T} = \sum_{\sigma \in T'_T} \beta_\sigma^{T'_T} v^h(\sigma).$$

The series of inequality (21–22) works, so that we get the stability and the error estimate (25). The proof is given in Appendix 2.

An extension of the LDA scheme (which has been tested and provide similar error curves, but the results are not reported here) is the following. Consider $T'_T \subset T$, define $k_\sigma^{T'_T} = \vec{\lambda} \cdot \vec{n}_\sigma^{T'_T}$ where $\vec{n}_\sigma^{T'_T}$ is the inward normal to T'_T opposite to σ. In (26), we replace $\beta_\sigma^{T'_T}$ by

$$\beta_\sigma^{T'_T} = -(k_\sigma^{T'_T})^+ \left(\sum_{\sigma' \in T'_T} (k_{\sigma'}^{T'_T})^+ \right)^{-1}.$$

This defines a high order LDA scheme. It is not monotone.

6. SOME NUMERICAL RESULTS

In each figure, all the degrees of freedom are ploted.

6.1. Linear Advection

We consider the problem

$$
\begin{aligned}
&-y\frac{\partial u}{\partial x}+x\frac{\partial u}{\partial y}=0 \qquad (x, y)\in[-1, 1]\times[0, 1]\\
&u(x, 0)= \begin{cases} 0 & \text{if } \ x\notin[0.1, 0.7]\\ \sin\left(\pi\frac{x-0.1}{0.6}\right) & x\in[0.1, 0.7]\end{cases}
\end{aligned}
\tag{29}
$$

The exact solution is, with $r=\sqrt{x^2+y^2}$, $u(x, y)=\sin(\pi\frac{r-0.1}{0.6})$ if $r\in[0.1, 0.7]$ and 0 elsewhere. The numerical solution is computed with a pseudo-time marching algorithm where the initial solution is $u_0=2$.

On each sub-triangle T'_T, we employ the scalar N scheme. The average velocity is defined by

$$(-\bar{y}, \bar{x})=\int_T (-y, x)\,dx\,dy.$$

Three schemes are compared: the standard second order PSI scheme, a third order scheme and a fourth order one. The high order schemes are constructed using P_3 and P_4 interpolation with the first limiter of Section 5. The mesh has 628 vertices and 1162 elements and is displayed on Fig. 5.

The isolines of the solutions are displayed on Fig. 6. The circular shape of the solution is perfectly well respected in each case. What is more interesting is a plot in the exit section, i.e., for $x=0$ and $y\leqslant -1$. This is given in Fig. 7 with a comparison with the exact one. We also provide a zoom of the solution around $(x, y)=(0, -0.5)$ where the solution is maximum. This figure shows clearly that an increase of the formal accuracy of the scheme does improve its effective accuracy. The fourth order scheme is almost perfect.

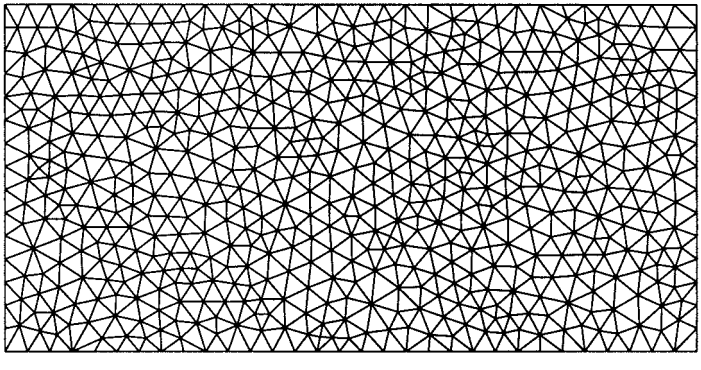


Fig. 5. Mesh for the linear circular advection problem.

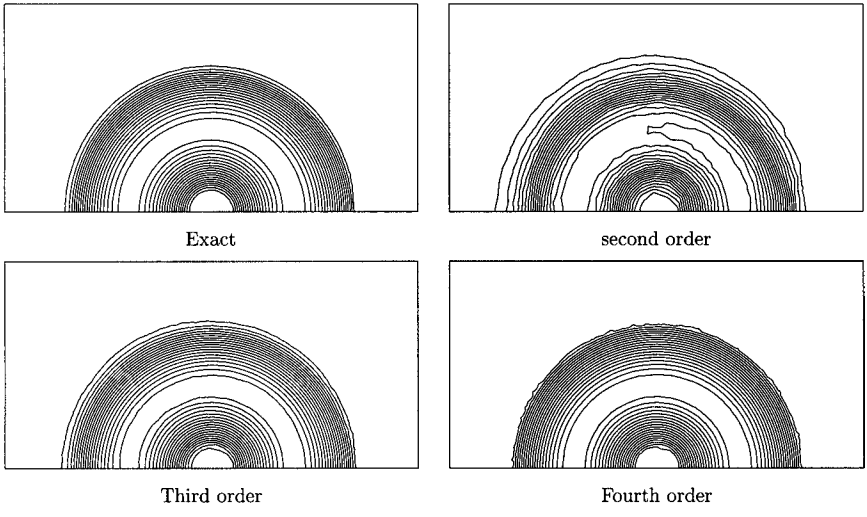


Fig. 6. Isolines of the exact and computed solutions.

Last we provide errors for the advection problem $\vec{\lambda} = (0, 1)^T$ and the boundary condition $u(x, y) = \cos \pi x$ on $[-1, 1] \times [0, 1]$. On Fig. 8, we provide the L^2 errors for the second order PSI scheme (Fig. 8a), the third order one (Fig. 8b) and the fourth order one (Fig. 8c). We see that the error slope of the r+1 th order scheme is between $r+1/2$ and $r+1$. This is not in contradiction with the expected theoretical $r+1/2$ order of accuracy. On Fig. 9, we display the L^∞ errors: the second order PSI scheme (Fig. 9a), the third order one (Fig. 9b) and the fourth order one (Fig. 9c). We see that the L^∞ errors are closer from the slope $r+1/2$, which seems to indicate that there is no clipping phenomena as in "standard" schemes. However, the fourth order schemes does not provide the expected accuracy, see Fig. 9c. This may be a

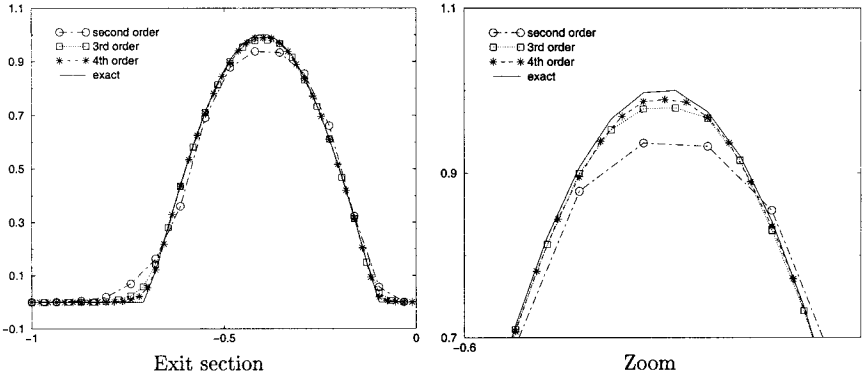


Fig. 7. Plot of the solutions in the interval $[-1, 0]$, and zoom around the maximum.

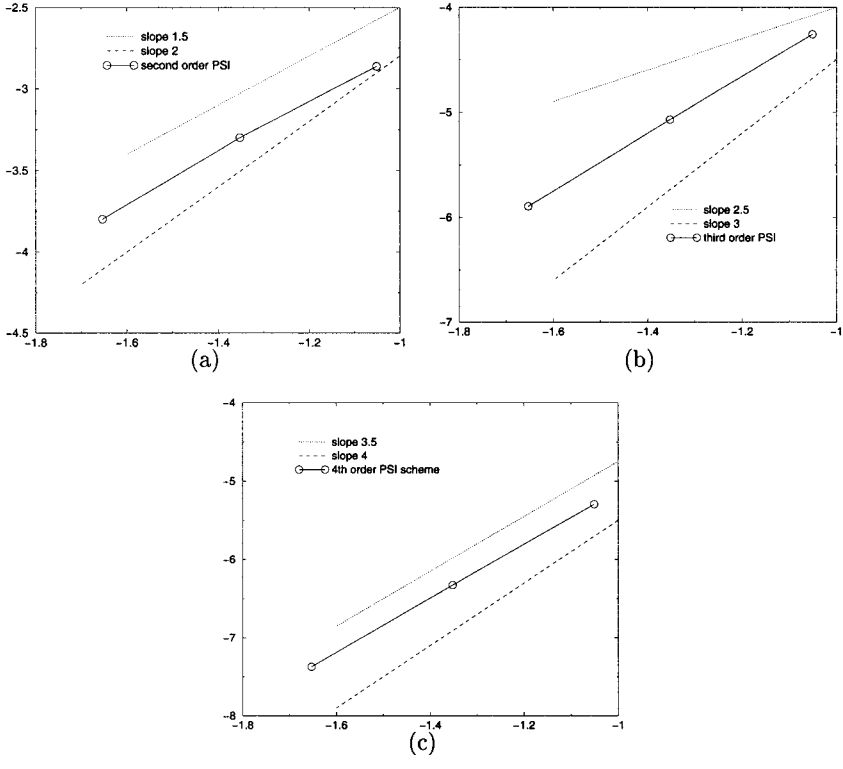


Fig. 8. L^2 error of the: (a) second order, (b) third order and (c) fourth order PSI schemes.

implementation error, or some roundoff problem (the computations have been performed on a 1Ghz Pentium IV processor with the Intel Fortran 90 compiler, and we have already experienced similar problems on this processor).

To obtain these results, a careful implementation of the limiter function

$$\hat{\beta}_i = \frac{\left(\frac{\Phi_i}{\Phi}\right)^+}{\sum_{j=1}^{3}\left(\frac{\Phi_j}{\Phi}\right)^+}$$

is needed. The first numerical results did not quite deliver the expected accuracy. It seems more important for the higher order schemes than for the the second order PSI scheme that $\sum_j \hat{\beta}_j = 1$ exactly. We have chosen to compute the revised weights (in pseudo fortran) as

$$\hat{\beta}_i = \frac{(1+\text{sign}(1., \Phi_i\Phi))\,\Phi_i + \epsilon}{\sum_{j=1}^{3}(1+\text{sign}(1., \Phi_j\Phi))\,\Phi_j + 3\epsilon}$$

with $\epsilon = 10^{-10}$ in double precision.

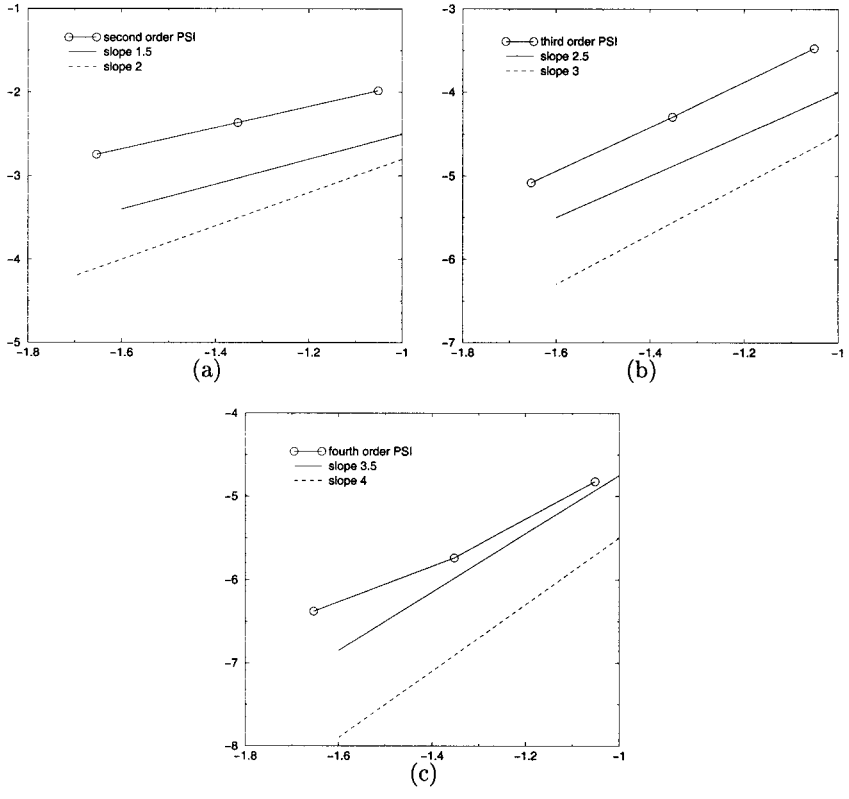


Fig. 9. L^{∞} error of the: (a) second order, (b) third order and (c) fourth order PSI schemes.

6.2. Burger Equation

We consider the Burger equation

$$\frac{1}{2}\frac{\partial u^2}{\partial x}+\frac{\partial u}{\partial y}=0, \qquad x\in[0,1]\times[0,1]$$

$$u(x,y)=1.5-2x \qquad \text{on the inflow boundary}$$

The exact solution is

$$u(x,y)=\begin{cases}-0.5 & \text{if} \quad y\leqslant 0.5 \quad \text{and} \quad -2(x-\frac{3}{4})+y-\frac{1}{2}\leqslant 0\\ 1.5- & \text{if} \quad y\leqslant 0.5 \quad \text{and} \quad -2(x-\frac{3}{4})+y-\frac{1}{2}\geqslant 0\\ \max\left(-0.5, \min\left(1.5, \dfrac{x-\frac{3}{4}}{y-\frac{1}{2}}\right)\right) & \text{else}\end{cases}$$

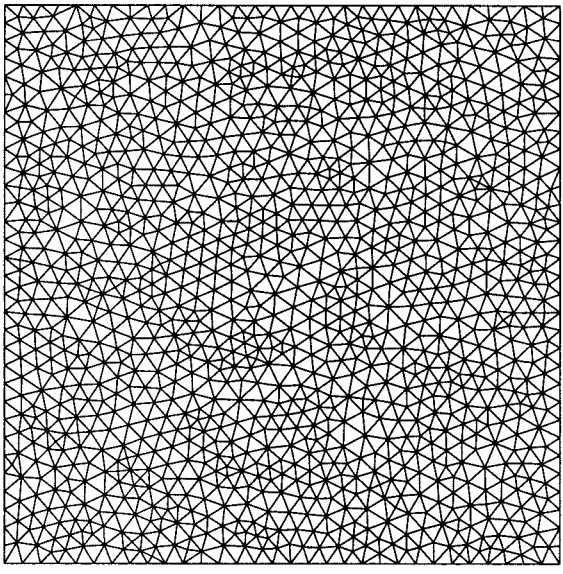


Fig. 10. Mesh for the Burger's equation.

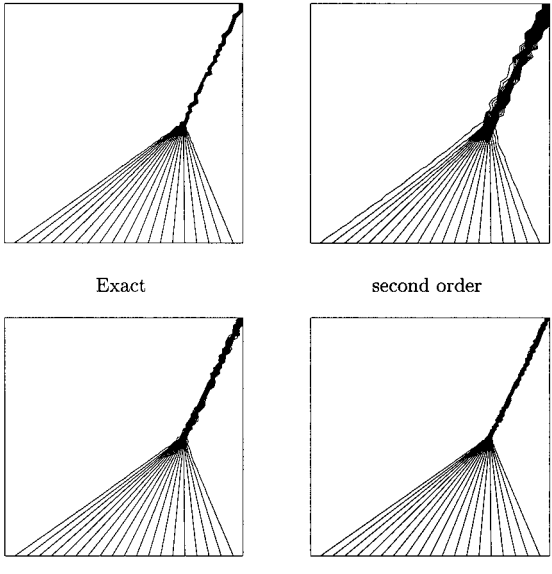


Fig. 11. Isolines of the exact and computed solutions.

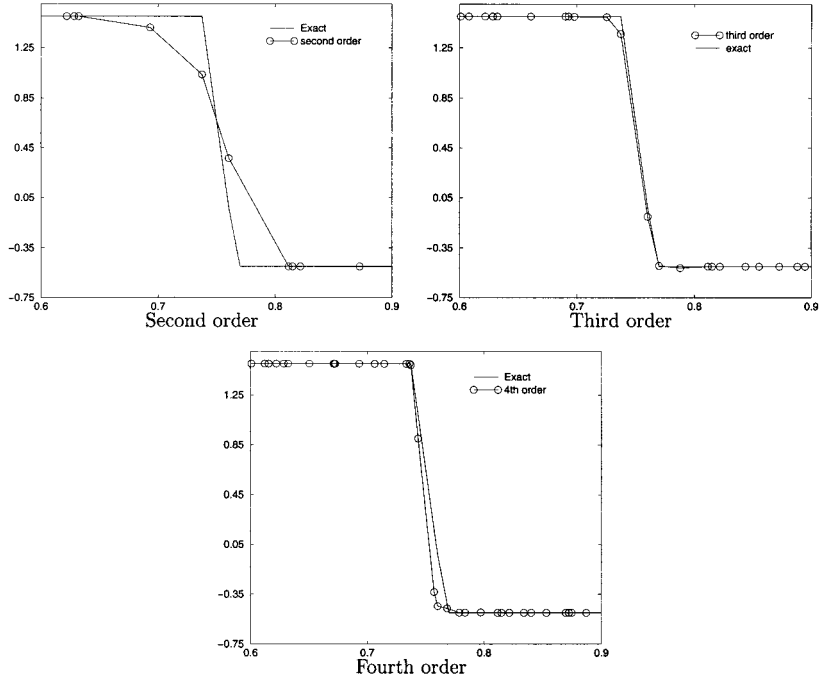


Fig. 12. Plot of the solutions across the shock, $y = 0.75$.

Once more, the solution is computed with a pseudo-time marching algorithm, the initial condition is set to $u = -0.5$. We represent the solutions computed for the mesh represented on Fig. 10 with 1041 vertices and 1960 triangles.

The isolines of the solutions are displayed on Fig. 11.

We also display cross sections for $y = 0.75$ across the discontinuity (Fig. 12) and $y = 0.25$ across the fan (Fig. 13).

Last, we give a zoom of the cross-section $y = 0.25$ around the corner of the fan on Fig. 14. We see a very clear improvement of the quality of the solution with the increase of the accuracy order.

7. CONCLUSIONS

We have presented and analyzed a class of schemes for hyperbolic problems, with a particular emphasis on scalar problems. We have provided a very general condition that ensures that the limit solution, if it exists, is a weak solution of the problem. Then we have discussed the construction of second-order monotone schemes from a new perspective that extends very naturally to higher-order. It is, however, not the only possible extension; it should be compared with proposals

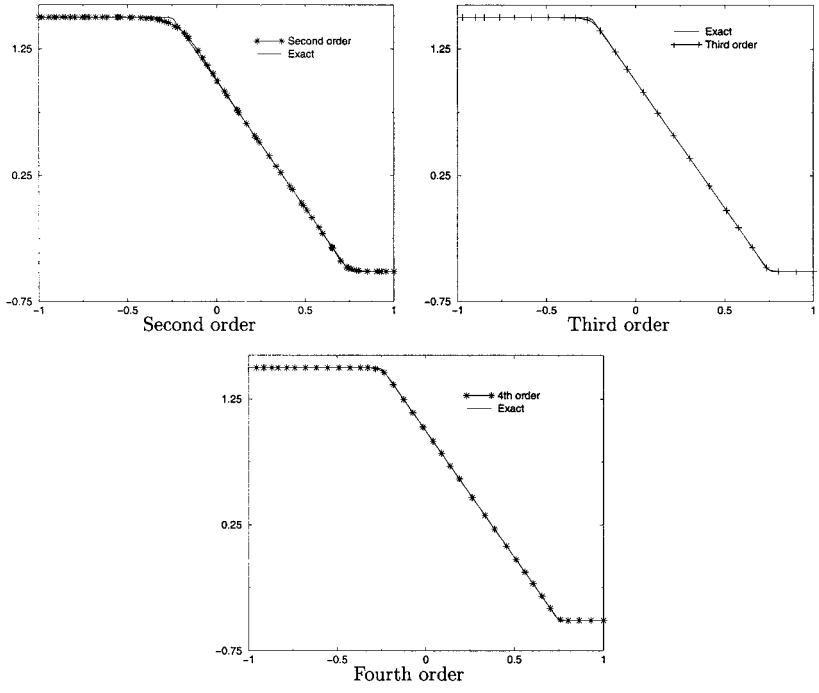


Fig. 13. Plot of the solutions across the fan, $y = 0.25$.

recently put forward by Caraeni [7] (see also [11]). It also has many points of similarity with ideas being developed in a finite-volume context by Wang [12]. Many variations of the present method are also possible, in particular the use of other function spaces for representing the solution. The extension to quadrangular meshes can be conducted by cutting every quadrangle into two triangles. Exploring these possibilities may lead to robust, high-order, methods that are less costly than the currently popular Discontinuous Galerkin methods.

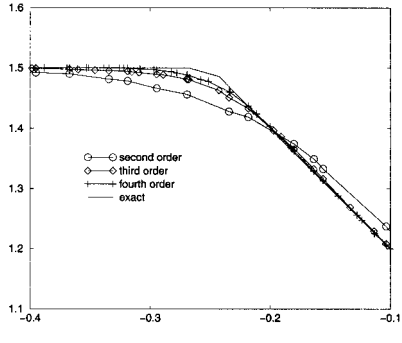


Fig. 14. Zoom across the fan, around the corner.

The extension to systems can proceed in several ways. The underlying N scheme can be extended formally by treating the coefficient k_j in Section 3 as matrix-valued [13]. Another method is to decompose a vector-valued residual into components associated with scalar advection problems and Cauchy-Riemann systems, treating each of these components independently, as discussed in [11].

APPENDIX A. PROOF OF THEOREM 2.2

For the sake of simplicity, we assume $d = 2$, the generalisation is immediate.

To begin with, we recall the notations of Section 2.2.1. If $\mathscr{T}_h$ is a mesh, T is a generic triangle, T'_T is any of the sub-triangles constructed from the degrees of freedom of the scheme (7)

$$u^{n+1}(\sigma) = u^n(\sigma) - \frac{\Delta t}{|C'_\sigma|} \sum_{T'_T \subset T,\, \sigma \in T'_T} \Phi_\sigma^{T'_T}$$

and $\mathscr{C}_h$ is the set of dual control volumes associated to them.

We also need the two spaces

- $V_h^k = \{v_h \in C^0(\mathbb{R}^2)^p;\ v_{h|T}$ polynomial of degree k, $\forall\, T'_T \subset T \in \mathscr{T}_h\}$
- $X^h = \{v_h;\ v_{h|C}$ constant $\in \mathbb{R}^p, \forall\, C \in \mathscr{C}_h\}$.

Last, π_h^k is the P^k interpolation defined on V_h^k.

We start by a lemma that is inspired from [15].

Lemma A.1. Let $\tau > 0$ and N the integer part of $\frac{\tau}{\Delta t}$. We consider $\mathscr{Q} \subset \mathbb{R}^2$, a bounded domain. Let $(u^h)_h$ be a sequence such that $u^h(.\,, t_n) \in (X_h)^p$ for any $n \leqslant N$. We assume there exists a constant C independant of h and $u \in L^2_{loc}(\mathscr{Q} \times [0, \tau])^p$ such that

$$\sup_h \sup_{x,y,t} |u^h(x, y, t)| \leqslant C, \qquad \lim_h \|u^h - u\|_{L^2(\mathscr{Q} \times [0,T])} = 0.$$

Then

1. $\lim_h \left(\sum_{n=0}^N \Delta t \sum_{T \subset \mathscr{Q}} \sum_{T'_T \subset T} |T'_T| \sum_{\sigma, \sigma' \in T'_T} \|u^h(\sigma) - u^h(\sigma')\|\right) = 0.$
2. $\tilde{u}^h = \pi_h^1 u^h$ satisfies $\lim_h h \|\nabla \tilde{u}^h\|_{L^2(\mathscr{Q} \times [0,T])} = 0.$

Proof. It is enough to prove the lemma for real valued functions v and $(v^h)_h$ that satisfies the assumptions.

First claim. For any sub-triangle T'_T, we define (ω being a cycle on the three vertices of T'_T)

$$v^h_{|T} = \sum_{\sigma \in T} v^h(\sigma)\, \chi_{C_\sigma \cap T}$$

$$w^h_{|T} = \sum_{\sigma \in T} v^h(\omega(\sigma))\, \chi_{C_\sigma \cap T}.$$

This enables to define two functions on $\mathbb{R}^2 \times \mathbb{R}^+$ that are bounded independantly of h. Moreover,

$$|T| \sum_{\sigma \neq \sigma'} |v^h(\sigma) - v^h(\sigma')| = 3 \int_T |v^h - w^h| \, dx.$$

The "3" comes form the definition of the dual cells,

We have

$$\begin{aligned}
&\sum_{n=0}^{N} \Delta t \sum_{T \subset \mathcal{Q}} \sum_{T'_T \subset T} |T'_T| \sum_{\sigma, \sigma' \in T'_T} |(v^h(\sigma))^n - (v^h(\sigma'))^n| \\
&\quad = 3 \int_0^{t_N} \int_{\cup T \subset \mathcal{Q}} |v^h - w^h| \, dx \, dt \\
&\quad \leqslant 3 \int_0^{\tau} \int_{\cup T \subset \mathcal{Q}} |v^h - w^h| \, dx \, dt.
\end{aligned}$$

Since (v^h) is bounded, there exists $v' \in L^\infty(\mathcal{Q} \times [0, \tau])$ such that $v^h \to v'$ for the weak star topology. Similarly, the exists $w \in L^\infty(\mathcal{Q} \times [0, \tau])$, with $w^h \to w$ for the weak star topology.

Since $v_h \to v$ in L^2_{loc}, we have $v' = v$ because $\mathcal{Q} \times [0, \tau]$ is bounded and $C_0^\infty(\mathcal{Q} \times [0, \tau])$ is dense in $L^1(\mathcal{Q} \times [0, \tau])$.

Let us show that $w = v$. Let be $\phi \in C_0^\infty(R^2 \times \mathbb{R}^+)^p$ We have

$$\begin{aligned}
\int_0^\tau \int_{\mathcal{Q}} v_h \phi \, dx \, dt &= \int_0^\tau \sum_{T, T \cap \mathcal{Q} \neq \varnothing} \sum_{T'_T \subset T} \sum_{\sigma \in T'_T} v_\sigma^h \int_{C_\sigma \cap T'_T} \phi(x, t) \, dx \, dt \\
&= \int_0^\tau \sum_{T, T \cap \mathcal{Q} \neq \varnothing} \sum_{T'_T \subset T} \sum_{\sigma \in T'_T} w_\sigma^h \int_{C_{\omega^{-1}(\sigma)} \cap T} \phi(x, t) \, dx \, dt \\
&= \int_0^\tau \int_{\mathcal{Q}} w_h \phi \, dx \, dt + \int_0^\tau \sum_{T, T \cap \mathcal{Q} \neq \varnothing} \sum_{T'_T \subset T} \sum_{\sigma \in T'_T} \\
&\quad \times w_\sigma^h \left(\int_{C_{\omega^{-1}(\sigma)} \cap T} \phi(x, t) \, dx \, dt - \int_{C_\sigma \cap T} \phi(x, t) \, dx \, dt \right)
\end{aligned}$$

We can also write

$$\int_{C_\sigma \cap T'_T} \phi \, dx \, dt = |C_\sigma \cap T'_T| \, \phi(\sigma), \qquad \int_{C_{\omega^{-1}(\sigma)} \cap T'_T} \phi \, dx \, dt = |C_\sigma \cap T'_T| \, \phi(\sigma')$$

for $\sigma \in C_\sigma \cap T'_T$ and $\sigma' \in C_{\omega^{-1}(\sigma) \cap T'_T}$ well chosen. Since $\nabla \phi$ is bounded on $\mathcal{Q} \times [0, \tau]$ and v^h is bounded, since $|C_\sigma \cap T'_T| = |C_{\omega^{-1}(\sigma)} \cap T'_T|$,[3] we have

$$\left| \int_0^\tau \int_{\mathcal{Q}} v_h \phi \, dx \, dt - \int_0^\tau \int_{\mathcal{Q}} w_h \phi \, dx \, dt \right| \leqslant Ch$$

where C is independant of h. Hence, $w = v' = v$.

[3] See Remark A.2

The same technique shows that (v_h^2) and (w_h^2) have the same weak star limit. Let us show it is v^2. By density of $C_0^\infty(\mathcal{Q}\times[0,\tau])$ in $L^1(\mathcal{Q}\times[0,\tau])$ and since v_h^2 is bounded independantly of h, we may choose test functions ϕ in $C_0^\infty(\mathcal{Q}\times[0,\tau])^p$. The function ϕ is bounded in $\mathcal{Q}\times[0,\tau]$. Hence, we have

$$\int_{\mathcal{Q}\times[0,\tau]} |v-v_h|^2\,\phi\,dx\,dt \to 0,$$

and then

$$\int_{\mathcal{Q}\times[0,\tau]} v^2\phi\,dx\,dt - 2\int_{\mathcal{Q}\times[0,\tau]} vv_h\phi\,dx\,dt + \int_{\mathcal{Q}\times[0,\tau]} v_h^2\phi\,dx\,dt \to 0.$$

By the Cauchy Schwarz inequality, $v\phi \in L^1(\mathcal{Q}\times[0,\tau])$; the second term tends towards

$$\int_{\mathcal{Q}\times[0,\tau]} v^2\phi\,dx\,dt.$$

and the $v_h^2 \to v^2$ in L^∞ weak star.

Last, by the same argument for $\phi = 1$, since $w_h^2 \to v^2$ in L^∞ weak star, we obtain

$$\int_{\mathcal{Q}\times[0,\tau]} |w_h-v|^2\,dx\,dt \to 0,$$

and finally,

$$\int_{\mathcal{Q}\times[0,\tau]} |w_h-v_h|^2\,dx\,dt \to 0.$$

The conclusion comes because $\mathcal{Q}\times[0,\tau]$ is bounded, hence $L^1(\mathcal{Q}\times[0,\tau]) \subset L^2(\mathcal{Q}\times[0,\tau])$.

Second claim. We have

$$\begin{aligned}\nabla\tilde{v}_h^{|T'_T} &= \frac{1}{2\,|T'_T|}\left(u_{\sigma_1}\vec{n}_{\sigma_1}^{T'_T} + u_{\sigma_2}\vec{n}_{\sigma_2}^{T'_T} + u_{\sigma_3}\vec{n}_{\sigma_3}^{T'_T}\right)\\ &= \frac{1}{2\,|T'_T|}\left((u_{\sigma_2}-u_{\sigma_1})\,\vec{n}_{\sigma_2}^{T'_T} + (u_{\sigma_3}-u_{\sigma_1})\,\vec{n}_{\sigma_3}^{T'_T}\right),\end{aligned}$$

thus

$$\begin{aligned}\Delta t\sum_0^N \int_{\mathcal{Q}} \|\nabla\tilde{v}_h^n\|^2\,dx\,dt &= \Delta t\sum_0^N \sum_{T,\,T\subset\mathcal{Q}} |T| \sum_{T'_T\subset T} \|(\nabla\tilde{v}_h^n)_{|T'_T}\|^2\\ &\leqslant C_1\Delta t\sum_0^N \sum_{T,\,T\subset\mathcal{Q}} \sum_{T'_T\subset T} \sum_{\sigma,\sigma'\in T'_T} |v_\sigma^n - v_{\sigma'}^n|^2\end{aligned}$$

because the triangulation is regular (Assumption 2.1). We conclude by the same arguments as before. □

Remark A.2. One can consider distributive schemes that are formulated on non triangular meshes. In this case, the arguments remain similar, as soon as we assume that the elements in the mesh have a number of edges that is bounded independantly of h, and that the area of the elements satisfies for any Ω bounded

$$\sum_{T_\sigma \cap \Omega \neq \varnothing} \sum_{(T'_T)_{\sigma_l} \text{ neighbor of } (T'_T)_{\sigma_i}} ||(T'_T)_{\sigma_i}| - |(T'_T)_{\sigma_{i_l}}|| \to 0$$

when $h \to 0$.

Lemma A.3. Let $\varphi \in C_0^{k+1}(\mathbb{R}^2 \times \mathbb{R}^+)$. With the assumptions of Theorem 2.2, one has

$$\sum_n \Delta t \sum_\sigma |C_\sigma| \, (u_\sigma^{n+1} - u_\sigma^n) \, \varphi(\sigma, t_n) + \int_{\mathbb{R}^2 \times \mathbb{R}^+} u \frac{\partial \varphi}{\partial t} \, dx \, dt + \int_{\mathbb{R}^2} u_0(x) \, \varphi(x, 0) \to 0$$

when $h \to 0$.

The proof is classical: it is the same as in the proof of the classical Lax Wendroff theorem.

Let us introduce ψ_σ the basis function at σ and Ψ_σ^T, $\Psi_\sigma^{T'}{}_T$ the signals

$$\Psi_\sigma^T = \int_T \psi_\sigma \operatorname{div} f^h(u^h),$$

$$\Psi_\sigma^{T'}{}_T = \int_{T'_T} \psi_\sigma \operatorname{div} f^h(u^h).$$

We have

Lemma A.4. Let $\varphi \in C_0^{k+1}(\mathbb{R}^2 \times \mathbb{R}^+)$, $f \in (C^1(\mathbb{R}^m))^d$, $d = 2$. Assume that f^h satisfies the Assumption 2.2(iii), and that u_h satisfies those of Theorem 2.2. Then

$$\Delta t \sum_{n,T} \sum_{\sigma \in T'_T \subset T} \varphi(\sigma, t_n) \, \Psi_\sigma^{T'_T}(u_h^n) + \int_{R^2 \times \mathbb{R}^+} f(u(x,t)) \operatorname{div} \varphi(x,t) \, dx \, dt \to 0$$

when $h \to 0$.

Proof. Let $\varphi \in C_0^1(\mathbb{R}^2 \times [0, +\infty[)$. Let Ω and τ such that $\operatorname{supp}(\varphi) \subset \Omega \times [0, \tau]$. Consider $T \in \mathcal{T}_h$ Now $\pi_h^k \varphi$ is the interpolation of φ of degree k. We have:

$$\Delta t \sum_{n,T} \sum_{\sigma \in T'_T} \varphi(\sigma, t_n) \, \Psi_\sigma^{T'_T}(u_h^n) = \sum_{n,T} \int_{t^n}^{t^{n+1}} \int_T \pi_h^k \varphi(x, t_n) \operatorname{div} f_T^h(u_h^n) \, dx \, dt$$

One can apply the Green formula,

$$\sum_T \int_T \pi_h^k \varphi(x, t^n) \operatorname{div} f_T^h(u_h^n) \, dx \, dt = -\sum_T \int_T \nabla(\pi_h^k \varphi)(x, t_n) . f_T^h(u_h^n) \, dx$$

$$+ \sum_T \int_{\partial T} \pi_h^k \varphi(x, t^n) \, f_T^h(u_h^n) . \vec{n} \, dx.$$

Since $\pi_h^k\varphi$ is continuous and from Assumption 2.2(iii), the second sum on the right hand side cancels.

Moreover, φ is C^{k+1} and the triangulation is regular: $\nabla(\pi_h^k\varphi)$ is uniformly bounded by a constant C independant of h. Thus we have

$$\left|\sum_{n,T}\int_{T\times[t_n,t_{n+1}]}\nabla(\pi_h^k\varphi)(x,t_n).f_T^h(u_h^n)\,dx\,dt-\sum_{n,T}\int_{T\times[t_n,t_{n+1}]}\nabla\varphi.f(u)\,dx\,dt\right|$$
$$\leqslant C\sum_{n,T}\int_{T\times[t^n,t^{n+1}]}\|f^h(u_h)-f(u)\|\,dx\,dt$$
$$+\sum_{n,T}\int_{T\times[t^n,t^{n+1}]}\|\nabla\pi_h^k\varphi-\nabla\varphi\|\,\|f(u_h)\|\,dx\,dt$$

The first sum is less than $\|f^h(u_h)-f(u)\|_{L^1(\Omega\times[0,\tau])}$ and tends to 0 because $\|u_h\|_\infty$ is bounded independantly of h, f is C^1 and $u_h\to u$ in L^2_{loc}.

Similarly, since u_h is bounded and f is continuous, $f(u_h)$ is bounded uniformly in h by a constant C. The second term of the right hand side sum is bounded by the L^1 norm of $\nabla\pi_h^k\varphi-\nabla\varphi$ that tends to 0 since the triangulation is uniform. □

Proof of Theorem 2.2. We multiply (7) by $\varphi(\sigma,t^n)\,|C_\sigma|$, where the test function φ belongs to $C_0^{k+1}(\mathbb{R}^2\times[0,+\infty[)$, such that $\mathrm{supp}(\varphi)\subset\Omega\times[0,\tau]$ for Ω and T well chosen ($n\Delta t\leqslant T$). Then we sum up on J the set of (σ,t_n) such that $(\sigma,t_n)\in\Omega$:

$$\sum_{(\sigma,t_n)\in J}|C_\sigma|\,(u_\sigma^{n+1}-u_\sigma^n)\,\varphi(\sigma,t^n)+\Delta t\sum_{(\sigma,t_n)\in J}\sum_{T_T';\,\sigma\in T_T'}\Phi_\sigma^{T_T'}(u_h^n)\,\varphi(\sigma,t^n)=0,$$

From Lemma A.3, we have

$$\sum_{(\sigma,t_n)\in J}|C_\sigma|\,(u_\sigma^{n+1}-u_\sigma^n)\,\varphi(\sigma,t^n)=-\int_{\mathbb{R}^2\times\mathbb{R}^+}u\frac{\partial\varphi}{\partial t}-\int_{\mathbb{R}^2}u_0\varphi(.,0)+o(1).$$

For the space term, we write:

$$\sum_{(\sigma,t_n)\in J}\sum_{T_T';\,\sigma\in T_T'}\Phi_\sigma^{T_t',n}\varphi(\sigma,t_n)$$
$$=\sum_{T\subset\Omega}\left(\sum_{\sigma\in T_T'\subset T}\varphi(\sigma,t_n)(\Phi_\sigma^{T,n}-\Psi_\sigma^{T_T',n})\right)+\sum_{T\subset\Omega}\sum_{\sigma\in T_T'\subset T}\varphi(\sigma,t_n)\,\Psi_\sigma^{T_T',n},$$

By the conservation relation, we have

$$\sum_{T_T'\subset T}\sum_{\sigma\in T_T'}(\Phi_\sigma^{T_T',n}-\Psi_\sigma^{T_T',n})=0,$$

so we have, whatever σ' degree of freedom in T,

$$\begin{aligned}
&\sum_{T'_T \subset T} \sum_{\sigma \in T'_T} \varphi(\sigma, t_n)(\Phi_\sigma^{T,n} - \Psi_\sigma^{T'_T,n}) \\
&\quad = \sum_{T'_T \subset T} \sum_{\sigma \in T'_T} \varphi(\sigma, t_n)(\Phi_\sigma^{T,n} - \Psi_\sigma^{T'_T,n}) - \varphi(\sigma', t_n) \sum_{T'_T \subset T} \sum_{\sigma \in T'_T} (\Phi_\sigma^{T'_T,n} - \Psi_\sigma^{T'_T,n}) \\
&\quad = \sum_{T'_T \subset T} \sum_{\sigma \in T'_T} (\varphi(\sigma, t_n) - \varphi(\sigma', t_n))(\Phi_\sigma^{T'_T,n} - \Psi_\sigma^{T'_T,n}).
\end{aligned}$$

Thanks to this relation, we can rewrite the second term of the previous sum as

$$\frac{k(k+1)}{2} \sum_{T \subset \Omega} \left(\sum_{\sigma, \sigma' \in T} (\varphi(\sigma, t_n) - \varphi(\sigma', t_n))(\Phi_\sigma^{T,n} - \Psi_\sigma^{T,n}) \right)$$

By Lemma A.4, we have

$$\Delta t \sum_{n=0,N} \sum_{\sigma \in J} \sum_{T;\sigma \in T} \Phi_\sigma^{T,n} \varphi(\sigma, t_n) = -\int_{\mathbb{R}^2 \times \mathbb{R}^+} f(u(x,t)) \operatorname{div} \varphi(x,t)\, dx\, dt + o(1).$$

Last, from Assumption 2.1 and Lemma A.1, setting $\Phi_\sigma^{T,n} - \Psi_\sigma^{T,n} = \mathscr{S}_\sigma^{T,n}$ to simplify, we have

$$\begin{aligned}
&\Delta t \left| \sum_{n,T} \sum_{\sigma \neq \sigma' \in T'_T} \mathscr{S}_\Sigma^{T,n} (\varphi(\sigma, t_n) - \varphi(\sigma', t_n)) \right| \\
&\qquad \leqslant Ch\, \Delta t \sum_{n=0,N} \sum_{T \subset \Omega} |\mathscr{S}_\sigma^T(u_h^n)| \\
&\qquad \leqslant Ch^2\, \Delta t \sum_{n=0,N} \sum_{T \subset \Omega} \sum_{\sigma, \sigma' \in T} |u^n(\sigma) - u^n(\sigma')| \to 0,
\end{aligned}$$

because both Ψ_σ and Φ_σ satisfy Assumption 2.3. This ends the proof. □

Remark A.5. If we assume only the continuity of the fluctuations and a regular mesh, the last part of the proof remains identical since it is sufficient to show that the translated of u_h^n which come in the computation of ϕ_σ^T converge to u at least for a subsequence and almost everywhere. As for the classical Lax–Wendroff theorem, this is proved via Lemma A.1.

APPENDIX B. STABILITY PROOF

Define V_h^k being the space of continuous functions made of piecewise polynomials of degree k on each triangle T with Dirichlet boundary conditions on the inflow boundary of Ω. We consider the following bilinear form defined on $V_h^k \times V_h^k$

$$a(v^h, u^h) = \int_\Omega \ell(v^h)\, \vec{\lambda} \cdot \nabla u^h\, dx,$$

where, following the notations of Section 5, we have defined

$$\ell_{T'_T}(v^h) = \sum_{\sigma \in T'_T} \beta_\sigma^{T'_T} v^h(\sigma).$$

We assume that for any degree of freedom, there is at least one sub-triangle T'_T for which $\beta_\sigma^{T'_T} = 0$ except for σ where $\beta_\sigma^{T'_T} = 1$. We note that ℓ maps V_h^k on the set of L^2 functions that are constant on each T'_T.

We know that there exists $c > 0$ such that

$$\int_\Omega \ell(v^h)^2 \, dx \geqslant c^2 \int_\Omega (v^h)^2.$$

In this appendix, we show that

Proposition B.1. With V_h^k defined as above, there exists a constant $C_k > 0$ such that

$$\inf_{u^h \in V_h^k} \sup_{v^h \in V_h^k} \frac{a(v^h, u^h)}{\sqrt{\int_\Omega (v^h)^2 \, dx} \sqrt{\int_\Omega (\vec{\lambda} \cdot \nabla v^h)^2}} \geqslant C_k > 0.$$

Proof. We denote by $\|.\|_{0,\Omega}$ the L^2 norm. First, we have

$$\sup_{v^h \in V_h^k} \frac{a(v^h, u^h)}{\|v^h\|_{0,\Omega}} = \sup_{v^h \in V_h^k} \left(\frac{a(v^h, u^h)}{\|\ell(v^h)\|_{0,\Omega}} \frac{\|\ell(v^h)\|_{0,\Omega}}{\|v^h\|_{0,\Omega}} \right) \geqslant c \sup_{v^h \in V_h^k} \frac{a(v^h, u^h)}{\|\ell(v^h)\|_{0,\Omega}} = \|\Pi_\ell(\vec{\lambda} \cdot \nabla u^h)\|$$

where Π_ℓ denote the L^2 projection on the space $\ell(V_h^k) \subset L^2(\Omega)$.

The problem is to estimate $\|\Pi_\ell(\vec{\lambda} \cdot \nabla u^h)\|$, that is

$$\sup_{v^h \in V_h^k} \frac{\int_\Omega \ell(v^h) \, \vec{\lambda} \cdot \nabla u^h}{\|\ell(v^h)\|_{0,\Omega}}.$$

We first note that for any $v^h \in V_h^k$, we have

$$\int_\Omega \ell(v^h) \, \vec{\lambda} \cdot \nabla u^h \, dx = \int_\Omega \ell(v^h) \, U \, dx$$

where U is the piecewise constant function defined on each T'_T by

$$U_{|T'_T} := \frac{\int_{T'_T} \vec{\lambda} \cdot \nabla u \, dx}{|T'_T|}. \tag{30}$$

The second thing is to note that

$$\int_\Omega \ell(v^h) \, U \, dx = \sum_T \sum_{T'_T \subset T} \int_{T'_T} \ell(v^h) \, U_{T'_T} \, dx = \sum_T \sum_{T'_T \subset T} \int_{T'_T} \ell(v_1^h) \, U_{T'_T} \, dx$$

where v_1^h is the continuous function, linear on each sub-triangle T'_T, that interpolate $\{v_h(\sigma)\}_{\sigma \in T'_T}$. The remark enables to reduce the problem to the case $k = 1$.

From now on, the symbol v_h denotes a continuous piecewise linear function defined on the conformal triangulation of Ω made by all the sub-triangles

$\{\{T'_T\}_{T'_T \subset T}\}_{T \subset \Omega}$. By abuse of notation, we still denote this space of functions as V_h^1. As a consequence, the degrees of freedom that define the functions of V_h^1 are the degrees of freedom σ.

Then, we show the existence of $C > 0$ such that for any v^h piecewise linear and continuous,

$$\sup_{v^h \in V_h^1} \frac{\int_\Omega \ell(v^h)\, U\, dx}{\|v^h\|_{0,\Omega}} \geqslant C \geqslant \alpha \sqrt{\sum_{T'_T} |T'_T|\, U_{T'_T}^2}$$

where U is the piecewise constant function defined by (30).

We have, for any U piecewise constant on each T,

$$\frac{\int_\Omega \ell(v^h)\, U}{\sqrt{\int_\Omega \ell(v^h)^2\, dx}} = \frac{\sum_{T'_T} |T'_T|\, \ell(v^h)_{|T'_T}\, U_{T'_T}}{\sqrt{\sum_{T'_T} |T'_T|\, \ell(v^h)_{|T'_T}^2}}.$$

We have also

$$\ell(v^h)_{|T'_T}^2 \leqslant \left(\sum_{\sigma \in T'_T} (\beta_\sigma^{T'_T})^2 \right) \left(\sum_{\sigma \in T'_T} v^h(\sigma) \right) \leqslant C \sum_{\sigma \in T'_T} v^h(\sigma)^2$$

for some $C > 0$ independant of h because $\beta_\sigma^{T'_T} \geqslant 0$. Thus we have

$$C \sum_{T'_T} |T'_T| \left(\sum_{\sigma \in T'_T} v^h(\sigma)^2 \right) \geqslant \int_\Omega \ell(v^h)^2\, dx,$$

and then

$$\frac{\int_\Omega \ell(v^h)\, U}{\sqrt{\int_\Omega \ell(v^h)^2\, dx}} \geqslant C^{-1} \frac{\sum_{T'_T} |T'_T| \left(\sum_\sigma (U_{T'_T} \beta_\sigma^{T'_T})\, v^h(\sigma) \right)}{\sum_{T'_T} |T'_T| \left(\sum_{\sigma \in T'_T} v^h(\sigma)^2 \right)}.$$

Taking the supremum of these inequalities, we get

$$\sup_{v^h \in V_h^1} \frac{a(v^h, u^h)}{\|v^h\|_{0,\Omega}} \geqslant C^{-1} \sqrt{\sum_{T'_T} |T'_T|\, U_{T'_T}^2 \left(\sum_\sigma (\beta_\sigma^{T'_T})^2 \right)}.$$

Since $\sum_{\sigma \in T'_T} \beta_\sigma^{T'_T} = 1$ and $\beta_\sigma^{T'_T} \geqslant 0$, there exists a constant $\alpha > 0$ such that

$$\sup_{v^h \in V_h^1} \frac{\int_\Omega \ell(v^h)\, U\, dx}{\|v^h\|_{0,\Omega}} \geqslant \alpha \sqrt{\sum_{T'_T} |T'_T|\, U_{T'_T}^2}.$$

Coming back to the original problem, we have shown the existence of $C > 0$ such that

$$C \sqrt{\sum_T \sum_{T'_T \subset T} \frac{\left(\int_{T'_T} \vec{\lambda} \cdot \nabla u^h \right)^2}{T'_T}} \leqslant \|\Pi_\ell(\vec{\lambda} \cdot \nabla u^h)\|_{0,\Omega}. \tag{31}$$

If $k = 1$, we have $U_T = \vec{\lambda} \cdot \nabla u^h$, so (31) is nothing more than

$$C\, \|\vec{\lambda} \cdot \nabla u^h\|_{0,\Omega} \leqslant \|\Pi_\ell(\vec{\lambda} \cdot \nabla u^h)\|_{0,\Omega}.$$

In all cases $k > 1$, the left hand side of (31) is quadratic in u^h, and it is easy to prove that if

$$\sum_{T'_T \subset T} \frac{(\int_{T'_T} \vec{\lambda} \cdot \nabla u^h)^2}{T'_T} = 0,$$

the $k^2 > (k+1)(k+2)/2 - 2$ terms[4] of the sum are all vanishing so that $\vec{\lambda} \cdot \nabla u^h = 0$ in the vector space of polynomials of degree k on T, denoted $P_k(T)$.

This shows that on the finite dimensional space $P_k(T)$, the two quadratic forms

$$\int_T (\vec{\lambda} \cdot \nabla u^h)^2$$

and

$$\sum_{T'_T \subset T} \frac{(\int_{T'_T} \vec{\lambda} \cdot \nabla u^h)^2}{T'_T}$$

are equivalent. Hence, there exists $c_k > 0$ such that

$$\sum_{T'_T \subset T} \frac{(\int_{T'_T} \vec{\lambda} \cdot \nabla u^h)^2}{T'_T} \geqslant c_k \int_T (\vec{\lambda} \cdot \nabla u^h)^2.$$

This ends the proof: there exists $C_k > 0$ such that

$$\inf_{u^h \in V_h^k} \sup_{v^h \in V_h^k} \frac{a(v^h, u^h)}{\sqrt{\int_\Omega (v^h)^2 \, dx} \sqrt{\int_\Omega (\vec{\lambda} \cdot \nabla v^h)^2}} \geqslant C_k > 0. \qquad \square$$

ACKNOWLEDGMENTS

This work was started during a visit to Bordeaux University, in June 2001, by the second author, who gratefully appreciated the warm hospitality shown on that occasion by members of the Département de Mathématiques Appliquées. Katherine Mer–Nkonga (CEA CESTA, France) is also acknowledged for her help in a very early draft of the paper [14]. The constructive criticisms of the referees are acknowledged and warmly thanked.

REFERENCES

1. Struijs, R., Deconinck, H., and Roe, P. L. Fluctuation splitting schemes for the 2d Euler equations, Technical report, VKI, 1991, VKI LS 1991-01, Computational Fluid Dynamics.
2. Deconinck, H., Struijs, R., and Roe, P. L. Compact advection schemes on unstructured grids, Technical report, VKI, 1993, VKI LS 1993-04, Computational Fluid Dynamics.
3. Hughes, T. J. R., and Mallet, M. (1986). A new finite element formulation for CFD: III. The generalized streamline operator for multidimensional advective-diffusive systems. *Comp. Meth. Appl. Mech. Engrg.* **58**, 305–328.

[4] The vector space of functions spanned by $\vec{\lambda} \cdot \nabla u^h$ with $u^h \in V_h^k$ is of dimension $\frac{(k+1)(k+2)}{2} - 2$.

4. Roe, P.L. (1982). Fluctuations and signals—a framework for numerical evolution problems. In *Numerical Methods for Fluid Dynamics*, Proc. Conf., Reading/U.K., pp. 219–257.
5. Abgrall, R. (2001). Toward the ultimate conservative scheme: following the quest. *J. Comput. Phys.* **167**(2), 277–315.
6. Abgrall, R., and Mezine, M. (2003). Construction of second order accurate monotone and stable residu al distributive schemes for unsteady flow problems. *J. Comput. Phys.*, in press.
7. Caraeni, D., Caraeni, M., and Fuchs, L. A parallel multidimensional upwind algorithm, AIAA Paper 201-2547, 2001, 15th AIAA CFD Meeting, Anaheim.
8. Leveque, R. (1992). *Numerical Methods for Conservation Laws*, Birkhäuser, Zürich.
9. Godlewski, E., and Raviart, P. A. (1995). *Hyperbolic Systems of Conservations Laws*, Applied Mathematical Sciences, Springer-Verlag.
10. Ern, A., and Guermond, J. L. (2002). *Eléments finis: théorie, applications, mise en oeuvre*, Vol. 36 of *Mathématiques et Applications*, Springer.
11. Nishikawa, H., Rad, M., and Roe, P. L. A third-order fluctuation splitting scheme that preserves potential flow, AIAA Paper 2001-2595, 2001, 15th AIAA CFD Meeting, Anaheim.
12. Wang, Z. J. Spectral (finite) volume method for conservation laws on unstructured grids: Basic formulation, *J. Comput. Phys.*, in press.
13. van der Weide, E., and Deconinck, H. (1996). Positive matrix distribution schemes for hyperbolic systems. In *Computational Fluid Dynamics '96*, Wiley, pp. 747–753.
14. Abgrall, R., and Mer, K. (March 1998). Un théorème de type Lax–Wendroff pour les schémas distributifs, Technical Report 98010, Mathématiques Appliquées de Bordeaux.
15. Kröner, D., Rokyta, M., and Wierse, M. (1996). A Lax–Wendroff type theorem for upwind finite volume schemes in 2-d. *East-West J. Numer. Math.* **4**(4), 279–292.