M2AN - Modélisation mathématique et analyse numérique

Jacques Baranger
Jean-François Maitre
Fabienne Oudin

Connection between finite volume and mixed finite element methods

M2AN - Modélisation mathématique et analyse numérique, tome 30, n° 4 (1996), p. 445-465
<http://www.numdam.org/item?id=M2AN_1996__30_4_445_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE

(Vol 30, n° 4, 1996, p 445 à 465)

CONNECTION BETWEEN FINITE VOLUME AND MIXED FINITE ELEMENT METHODS (*)

by Jacques BARANGER (1), Jean-François MAITRE (2) et Fabienne OUDIN (1)

Abstract — *For the model problem with Laplacian operator, we show how to produce cell-centered finite volume schemes, starting from the mixed dual formulation discretized with the Raviart-Thomas element of lowest order*

The method is based on the use of an appropriate integration formula (mass lumping) allowing an explicit elimination of the vector variables The analysis of the finite volume scheme (wellposed-ness and error bounds) is directly deduced from classical results of mixed finite element theory, which is the main interest of the method

We emphasize existence and properties of the diagonalizing integration formulas, specially in the case of N-dimensional simplicial elements

Résumé — *Pour le problème modèle avec l'opérateur laplacien, nous montrons comment construire un schéma volumes finis, en partant de la formulation mixte duale discrétisée avec l'élément de Raviart-Thomas de plus bas degré*

La méthode repose sur l'utilisation d'une formule d'intégration numérique appropriée (condensation de masse) permettant l'élimination explicite du champ vectoriel L'analyse du schéma volumes finis (existence-unicité, borne d'erreur) se déduit directement des résultats classiques de la théorie des éléments finis mixtes, ce qui constitue l'intérêt principal de la démarche

Un intérêt particulier est porté à l'existence et aux propriétés de formules d'intégration diagonalisantes, cela particulièrement pour le cas des éléments simpliciaux en dimension N quelconque

1. INTRODUCTION

Let Ω be an open bounded polygonal set of $\mathbb{R}^N$ (where $N = 1, 2$ or 3), and consider the Dirichlet model problem

$$- \operatorname{div} (\underline{\operatorname{grad}}\, u) = f \quad \text{in } \Omega, \quad u = 0 \text{ on } \Gamma = \delta\Omega \tag{1}$$

(*) Manuscript received February 9, 1995 , revised July 6, 1995

(1) Université de Lyon 1, Laboratoire d'Analyse Numérique, Bât 101, 69622 Villeurbanne Cedex

(2) Département M I S , École Centrale de Lyon, BP 163, 69131 Ecully Cedex

and its mixed dual formulation

$$
(P)\begin{cases}
\text{find } (\underline{p}, u) \in H(\operatorname{div}; \Omega) \times L^2(\Omega) \text{ such that} \\
\displaystyle\int_\Omega \underline{p} \cdot \underline{q}\, dx + \int_\Omega \operatorname{div} \underline{q} \cdot u\, dx = 0 \quad \forall \underline{q} \in H(\operatorname{div}; \Omega) = Q \\
\displaystyle\int_\Omega \operatorname{div} \underline{p} \cdot v\, dx = -\int_\Omega f \cdot v\, dx \quad \forall v \in L^2(\Omega) = V .
\end{cases}
$$

Using a mesh Θ_h of Ω $\left(\overline{\Omega} = \bigcup_{K \in \Theta_h} K\right)$, we consider the approximation of (P)

$$
(P_h)\begin{cases}
\text{find } (\underline{p}_h, u_h) \in Q_h \times V_h \text{ such that} \\
\displaystyle\int_\Omega \underline{p}_h \cdot \underline{q}_h\, dx + \int_\Omega \operatorname{div} \underline{q}_h \cdot u_h\, dx = 0 \quad \forall \underline{q}_h \in Q_h \\
\displaystyle\int_\Omega \operatorname{div} \underline{p}_h \cdot v_h\, dx = -\int_\Omega f \cdot v_h\, dx \quad \forall v_h \in V_h
\end{cases}
$$

with :

$$V_h = \{ v_h \in L^2(\Omega) ; v_h|_K \in P_0(K), \forall K \in \Theta_h \} ,$$

$$Q_h = \{ \underline{q}_h \in H(\operatorname{div}; \Omega) ; \underline{q}_h|_K \in RT_1(K), \forall K \in \Theta_h \} ,$$

where $RT_1(K)$ is the Raviart-Thomas space of smallest order.

It is well known that (P) (resp. (P_h)) has a unique solution $(\underline{p}, u)$ (resp. $(\underline{p}_h, u_h)$) satisfying (under regularity assumptions) :

$$\| u - u_h \|_{L^2(\Omega)} + \| \underline{p} - \underline{p}_h \|_{H(\operatorname{div}; \Omega)} = O(h) \tag{2}$$

(see Brezzi-Fortin [3], Raviart-Thomas [14]).

Recent works on finite volume methods applied to the same problem are : W. Hackbusch [9] for the box methods and Morton-Suli [12] for the cell vertex methods. P. Emonot [1] has studied finite volumes using polynomials with degree greater than 1.

(P_h) presents implicitly a finite volume aspect : in equation (1), we can consider as test function the characteristic function of element K, which gives the integration on K of the equation of conservation $\operatorname{div} p + f = 0$. But finite volume schemes for (1) are usually expressed with the only unknowns u_i approximating u.

The aim here is to show that, from the mixed finite element scheme of (1), we can obtain a finite volume scheme with the unknows $(u_K)_K$. This needs the elimination of the unknows $(p_i)_i$. Using a technique similar to that of Haugazeau-Lacoste [10], we show that there exists a numerical integration formula on K of the form $\sum_f \alpha_f \varphi_f(\underline{p}_h)\, \varphi_f(\underline{q}_h)$ (where $\varphi_f(\underline{p}_h)$ is the flux of $\underline{p}_h$ through face f), exact for $\underline{p}_h$ and $\underline{q}_h$ piecewise constant. We consider the approximate discretisation corresponding to this integration formula, that is :

$$(\tilde{P}_h)\begin{cases}\text{find } (\underline{\tilde{p}}_h, \tilde{u}_h) \in Q_h \times V_h \text{ such that} \\ \sum_{f \in F_h} \alpha_f \varphi_f(\underline{\tilde{p}}_h) \cdot \varphi_f(\underline{q}_h) + \int_\Omega \operatorname{div} \underline{q}_h \cdot \tilde{u}_h \, dx = 0 \quad \forall \underline{q}_h \in Q_h \\ \qquad \int_\Omega \operatorname{div} \underline{\tilde{p}}_h \cdot v_h \, dx = -\int_\Omega g \cdot v_h \, dx \quad \forall v_h \in V_h \end{cases}$$

where F_h denotes the set of all faces of the mesh Θ_h.

The approximate mass matrix being diagonal, we can eliminate the gradient unknowns to obtain a finite volume scheme with the approximate values of u as the only unknowns.

The interest of this method is that, thanks to general results of mixed finite element methods, we can obtain results of existence and unicity, and also error estimations from that of type (2) (see [3] or [15]). Concerning this method, we have been aware recently, after having achieved this present study, that it had been already proposed by A. Weiser and M. F. Wheeler [16], but for the only case of the rectangular mesh, and apparently without detailed proof.

For our proof, we shall use a Theorem of Roberts-Thomas ([15]), which compares the solution $(\varphi, \lambda) \in W \times M$ of the problem (P_1) :

$$(P_1) \quad \begin{cases} a(\varphi, \psi) + b(\psi, \lambda) = f(\psi), & \forall \psi \in W \\ b(\varphi, \mu) = g(\mu), & \forall \mu \in M, \end{cases}$$

and the solution $(\varphi_h^*, \lambda_h^*) \in W_h \times M_h$ of the discrete problem $(P_{1,h}^*)$ using approximated linear and bilinear forms, with $W_h \subset W$ and $M_h \subset M$:

$$(P_{1,h}^*) \quad \begin{cases} a_h(\varphi_h^*, \psi_h) + b_h(\psi_h, \lambda_h^*) = f_h(\psi_h), & \forall \psi_h \in W_h \\ b_h(\varphi_h^*, \mu_h) = g_h(\mu_h), & \forall \mu_h \in M_h, \end{cases}$$

This Theorem is the following :

THEOREM 1 : *Suppose that $a(\,.\,,\,.\,)$ (resp. $b(\,.\,,\,.\,)$) is a continuous bilinear form on $W \times W$ (resp. $W \times M$) with A (resp. B) as constant of continuity. Let A_h be a number such that :*

$$a_h(\varphi_h, \psi_h) \leq A_h \|\varphi_h\|_W \|\psi_h\|_W, \quad \forall \varphi_h \in W_h, \forall \psi_h \in W_h .$$

Suppose further that there exists numbers $\alpha_h > 0$ and $\beta_h > 0$ such that :

$$a_h(v_h, v_h) \geq \alpha_h \|v_h\|_W^2, \quad \forall v_h \in \{w_h \in W_h, b_h(w_h, \mu_h) = 0, \forall \mu_h \in M_h\}$$

and

$$\inf_{\{\mu_h \in M_h,\ \|\mu_h\|_M = 1\}} \ \sup_{\{\psi_h \in W_h,\ \|\psi_h\|_M = 1\}} b_h(\psi_h, \mu_h) \geq \beta_h .$$

Then there exists a constant C dependent only on A, B, A_h, α_h and β_h such that (φ, λ) and $(\varphi_h^*, \lambda_h^*)$ the solutions of (P_1) and $(P_{1,h}^*)$ satisfy :

$$\|\varphi - \varphi_h^*\|_W + \|\lambda - \lambda_h^*\|_M$$

$$\leq C \left\{ \inf_{\psi_h \in W_h} \left(\|\varphi - \psi_h\|_W + \sup_{\eta_h \in W_h} \frac{a(\psi_h, \eta_h) - a_h(\psi_h, \eta_h)}{\|\eta_h\|_W} \right) \right.$$

$$+ \inf_{\mu_h \in M_h} \left(\|\lambda - \mu_h\|_M + \sup_{\eta_h \in W_h} \frac{b(\eta_h, \mu_h) - b_h(\eta_h, \mu_h)}{\|\eta_h\|_W} \right)$$

$$\left. + \sup_{\eta_h \in W_h} \frac{f(\eta_h) - f_h(\eta_h)}{\|\eta_h\|_W} + \sup_{\nu_h \in M_h} \frac{g(\nu_h) - g_h(\nu_h)}{\|\nu_h\|_M} \right\} .$$

Remark : in this study, being interested in the numerical integration for a, that is the error $a(\,.\,,\,.\,) - a_h(\,.\,,\,.\,)$, we shall assume $b = b_h$, $f = f_h$, $g = g_h$.

In Section 2, we study briefly the problem in the one-dimensional case ; we find, using the trapezoidal rule, a scheme obtained with a finite difference method (see [6]) with an error estimation of type (2). We can remark that, for this simple 1.D case, it is possible to eliminate directly the gradient unknowns, without using numerical integration, obtaining for the Laplacian an other scheme of finite difference type, but on a staggered grid.

In Section 3, we study the problem in two-dimension with a triangular mesh. We give a method of construction of the numerical integration, exact for constant fields, to obtain an error estimation of type (2). The resulting scheme is exactly the finite volume scheme given by R. Herbin [11]. As a remark, we give the result for a rectangular mesh.

In Section 4, we consider the case of higher dimensions ($N \geq 3$). We show that the process used for the two-dimensional case does not work in general. In the example of a tetrahedral 3-D mesh, the construction of an integration formula exact for constant fields and diagonalizing the mass matrix is only possible in the case of special tetrahedra, among which regular ones.

2. ONE-DIMENSIONAL CASE

In this case, the Raviart-Thomas element of smallest order used the gradient space $RT_1(I) = P_1(I)$ (where I is the interval of length h_I), and its degrees of freedom are the values at vertices of the mesh I. We note by $a^I(\,.\,,\,.\,)$ the bilinear form associated to the elementary mass matrix, that is to say: $a^I(\underline{p}, \underline{q}) = \int_I \underline{p} \cdot \underline{q}\, dx$.

PROPOSITION 1 : *There exists a diagonal matrix A_h^I with associated bilinear form $a_h^I(\,.\,,\,.\,)$ satisfying for all $\underline{p}_0 \in P_0(I)$: $a^I(\underline{p}_0, \underline{p}_0) = a_h^I(\underline{p}_0, \underline{p}_0)$. Moreover, for all $\underline{p}_h$ and all $\underline{q}_h$ in $RT_1(\bar{I})$:*

$$|a^I(\underline{p}_h, \underline{q}_h) - a_h^I(\underline{p}_h, \underline{q}_h)| = \frac{1}{6} h_I^2 |\underline{p}_h|_{H^1(I)} |\underline{q}_h|_{H^1(I)} ,$$

where $|\,.\,|_{H^1(I)}$ denotes the semi norm of $H^1(I)$.

Proof: We can choose the numerical integration corresponding to the trapezoidal rule, which is exact for the constants and uses values at the vertices of I, that is to say the degrees of freedom of $RT_1(I)$; the elementary matrix is then given by : $A_h^I = \text{diag}\left(\frac{1}{2} h_I, \frac{1}{2} h_I\right)$ and the result is easy to obtain since $\underline{p}_h$ and $\underline{q}_h$ are in $P_1(I)$.

Denoting by $(u, \underline{p})$ the solution of problem (P) and $(\tilde{u}_h, \tilde{\underline{p}}_h)$ the solution of problem $(\tilde{P}_h)$, where the numerical integration is that corresponding to the matrix A_h^I, we have the following result :

PROPOSITION 2 : *Problem $(\tilde{P}_h)$ has a unique solution, and there exists a constant C such that, if $(u, \underline{p}) \in H^1(\Omega) \times (H^2(\Omega))^2$, then :*

$$\|\underline{p} - \tilde{\underline{p}}_h\|_{H^1(\Omega)} + \|u - \tilde{u}_h\|_{L^2(\Omega)} \leq Ch(\|\underline{p}\|_{H^2(\Omega)} + \|u\|_{H^1(\Omega)}) .$$

Proof: It is a direct consequence of Theorem 1 of Roberts-Thomas [15], the hypotheses of which are easily verified, and of the former proposition.

Remark : defining $I_k = [x_k, x_{k+1}]$ (where $[0, 1] = \bigcup_{k=1}^{n} I_k$), $h_k = x_{k+1} - x_k$, and denoting $u_{k+1/2}$, the unknown in I_k, the scheme obtained after the elimination of gradients $\underline{p}_h$ is :

$$\begin{cases} \dfrac{u_{k+3/2} - u_{k+1/2}}{h_{k+1/2}} - \dfrac{u_{k+1/2} - u_{k-1/2}}{h_{k-1/2}} = - h_k \bar{f}_k, \quad \forall k \in \{1, \dots, n\} \\ u_{1/2} = u_{n+3/2} = 0 \end{cases} \tag{3}$$

where

$$h_{k+1/2} = 1/2(h_k + h_{k+1}) \quad \text{and} \quad h_k \bar{f}_k = \int_{x_k}^{x_{k+1}} f(x)\, dx .$$

This scheme is in fact the classical three points scheme obtained by cell centered finite differences method (see for example [6])

3. TWO-DIMENSIONAL CASE

We have studied both cases of rectangular and triangular meshes, but we will detail only the study of the last one.

For rectangular meshes, different diagonalising integration formulas are exact for constant elements of RT_1, among which the trapezoidal rule. The corresponding finite difference scheme on each rectangle is a 5 points one, the approximation of the fluxes being the natural ones. The error bound $O(h)$ is still valid, but cannot be obtained by Theorem 1.

Let us study the case of a triangular mesh of Ω We have : $\bar{\Omega} = \bigcup_{T \in \Theta_h} T$, where the T's are triangles. We suppose that Θ_h is in a regular family of triangulations of $\bar{\Omega}$, in the sense that there exists a constant $\sigma > 0$ independent of h such that :

$$\max_{T \in \Theta_h} \frac{h_T}{\delta_T} \leq \sigma ,$$

where h_T is the diameter of T, and δ_T is the diameter of the inscribed circle of T.

For each T, we shall use the following notations :

- $|T|$: area of T ;
- a_i : i^{th} vertex of T, with coordinates (x_i, y_i) ($i \in \{1, 2, 3\}$) ,
- f_i : face opposite to vertex a_i, with length $|f_i|$,

- $\underline{n}_i$: unit exterior normal to face f_i ;
- $\overline{\theta}_i$: angle at vertex a_i ;

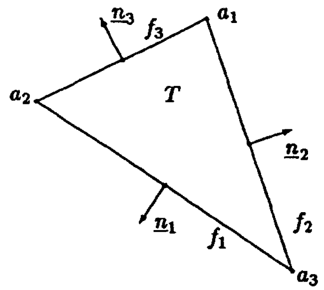

Figure 1.

For any triangle T, the Raviart-Thomas space of smallest order is defined by :

$$RT_1(T) = (P_0(T))^2 \oplus P_0(T) \begin{pmatrix} x \\ y \end{pmatrix}.$$

The local shape functions $\{\underline{p}_i\}_{i \in \{1,2,3\}}$ of any triangle T, associated to the fluxes $\{\varphi_{f_i}\}_{i \in \{1,2,3\}}$ through faces $\{f_i\}_{i \in \{1,2,3\}}$, are defined by :

$$\underline{p}_i(x, y) = \frac{1}{2|T|}(x - x_i, y - y_i), \quad \forall (x, y) \in T. \tag{4}$$

We denote $a^T(\cdot , \cdot)$, the bilinear form on $Q \times Q$ defined by :

$$a^T(\underline{p}, \underline{q}) = \int_T \underline{p}(x, y) \cdot \underline{q}(x \cdot y)\, dx\, dy.$$

We have the following results :

PROPOSITION 3 : *There exists a unique diagonal matrix, denoted A_h^T, such that $a_h^T(\cdot , \cdot)$, its associated bilinear form, satisfies :*

$$\textit{for all } \underline{p}_0 \in (P_0(T))^2, \quad a^T(\underline{p}_0, \underline{p}_0) = a_h^T(\underline{p}_0, \underline{p}_0). \tag{5}$$

This matrix is given by $(A_h^T)_{ii} = \frac{1}{2}\cot(\theta_i)$, *where* θ_i *is the angle at vertex* a_i, *and* $a_h^T(\ .\ ,\ .\ )$ *is given by :*

$$a_h^T(\underline{p}_h, \underline{q}_h) = \frac{1}{2}\sum_{i=1}^{3} c_i\, \varphi_{f_i}(\underline{p}_h)\, \varphi_{f_i}(\underline{q}_h)\,, \tag{6}$$

with $c_i = \cot(\theta_i)$.

Proof : see Section 4.

Remark : relation (5) implies the more general one :

$$\forall(\underline{p}_0, \underline{q}_0) \in (P_0(T))^2 \times (P_0(T))^2, \quad a^T(\underline{p}_0, \underline{q}_0) = a_h^T(\underline{p}_0, \underline{q}_0)\,.$$

PROPOSITION 4 : *The bilinear form* a_h^T *of Proposition 3 satisfies :*

$$|a^T(\underline{p}_h, \underline{q}_h) - a_h^T(\underline{p}_h, \underline{q}_h)| \leqslant \varepsilon(h_T)\ \|\underline{p}_h\|_{H(\mathrm{div}\,,T)} \|\underline{q}_h\|_{H(\mathrm{div}\,,T)}\,, \tag{7}$$

for all $\underline{p}_h$ *and* $\underline{q}_h$ *in* $RT_1(T)$, *with* $\varepsilon(h_T) = h_T^2/48$ *if* T *is equilateral* $\left(h_T^2 / \delta_T^2 = 3\right)$ *and* $\varepsilon(h_T) = \frac{h_T}{3}\sqrt{\sigma^2 - 3} + O(h_T^2)$ *else, where* σ *is the constant of the regular family of meshes* $(h_T/\delta_T \leqslant \sigma)$.

Proof : For $\underline{p}_h$ in $RT_1(T)$, we can write :

$$\underline{p}_h = \underline{p}_0 + \beta(\underline{p}_1 + \underline{p}_2 + \underline{p}_3), \quad \text{with} \quad \beta = \frac{|T|}{3}\,\mathrm{div}\,(\underline{p}_h) \tag{8}$$

and $\quad \underline{p}_0 = \alpha_1\, \underline{p}_1 + \alpha_2\, \underline{p}_2 + \alpha_3\, \underline{p}_3$ with $\sum_{i=1}^{3} \alpha_i = 0$.

For $\underline{q}_h$ in $RT_1(T)$, we can write the same relations, replacing β by γ and (α_i) by (δ_i).

It is easy to verify the following equalities :

$$a^T(\underline{p}_h, \underline{q}_h) = a^T(\underline{p}_0, \underline{q}_0) + \gamma\beta|\underline{p}_1 + \underline{p}_2 + \underline{p}_3|_{0,T}^2\,,$$

$$a_h^T(\underline{p}_h, \underline{q}_h) = a_h^T(\underline{p}_0, \underline{q}_0) + \gamma e^T A_h^T \alpha + \beta e^T A_h^T \delta + \gamma\beta e^T A_h^T e\,,$$

where $e^T = (1, 1, 1)$, and, introducing the barycenter g_T and the gyration radius ρ_T of T, to obtain :

$$|\underline{p}_1 + \underline{p}_2 + \underline{p}_3|_{0,T}^2 = \frac{9}{4|T|^2}\int_T (x - g_T)^2\, dx = \frac{9}{4|T|}\,\rho_T^2\,.$$

It can be shown that $\sum_{i=1}^{3} c_i = \frac{9}{|T|} \rho_T^2$, which implies :

$$e^T A_h^T e = \frac{1}{2} \sum_{i=1}^{3} c_i = \frac{9}{2|T|} \rho_T^2 .$$

Finally, we have :

$$a_h^T(\underline{p}_h, \underline{q}_h) - a^T(\underline{p}_h, \underline{q}_h) = \gamma\beta\left(\frac{9}{4|T|} \rho_T^2\right) + \gamma e^T A_h^T \alpha + \beta e^T A_h^T \delta \tag{9}$$

since $a_h^T(\underline{p}_0, \underline{q}_0) = a^T(\underline{p}_0, \underline{q}_0)$.

For the $H(\operatorname{div}; T)$-norms, we have :

$$\|\underline{p}_h\|^2_{H(\operatorname{div}; T)} = a^T(\underline{p}_0, \underline{p}_0) + \beta^2 |\underline{p}_1 + \underline{p}_2 + \underline{p}_3|^2_{0,T} + |T| (\operatorname{div} \underline{p}_h)^2$$

that is :

$$\|\underline{p}_h\|^2_{H(\operatorname{div}; T)} = \alpha^T A_h^T \alpha + \frac{9}{|T|} \left(1 + \frac{1}{4} \rho_T^2\right) \beta^2 \tag{10}$$

and, similarly :

$$\|\underline{q}_h\|^2_{H(\operatorname{div}; T)} = \delta^T A_h^T \delta + \frac{9}{|T|} \left(1 + \frac{1}{4} \rho_T^2\right) \gamma^2 . \tag{11}$$

Thanks to (9), (10), (11), and to prove (7), we have to solve or bound the following supremum :

$$\sup_{\{\beta, \alpha\}, \{\gamma, \delta\}} \frac{\left|\beta\gamma\left(\frac{9}{4|T|} \rho_T^2\right) + \beta e^T A_h^T \delta + \gamma e^T A_h^T \alpha\right|}{\left(\alpha^T A_h^T \alpha + \frac{9}{|T|}\beta^2\right)^{1/2} \left(\delta^T A_h^T \delta + \frac{9}{|T|} \gamma^2\right)^{1/2}} . \tag{12}$$

To eliminate the constraints $\sum_{i=1}^{3} \alpha_i = 0$, $\sum_{i=1}^{3} \delta_i = 0$, we write

$$\alpha = (\tilde{\alpha}_1, \tilde{\alpha}_2, -\tilde{\alpha}_1 - \tilde{\alpha}_2), \quad \delta = (\tilde{\delta}_1, \tilde{\delta}_2, -\tilde{\delta}_1 - \tilde{\delta}_2),$$

with $\tilde{\alpha}$, $\tilde{\delta}$ frees in $\mathbb{R}^2$, and obtain :

$$e^T A_h^T \alpha = \frac{1}{2}(c_1 - c_3, c_2 - c_3)\,\tilde{\alpha}\,, \quad e^T A_h^T \delta = \frac{1}{2}(c_1 - c_3, c_2 - c_3)\,\tilde{\delta}\,,$$

$$\alpha^T A_h^T \alpha = \frac{1}{2}\tilde{\alpha}^T C\tilde{\alpha}\,, \quad \delta^T A_h^T \delta = \frac{1}{2}\tilde{\delta}^T C\tilde{\delta}\,,$$

with

$$C = \begin{pmatrix} c_1 + c_3 & c_3 \\ c_3 & c_1 + c_3 \end{pmatrix} \quad \text{and} \quad C^{-1} = \begin{pmatrix} c_2 + c_3 & -c_3 \\ -c_3 & c_1 + c_3 \end{pmatrix},$$

since $\quad c_1 c_2 + c_2 c_3 + c_3 c_1 = 1$.

Then the solution of (12) is given by the spectral radius of the following 3×3 matrix :

$$\begin{bmatrix} \frac{|T|}{9} & 0 & 0 \\ 0 & & \\ & 2 & C^{-1} \\ 0 & & \end{bmatrix} \begin{bmatrix} \frac{9}{4|T|}\rho_T^2 & \frac{1}{2}(c_1 - c_3) & \frac{1}{2}(c_2 - c_3) \\ \frac{1}{2}(c_1 - c_3) & & \\ & & 0 \\ \frac{1}{2}(c_2 - c_3) & & \end{bmatrix} \tag{13}$$

which eigenvalues are 0 and the roots of

$$\lambda^2 - \frac{\rho_T^2}{4}\lambda - \frac{\rho_T^2}{2}\left(1 - 9\frac{c_1 c_2 c_3}{c}\right) = 0\,.$$

The spectral radius of (13) is bounded by :

$$\rho_T \frac{\sqrt{2}}{2}\sqrt{\frac{8}{3}\left(\frac{h_T^2}{\delta_T^2} - 3\right) + \frac{\rho_T^2}{32}} + \frac{\rho_T^2}{8}$$

since $1 - 9\frac{c_1 c_2 c_3}{c} \leqslant \frac{8}{3}\left(\frac{h_T^2}{\delta_T^2} - 3\right)$, that is by :

$$\frac{h_T}{\sqrt{24}}\sqrt{\frac{8}{3}\left(\frac{h_T^2}{\delta_T^2} - 3\right) + \frac{h_T^2}{12.32}} + \frac{h_T^2}{96}$$

since $\rho_T^2 \leqslant h_T^2/12$.

This bound implies (7), proving Proposition 4.

We consider now the approximate problem $(\tilde{P}_h)$, with :

$$a_h(\underline{p}_h, \underline{q}_h) = \sum_{T \in \Theta_h} a_h^T(\underline{p}_h, \underline{q}_h)$$

and $a_h^T(. , .)$ constructed in Proposition 3.

PROPOSITION 5 : *Problem $(\tilde{P}_h)$ has a unique solution $(\underline{\tilde{p}}_h, \tilde{u}_h)$; moreover, for a regular family of meshes, there exists a constant C such that :*

$$\|\underline{p} - \underline{\tilde{p}}_h\|_{H(\mathrm{div}\,;\,\Omega)} + \|u - \tilde{u}_h\|_{L^2(\Omega)} \leq Ch(|u|_{1,\Omega} + \|\underline{p}\|_{1,\Omega} + \|\mathrm{div}\,(\underline{p})\|_{1,\Omega})$$

where $(\underline{p}, u)$, the solution to (P) is assumed to verify :

$$(\underline{p}, u) \in (H^1(\Omega))^2 \times H^1(\Omega) \textit{ and } \mathrm{div}\,(\underline{p}) \in H^1(\Omega)\,.$$

Proof : We use again Theorem 1 of Roberts-Thomas [15] and begin with verifying its hypotheses, that is the existence of constants, independent of h, for continuity and ellipticity of a_h.

For the continuity of $a_h(. , .)$, we deduce from Proposition 4 :

$$\begin{aligned} |a_h^T(\underline{p}_h, \underline{q}_h)| &\leq |a^T(\underline{p}_h, \underline{q}_h)| + C_1\, h_T \|\underline{p}_h\|_{H(\mathrm{div},T)} \|\underline{q}_h\|_{H(\mathrm{div},T)} \\ &\leq (1 + C_1\, h_T) \|\underline{p}_h\|_{H(\mathrm{div},T)} \|\underline{q}_h\|_{H(\mathrm{div},T)}\,, \end{aligned}$$

and by summation on all triangles T, that the constant of continuity is bounded uniformly on h.

For the ellipticity of $a_h(. , .)$ on

$$Q_h^0 = \{\underline{q}_h \in Q_h\,;\, b(\underline{q}_h, v_h) = 0, \forall v_h \in V_h\},$$

we have for $\underline{q}_{0h}$ element of Q_h^0 :

$$b(\underline{q}_{0h}, v_h) = \int_\Omega \mathrm{div}\,(\underline{q}_{0h})\,.\,v_h\,dx\,dy = 0, \quad \forall v_h \in V_h$$

or equivalently :

$$\int_T \mathrm{div}\,(\underline{q}_{0h})\,dx\,dy = 0, \quad \forall T \in \Theta_h\,,$$

$$\mathrm{div}\,(\underline{q}_{0h})|_T = 0, \quad \forall T \in \Theta_h\,,$$
$$\underline{q}_{0h}|_T \in (P_0(T))^2\,, \ \forall T \in \Theta_h\,,$$

That gives, by Proposition 3 :

$$a_h^T(\underline{q}_{0h}, \underline{q}_{0h}) = a^T(\underline{q}_{0h}, \underline{q}_{0h}) = |\underline{q}_{0h}|_{0,T}^2 = \|\underline{q}_{0h}\|_{H(\mathrm{div}\,;T)}^2$$

which gives 1 for the constant of ellipticity of $a_h(\,.\,,.\,)$.

Finally, we have to check the inf-sup condition on $b(\,.\,,.\,)$, which is easy to do according to Theorem 4 of Raviart-Thomas [14].

Then, Theorem 1 gives a result of existence and unicity of the solution $(\tilde{\underline{p}}_h, \tilde{u}_h)$ of problem $(\tilde{P}_h)$, with an error estimation of type :

$$\|\underline{p} - \tilde{\underline{p}}_h\|_Q + \|u - \tilde{u}_h\|_V \leq C\Big[\inf_{v_h \in V_h} \|u - v_h\|_V + $$

$$+ \inf_{\underline{q}_h \in Q_h} \left\{ \|\underline{p} - \underline{q}_h\|_Q + \sup_{\tilde{\underline{q}}_h \in Q_h} \frac{|a(\underline{q}_h, \tilde{\underline{q}}_h) - a_h(\underline{q}_h, \tilde{\underline{q}}_h)|}{\|\tilde{\underline{q}}_h\|_Q} \right\}\Big].$$

By Proposition 4, after summation on all triangles T, we have :

$$\sup_{\tilde{\underline{q}}_h \in Q_h} \frac{|a(\underline{q}_h, \tilde{\underline{q}}_h) - a_h(\underline{q}_h, \tilde{\underline{q}}_h)|}{\|\tilde{\underline{q}}_h\|_Q} \leq C_2\, h \|\underline{q}_h\|_Q, \quad \forall \underline{q}_h \in Q_h$$

where $\quad h = \sup_T h_T$.

According to a result of Raviart-Thomas [14] for elements of $RT_1(T)$, there exists a constant $C_3 > 0$, independent of h such that, if $\underline{p} \in (H^1(\Omega))^2$ with $\mathrm{div}\,\underline{p} \in H^1(\Omega)$, then :

$$\inf_{\underline{q}_h \in Q_h} \|\underline{p} - \underline{q}_h\|_{H(\mathrm{div}\,;\Omega)} \leq C_3\, h(|\underline{p}|_{1,\Omega} + |\mathrm{div}\,\underline{p}|_{1,\Omega}).$$

On the other hand, an application of a result of Ciarlet-Raviart [4] (Theorem 5) gives for some constant C_4 independent of h :

$$\inf_{v_h \in V_h} \|u - v_h\|_{0,\Omega} \leq C_4\, h \|u\|_{1,\Omega}.$$

These three last results imply the error estimation given in Proposition 5.

Finally, we will emphasize the explicit scheme corresponding to the unique « mass lumping » of Proposition 3. To describe the finite difference equation in u associated with triangle T, we introduce the following notations :

- T_i is the triangle sharing face f_i with T ;
- C (resp. C_i) is the center of the circumscribed circle to T (resp. T_i) ;
- d_i (resp. d_i^i) is the distance between center C (resp. C_i) and middle m_i of face f_i ;

• c_i (resp. c_i^i) is the cotangent of the angle of T (resp. T_i) opposite of f_i ;

• D_i is the distance between C and C_i.

By simple geometrical properties, we have the following relation :

$$\overrightarrow{CC_i} = \frac{c_i + c_i^i}{2} |f_i| \overrightarrow{n}_i$$

that is $D_i = \frac{1}{2} |c_i + c_i^i| |f_i|$.

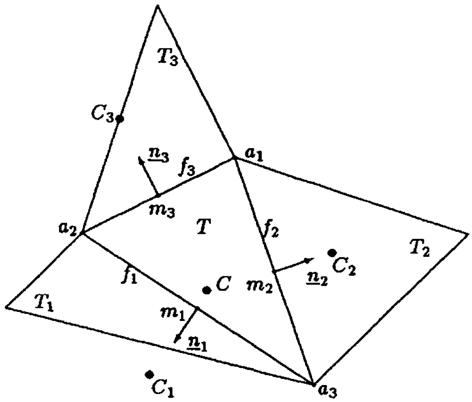

Figure 2.

For problem $(\tilde{P}_h)$, with the chosen orientation of the normals, the equation associated with T can be written :

$$\varphi_{f_1} + \varphi_{f_2} + \varphi_{f_3} = - \int_T f \, dx \, dy \tag{14}$$

where each flux φ_{f_i} (though face f_i) satisfies the simple equation :

$$\frac{1}{2} (c_i + c_i^i) \varphi_{f_i} + u_T - u_{T_i} = 0 , \tag{15}$$

thanks to the mass-lumping of Proposition 3.

Depending on the sign of $c_i + c_i^i$, three cases are to be considered for the approximation of φ_{f_i} given by (15) :

• Case 1 : $c_i + c_i^i > 0$.

That means that the pair (T, T_i) verifies Delaunay property (T_i is not included in the circumscibed circle of T, or the sum of the angles opposite of f_i is strictly smaller than π). Here (15) gives :

$$\varphi_{f_i} = \frac{u_{T_i} - u_T}{D_i} |f_i|$$

which is the natural finite difference approximation associating values u_T (resp. u_{T_i}) to point C (resp. C_i).

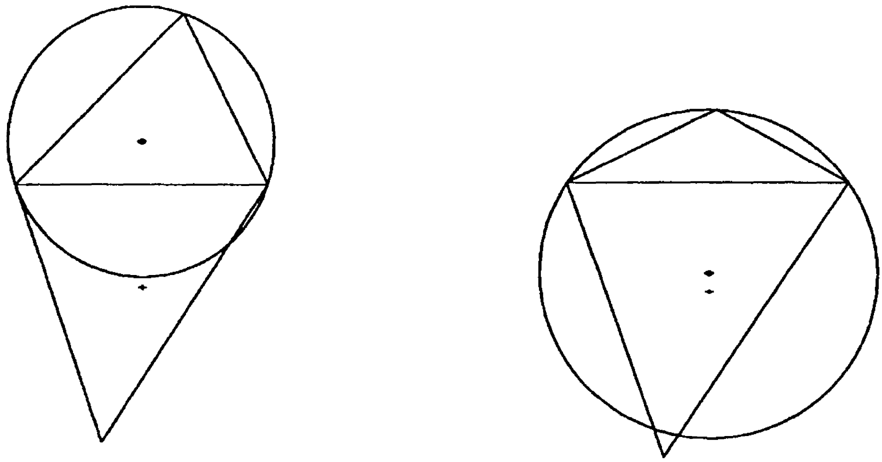

Figure 3.

• Case 2 : $c_i + c_i^i = 0.$

That means that T and T_i have the same circumscribed circle ($C_i = C$), or that the sum of the angles opposite to f_i is exactly π. Here (15) implies $u_{T_i} = u_T$, that is only one value for the quadrangular cell $T \cup T_i$ to which corresponds the conservation equation obtained by summing those of T and T_i *(voir figure 4).*

• Case 3 : $c_i + c_i^i < 0.$

That means that T_i is included in the interior of the circumscribed circle of T, or the sum of the angles opposite to f_i is greater than π. Here (15) gives :

$$\varphi_{f_i} - \frac{u_{T_i} - u_T}{D_i} |f_i|$$

which has the opposite sign of a natural approximation ! *(voir figure 5).*

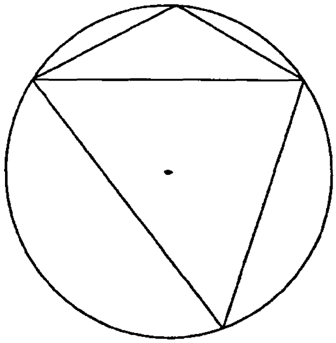

Figure 4.

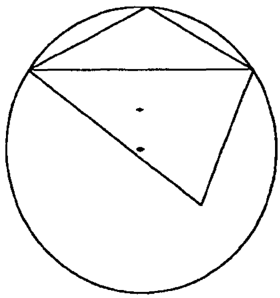

Figure 5.

Remark : in the three cases, it may happen that an angle opposite to f_i be obtuse, which implies that one center, at least, is outside the corresponding triangle. In this situation, it is not natural to affect the value of u in the triangle to that point... A remedy to avoid such situation is to use only acute triangles.

We must emphasize the validity of Proposition 5 (that is wellposedness of the problem $(\tilde{P}_h)$ and error bound $O(h)$), even when exotic situations (listed above) happen for some faces.

4. EXTENSIONS TO N-DIMENSIONAL CASE $(N \geq 3)$

We want to extend our method to N-dimensional simplicial elements K that is to search for a numerical integration formula diagonalising the mass matrix and being exact on $(P_0(K))^N$. This includes the case of triangles (proof of Proposition 3 of § 2) and 3-D tetrahedra.

We consider a simplex K, of measure $|K|$, vertices a_i with opposite face f_i, and barycenter $g = \frac{1}{N+1} \sum_{i=1}^{N+1} a_i$.

Shape functions $(\underline{p}_i)_i$ of space $RT_1(K)$, corresponding to the fluxes through faces $(f_i)_i$ of K, are given by (see J. C. Nedelec [13]) :

$$\underline{p}_i = \frac{1}{N|K|}(x - a_i), \quad \forall i \in \{1, \dots, N+1\}.$$

For $\underline{p}_0 \in (P_0(K))^N$, we have :

$$\underline{p}_0 = \sum_{i=1}^{N+1} \alpha_i \underline{p}_i = \frac{1}{N|K|}\left[\sum_{i=1}^{N+1} \alpha_i x - \sum_{i=1}^{N+1} \alpha_i a_i\right]$$

with :

$$\sum_{i=1}^{N+1} \alpha_i = 0. \tag{16}$$

We have $\int_K |\underline{p}_0|^2 \, dx = \alpha^T M\alpha$, with $M_{ij} = (\underline{p}_i, \underline{p}_j)_{(L^2(K))^N}$ and $\alpha^T = (\alpha_1, \dots, \alpha_{N+1})$. Our aim is here to search for $\mu = (\mu_1, \dots, \mu_{N+1})$ such that :

$$\int_K |\underline{p}_0|^2 \, dx = \alpha^T D_\mu \, \alpha$$

with $D_\mu = \operatorname{diag}(\mu_1, \dots, \mu_{N+1})$, that is to say :

$$\alpha^T M\alpha = \alpha^T D_\mu \, \alpha \tag{17}$$

with $\alpha \in \mathbb{R}^{N+1}$ satisfying (16). But condition (16) can also be written $e^T \alpha = 0$ with $e^T = (1, \dots, 1)$ and $e \in \mathbb{R}^{N+1}$, or also

$$\alpha = \left(I - \frac{1}{N+1} ee^T\right) \beta$$

with $\beta \in \mathbb{R}^{N+1}$. We notice that $Q = I - \frac{1}{N+1} ee^T$ is a projector which satisfies :

$$Qe = 0 \quad \text{and} \quad Q^T = Q . \tag{18}$$

Hence, relation (17) is equivalent to the matricial equality :

$$QMQ = QD_\mu Q . \tag{19}$$

We first calculate the exact mass matrix M (where g denotes the barycenter of K) :

$$M_{ij} = \int_K \underline{p}_i \cdot \underline{p}_j \, dx = \frac{1}{N^2 |K|^2} \int_K (x - a_i, x - a_j) \, dx$$

$$= \frac{1}{N^2 |K|^2} \int_K (x - g + g - a_i, x - g + g - a_j) \, dx$$

$$= \frac{1}{N^2 |K|^2} \left[\int_K (x - g)^2 \, dx + \int_K (g - a_i, g - a_j) \, dx \right]$$

since $\int_K (x - g) \, dx = 0.$

We obtain finally :

$$M_{ij} = \frac{1}{N^2 |K|^2} [|K| \rho_K^2 + |K| (g - a_i, g - a_j)] ,$$

that is $M = \frac{1}{N^2 |K|} (\rho_K^2 ee^T + GG^T)$ with $G^T = (g - a_1, \dots, g - a_{N+1})$.

From (18), we get $QM = \frac{1}{N^2 |K|} QGG^T$, that is to say :

$$QMQ = \frac{1}{N^2 |K|} QGG^T Q . \tag{20}$$

We have moreover :

$$G^T Q = (g - a_1, \dots, g - a_{N+1})\Big(I - \frac{1}{N+1} ee^T\Big)$$

$$= G^T - \frac{1}{N+1} (g - a_1, \dots, g - a_{N+1})\, ee^T$$

$$= G^T - \frac{1}{N+1} \left[\sum_{i=1}^{N+1} (g - a_i) \right] e^T$$

$$= G^T$$

and (20) reduces to $QMQ = \frac{1}{N^2|K|} GG^T$ and (19) can be rewriten :

$$\frac{1}{N^2|K|} GG^T = QD_\mu Q\,. \tag{21}$$

On the other hand :

$$QD_\mu Q = \Big(I - \frac{1}{N+1} ee^T\Big) D_\mu \Big(I - \frac{1}{N+1} ee^T\Big)$$

$$= D_\mu - \frac{1}{N+1} e\mu^T - \frac{1}{N+1} \mu e^T + \frac{1}{(N+1)^2} e\mu^T ee^T\,,$$

and (21) becomes finally :

$$\frac{1}{N^2|K|} (g - a_i, g - a_j) = \mu_i \delta_{ij} - \frac{1}{N+1} \mu_j - \frac{1}{N+1} \mu_i + \frac{1}{(N+1)^2} \sum_{k=1}^{N+1} \mu_k\,. \tag{22}$$

- For $i = j$ $(i = 1 \text{ to } N+1)$:
  we obtain, after summation on all i :

$$\sum_{i=1}^{N+1} \mu_i = \frac{N+1}{N^3|K|} \sum_{i=1}^{N+1} |g - a_i|^2\,.$$

Reporting this result in (22), we have the following relation :

$$\mu_i = \frac{N+1}{N^2(N-1)|K|} \left[|g - a_i|^2 - \frac{1}{N(N+1)} \sum_{i=1}^{N+1} |g - a_i|^2 \right]. \tag{23}$$

- For $i \neq j$:
  with (23) and (22), we have to satisfy :

$$\frac{1}{N^2|K|}(g - a_i, g - a_j) + \frac{1}{N+1}(\mu_i + \mu_j) - \frac{1}{(N+1)^2}\sum_{k=1}^{N+1} \mu_k = 0$$

or :

$$(g - a_i, g - a_j) + \frac{1}{N-1}\left[|g - a_i|^2 + |g - a_j|^2 - \frac{1}{N}\sum_{k=1}^{N+1} |g - a_k|^2 \right] = 0 . \quad (24)$$

For $N = 2$, condition (24) is satisfied for all $(i,j) \in (1, ..., N+1)^2$; moreover, we find for $(\mu_i)_i$, the unique values for a triangular element :

$$\mu_1 = \frac{1}{2}c_1, \quad \mu_2 = \frac{1}{2}c_2, \quad \mu_3 = \frac{1}{2}c_3,$$

proving Proposition 3 of § 2.

For $N = 3$, condition (24) is equivalent to the following relations :

$$\left|g - \frac{1}{2}(a_2 + a_3)\right|^2 = \left|g - \frac{1}{2}(a_2 + a_1)\right|^2 = \left|g - \frac{1}{2}(a_2 + a_4)\right|^2$$

$$= \left|g - \frac{1}{2}(a_1 + a_3)\right|^2 = \left|g - \frac{1}{2}(a_1 + a_4)\right|^2 = \left|g - \frac{1}{2}(a_3 + a_4)\right|^2$$

$$= \frac{1}{12}\sum_{i=1}^{4} |g - a_i|^2$$

or also :

$$|a_3 - a_1|^2 + |a_4 - a_2|^2 = |a_2 - a_1|^2 + |a_3 - a_4|^2 = |a_2 - a_3|^2 + |a_4 - a_1|^2 .$$

This last relation is a necessary and sufficient condition for K to be an orthocentered tetrahedron, which means that its four heights are concurrent ; in particular, regular tetrahedra satisfy this condition.

With this condition, coefficients $(\mu_i)_i$ are given by :

$$\mu_i = \frac{d_i}{|T_i|}$$

where :

- $|T_i|$ is the area of the face opposite to vertex a_i ;

• d_i is the distance between the barycenter of T_i and I, point of convergence of straight line perpendicular to each barycenter of the four faces (in case of regular tetrahedron, I is the barycenter of K).

For $N > 3$, (24) is verified by regular simplices, but the whole set of simplices verifying (24) (extension to N of orthocentered tetrahedra) is less evident to characterize geometrically.

5. CONCLUDING REMARKS

Essentially for a general 2-D triangular mesh, we have produced a 4-points cell-centered finite volume scheme in the variable u, after elimination of the variable $\underline{p}$ from the mixed finite element system, and that, thanks to an appropriate « mass-lumping ». We emphasize the fact that the error bound of Proposition 5 is $O(h)$ for an approximate $H^2(\Omega)$-norm of the error $u - \tilde{u}_h$, since $\underline{p} - \underline{\tilde{p}}_h = \nabla u - \nabla_h \tilde{u}_h$ and $\operatorname{div}(\underline{p} - \underline{\tilde{p}}_h) = \Delta u - \Delta_h \tilde{u}_h$.

Concerning the mass-lumping, it must be noticed that we have restricted ourselves to integration formulas being exact for constant fields, that is satisfying (5). It is with this restriction that we have obtained the unique formula of Proposition 3 for a general triangle, and in $N - D$ $(N \geqslant 3)$ existence only for special simplices, among which the regular ones. More general diagonalizing formulas have only to verify an inequality of the type of (7) with $\varepsilon(h_T) = O(h_T)$, implying a finite volume scheme verifying Proposition 5, that is with a $O(h)$ error bound.

Although our presentation is limited to the model problem, we can extend the present technique to the operator $-\operatorname{div}(\mathscr{A}\, \underline{\operatorname{grad}}(u))$, where $\mathscr{A}$ is an appropriate variable matrix, and to other problems such as those of convection-diffusion and elasticity. Extension to more general diagonalizing formulas is in progress.

REFERENCES

[1] J. BARANGER, P. EMONOT, J.-F. MAITRE, 1992, High order lagrangian finite volume elements for elliptic boundary value problems, *Numerical Methods in Engineering 92,* Ch. Hirsch *et al.* ed., Elsevier Science Publishers, pp. 709-713.

[2] J. BARANGER, J.-F. MAITRE, F. OUDIN, 1994, Application de la théorie des éléments finis mixtes à l'étude d'une classe de schémas aux volumes différences finis pour les problèmes elliptiques, *C.R. Acad. Sci. Paris,* **319,** Série I, pp. 401-404.

[3] F. BREZZI, M. FORTIN, 1991, Mixed and Hybrid Finite Element Methods, *Springer Series in Computational Mathematics,* **15,** Springer-Verlag, New-York.

[4] P. G. CIARLET, P. A. RAVIART, 1972, General Lagrange and Hermite interpolation in $\mathbb{R}^n$ with applications to finite element methods, *Arch. Rat. Mech. Anal.*, **46,** pp. 177-199.

[5] I. FAILLE, 1992, A control volume method to solve an elliptic equation on a two-dimensional irregular mesh, *Comput. Methods Appl. Mech. Engrg.*, **100,** pp. 275-290.

[6] I. FAILLE, T. GALLOUET, R. HERBIN, 1991, Des mathématiciens découvrent les volumes finis, S.M.A.I., *Matapli, Bulletin de liaison,* **28.**

[7] M. FARHLOUL, 1991, Méthodes d'éléments finis mixtes et volumes finis, Thèse, Université Laval, Québec.

[8] D. A. FORSYTH Jr, P. H. SAMMON, 1988, Quadratic convergence for cell-centered grids, *Applied Numerical Mathematics,* **4,** pp. 377-394.

[9] W. HACKBUSCH, 1989, On first and second order box schemes, *Computing,* **41,** pp. 277-296.

[10] Y. HAUGAZEAU, P. LACOSTE, 1993, Condensation de la matrice de masse pour les éléments finis mixtes de $H(\mathrm{rot})$, *C.R. Acad. Sci. Paris,* **316,** Série I, pp. 509-512.

[11] R. HERBIN, An error estimate for a finite volume scheme for a diffusion convection problem on a triangular mesh, *accepted for publication in Num. Meth. P.D.E.*

[12] K. W. MORTON, E. SULI, 1991, Finite volume methods and their analysis, *IMA Journal of Num. Anal.,* **11,** pp. 241-260.

[13] J. C. NEDELEC, 1991, Notions sur les techniques d'éléments finis, *Mathématiques et Applications,* **7,** Ellipses-Edition Marketing.

[14] P. A. RAVIART, J. M. THOMAS, 1977, A mixed finite element method for second order elliptic problems, *Mathematical Aspects of the Finite Element Methods* (I. Galligani, E. Magenes, eds), *Lectures Notes in Math.,* **606,** Springer-Verlag, New-York.

[15] J. E. ROBERTS, J. M. THOMAS, 1989, Mixed and hybrid methods, in *Handbook of Numerical Analysis,* (P. G. Ciarlet and J. L. Lions, eds.), Vol. II, Finite Element Methods (Part 1), North-Holland, Amsterdam.

[16] A. WEISER, M. F. WHEELER, 1988, On convergence of block-centered finite differences for elliptic problems, *SIAM J. Numer. Anal.,* **25,** pp. 351-375.