

EXPLICIT ERROR BOUNDS IN A CONFORMING FINITE ELEMENT METHOD

PHILIPPE DESTUYNDER AND BRIGITTE MÉTIVET

ABSTRACT. The goal of this paper is to define a procedure for bounding the error in a conforming finite element method. The new point is that this upper bound is fully explicit and can be computed locally. Numerical tests prove the efficiency of the method. It is presented here for the case of the Poisson equation and a first order finite element approximation.

1. INTRODUCTION

Let us consider the following problem:

$$(1) \quad \begin{cases} \text{find } u \in H_0^1(\Omega) \text{ such that, for all } v \in H_0^1(\Omega), \\ \int_{\Omega} \nabla u \bullet \nabla v = \int_{\Omega} f v, \end{cases}$$

where f is a function in the space $L^2(\Omega)$ and ∇ denotes the gradient of a function. It is well known that (1) has a unique solution. Furthermore, under classical assumptions, one can prove that u is an element of the space $H^2(\Omega) \cap H_0^1(\Omega)$ (no re-entrant angle* on the boundary, which should be piecewise smooth enough).

Let us now consider a family of triangulations of Ω , assumed to be uniformly regular (see Girault and Raviart [16]). One triangulation is denoted by \mathcal{T}^h , where h denotes the size of the mesh. The approximation space of $H_0^1(\Omega)$, based on the triangulation \mathcal{T}^h , is denoted by V^h and is, for instance, defined by

$$(2) \quad V^h = \{v \in H_0^1(\Omega), \forall K \in \mathcal{T}^h, v|_K \in P_1(K)\},$$

where $P_1(K)$ is the first degree polynomial space. Then the approximation of u , denoted by u^h , is defined by

$$(3) \quad \begin{cases} \text{find } u^h \in V^h \text{ such that for all } v \in V^h \\ \int_{\Omega} \nabla u^h \bullet \nabla v = \int_{\Omega} f v. \end{cases}$$

The classical error estimate between u and u^h is derived from the a priori inequality

$$(4) \quad |u - u^h|_{1,\Omega} \leq \inf_{v \in V^h} |v - u|_{1,\Omega},$$

where $|\cdot|_{1,\Omega}$ is the classical H^1 seminorm. From the interpolation results (see Ciarlet [8] or Raviart and Thomas [22]), one can deduce that there exists a constant c which

Received by the editor June 5, 1996 and, in revised form, February 19, 1998.
1991 *Mathematics Subject Classification*. Primary 65N30, 65R20, 73C50.

*As a matter of fact, the internal angles on the boundary should be smaller than π .

is independent of both h and u , and such that

$$(5) \quad |u - u^h|_{1,\Omega} \leq ch|u|_{2,\Omega}.$$

But, unfortunately, this estimate is not explicit because $|u|_{2,\Omega}$ is not.

Let us explain another way to derive an error bound, which was introduced by P. Ladevèze in his thesis [17] and which is a particular implication of Prager and Synge's identity. Let us first introduce the following set of vector fields in Ω :

$$(6) \quad H_f(\text{div}, \Omega) = \{p \in (L^2(\Omega))^2, \text{div } p + f = 0 \text{ in } \Omega\}.$$

Then one has the inequality, used first by Ladevèze [17],

$$(7) \quad |u - u^h|_{1,\Omega} \leq \inf_{p \in H_f(\text{div}, \Omega)} \|p - \nabla u^h\|_{0,\Omega},$$

the proof of which is a straightforward consequence of the following identity (Prager and Synge [20]). Let p and v be arbitrary elements in the sets $H_f(\text{div}, \Omega)$ and V^h , respectively. Then, if u is a solution of the Poisson model, one has

$$|u - v|_{1,\Omega}^2 + \|p - \nabla u\|_{0,\Omega}^2 = \|p - \nabla v\|_{0,\Omega}^2.$$

Because of its simplicity the proof is left to the reader. The inequality (7) is obtained with $v = u^h$.

The goal of the method that we develop is then to define an element p in the set $H_f(\text{div}, \Omega)$ such that $p - \nabla u^h$ is as small as possible. In this paper we suggest a choice for p , and we prove that the term $\|p - \nabla u^h\|_{0,\Omega}$ is $O(h)$, provided that u is in the space $H^2(\Omega)$ and that the mesh family satisfies a uniform regularity assumption.

Let us point out the differences between (4) and (7). The first is a so-called a priori estimate, and the second is a posteriori. In the first case the exact solution u is involved, but in the second case only u^h is necessary. The error bound deduced from (4) requires us to define an element v in the case V^h such that $|\nabla v - \nabla u|_{1,\Omega}$ will be as small as possible. The space V^h is a conforming approximation of $H_0^1(\Omega)$. When (7) is used, the infimum is taken over vector fields chosen in the admissible set $H_f(\text{div}, \Omega)$ for the problem dual to (1). Let us recall that this dual problem consists in minimizing in $H_f(\text{div}, \Omega)$ the function

$$(8) \quad p \rightarrow \frac{1}{2} \int_{\Omega} |p|^2.$$

The numerical approximation of this problem is very difficult, and one prefers to use a mixed formulation. It enables one to avoid requiring exact satisfaction of the condition

$$(9) \quad \text{div } p + f = 0 \quad \text{in } \Omega.$$

As a matter of fact it is, for instance, replaced by

$$\text{div } p^h + \frac{1}{|K|} \int_K f = 0, \quad \forall K \in \mathcal{T}^h$$

(p^h being the solution of a first order mixed finite element).

Such an element p^h cannot be used in (7) because $p^h \notin H_f(\text{div}, \Omega)$. Further details concerning mixed finite elements can be found in Roberts and Thomas [23].

Hence we add a complementary element, call it δp_K , defined by

$$(10) \quad \begin{cases} \begin{cases} \delta p_K = \nabla \delta u_K, & \text{in } K, \\ -\Delta \delta u_K = f - \frac{1}{|K|} \int_K f, & \text{in } K, \end{cases} \\ \frac{\partial \delta u_K}{\partial \nu} = 0 \text{ on } \partial K \quad \text{and} \quad \int_K \delta u_K = 0. \end{cases}$$

The term $p = p^h + \delta p_K$ is then in $H_f(\text{div}, \Omega)$. It can be proved that, with this choice, $p - \nabla u^h$ is $O(h)$ in the $L^2(\Omega)$ norm. But unfortunately the computation of p^h is not local and requires the solution of a global linear system (over the whole triangulation \mathcal{T}^h). Hence our goal is to construct a local approximation of p^h that does not require many computations. One application of the method is the adaptive mesh refinement, but let us point out that the true new point is that the error bound is explicit. Then we also discuss the asymptotic exactness of our error estimator.

There is a well-developed literature on a posteriori error estimates and adaptive mesh refinement for the elliptic equations. It seems quite impossible to list each contribution in a single paper. But let us try to mention some of the papers that are closest to our formulation.

The closest idea is due to Ladevèze [17]. But it appears that this author did not use an exact construction of the dual variable, which we need in our formulation. From a mathematical point of view, Ainsworth and Oden [1] have underlined the interest in a coupling between a conformal finite element approximation and a hybrid one. They suggest using the Lagrange multiplier, which is defined in order to prescribe the inter-element continuity, in order to construct an error estimator by solving a local (i.e., element by element) problem. The way they do it is close to but different from the one we suggest in this paper. The idea of comparing the finite element solution with the dual problem is also the origin of the method developed by Zienkiewicz and Zhu in [26] and [27], but they did not require the dual variable (the stress field in their mechanical applications) to satisfy the equilibrium equation. Moreover, they used a whole continuity of this dual variable at the inter-element, instead of only the one of the normal component. Therefore, the strategy seems to be hazardous in case of singularities like a discontinuity of coefficients in the operator (bimaterial).

From the mathematical point of view, let us mention three other strategies which are well founded and seem to be very promising. To our best knowledge the first is due to Babuška and Rheinboldt [4]. The basic trick consists in bounding the error between the exact and approximate solution by a constant times the so-called residual terms. There are two. One of them is the jump between the normal derivatives of the finite element solution across the inter-elements, and the other one is the lack of equilibrium inside the elements. Then Verfürth [25] and Bernardi, Métivet and Verfürth [7] proved that this error bound is also—up to a multiplicative constant—a lower bound on the error. The method, which can be extended to elasticity and the Stokes model [25], [7], seems to be very efficient in numerical applications. This is why we used it to compare with our formulation in the numerical tests in the last section of this paper.

The second strategy is certainly the most promising for the near future. It is based on superconvergence results. For particular meshes this strategy was developed by Babuška and Rodriguez [5], but the most important step for defining superconvergence points was achieved by Schatz, Sloan and Walkbin [24]. The mesh refinement should be defined using these points. The advantage is that the method would lead to a local contribution to the error.

The third method is quite close to ours from a theoretical point of view. It was developed by Bank and Weiser [6]. The basic point seems to be to solve a local Neumann problem in order to construct an a posteriori error estimator. The main advantage of the method, compared to others, is that it gives (in an appropriate norm) an asymptotically exact estimate of the error as the mesh size tends to zero.

2. ORGANISATION OF THE PAPER

First we will recall a few properties of the approximation model. Then we find an element, call it p^h , whose construction can be performed locally (i.e., in the vicinity of one vertex of the mesh). This element satisfies

$$\operatorname{div} p^h + \frac{1}{|K|} \int_K f = 0 \quad \forall K \in \mathcal{T}^h,$$

and $p^h \in H(\operatorname{div}, \Omega)$. The next step consists in finding a solution δu_K of (10) and in proving that δu_K can be small if f is smooth enough. Using the Green kernel, this term is explicit as the solution of a local boundary integral equation. The numerical solution can then be found with a predefined accuracy.

The last step, but not the least, is to prove that the error bound $p - \nabla u^h$, where $p = p^h + \nabla \delta u_K$ in each element K of \mathcal{T}^h , is itself bounded by $O(h)$.

3. PROPERTIES OF THE CONFORMING FINITE ELEMENT SOLUTION

Let us denote by \mathcal{S}^h the set of all the internal vertices of \mathcal{T}^h . For each vertex S_i we introduce the basis function λ_i of V^h , which is equal to 1 at S_i and 0 at all the other vertices. From the definition of u^h we have

$$(11) \quad \int_{\mathcal{C}_i^h} \nabla u^h \bullet \nabla \lambda_i - \int_{\mathcal{C}_i^h} f \lambda_i = 0,$$

where \mathcal{C}_i^h is the so-called “cluster” around S_i —the collection of elements K of \mathcal{T}^h which have S_i as a vertex (see Figure 1). We denote by γ_i^k the sides of \mathcal{C}_i^h that have S_i as one of their two extremities. The number of elements K in \mathcal{C}_i^h is κ .

Using the Stokes formula, one can transform (11) into the following relationship:

$$(12) \quad \sum_{k=1}^{\kappa} \left\{ \frac{\operatorname{meas}(\gamma_i^k)}{2} \left[\frac{\partial u^h}{\partial \nu} \right]_i^k - \int_{K^k} f \lambda_i \right\} = 0,$$

where $[\cdot]_i^k$ is the jump of a quantity across the side γ_i^k . This relation can be interpreted from a mechanical (for instance) point of view. The first terms represent the moment of ∇u^h at the vertex S_i and along the side γ_i^k . The second term is the moment of the external forces acting in K and expressed at the vertex S_i . Then (12) gives a global equilibrium of these moments at S_i .

The basic idea in the construction of p^h mentioned in the introduction is to equilibrate separately on each element K of \mathcal{C}_i^h the moments at the vertex S_i and

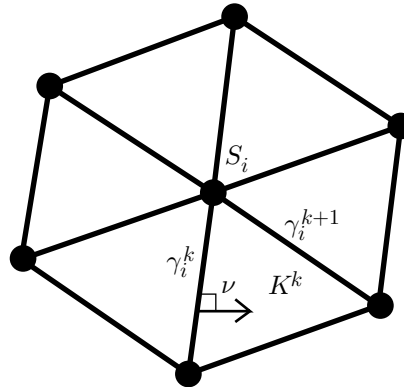


FIGURE 1. Cluster \mathcal{C}_i^h , $S_i \in \mathcal{S}^h$

to ensure the equilibrium between two neighbour elements. Hence we look for p^h such that

$$(13) \quad \begin{cases} p_i^h \in H(\text{div}, \mathcal{C}_i^h), & p_i^h \cdot \nu = 0 \quad \text{on } \partial K \cap \partial \mathcal{C}_i^h, \\ \text{div } p_i^h = \frac{1}{|K|} \left\{ \int_K \nabla u^h \cdot \nabla \lambda_i - \int_K f \lambda_i \right\}, \end{cases}$$

which implies that an assumed solution satisfies, for all $K \in \mathcal{C}_i^h$,

$$\begin{aligned} \int_{\partial K} p_i^h \cdot \nu &= \int_K \nabla u^h \cdot \nabla \lambda_i - \int_K f \lambda_i \\ &= \int_{\gamma_i^1} \frac{\partial u^h}{\partial \nu} \lambda_i + \int_{\gamma_i^2} \frac{\partial u^h}{\partial \nu} \lambda_i - \int_K f \lambda_i \\ &= \frac{\text{meas}(\gamma_i^1)}{2} \left(\frac{\partial u^h}{\partial \nu} \right) + \frac{\text{meas}(\gamma_i^2)}{2} \left(\frac{\partial u^h}{\partial \nu} \right) - \int_K f \lambda_i \end{aligned}$$

(γ_i^1 and γ_i^2 are the two sides of K which have S_i as an extremity). Hence the term p_i^h appears as the complementary system of forces which could be applied in order to equilibrate separately each triangle of the cluster \mathcal{C}_i^h for the test variable λ_i . The existence of p_i^h is proved in the next section.

4. DEFINITION OF AN EQUILIBRIUM VECTOR FIELD ON Ω

Let us consider an arbitrary vertex S_i of the triangulation \mathcal{T}^h . We associate to S_i the cluster \mathcal{C}_i^h , which is the set of elements K of \mathcal{T}^h such that S_i is a vertex of K . But S_i can be a point on the boundary of Ω . In both cases the cluster \mathcal{C}_i^h can be defined as shown in Figure 2.

It is worth noting that the definition of the boundary $\partial \mathcal{C}_i^h$ in this second situation does not include the vertices which belong to the boundary $\partial \Omega$ of Ω .

In order to approach the vector fields of the space $H(\text{div}, \Omega)$, we make use of the finite elements introduced by Raviart and Thomas in [21]. Their restriction to the

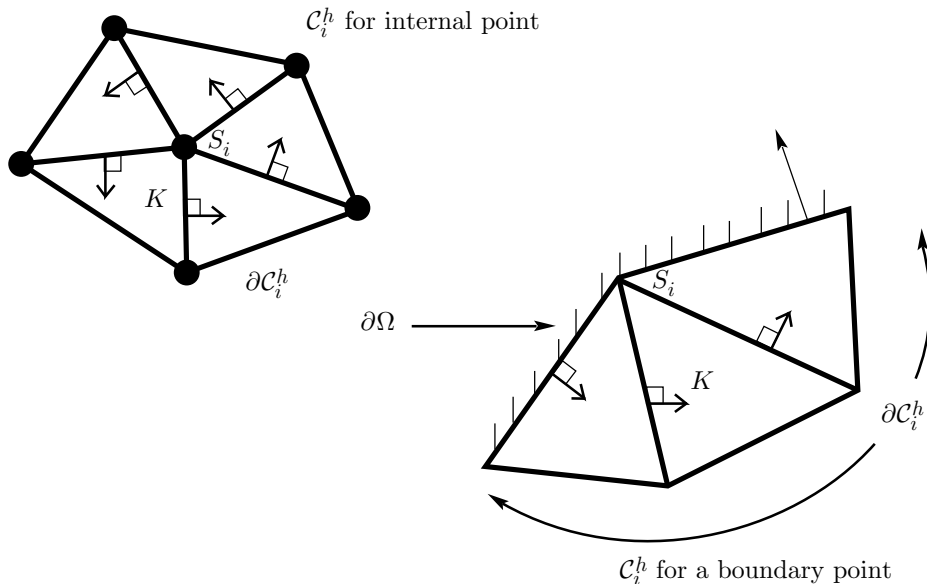


FIGURE 2. The two types of clusters

cluster \mathcal{C}_i^h is denoted by $H_{RT1}(\text{div}, \mathcal{C}_i^h)$, and we use the definition

$$H_{RT1}(\text{div}, \mathcal{C}_i^h) = \left\{ p \in (L^2(\mathcal{C}_i^h))^2, p \bullet \nu = 0 \text{ on } \partial \mathcal{C}_i^h, \right. \\ \left. \forall K \in \mathcal{C}_i^h, p|_K = \begin{Bmatrix} a_K + b_K x \\ c_K + b_K y \end{Bmatrix} \right\}.$$

The index RT1 means Raviart-Thomas, degree 1. One remarkable property of the vectors in the space $H_{RT1}(\text{div}, \mathcal{C}_i^h)$ is that $p \bullet \nu$ is constant and continuous across the sides of the elements, because p is in the space $H(\text{div}, \mathcal{C}_i^h)$. Then we introduce the following problem:

$$(14) \quad \begin{cases} \text{Find } p_i^h \in H_{RT1}(\text{div}, \mathcal{C}_i^h) \text{ such that for all } K \in \mathcal{C}_i^h, \\ \text{div } p_i^h = \frac{1}{|K|} \left\{ \int_K \nabla u^h \bullet \nabla \lambda_i - \int_K f \lambda_i \right\}, \text{ in } K, \end{cases}$$

and we prove the next result.

Theorem 1. *There exists a solution to (14) that is defined up to an element of the kernel of the linear system. More precisely,*

$$p_i^h = \mathbf{p}_i^h + \alpha_i \text{rot } \lambda_i, \quad \alpha_i \in \mathbb{R},$$

where \mathbf{p}_i^h is a particular solution of (14)

Proof. a) Let us begin with the case where S_i is an internal vertex of the triangulation \mathcal{T}^h . Then the linear system (14) has the same number of unknowns and of equations (there are as many sides from S_i as elements in \mathcal{C}_i^h). Let us therefore analyze the homogeneous system associated to (14). This problem consists in finding an element δ_i^h in $H_{RT1}(\text{div}, \mathcal{C}_i^h)$ such that for all $K \in \mathcal{C}_i^h$,

$$(15) \quad \text{div } \delta_i^h = 0 \text{ in } K.$$

Because the cluster \mathcal{C}_i^h is a simply connected open set, there exists a function ψ_i such that

$$\delta_i^h = \text{rot } \psi_i, \quad \psi_i \in H^1(\mathcal{C}_i^h).$$

Furthermore, the condition $\delta_i^h \bullet \nu = 0$ on $\partial\mathcal{C}_i^h$ implies that ψ_i is constant along $\partial\mathcal{C}_i^h$. As ψ_i is defined up to a constant, without any loss of generality we can choose $\psi_i = 0$ on $\partial\mathcal{C}_i^h$. But δ_i^h is in the space $H_{\text{RT1}}(\text{div}, \partial\mathcal{C}_i^h)$. Hence on each element K of \mathcal{C}_i^h one has

$$\text{rot } \psi_i = \begin{vmatrix} a_K + b_K x \\ c_K + b_K y \end{vmatrix}_{c_K}^{a_K}$$

(b_K is zero because $\text{div } \delta_i^h = 0$). Finally, ψ_i is piecewise linear and therefore is proportional to the basis function λ_i . In this situation (internal vertex), the kernel of (14), which is one dimensional, is generated by the vector $\text{rot } \lambda_i$.

b) If now we consider S_i on the boundary of Ω , everything we did in the previous situation is still valid concerning the kernel.

It has been proved that the linear system (14) is singular and that the kernel is one dimensional. When the vertex S_i is internal to the triangulation \mathcal{T}^h , the matrix of the linear system (14) is a square matrix and therefore the right-hand side must be orthogonal to the cokernel (i.e., the kernel of the transposed matrix). When S_i is on the boundary $\partial\Omega$, there is no compatibility requirement because the matrix associated to (14) is rectangular and we have one more unknown than equations.

Let us characterize the cokernel of the matrix associated to the linear system (14). We already know that it is one dimensional.

An element $X = (X_i)$, $i = 1, \dots, \kappa$, of the cokernel satisfies

$$\forall q \in H_{\text{RT1}}(\text{div}, \mathcal{C}_i^h), \quad \sum_{j=1}^{\kappa} X_j (\text{div } q)|_{K_j} = 0 \quad (K_j \in \mathcal{C}_i^h),$$

but as $X_j (\text{div } q)|_{K_j}$ is constant on each triangle K_j of \mathcal{C}_i^h , one has

$$\sum_{j=1}^{\kappa} \frac{1}{|K_j|} \int_{K_j} X_j (\text{div } q)|_{K_j} = 0,$$

or else

$$\sum_{j=1}^{\kappa} \int_{\gamma_i^j} \left(\frac{X_j}{|K_j|} - \frac{X_{j-1}}{|K_{j-1}|} \right) q \bullet \nu|_{\gamma_i^j} = 0,$$

where γ_i^j denotes the sides of \mathcal{C}_i^h that have S_i as an extremity, as shown in Figure 2. Thus the quantity

$$\xi = X_j / |K_j|, \quad j = 1, \dots, \kappa,$$

is constant for any j . The cokernel is finally spanned by the vector $X = (X_j) = (|K_j|)$ (measure of K_j).

The compatibility condition for the system (14) can then be formulated as

$$\sum_{j=1}^{\kappa} \left[\int_{K_j} \nabla u^h \cdot \nabla \lambda_i - \int_{K_j} f \lambda_i \right] = 0,$$

or else

$$(16) \quad \int_{C_i^h} \nabla u^h \cdot \nabla \lambda_i = \int_{C_i^h} f \lambda_i,$$

which is precisely (for internal vertices) one of the equations characterizing u^h . Therefore the right-hand side of (14) is orthogonal to the cokernel, and Theorem 1 is proved. \square

From the elements p_i^h defined in Theorem 1 we introduce the term

$$(17) \quad p^h = \sum_{i \in \bar{S}^h} p_i^h,$$

where \bar{S}^h is the set of all vertices of triangulation \mathcal{T}^h , including those on the boundary of Ω (S^h is restricted to the internal points). From the definition of p_i^h , and because $p_i^h = 0$ on $\Omega - C_i^h$, one has

$$\operatorname{div} p^h = \sum_{i \in \bar{S}^h} \operatorname{div} p_i^h = \sum_{i \in \bar{S}^h} \sum_{K \in C_i^h} \frac{1}{|K|} \left\{ \int_K \nabla u^h \cdot \nabla \lambda_i - \int_K f \lambda_i \right\},$$

and because on each triangle K we have $\sum_{i \in \bar{S}^h} \lambda_i = 1$, we conclude that, for all $K \in \mathcal{T}^h$,

$$(18) \quad \operatorname{div} p^h + \frac{1}{|K|} \int_K f = 0.$$

As the element p^h defined in (18) is not in the set $H_f(\operatorname{div}, \Omega)$, we add a local term δu_K (defined on each triangle of \mathcal{T}^h) such that

$$(19) \quad \begin{cases} \delta p_K = \nabla \delta u_K & \text{on } K, \quad \text{with } \delta u_K \in H^1(K), \\ -\Delta \delta u_K = f - \frac{1}{|K|} \int_K f & \text{on } K, \\ \frac{\partial \delta u_K}{\partial \nu} = 0 \text{ on } \partial K & \text{and } \int_K \delta u_K = 0. \end{cases}$$

The existence and uniqueness of a solution to (19) is very classical, and finally we set, on each K of \mathcal{T}^h ,

$$(20) \quad p = p^h + \delta p_K.$$

It is worth noting that $\delta p_K \in H(\operatorname{div}, \Omega)$, because of the homogeneous Neumann boundary condition that we chose on ∂K . Then a simple compilation of the previous results shows that

$$(21) \quad \operatorname{div} p + f = 0 \quad \text{on } \Omega.$$

As a matter of fact, the term δu_K is only dependent of the right-hand side f of the problem (1). It is obvious that $\delta u_K = 0$ if $f = 0$ on K . More precisely, we can upper bound δu_K depending on the regularity (local) of f .

The result is made explicit in the following theorem.

Theorem 2. *Assume that f is in $L^2(\Omega)$ and that the triangulation family is regular. Then there exists a constant c , independent of both f and h , such that*

$$|\delta u_K|_{1,K} \leq ch \|f\|_{0,K}.$$

Furthermore, if f is in $H^1(K)$, then, under the same assumptions,

$$|\delta u_K|_{1,K} \leq ch^2 \|f\|_{1,K}.$$

Proof. From the definition of δu_K , and letting Π_0 denote the $L^2(K)$ projection onto the constants, one obtains

$$\begin{aligned} |\delta u_K|_{1,K}^2 &= - \int_K \Delta(\delta u_K) \delta u_K = \int_K \left(f - \frac{1}{|K|} \int_K f \right) \delta u_K \\ &\leq \|f - \Pi_0 f\|_{0,K} \|\delta u_K\|_{0,K} \end{aligned}$$

and, by Lemmas 1 and 2 (see the Appendix),

$$\begin{aligned} |\delta u_K|_{1,K}^2 &\leq ch \|f - \Pi_0 f\|_{0,K} |\delta u_K|_{1,K} \\ &\leq \begin{cases} ch \|f\|_{0,K} |\delta u_K|_{1,K} & \text{if } f \in L^2(K), \\ ch^2 \|f\|_{1,K} |\delta u_K|_{1,K} & \text{if } f \in H^1(K). \end{cases} \end{aligned}$$

This completes the proof of Theorem 2. □

Remark. In the definition of p_i^h the coefficient α_i (see Theorem 1) is not yet defined. Let us mention one possibility. Consider one side of a cluster with the center S_i as an extremity. Then on this side, call it γ_i^1 , one has

$$\begin{aligned} \alpha_i &\int_{\gamma_i^1} \text{rot } \lambda_i \bullet \nu + \int_{\gamma_i^1} \mathbf{p}_i^h \bullet \nu - \int_{\gamma_i^1} \frac{1}{2} \frac{\partial u^h}{\partial \nu} \\ &= \alpha_i + \int_{\gamma_i^1} \left(\mathbf{p}_i^h \bullet \nu - \frac{1}{2} \frac{\partial u^h}{\partial \nu} \right), \end{aligned}$$

and we can choose α_i such that this quantity is zero. Hence

$$\begin{aligned} \alpha_i &= - \int_{\gamma_i^1} \left(\mathbf{p}_i^h \bullet \nu - \frac{1}{2} \frac{\partial u^h}{\partial \nu} \right) \\ (22) \quad &= - \int_{\gamma_i^1} \left(\mathbf{p}_i^h \nu - \frac{\partial u^h}{\partial \nu} \lambda_i \right) \\ &= - \text{meas}(\gamma_i^1) \left[\mathbf{p}_i^h \bullet \nu - \frac{1}{2} \frac{\partial u^h}{\partial \nu} \right]_{|\gamma_i^1} \end{aligned}$$

We shall prove below that such a choice leads to a consistent error bound. □

5. ASYMPTOTIC BEHAVIOUR OF THE EXPLICIT ERROR BOUND BETWEEN u AND u^h WHEN h TENDS TO ZERO

Let us consider the element p defined in (20). From the a posteriori inequality, we have

$$|u - u^h|_{1,\Omega} \leq \|p - \nabla u^h\|_{0,\Omega} = \varepsilon.$$

The main result of this section is to prove that ε is bounded by $O(h)$. This will justify that this explicit error bound is consistent with respect to the classical results known in finite element methods.

Theorem 3. *Assume that f is in $L^2(\Omega)$ and that the triangulation family is uniformly regular. Then there exists a constant c , independent of both h (the mesh size) and f , such that*

$$\|p - \nabla u^h\|_{0,\Omega} \leq ch[\|f\|_{0,\Omega} + |u|_{2,\Omega}].$$

Proof. First of all, on each element K of the triangulation \mathcal{T}^h we set

$$k = p^h + \nabla \delta u_K - \nabla u^h.$$

Note that on K we have $\text{curl } k = 0$ and $\text{div } k + f = 0$. Furthermore, on the boundary ∂K of K , k satisfies

$$k \bullet \nu = p^h \nu \bullet - \frac{\partial u^h}{\partial \nu} \quad \left(\text{because } \frac{\partial \delta u_K}{\partial \nu} = 0 \text{ on } \partial K \right).$$

Therefore, we can deduce that there exists a function φ_K such that

$$\begin{cases} k = \nabla \varphi_K & \text{and} & \int_K \varphi_K = 0, \\ \varphi_K \in H^1(K); \end{cases}$$

in addition φ_K is a solution of

$$(23) \quad \begin{cases} -\Delta \varphi_K = f & \text{in } K, & \int_K \varphi_K = 0, \\ \frac{\partial \varphi_K}{\partial \nu} = p^h \bullet \nu - \frac{\partial u^h}{\partial \nu} & \text{on } \partial K, & \varphi_K \in H^1(K). \end{cases}$$

The previous model defines φ_K uniquely. But one also has

$$(24) \quad \varepsilon_K^2 \stackrel{\text{def}}{=} \int_K |p^h + \nabla \delta u_K - \nabla u^h|^2 = |\varphi_K|_{1,K}^2$$

and, from (23),

$$\varepsilon_K^2 = \int_K f \varphi_K + \int_{\partial K} \left(p^h \bullet \nu - \frac{\partial u^h}{\partial \nu} \right) \varphi_K.$$

Our goal is now to prove that ε_K is $O(h)$. First of all,

$$\varepsilon_K^2 = \int_K f \varphi_K + \sum_{i \in \overline{S}^h} \int_{\partial K} p_i^h \bullet \nu \varphi_K - \int_{\partial K} \frac{\partial u^h}{\partial \nu} \varphi_K.$$

But the summation over the index i here is restricted to the three vertices of K . Let us introduce the element \hat{q} in the space $H_{RT1}(\text{div}, K)$ defined (see Raviart and Thomas in [21]) by

$$\forall \gamma \in \partial K, \quad \hat{q} \bullet \nu|_\gamma = \frac{1}{\text{meas}(\gamma)} \int_\gamma \frac{\partial u}{\partial \nu},$$

where γ is a side of K and u is the solution of the initial problem, and we assume that u is in $H^2(\Omega)$. The error estimates proved by Raviart and Thomas [21] lead to (the triangulation family is assumed to be regular)

$$(25) \quad \begin{cases} \|\hat{q} - \nabla u\|_{0,K} \leq ch|u|_{2,K}, \\ \text{and} \\ \|\text{div } \hat{q} - \text{div}(\nabla u)\|_{0,K} = \|\text{div } \hat{q} + f\|_{0,K} \leq ch[\|f\|_{0,K} + |u|_{2,K}]. \end{cases}$$

Then one obtains the following equality:

$$\begin{aligned} \varepsilon_K^2 &= \int_K f \varphi_K + \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \hat{q} \cdot \nu) \varphi_K \\ &\quad + \int_{\partial K} \left(\hat{q} \cdot \nu - \frac{\partial u^h}{\partial \nu} \right) \varphi_K, \end{aligned}$$

or else, using the Stokes formula (i_1, i_2 and i_3 are the three vertices of K),

$$\begin{aligned} \varepsilon_K^2 &= \int_K f \varphi_K + \int_K (\hat{q} - \nabla u^h) \nabla \varphi_K + \int_K \operatorname{div} \hat{q} \varphi_K \\ &\quad + \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \hat{q} \cdot \nu) \varphi_K. \end{aligned}$$

Hence

$$\begin{aligned} \varepsilon_K^2 &\leq \|f + \operatorname{div} \hat{q}\|_{0,K} \|\varphi_K\|_{0,K} + \|\hat{q} - \nabla u^h\|_{0,K} |\varphi_K|_{1,K} \\ (26) \quad &\quad + \left| \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \hat{q} \cdot \nu) \varphi_K \right|. \end{aligned}$$

From the triangular inequality and Lemma 1 in the Appendix, we deduce that

$$\begin{aligned} \varepsilon_K^2 &\leq ch[\|f\|_{0,K} + |u|_{2,K} + |u - u^h|_{1,K}] |\varphi_K|_{1,K} \\ (27) \quad &\quad + \left| \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \hat{q} \cdot \nu) \varphi_K \right|. \end{aligned}$$

But on each side γ of ∂K only two terms p_i^h (for $i = i_1$ and i_2 , for instance) are different from zero. Therefore

$$\begin{aligned} &\left| \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \hat{q} \cdot \nu) \varphi_K \right| \\ &= \sum_{\gamma \subset \partial K} \left\{ \sum_{i \in \{i_1, i_2\}} \int_{\gamma} (p_i^h \cdot \nu - \frac{1}{2} \hat{q} \cdot \nu) \varphi_K \right\} \\ &\leq 2 \sum_{\gamma \subset \partial K} \left\{ |p_i^h \nu - \frac{1}{2} \hat{q} \cdot \nu|_{|\gamma} \sqrt{\operatorname{meas}(\gamma)} \|\varphi_K\|_{0,\gamma} \right\}, \end{aligned}$$

where γ_i^k is one side of the cluster C_i^h with the center S_i as an extremity. First of all, if we define

$$X_i^k \stackrel{\text{def}}{=} \int_{\gamma_i^k} (p_i^h \cdot \nu - \lambda_i \hat{q} \cdot \nu) = \operatorname{meas}(\gamma_i^k) [r_i^h \cdot \nu - \frac{1}{2} \hat{q} \cdot \nu]_{|\gamma_i^k},$$

then

$$\begin{aligned} (28) \quad X_i^{k+1} - X_i^k &= \int_{\gamma_i^{k+1}} (p_i^h \cdot \nu - \lambda_i \hat{q} \cdot \nu) - \int_{\gamma_i^k} (p_i^h \cdot \nu - \lambda_i \hat{q} \cdot \nu) \\ &= \int_{K^k} \operatorname{div} p_i^h - \int_{K^k} \operatorname{div} \hat{q} \lambda_i - \int_{K^k} \hat{q} \cdot \nabla \lambda_i \end{aligned}$$

and, because of the definition p_i^h (see Figure 3),

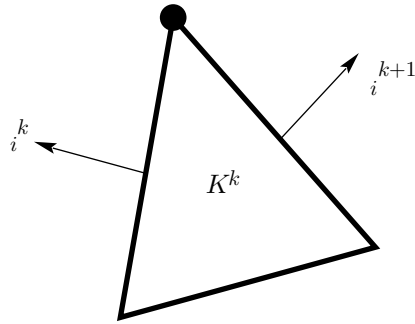


FIGURE 3. A triangle K^k of \mathcal{C}_i^h

$$X_i^{k+1} - X_i^k = \int_{K^k} (\nabla u^h - \hat{q}) \bullet \nabla \lambda_i - \int_{K^k} (f + \operatorname{div} \hat{q}) \lambda_i,$$

which is bounded by

$$c[h|u|_{2,K^k} + |u - u^h|_{1,K^k} + [\|f\|_{0,K^k} + |u|_{2K^k}] \|\lambda_i\|_{0,K^k}]$$

or else

$$(29) \quad X_i^{k+1} - X_i^k \leq ch(\|f\|_{0,K^k} + |u|_{2,K^k}) + |u - u^h|_{1,K^k}.$$

But from the definition of the coefficient α_i in the expression of p_i^h one has (see (22)):

$$\begin{aligned} X_i^1 &= \int_{\gamma_i^1} (p_i^h \bullet \nu - \lambda_i \hat{q} \bullet \nu) \\ &= \alpha_i + \int_{\gamma_i^1} p_i^h \bullet \nu - \frac{1}{2} \int_{\gamma_i^1} \frac{\partial u^h}{\partial \nu} + \frac{1}{2} \int_{\gamma_i^1} \left(\frac{\partial u^h}{\partial \nu} - \hat{q} \bullet \nu \right) \\ &= \frac{1}{2} \int_{\gamma_i^1} \left(\frac{\partial u^h}{\partial \nu} - \hat{q} \bullet \nu \right). \end{aligned}$$

Let us denote by ξ the second degree polynomial function on the triangle K_i^1 equal to 1 on the middle of γ_i^1 and zero on the two other sides. Then, setting $(\int_{\gamma_i^1} \xi = \operatorname{meas}(\gamma_i^1) \frac{2}{3})$, we have

$$X_i^1 = \frac{3}{4} \int_{\gamma_i^1} \left(\frac{\partial u^h}{\partial \nu} - \hat{q} \bullet \nu \right) \xi,$$

and, from the Stokes formula,

$$\begin{aligned} X_i^1 &= \int_{K^1} (\nabla u^h - \hat{q}) \bullet \nabla \xi - \int_{K^1} \operatorname{div} \hat{q} \xi \\ &\leq c(|u^h - u|_{1,K^1} + \|\nabla u - \hat{q}\|_{0,K^1} + h\|\operatorname{div} \hat{q}\|_{0,K^1}). \end{aligned}$$

Finally, the inequalities

$$\|\nabla u - \hat{q}\|_{0,K^1} \leq ch|u|_{2,K^1}, \quad \|\operatorname{div} \hat{q}\|_{0,K^1} \leq c|u|_{2,K^1}$$

enable one to obtain the estimate

$$(30) \quad |X_i^1| \leq ch[\|f\|_{0,K^1} + |u|_{2,K^1}] + |u - u^h|_{1,K^1},$$

and from (29) we obtain

$$(31) \quad |X_i^k| \leq ch[\|f\|_{0,C_i^h} + |u|_{2,C_i^h}] + c|u - u^h|_{1,C_i^h}$$

for all $k = 1, \dots, \kappa$ (all the sides γ_i^k of the cluster C_i^h with S_i as an extremity).

The proof of Theorem 3 is then a consequence of (5), (27), (31), and Lemmas 1 and 3 in the Appendix. Obviously it requires that the number of triangles in a cluster must be bounded above. \square

Remark. As we proved in Theorem 2 that

$$|\delta u_K|_{1,K} \leq ch^2 \|f\|_{1,K},$$

it can be suggested that if f is smooth enough, this term can be omitted. \square

Remark. When δu_K must be computed, it is interesting to use a subgrid on K . We point out that this computation is highly parallel, or can even be vectorized. Therefore, the computational time is very much reduced. \square

6. NUMERICAL TESTS

Let us now suppose that u^h is a solution of the classical conforming finite element method defined by (3). Let us recall that in order to apply the Prager-Syngé relation, one has to construct a vector field p lying in the set $H_f(\text{div}, \Omega)$. The Raviart-Thomas finite element is used. For clarity, we recall briefly the basic principles of our strategy.

At each node S_i of a mesh \mathcal{T}^h , we define the cluster of elements C_i^h , which is the union of elements having S_i as a vertex.

Then for each S_i (even on the boundary of Ω), we set

$$(32) \quad \begin{cases} p_i^h \in H_{\text{RT1}}(\text{div}, C_i^h), \\ \text{div } p_i^h = \frac{1}{|K|} \left\{ \int_K \nabla u^h \bullet \nabla \lambda_i - \int_K f \lambda_i \right\}, \\ \forall K \in C_i^h \end{cases}$$

(λ_i is the continuous piecewise linear function equal to one at S_i and 0 at all the other nodes). Here we have put

$$H_{\text{RT1}}(\text{div}, C_i^h) = \left\{ p \in H(\text{div}, C_i^h), p \bullet \nu = 0 \text{ on } \partial C_i^h \right. \\ \left. \text{and } \forall K \in C_i^h \ p|_K = \begin{pmatrix} a_K + b_K x \\ c_K + b_K y \end{pmatrix} \right\}.$$

The existence of a solution to (32) has been proved in §4. But the solution is not unique. More precisely, we proved that the general solution is

$$p_i^h = \mathbf{p}_i^h + \alpha_i \text{rot } \lambda_i,$$

where \mathbf{p}_i^h is a particular solution of (32) and α_i is an arbitrary constant. We set

$$(33) \quad p^h = \sum_{i \in \mathcal{S}^h} p_i^h = \sum_{i \in \mathcal{S}^h} \mathbf{p}_i^h + \sum_{i \in \mathcal{S}^h} \alpha_i \text{rot } \lambda_i,$$

where \mathcal{S}^h denotes the set of all the nodes of \mathcal{T}^h (including those on the boundary).

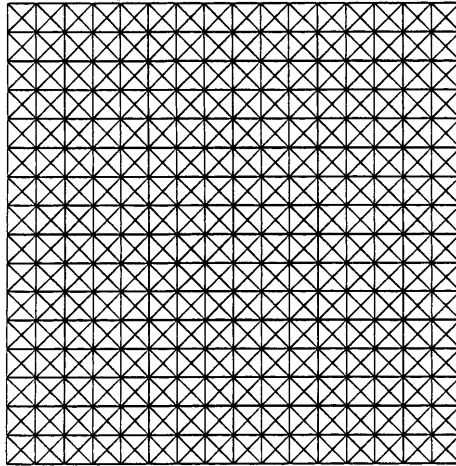


FIGURE 4.1. Examples of a regular mesh

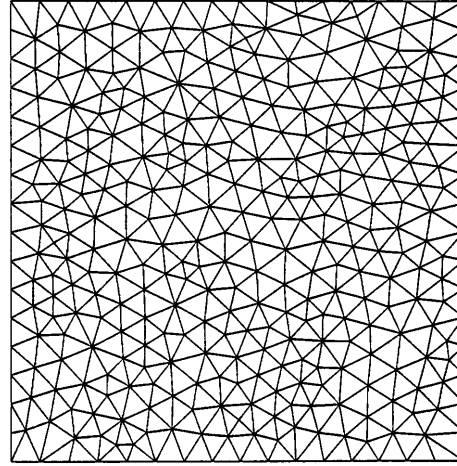


FIGURE 4.2. Examples of a mesh obtained by a mesh-generator

A nice choice for the coefficients α_i is obtained by minimizing the error bound (assuming that $\operatorname{div} p^h + f = 0$, or else that f is piecewise constant, for simplicity):

$$(34) \quad \alpha \in \mathbb{R}^L \rightarrow \left\| p^h + \sum_{i \in \mathcal{S}^h} \alpha_i \operatorname{rot} \lambda_i - \nabla u^h \right\|_{0,\Omega},$$

where $L = \operatorname{card}(\mathcal{S}^h)$.

Two strategies can be then discussed. One consists in replacing (34) by a local minimization (one iteration of the Jacobi algorithm, for instance, even if the matrix is not diagonal dominant). The second one is more well founded, and it consists in adding to the former one iteration of the SSOR algorithm. These two strategies have been checked on the test model presented in this paper. One can see that the second one is more reliable for irregular meshes (see Figures 5 and 6). As the additional cost is negligible, it has to be recommended for general applications.

The open set used is a square and two different kinds of meshes are used. They are represented in Figures 4.1 and 4.2.

In order to compare the method described here and the error indicator strategy of Bernardi, Métivet and Verfürth [7], we have plotted this quantity (denoted by B^h) in Figures 5 and 6. Let us recall that it is defined by

$$(35) \quad B^h = \left[\sum_{K \in \mathcal{T}^h} \left\{ h_K \|f + \Delta u^h\|_{0,K} + \frac{1}{2} \sum_{\gamma \subset \partial K} |\gamma|^{1/2} \left\| \left[\frac{\partial u^h}{\partial \nu} \right] \right\|_{0,\gamma} \right\}^2 \right]^{1/2}.$$

The indicator B^h is larger (a ratio of 6 with the exact error instead of 1.3 for the method developed here). Another advantage of our error bound is that one can improve the approximation of u by a local minimization problem, for instance, by adding degrees of freedom on the sides between elements. For example, we have added degrees of freedom at the midpoint of the inner edges of the meshes. The

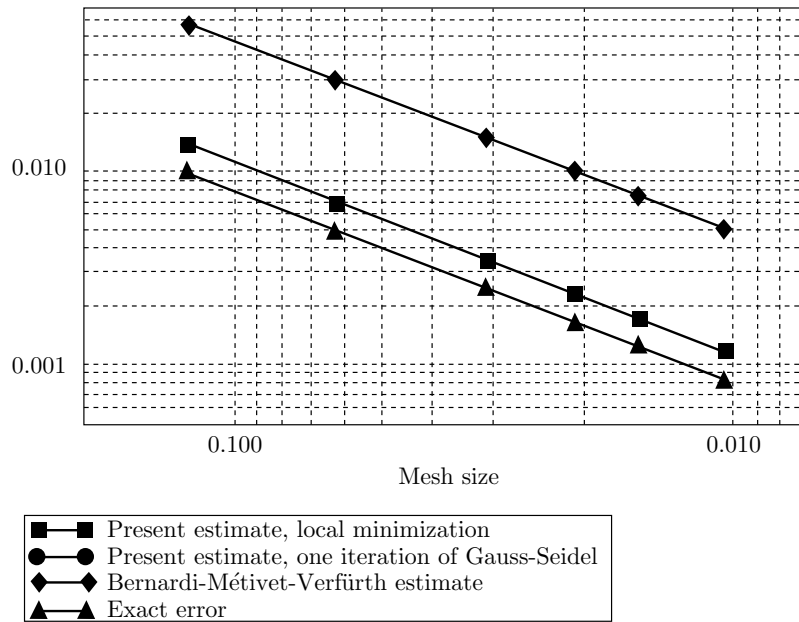


FIGURE 5. The error bounds, the exact errors, and the error indicators for regular meshes

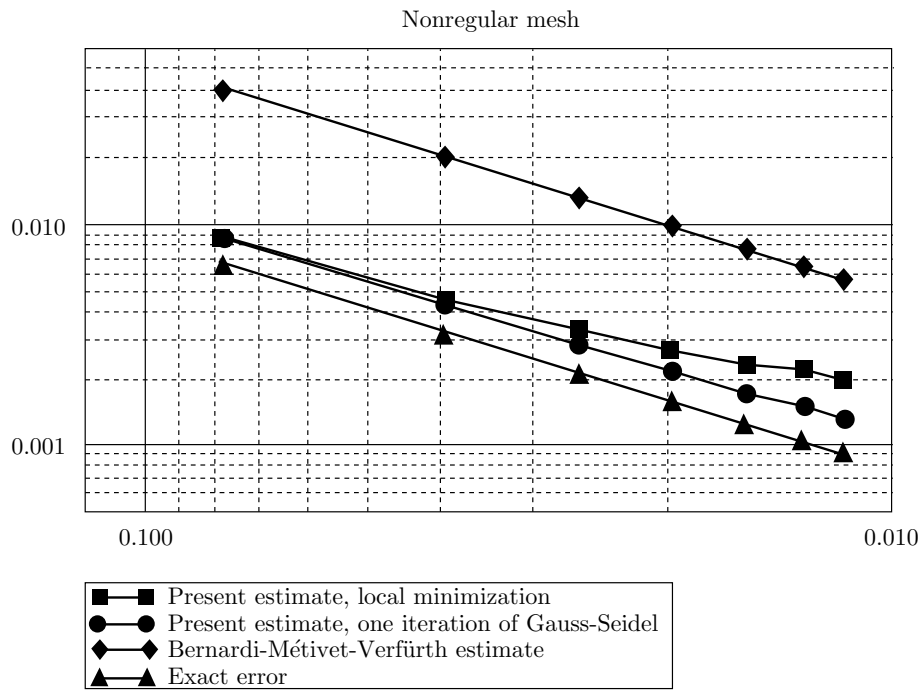


FIGURE 6. The error bounds, the exact errors, and the error indicators for irregular meshes

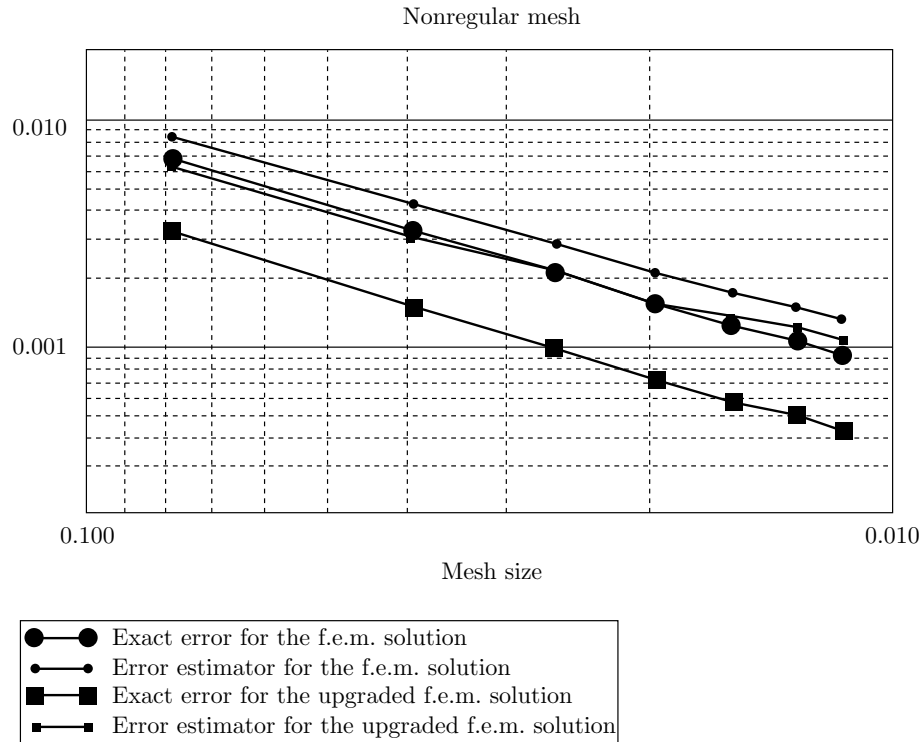


FIGURE 7. Local upgrade of the finite element solution using mid-point degrees of freedom

error estimator obtained with this upgraded solution has been plotted in Figure 7. We have also represented the exact error between the solution u and this new term. The former results have been recalled in order to evaluate the improvement due to this trick.

7. CONCLUSION

The method that we have developed in this paper is a new strategy for explicitly bounding the error in a finite element method using conforming approximation. The extension to elliptical problems does not require any new tricks. For instance, the case of 2D-elasticity can be handled. The difficulty is then to construct a symmetrical equilibrium finite element. The way we know consists in replacing a Raviart-Thomas element by one of the family suggested by Arnold, Douglas and Gupta in [3]. In another respect the contribution to the error ε_K defined at (24) can be used as an error indicator in an automatic mesh refinement algorithm. The restricted numerical discussion given in §6 enables one to observe a few of the advantages of the method developed in this paper. A more extensive presentation of the numerical tests is given by Destuynder, Collot and Salaün in [13], and also in M. Collot's thesis [9], where the extension of the method to adaptive mesh refinements is discussed.

APPENDIX

In this paper we used several classical results which are quite well known. They are recalled here.

Lemma 1. *Let K be a triangle of a mesh family assumed to be regular (see Ciarlet [8], and Girault and Raviart [16]). Then let h be the maximum length of the sides of K . For any function $\varphi \in H^1(K)$, satisfying*

$$\int_K \varphi = 0,$$

there exists a constant (say c) which is independent of both h and φ and such that

$$\text{i) } \|\varphi\|_{0,K} \leq ch|\varphi|_{1,K}. \quad \square$$

Lemma 2 (Same hypothesis as in Lemma 1). *For all $\varphi \in L^2(K)$ and $K \in \mathcal{T}^h$, let $\Pi_0\varphi$ be defined by*

$$\Pi_0\varphi = \frac{1}{\text{meas}(K)} \int_K \varphi,$$

then there exists a constant c which is independent of both h and φ and such that

$$\|\varphi - \Pi_0\varphi\|_{0,K} \leq ch|\varphi|_{1,K}. \quad \square$$

Lemma 3. (Same hypothesis as in Lemma 1 but we also assume that the family of triangulations is uniformly regular, as described by Girault and Raviart in [16]). *There exists a constant c which is independent of both h and φ and such that*

$$\|\varphi\|_{0,\partial K} \leq c\sqrt{h}|\varphi|_{1,K}. \quad \square$$

REFERENCES

1. M. Ainsworth and J. T. Oden [1993], *A unified approach to a posteriori error estimation using element residual methods*, Numer. Math. **65**, 23–50. MR **95a**:65185
2. R. Arcangéli and J. L. Gout [1976], *Sur l'évaluation de l'erreur d'interpolation de Lagrange dans un ouvert de \mathbb{R}^n* , Rev. Française Automat. Inform. Rech. Opér. Sér. Rouge Anal. Numér. **10**, 5–27. MR **58**:29586 (preprint)
3. D. Arnold, J. Douglas, and C. P. Gupta [1984], *A family of higher order mixed finite element methods for plane elasticity*, Numer. Math. **45**, 1–22. MR **86a**:65112
4. I. Babuška and W. C. Rheinboldt [1978], *Error estimates for adaptive finite element computation*, SIAM J. Numer. Anal. **15**, 736–754. MR **58**:3400
5. I. Babuška and R. Rodriguez [1993], *The problem of the selection of an a posteriori error indicator based on smoothing techniques*, Internat. Numer. Methods Engrg. **36**, 539–567. MR **93k**:65089
6. R. E. Bank and A. Weiser [1985], *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comp. **44**, 283–301. MR **86g**:65207
7. C. Bernardi, B. Métivet and R. Verfürth [1993], *Analyse numérique d'indicateurs d'erreur*, note EDF-M172/93062.
8. P. G. Ciarlet [1978], *The finite element method for elliptic problems*, North-Holland, Amsterdam. MR **58**:25001
9. M. Collot [1998], *Analysis and implementation of explicit error bounds for finite element methods; Static and dynamic aspects*, CNAM, Paris.
10. Ph. Destuynder and B. Métivet [1998], *Explicit error bounds for a non-conforming finite element method*, Finite Element Methods (Jyväskylä, 1997), Lecture Notes Pure Appl. Math., vol. 196, Marcel Dekker, New York, 1998, pp. 95–111. CMP 98:08
11. ——— [1996a], *Estimation de l'erreur explicite pour une méthode d'éléments finis non conformes*, C. R. Acad. Sci. Paris Sér. I. Math. **322**, 1081–1086. MR **97e**:65110
12. ——— [1996b], *Estimation de l'erreur explicite pour une méthode d'éléments finis conformes*, C. R. Acad. Sci. Paris Sér. I. Math. **323**, 679–684. MR **97h**:65138

13. Ph. Destuynder, M. Collot, and M. Salaün [1997], *A mathematical framework for the P. Ladevèze a posteriori error bounds in finite element methods*, New Advances in Adaptive Computational Methods in Mechanics (T. Oden and P. Ladevèze, eds.), Elsevier, Amsterdam. (to appear)
14. K. Eriksson and C. Johnson [1988], *An adaptive finite element method for linear elliptic problems*, Math. Comp. **50** 361–383. MR **89c**:65119
15. L. Gallimard, P. Ladeveze, and J. P. Pelle [1995], *La méthode des erreurs en relation de comportement appliquée au contrôle des calculs non linéaires*. École CEA EDF-INRIA. Cours du 18–21 septembre 1995 sur le calcul d’erreur a posteriori et adaptation de maillage.
16. V. Girault and P. A. Raviart [1986], *Finite element methods for Navier-Stokes equations. Theory and algorithms*, Springer-Verlag, Berlin. MR **88b**:65129
17. P. Ladevèze [1975], *Comparaison de modèles de mécanique des milieux continus*. Thèse d’état, Université Paris VI, Paris.
18. P. Ladevèze and D. Leguillon [1983], *Error estimate procedure in the finite element method and applications*, SIAM J. Numer. Anal. **20**, 485–509. MR **84g**:65150
19. J. T. Oden, L. Demkowicz, W. Rachowicz, and T. A. Westermann [1989], *Toward a universal h-p adaptive finite element strategy. Part 2. A posteriori error estimation*. Comp. Methods Appl. Mech. Engrg. **77**, 113–180. MR **91b**:65123b
20. W. Prager and J. L. Synge [1947], *Approximations in elasticity based on the concept of function space*, Quart. Appl. Math. **5**, 241–269. MR **10**:81b
21. P. A. Raviart and J. M. Thomas [1975], *A mixed finite element method for second order elliptic problems*, Mathematical Aspects of Finite Element Methods (Proc. Conf., C.N.R., Rome, 1975) Lecture Notes in Math., vol. 60b, Springer-Verlag, Berlin, 1977, pp. 292–315. MR **58**:3547
22. ——— [1983], *Introduction à l’analyse numérique des équations aux dérivées partielles*, Masson, Paris. MR **87a**:65001a
23. J. E. Roberts and J. M. Thomas [1987], *Mixed and hybrid finite element methods*, Rapport INRIA n°737.
24. A. H. Schatz, L. H. Sloan, and L. B. Wahlbin [1996], *Superconvergence in finite element methods and meshes that are locally symmetric with respect to a point*, SIAM J. Numer. Anal. **33**, 505–521. MR **98f**:65112
25. R. Verfürth [1994], *A posteriori error estimation and adaptive mesh-refinement techniques*, J. Comput. Appl. Math. **50**, 67–83. MR **95c**:65171
26. O. C. Zienkiewicz and J. Z. Zhu [1987], *A simple error estimator and adaptive procedure for practical engineering analysis*, Internat. J. Numer. Methods Engrg. **24**, 337–357. MR **87m**:73055
27. ——— [1992], *The super convergent patch recovery and a posteriori error estimates. Part 1: The recovery technique*, Internat. J. Numer. Methods Engrg. **33** 1331–1364. MR **93c**:73098

CNAM/IAT, 15 RUE MARAT, 78210 SAINT-CYR-L’ÉCOLE, FRANCE

E-mail address: `destuynd@cnam.fr`

1 AVENUE DU GÉNÉRAL DE GAULLE, 92141 CLAMART, FRANCE

E-mail address: `brigitte.metivet@der.edfgdf.fr`