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:

\[
\begin{aligned}
\text{find } u \in H^1_0(\Omega) \text{ such that, for all } v \in H^1_0(\Omega), \\
\int_\Omega \nabla u \cdot \nabla v = \int_\Omega f v,
\end{aligned}
\]

where \( f \) is a function in the space \( L^2(\Omega) \) and \( \nabla \) 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 \( H^2(\Omega) \cap H^1_0(\Omega) \) (no re-entrant angle\(^\ast\) on the boundary, which should be piecewise smooth enough).

Let us now consider a family of triangulations of \( \Omega \), assumed to be uniformly regular (see Girault and Raviart [16]). One triangulation is denoted by \( T^h \), where \( h \) denotes the size of the mesh. The approximation space of \( H^1_0(\Omega) \), based on the triangulation \( T^h \), is denoted by \( V^h \) and is, for instance, defined by

\[
V^h = \{ v \in H^1_0(\Omega), \forall K \in T^h, v|_K \in P_1(K) \},
\]

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

\[
\begin{aligned}
\text{find } u^h \in V^h \text{ such that for all } v \in V^h \\
\int_\Omega \nabla u^h \cdot \nabla v = \int_\Omega f v.
\end{aligned}
\]

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

\[
|u - u^h|_{1,\Omega} \leq \inf_{v \in V^h} |v - u|_{1,\Omega},
\]

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

\( \ast \) As a matter of fact, the internal angles on the boundary should be smaller than \( \pi \).
is independent of both $h$ and $u$, and such that
\begin{equation}
|u - u^h|_{1,\Omega} \leq ch|u|_{2,\Omega}.
\end{equation}

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 $\Omega$:
\begin{equation}
H_f(\text{div}, \Omega) = \{ p \in (L^2(\Omega))^2, \text{div} \ p + f = 0 \text{ in } \Omega \}.
\end{equation}
Then one has the inequality, used first by Ladevèze [17],
\begin{equation}
|u - u^h|_{1,\Omega} \leq \inf_{p \in H_f(\text{div}, \Omega)} \| p - \nabla u^h \|_{0,\Omega},
\end{equation}
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
\begin{equation}
|u - v|^2_{1,\Omega} + \| p - \nabla u^h \|^2_{0,\Omega} = \| p - \nabla v \|^2_{0,\Omega}.
\end{equation}
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 $|\nabla v - \nabla u^h|_{1,\Omega}$ will be as small as possible. The space $V^h$ is a conforming approximation of $H^1_0(\Omega)$. 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
\begin{equation}
p \to \frac{1}{2} \int_{\Omega} |p|^2.
\end{equation}
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
\begin{equation}
\text{div} \ p + f = 0 \text{ in } \Omega.
\end{equation}
As a matter of fact it is, for instance, replaced by
\begin{equation}
\text{div} \ p^h + \frac{1}{|K|} \int_K f = 0, \quad \forall K \in T^h
\end{equation}(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 $\delta p_K$, defined by

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

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 Walhbin [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

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

and \( p^h \in H(\text{div}, \Omega) \). The next step consists in finding a solution \( \delta u_K \) of (10) and in proving that \( \delta u_K \) 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 \( T^h \), is itself bounded by \( O(h) \).

3. Properties of the conforming finite element solution

Let us denote by \( S^h \) the set of all the internal vertices of \( T^h \). For each vertex \( S_i \) we introduce the basis function \( \lambda_i \) 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

\[
\int_{C^h_i} \nabla u^h \cdot \nabla \lambda_i - \int_{C^h_i} f \lambda_i = 0,
\]

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

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

\[
\sum_{k=1}^{\kappa} \frac{\text{meas}(\gamma^k_i)}{2} \left[ \left( \frac{\partial u^h}{\partial n} \right)^k - \int_{K^k} f \lambda_i \right] = 0,
\]

where \([ \cdot ]^k_i \) is the jump of a quantity across the side \( \gamma^k_i \). This relation can be interpreted from a mechanical (for instance) point of view. The first terms represent the moment of \( \nabla u^h \) at the vertex \( S_i \) and along the side \( \gamma^k_i \). 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 \( C^h_i \) the moments at the vertex \( S_i \) and...
to ensure the equilibrium between two neighbour elements. Hence we look for \( p^h \) such that

\[
\begin{align*}
  \left\{ \begin{array}{l}
    p_i^h \in H(\text{div}, C_i^h), \\
    \text{div } p_i^h = \frac{1}{|K|} \left\{ \int_K \nabla u^h \cdot \nabla \lambda_i - \int_K f \lambda_i \right\}, \\
    p_i^h \cdot \nu = 0 \text{ on } \partial K \cap \partial C_i^h,
  \end{array} \right.
\end{align*}
\]

which implies that an assumed solution satisfies, for all \( K \in C_i^h \),

\[
\int_{\partial K} p_i^h \cdot \nu = \int_K \nabla u^h \cdot \nabla \lambda_i - \int_K f \lambda_i
\]

\[
= \int_{\gamma_i^1} \frac{\partial u^h}{\partial \nu} \lambda_i + \int_{\gamma_i^2} \frac{\partial u^h}{\partial \nu} \lambda_i - \int_K f \lambda_i
\]

\[
= \frac{\text{meas}(\gamma_i^1)}{2} \left( \frac{\partial u^h}{\partial \nu} \right) + \frac{\text{meas}(\gamma_i^2)}{2} \left( \frac{\partial u^h}{\partial \nu} \right) - \int_K f \lambda_i
\]

(\( \gamma_i^1 \) and \( \gamma_i^2 \) 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 \( C_i^h \) for the test variable \( \lambda_i \). The existence of \( p_i^h \) is proved in the next section.

4. Definition of an equilibrium vector field on \( \Omega \)

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

It is worth noting that the definition of the boundary \( \partial C_i^h \) in this second situation does not include the vertices which belong to the boundary \( \partial \Omega \) of \( \Omega \).

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

$$H_{RT1}(\text{div},C^h_i) = \left\{ p \in (L^2(C^h_i))^2, \; p \bullet \nu = 0 \; \text{on} \; \partial C^h_i, \; \forall K \in C^h_i, p|_K = \frac{1}{|K|} \int_K \nabla u^h \bullet \nabla \lambda_i - \int_K f \lambda_i \right\},$$

for internal point $h$.

$$C^h_i$$

for a boundary point $h$.

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

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

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,

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

where $p^h_i$ is a particular solution of (14).

**Proof.** a) Let us begin with the case where $S_i$ is an internal vertex of the triangulation $T^h$. 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 $C^h_i$). Let us therefore analyze the homogeneous system associated to (14). This problem consists in finding an element $\delta^h_i$ in $H_{RT1}(\text{div},C^h_i)$ such that for all $K \in C^h_i$,

$$\text{div } \delta^h_i = 0 \; \text{in } K. \tag{15}$$
Because the cluster $\mathcal{C}^h_i$ is a simply connected open set, there exists a function $\psi_i$ such that

$$\delta^h_i = \text{rot} \psi_i, \quad \psi_i \in H^1(\mathcal{C}^h_i).$$

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

$$\text{rot} \psi_i = \left| \begin{array}{c} a_K + b_K x \\ c_K + b_K y \end{array} \right| = \left| \begin{array}{c} a_K \\ c_K \end{array} \right|,$$

($b_K$ is zero because $\text{div} \delta^h_i = 0$). Finally, $\psi_i$ is piecewise linear and therefore is proportional to the basis function $\lambda_i$. 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 $\Omega$, 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 $X = (X_j), i = 1, \ldots, \kappa,$ of the cokernel satisfies

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

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

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

or else

$$\sum_{j=1}^\kappa \int_{\gamma_j} \left( \frac{X_j}{|K_j|} - \frac{X_{j-1}}{|K_{j-1}|} \right) q \cdot \nu_{|\gamma_j} = 0,$$

where $\gamma_j$ denotes the sides of $\mathcal{C}^h_i$ that have $S_i$ as an extremity, as shown in Figure 2. Thus the quantity

$$\xi = \frac{X_j}{|K_j|}, \quad j = 1, \ldots, \kappa,$$

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
\[ \int_{C_h} \nabla u_h \cdot \nabla \lambda_i = \int_{C_h} f \lambda_i, \]
(16)
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 \( p^h_i \) defined in Theorem 1 we introduce the term
\[ p^h = \sum_{i \in \mathcal{S}^h} p^h_i, \]
(17)
where \( \mathcal{S}^h \) is the set of all vertices of triangulation \( T^h \), including those on the boundary of \( \Omega \) (\( \mathcal{S}^h \) is restricted to the internal points). From the definition of \( p^h_i \), and because \( p^h_i = 0 \) on \( \Omega - C^h \), one has
\[ \text{div} p^h = \sum_{i \in \mathcal{S}^h} \text{div} p^h_i = \sum_{i \in \mathcal{S}^h} \sum_{K \in C^h_i} \frac{1}{|K|} \left\{ \int_K \nabla u_h \cdot \nabla \lambda_i - \int_K f \lambda_i \right\}, \]
and because on each triangle \( K \) we have \( \sum_{i \in \mathcal{S}^h} \lambda_i = 1 \), we conclude that, for all \( K \in T^h \),
\[ \text{div} p^h + \frac{1}{|K|} \int_K f = 0. \]
(18)
As the element \( p^h \) defined in (18) is not in the set \( H_f(\text{div}, \Omega) \), we add a local term \( \delta u_K \) (defined on each triangle of \( T^h \)) such that
\[ \begin{cases} 
\delta p_K = \nabla \delta u_K & \text{on } K, \\
-\Delta \delta u_K = f - \frac{1}{|K|} \int_K f & \text{on } K, \\
\frac{\partial \delta u_K}{\partial \nu} = 0 & \text{on } \partial K \quad \text{and} \quad \int_K \delta u_K = 0.
\end{cases} \]
(19)
The existence and uniqueness of a solution to (19) is very classical, and finally we set, on each \( K \) of \( T^h \),
\[ p = p^h + \delta p_K. \]
(20)
It is worth noting that \( \delta p_K \in H(\text{div}, \Omega) \), because of the homogeneous Neumann boundary condition that we chose on \( \partial K \). Then a simple compilation of the previous results shows that
\[ \text{div} p + f = 0 \quad \text{on } \Omega. \]
(21)
As a matter of fact, the term \( \delta u_K \) 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 \( \delta u_K \) 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,

\[
|\delta u_K|_{1,K} \leq c h^2\|f\|_{1,K}.
\]

**Proof.** From the definition of \( \delta u_K \), and letting \( \Pi_0 \) denote the \( L^2(K) \) projection onto the constants, one obtains

\[
|\delta u_K|_{1,K}^2 = - \int_K \Delta(\delta u_K)\delta u_K = \int_K \left( f - \frac{1}{|K|} \int_K f \right) \delta u_K 
\]

\[
\leq \|f - \Pi_0 f\|_{0,K}\|\delta u_K\|_{0,K},
\]

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

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

This completes the proof of Theorem 2. \( \square \)

**Remark.** In the definition of \( p_h^i \) the coefficient \( \alpha_i \) (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 \( \gamma_i^1 \), one has

\[
\alpha_i \int_{\gamma_i^1} \text{rot} \lambda_i \cdot \nu + \int_{\gamma_i^1} p_h^i \cdot \nu - \int_{\gamma_i^1} \frac{1}{2} \frac{\partial u_h}{\partial \nu} 
\]

\[
= \alpha_i + \int_{\gamma_i^1} \left( p_h^i \cdot \nu - \frac{1}{2} \frac{\partial u_h}{\partial \nu} \right),
\]

and we can choose \( \alpha_i \) such that this quantity is zero. Hence

\[
\alpha_i = - \int_{\gamma_i^1} \left( p_h^i \cdot \nu - \frac{1}{2} \frac{\partial u_h}{\partial \nu} \right)
\]

\[
= - \int_{\gamma_i^1} \left( p_h^i \nu - \frac{\partial u_h}{\partial \nu} \lambda_i \right)
\]

\[
= - \text{meas}(\gamma_i^1) \left[ p_h^i \cdot \nu - \frac{1}{2} \frac{\partial u_h}{\partial \nu} \right]_{\gamma_i^1}.
\]

We shall prove below that such a choice leads to a consistent error bound. \( \square \)

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 \( \varepsilon \) 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

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

Proof. First of all, on each element $K$ of the triangulation $T^h$ 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 $\partial K$ of $K$, $k$ satisfies

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

Therefore, we can deduce that there exists a function $\varphi_K$ 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 $\varphi_K$ is a solution of

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

(23)

The previous model defines $\varphi_K$ uniquely. But one also has

$$\varepsilon^2_K \overset{\text{def}}{=} \int_K |p^h + \nabla \delta u_K - \nabla u^h|^2 = |\varphi_K|_{1,K}^2$$

and, from (23),

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

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

$$\varepsilon^2_K = \int_K f \varphi_K + \sum_{i \in S_K} \int_{\partial K} p^h_i \nu \varphi_K - \int_{\partial K} \frac{\partial u}{\partial \nu} \varphi_K.$$  

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

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

where $\gamma$ 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)

$$\begin{cases}
\|\hat{q} - \nabla u\|_{0,K} \leq c h |u|_{2,K}, & \\
\|\text{div } \hat{q} - \text{div}(\nabla u)\|_{0,K} = \|\text{div } \hat{q} + f\|_{0,K} \leq c h \|f\|_{0,K} + |u|_{2,K}.
\end{cases}$$

(25)
Then one obtains the following equality:
\[
\varepsilon_K^2 = \int_K f \phi_K + \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \varphi_i \hat{q} \cdot \nu) \phi_K
\]
\[+ \int_{\partial K} \left( \hat{q} \cdot \nu - \frac{\partial u^h}{\partial \nu} \right) \phi_K,
\]
or else, using the Stokes formula \((i_1, i_2, i_3)\) are the three vertices of \(K\),
\[
\varepsilon_K^2 = \int_K f \phi_K + \int_K (\hat{q} - \nabla u^h) \nabla \phi_K + \int_K \text{div} \, \hat{q} \phi_K
\]
\[+ \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \varphi_i \hat{q} \cdot \nu) \phi_K.
\]
Hence
\[
\varepsilon_K^2 \leq \|f + \text{div} \, \hat{q}\|_{0,K} \|\phi_K\|_{0,K} + \|\hat{q} - \nabla u^h\|_{0,K} \|\phi_K\|_{1,K}
\]
\[+ \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \varphi_i \hat{q} \cdot \nu) \phi_K.
\]
From the triangular inequality and Lemma 1 in the Appendix, we deduce that
\[
\varepsilon_K^2 \leq ch \|f\|_{0,K} + \|u\|_{1,K} + \|u - u^h\|_{1,K} \|\phi_K\|_{1,K}
\]
\[+ \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \varphi_i \hat{q} \cdot \nu) \phi_K.
\]
But on each side \(\gamma\) of \(\partial K\) only two terms \(p_i^h\) (for \(i = i_1\) and \(i_2\), for instance) are different from zero. Therefore
\[
\left| \sum_{i \in \{i_1, i_2, i_3\}} \int_{\partial K} (p_i^h \cdot \nu - \lambda_i \varphi_i \hat{q} \cdot \nu) \phi_K \right|
\]
\[= \sum_{\gamma \subseteq \partial K} \left\{ \sum_{i \in \{i_1, i_2\}} \int_{\gamma} (p_i^h \cdot \nu - \frac{1}{2} \hat{q} \cdot \nu) \phi_K \right\}
\]
\[\leq 2 \sum_{\gamma \subseteq \partial K} \left\{ |p_i^h \nu - \frac{1}{2} \hat{q} \cdot \nu|_{\gamma} \sqrt{\text{meas}(\gamma)} \|\phi_K\|_{0,\gamma} \right\},
\]
where \(\gamma_i^h\) is one side of the cluster \(C_i^h\) with the center \(S_i\) as an extremity. First of all, if we define
\[
X_i^k \triangleq \int_{\gamma_i^h} (p_i^h \cdot \nu - \lambda_i \varphi_i \hat{q} \cdot \nu) = \text{meas}(\gamma_i^h)[r_i^h \cdot \nu - \frac{1}{2} \hat{q} \cdot \nu]|_{\gamma_i^h},
\]
then
\[
X_i^{k+1} - X_i^k = \int_{\gamma_i^{k+1}} (p_i^h \cdot \nu - \lambda_i \varphi_i \hat{q} \cdot \nu) - \int_{\gamma_i^k} (p_i^h \cdot \nu - \lambda_i \varphi_i \hat{q} \cdot \nu)
\]
\[= \int_{K^h} \text{div} \, p_i^h - \int_{K^h} \text{div} \, \hat{q} \lambda_i - \int_{K^h} \hat{q} \cdot \nabla \lambda_i
\]
and, because of the definition \(p_i^h\) (see Figure 3),
Figure 3. A triangle $K^k$ of $C_i^h$

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

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

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

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

$$X_1^1 = \int_{\gamma_1^1} (p_1^h \cdot \nu - \lambda_i \hat{q} \cdot \nu)$$

$$= \alpha_i + \int_{\gamma_1^1} p_1^h \cdot \nu - \frac{1}{2} \int_{\gamma_1^1} \frac{\partial u^h}{\partial \nu} + \frac{1}{2} \int_{\gamma_1^1} \left( \frac{\partial u^h}{\partial \nu} - \hat{q} \cdot \nu \right)$$

$$= \frac{1}{2} \int_{\gamma_1^1} \left( \frac{\partial u^h}{\partial \nu} - \hat{q} \cdot \nu \right).$$

Let us denote by $\xi$ the second degree polynomial function on the triangle $K_1^1$ equal to 1 on the middle of $\gamma_1^1$ and zero on the two other sides. Then, setting $(f, \xi = \text{mes}(\gamma_1^1)^\frac{1}{2})$, we have

$$X_1^1 = \frac{3}{4} \int_{\gamma_1^1} \left( \frac{\partial u^h}{\partial \nu} - \hat{q} \cdot \nu \right) \xi,$$

and, from the Stokes formula,

$$X_1^1 = \int_{K^1} (\nabla u^h - \hat{q}) \cdot \nabla \xi - \int_{K^1} \text{div} \hat{q} \xi$$

$$\leq c(|u^h - u|_{1,K^1} + ||\nabla u - \hat{q}||_{0,K^1} + h||\text{div} \hat{q}||_{0,K^1}).$$

Finally, the inequalities

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

enable one to obtain the estimate

$$(30) \quad |X_1^1| \leq ch(||f||_{0,K^1} + |u|_{2,K^1}) + |u - u^h|_{1,K^1},$$
and from (29) we obtain

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

for all \(k = 1, \ldots, \kappa\) (all the sides \(\gamma^k_i\) 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.

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.

Remark. When \(\delta u_K\) 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.

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-Synge 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 \(T^h\), 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 \(\Omega\)), we set

\[\begin{cases} p_i^h \in H_{\text{RT}1}(\text{div}, C_i^h), \\
\text{div} p_i^h = \frac{1}{|K|} \left\{ \int_K \nabla u^h \cdot \nabla \lambda_i - \int_K f \lambda_i \right\}, \\
\forall K \in C_i^h \end{cases}\]

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

\[H_{\text{RT}1}(\text{div}, C_i^h) = \left\{ p \in H(\text{div}, C_i^h), p \bullet \nu = 0 \quad \text{on } \partial C_i^h \right\} \quad \text{and} \quad \forall K \in C_i^h, p_{i|K} = \begin{bmatrix} a_K + b_{Kx} \\ c_K + b_{Ky} \end{bmatrix}.\]

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

\[p_i^h = p_i^h + \alpha_i \text{rot} \lambda_i,\]

where \(p_i^h\) is a particular solution of (32) and \(\alpha_i\) is an arbitrary constant. We set

\[p^h = \sum_{i \in S^h} p_i^h = \sum_{i \in S^h} p_i^h + \sum_{i \in S^h} \alpha_i \text{rot} \lambda_i,\]

where \(S^h\) denotes the set of all the nodes of \(T^h\) (including those on the boundary).
A nice choice for the coefficients $\alpha_i$ is obtained by minimizing the error bound (assuming that $\text{div} p^h + f = 0$, or else that $f$ is piecewise constant, for simplicity):

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

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

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, Mètivet 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

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

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.
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
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 $\varepsilon_K$ 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 Salain 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 $\Delta K$ 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 $\Delta K$. For any function $\varphi \in H^1(\Delta K)$, satisfying

$$\int_{\Delta K} \varphi = 0,$$

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

1) $\|\varphi\|_{0,\Delta K} \leq ch|\varphi|_{1,\Delta K}.$

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

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

then there exists a constant $c$ which is independent of both $h$ and $\varphi$ and such that

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

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 $\varphi$ and such that

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

REFERENCES


