PH. CLÉMENT

Approximation by finite element functions using local regularization

Revue française d’automatique, informatique, recherche opérationnelle. Analyse numérique, tome 9, n° R2 (1975), p. 77-84

<http://www.numdam.org/item?id=M2AN_1975__9_2_77_0>
Approximation by Finite Element Functions Using Local Regularization (*)

par Ph. Clément (1)

Communicated by P G Ciarlet

Abstract — The aim of this paper is to give an elementary proof of a theorem of approximation of Sobolev spaces $H^q(\Omega)$ by finite elements without using classical interpolation. The construction which we give here allows us in some cases to fit boundary conditions.

1. INTRODUCTION

The mathematical problem of approximation by finite element functions has been first studied by Goël ([2]) and Zlamal ([3]). In [4], Bramble and Zlamal give estimates for the error in terms of Sobolev norms, however their results are based on the existence and continuity of the interpolate; but the interpolate may not exist; for example if $u \in H^1(\Omega)$ where $\Omega$ is a two-dimensional domain, by Sobolev's imbedding theorem, the pointwise values of $u$ cannot be defined and consequently no interpolation is possible. In [5], Strang defines an approximation by considering interpolates of regularized functions.

The purpose of this paper is to give an elementary construction of an approximation based on «local» regularization; the results are slightly less restrictive than Strang's ones. As shown by an example, the construction may be modified to fit boundary conditions. Most of the basic tools are known in the literature and their proofs will not be reproduced here; however all details are contained in [1].

(*) This paper is drawn from a thesis presented at the Federal Institute of Technology Lausanne. The author wants to express his gratitude to Prof J Descloux for his suggestions and helpful assistance.

(1) Ecole polytechnique fédérale de Lausanne, Département de Mathématiques

Revue Française d'Automatique, Informatique et Recherche Opérationnelle n° août 1975, R-2
2. RESULTS

For the sake of concreteness and simplicity, we shall restrict ourselves to the case of triangular finite element subspaces described by the following situation.

Let $\Lambda$ be a closed bounded two-dimensional domain with polygonal boundary $\Gamma$. One considers a set $D$ of decompositions $D = \{ T_1, \ldots, T_n \}$ of $\Lambda$ in closed triangles $T_1$, $T_2$, $\ldots$, $T_n$ called « elements » such that: 1) $\Lambda = \bigcup_{j=1}^{n} T_j$, 2) two triangles $T_i$ and $T_j \in D$ are either disjoint, or have a vertex in common or have a side in common, 3) $n$ depends on $D$. To each $D = \{T_1, \ldots, T_n\} \in D$ is associated a set $\{ \varphi_1, \varphi_2, \ldots, \varphi_m \}$ of independent real functions defined on $\Lambda$ ($m$ depends on $D$) and to each $\varphi_i$ is associated a point $Q_i \in \Lambda$ called « node »; the $Q_i$'s are not necessarily distinct; let $S_1, \ldots, S_m$ be the supports of $\varphi_1, \ldots, \varphi_m$. $S_i$ is connected to $Q_i$ by the relation: $S_i = \bigcup_{Q_i \in T_j} T_j$. Furthermore to each $\varphi_i$ is associated a functional $\gamma_i : C^\omega(\Lambda) \to \mathbb{R}$ of the form $\sum_{|s|=i} a_{is} D^2 f(Q_i)$; $l_i$ is called the « order » of $\gamma_i$.

$$V = \left\{ \sum_{i=1}^{m} a_i \varphi_i : a_i \in \mathbb{R} \right\}$$

is the finite element space associated to $D$. For specific examples see [2], [3], [4], [5], [6], [7].

Let $\Phi \subset \Lambda$; $d(\Phi)$ is the diameter of $\Phi$; $\mu(\Phi)$ is the measure of $\Phi$; $H^q(\Phi)$ denotes the Sobolev space of square integrable functions on $\Phi$ possessing square integrable derivatives of order $\leq q$. For $u, v \in H^q(\Phi)$ and $w \in C^\omega(\Lambda)$ we define the scalar product and seminorms:

$$\langle u, v \rangle_\Phi = \int u v \quad |u|_{k,\Phi}^2 = \sum_{|s|=k} (D^k u, D^k u)_\Phi \quad k \leq q,$$

$$|w|_{k,\infty,\Phi} = \max_{|s|=k} \sup_{x \in \Phi} |D^k w(x)| ;$$

for a function $u \in H^q(T_j), j = 1, 2, \ldots, n$ we set $|u|_{k,\Lambda}^2 = \sum_{j=1}^{n} |u|_{k,T_j}^2 k \leq q$.

In the following $c$ will denote a generic constant independent of $D \in D$. We introduce the following hypotheses.

**H1.** For any $p \in \mathcal{P}_\rho$ (polynomials of degree $\leq \rho$), where $\rho$ is independent of $D \in D$, one has for each $T_j \in D$:

$$p(x) = \sum_{Q_i \in T_j} \gamma_i(p) \varphi_i(x) \quad x \in T_j.$$
The functions \( \phi_i \) are the basis functions associated with functionals \( \gamma_i \).

**H2.** If \( f \in C^\infty(\Lambda) \), for any \( \phi_i \) with \( l_i \leq \rho \) and any \( T_j \in D \), one has
\[
|\gamma_i(f)\phi_i|_{k,T_j} \leq c(d(T_j))^{l_i-k}(\mu(T_j))^{\frac{1}{2}} |f|_{l_i,\infty,q}, k \leq \rho + 1
\]

**H3.** For any \( T_i \) the number of \( S_i \supseteq T_j \) is bounded by \( c \).

**H4.** For any \( D \in \mathcal{D} \), all the angles of the \( T'_i \)'s are \( \geq c > 0 \).

Now we define the linear mapping \( \Pi : H^0(\Lambda) \to V \) by the following construction. Let \( u \in H^0(\Lambda) \); to each \( S_i \) we associate the polynomial \( p_i \in \mathcal{P}_\rho \) which is the best approximation of \( u \) with respect to the norm \( \cdot \|_0, s \), i.e.
\[
(u - p_i, p)_{S_i} = 0 \text{ for all } p \in \mathcal{P}_\rho; \text{ we set}
\]
\[
\Pi u = \sum_{i=1}^m \gamma_i(p_i)\phi_i.
\]

Let \( h = h(D) = \max_{j=1,2,\ldots,n} d(T_j) \). Under the above situation and hypotheses we shall prove:

**Theorem 1**

For \( u \in H^q(\Lambda), q \leq \rho + 1 \), one has
\[
|u - \Pi u|_{k,\Lambda} \leq c h^{q-k}|u|_{q,\Lambda}, \quad k = 0, 1, \ldots, q; \tag{1}
\]

furthermore if \( q \leq \rho \) one has also
\[
\lim_{h \to 0} |u - \Pi u|_{q,\Lambda} = 0. \tag{2}
\]

As mentioned in the introduction, \( \Pi \) can be modified in order to fit boundary condition; we restrict ourselves to the case where the function \( u \) to approximate takes the value 0 on the boundary \( \Gamma \) of \( \Lambda \) i.e. \( u \in \tilde{H}^1(\Lambda) \).

We define \( \tilde{\Pi} : H^0(\Lambda) \to V \) by
\[
\tilde{\Pi} u = \sum_{i=1}^m \gamma_i(p_i)\phi_i
\]

where \( \Sigma' \) means that we omit in the sum the terms relative to indices \( i \) for which \( Q_i \in \Gamma \) and \( l_i = 0 \). Then under the above hypotheses we shall prove:

**Theorem 2**

For \( u \in \tilde{H}^1(\Lambda) \cap H^q(\Lambda), q \leq \rho + 1 \) one has
\[
|u - \tilde{\Pi} u|_{k,\Lambda} \leq c h^{q-k}|u|_{q,\Lambda}, \quad k = 0, 1, \ldots, q; \tag{3}
\]

furthermore if \( q \leq \rho \) one has also
\[
\lim_{h \to 0} |u - \tilde{\Pi} u|_{q,\Lambda} = 0. \tag{4}
\]
Remarks

1. \( \tilde{\Pi} u \) will not automatically belong to \( \tilde{H}^1(\Lambda) \); however if
   
a) \( \varphi_i \in H^1(\Lambda) \) \( i = 1, 2, \ldots, m \),
   
b) for all \( Q_i \in \Gamma \) one has \( l_i = 0 \), then specific examples show that
   \( \tilde{\Pi} u \in \tilde{H}^1(\Lambda) \).

2. For one-dimensional finite elements \( H^4 \) has to be replaced by :
   if \( T_i \cap T_j \neq \emptyset \) then \( d(T_i)/d(T_j) \geq c > 0 \); here the \( T_i \)'s are segments and \( d(T_i) \)
   is the length of \( T_i \).

3. The local character of the definition of \( \Pi \) clearly implies the following
   property. Let \( u \in H^0(\Lambda) \), \( \Phi \subset \Omega \subset \Lambda \) where \( \Phi \) is closed and \( \Omega \) open,
   \( u \in H^q(\Omega) \); then for \( m \leq \rho + 1 \) one has \( |u - \Pi u|_{k, \Phi} \leq c h^{q-k}|u|_{q, \Omega} \) and
   for \( q \leq \rho \) one has \( \lim_{h \to 0} |u - \Pi u|_{q, \Phi} = 0 \).

4. There is a great number of alternative possibilities of defining \( \Pi \) in the
   same spirit.

5. One can without any difficulty give the same results for \( \Lambda \subset \mathbb{R}^n \).

3. PROOFS

In this section we use all the definitions, notations, hypotheses introduced
in section 2.

Lemma 1

Let \( S \) be any of the supports \( S_1, S_2, \ldots, S_m, u \in H^q(S) \), \( q \leq \rho + 1 \), \( t \in \mathbb{F}_p \)
such that \( (u - t, p)_S = 0 \) for all \( p \in \mathbb{F}_p \). Then

\[ |u - t|_{k, S} \leq c (d(S))^{q-k}|u|_{q, S} \quad 0 \leq k \leq q. \]

Lemma 1, which supposes \( H^4 \), has analogue in the literature, see [3],
[5], [9]; however, we give below a sketch of the proof.

We restrict ourselves to the case where \( S \) is formed by two adjacents
triangles; the other cases are treated similarly. Let us consider first the case
where \( d(S) = 1 \). Let \( \Delta_1, \Delta_2 \) be triangles in the \((\xi, \eta)\) plane with vertices
at \((0, 0), (0, 1), (1, 0)\) and \((0, 0), (0, -1), (1, 0)\) respectively. Let \( \Delta = \Delta_1 \cup \Delta_2 \).
Let in the \((x, y)\) plane \( S \) be the domain formed by two triangles \( T_1 \) and \( T_2 \) having
a common side. \( T_i \) is the image of \( \Delta_i \) by the application \( \varphi_i : \)

\[ (\xi) \mapsto A_i(\xi) + b_i = (\gamma) \quad i = 1, 2. \]

\( \varphi \) is the application of \( \Delta \to S \) such that \( \varphi |_{\Delta_i} = \varphi_i \). We consider a set \( \mathcal{E} \)
of domains \( S \) of this type satisfying the relations :

\[ \|A_i\| \leq c \quad \text{and} \quad \|A_i^{-1}\| \leq c \]
where \( c \) is a generic constant which doesn’t depend on \( S \in \mathcal{E} \). Let us show that \( (u, 1)_S = 0 \) implies \( |u|_{0, S} \leq c |u|_{1, S} \) if \( u \in H^1(S) \).

Revue Française d'Automatique, Informatique et Recherche Opérationnelle
Let \( v = u \circ \phi : \Delta \to R \). We verify that \( v \) belongs to \( H^1(\Delta) \) and if we put
\[
\alpha = \frac{(v, 1)_\Delta}{(1, 1)_\Delta},
\]
we get \( |v - \alpha|_{0, \Delta} \leq c|v|_{1, \Delta} \). The hypothesis on the matrices \( A_i \) imply that
\[
|u - \alpha|_{0, \tau} \leq c|v - \alpha|_{0, \Delta} \quad \text{and} \quad |v|_{1, \Delta} \leq c|u|_{1, \tau}.
\]
So
\[
|u_{0, s}^2 \leq |u_{0, s}^2 + |\alpha|_{0, s}^2 = |u - \alpha|_{0, \tau}^2 + |u - \alpha|_{0, \tau}^2 \leq c(|v - \alpha|_{0, \Delta}^2 + |v - \alpha|_{0, \Delta}^2) = c|v - \alpha|_{0, \Delta} \leq |v|_{1, \Delta}^2 \leq c|u|_{1, \tau}^2.
\]
This allows us to prove the lemma for \( k = 0 \). Indeed, let \( u \in H^q(S), q \leq \rho + 1 \)
\( t \in P_\rho \) such that \( (u - t, p)_S = 0 \) for all \( p \in P_\rho \) Let \( \bar{p} \) the unique polynomial
belonging to \( P_\rho \) such that \( D^s(u - \bar{p}), 1)_S = 0 \) for \( 0 \leq s \leq q - 1 \); then by
applying the preceding result to \( D^s(u - \bar{p}) \) for \( s = q - 1, q - 2, ..., 0 \) we
get : \( |u - \bar{p}|_{0, s} \leq c|u - \bar{p}|_{q, s} = c|u|_{q, s} \); then \( |u|_{0, s} \leq |u - \bar{p}|_{0, s} \) hence we are
done. We obtain the general case from the interpolation formula
\( |u|_{k, s} \leq c(|u|_{0, s} + |u|_{q, s}) \) \( \forall u \in H^q(S) \) and from the fact that \( |p|_{k, s} \leq c|p|_{q, s} \)
\( \forall p \in P_\rho \). (These relations can be established by returning to the fundamental
domain \( \Delta \) by the application \( \phi \). Indeed, for \( \rho = q - 1, 0 \leq k \leq q \) we get :
\( |u - t|_{k, s} \leq c(|u - t|_{0, s} + |u - t|_{q, s}) \leq c|u|_{q, s} \). For \( q < \rho + 1 \), let \( \bar{\tau} \in P_{q-1} \)
such that \( (u - \bar{\tau}, p)_S = 0 \) \( \forall p \in P_{q-1} \); then \( |u - \bar{\tau}|_{k, s} \leq |u - \bar{\tau}|_{k, s} + |\bar{\tau} - t|_{k, s} \)
\( |u - t|_{k, s} \leq c|u|_{q, s} \) and \( |\bar{\tau} - t|_{k, s} \leq c|\bar{\tau} - t|_{0, s} \leq c(|u - \bar{\tau}|_{0, s} + |u - t|_{0, s}) \leq c|u|_{q, s} \). We
obtain the case \( d(S) \neq 1 \) by a dilatation.

**Lemma 2**

Let \( T \in D, p \in \mathfrak{F}_\rho \); then
\[
|p|_{k, \omega, \tau} \leq c(\mu(T))^{-\frac{1}{2}}(d(T))^{-k - 1} |p|_{0, \tau} \quad k = 0, 1, 2, ...
\]

Lemma 2, which supposes H4, is an elementary property based on the
equivalence of all the norms for a finite dimensional space (see [8]).

**Lemma 3**

\( a) \) If \( T_j \subset S_i \) then \( d(S_i) \leq c d(T_j) \).

\( b) \) The number of elements \( T_j \) contained in any support \( S_i \) is \( \leq c \).

Lemma 3 is a consequence of H3 and H4.

**Proof of theorem 1**

Let \( T \) be a particular element of \( D, Q_1, Q_2, ..., Q_s \) the nodes belonging
to \( T, \phi_1, ..., \phi_\alpha, \gamma_1, ..., \gamma_\alpha, l_1, l_2, ..., l_\alpha, p_1, ..., p_\alpha \) the corresponding basic
functions, functionals, orders and polynomials. Let \( u \in H^q(\Lambda), q \leq \rho + 1 \). One
has on \( T \):
\[
\Pi u = \sum_{i=1}^\alpha \gamma_i(p_i) \phi_i = \sum_{i=1}^\alpha \gamma_i(p_i) \phi_i + \sum_{i=2}^\alpha \gamma_i(p_i - p_1) \phi_i;
\]
n° août 1975, R-2.
because of $H_1$, the first sum on the right side is equal to $p_1$; by lemmas 1 and 3, one gets for $0 \leq k \leq q$:

$$|\Pi u - u|_{k,T} \leq ch^{q-k}|u|_{q,S_1} + \sum_{i=2}^{q} |\gamma_i(p_i - p_1)| |\varphi_{i,k,T}|; \quad (5)$$

in order to estimate the second term of the right side, one remarks that one has by lemmas 1 and 3:

$$|p_i - p_1|_{0,T} \leq |u - p_1|_{0,T} + |u - p_1|_{0,S_1} + |u - p_1|_{0,S_1} \leq c(d(S_i))^q|u|_{q,S_i} + c(d(S_i))^q|u|_{q,S_1} \leq c(d(T))^q(|u|_{q,S_i} + |u|_{q,S_1});$$

by lemma 2 and $H_2$, one gets:

$$|p_i - p_1|_{0,T} \leq c(\mu(T))^{-\frac{1}{2}}(d(T))^{q-1}|u|_{q,S_i} + c|u|_{q,S_1};$$

introducing this last inequality in (5) one gets by $H_3$ ($\alpha \leq c$):

$$|\Pi u - u|^2_{k,T} \leq ch^{2(q-k)} \sum_{i=1}^{q} |u|^2_{q,S_i};$$

this relation is valid for any $T \in D$; using again $H_3$ and lemma 3, one gets by summing for $i = 1, 2, ..., n$ precisely relation (1). Now suppose $q \leq \rho$; for any $v \in H^{q+1}(\Lambda)$ one has by (1):

$$|u - \Pi u|_{q,\Lambda} \leq |v - \Pi v|_{q,\Lambda} + |(u - v) - \Pi(u - v)|_{q,\Lambda} \leq ch |v|_{q+1,\Lambda} + c |u - v|_{q,\Lambda};$$

let $\varepsilon > 0$; one first chooses $v$ such that $c |u - v|_{q,\Lambda} < \varepsilon/2$; let $h_0 > 0$ be such that $ch_0 |v|_{q+1,\Lambda} < \varepsilon/2$; then for $h < h_0, |u - \Pi u|_{q,\Lambda} < \varepsilon$ which proves relation (2).

**Lemma 4**

Let $T \in D$, $\tau$ a side of $T$, $u \in H_1(T)$; then

$$d(T)|u|^2_{0,\tau} \leq c \{ |u|^2_{0,T} + (d(T))^2 |u|^2_{1,T} \}.$$  

Lemma 4 is a consequence of the trace theorem (see [10]); for a detailed proof, see [1].

**Proof of theorem 2**

Let $u \in \hat{H}^1(\Lambda) \cap H^q(\Lambda)$, $1 \leq q \leq \rho + 1$. Using the notations and arguments of the proof of theorem 1, one remarks (see [6]) that for proving (3) it suffices to show that for $T \in D$ one has

$$|\Pi u - \tilde{\Pi} u|^2_{k,T} \leq ch^{2(q-k)} \sum_{i=1}^{q} |v|^2_{q,S_i}, \quad 0 \leq k \leq q.$$  

(7)
Let $Q_i \in \Gamma$ and $l_i = 0$ for $i = 1, 2, \ldots, \beta$ and for $\beta < i \leq \alpha$, $Q_i \notin \Gamma$ or $l_i > 0$; if $\beta = 0$, $\Pi u = \tilde{\Pi} u$ on $T$ so that we can suppose $\beta > 0$; by H2, one has

$$|\Pi u - \tilde{\Pi} u|_{k,T} = \left| \sum_{i=1}^{\beta} \gamma_i(p_i) \phi_i \right|_{k,T} \leq c(d(T))^{-k}(\mu(T))^{\frac{1}{2}} \sum_{i=1}^{\beta} |p_i(Q_i)|;$$

let $\tau_i \subset \Gamma$ be a side of an element $G_i$, satisfying the relation $Q_i \in \tau_i \subset G_i \subset S_i$, $i = 1, \ldots, \beta$. Since $u \in \tilde{H}^1(\Lambda), |u|_{0,\tau} = 0$ and $|p_i|_{0,\tau} = |u - p_i|_{0,\tau}$, from lemmas 4 and 1, one gets:

$$d(G_i)|p_i|_{0,\tau} \leq c \{ |u - p_i|^2_{0,S} + (d(S_i))^2 |u - p_i|^2_{1,S_i} \} \leq c(d(S_i))^{2q} |u|_{q,S_i}^2; \quad (9)$$

a one-dimensional version of lemma 2 allows to write

$$|p_i(Q_i)| \leq c(L(\tau_i))^{-\frac{1}{2}} |p_i|_{0,\tau}, \quad (10)$$

where $L(\tau)$ is the length of $\tau$; by H4 and lemma 3, (9) and (10) imply:

$$(d(T))^{-2k}\mu(T)(p_i(Q_i))^2 \leq c(d(T))^{-2k}\mu(T)(L(\tau))^{-1}(d(G_i))^{-1}(d(S_i))^{2q} |u|_{q,S_i}^2 \leq c h^{2(q-k)} |u|_{q,S_i}^2;$$

replacing in (8) and using H3 one obtains (7). It remains to prove (4); let $u \in \tilde{H}^1(\Lambda) \cap H^q(\Lambda), q \leq \rho$; by theorem 1, it suffices to verify that $\lim_{h \to 0} |\Pi u - \tilde{\Pi} u|_{q,\Lambda} = 0$; let $J = \{ j: T_j \cap \Gamma \neq \emptyset \}$ and $\emptyset$ be the union of all $S_i$ containing a $T_j$ with $j \in J$; by (7) H3, and lemma 3, one has

$$|\Pi u - \tilde{\Pi} u|_{q,\Lambda}^2 = \sum_{j \in J} |\Pi u - \tilde{\Pi} u|_{q,T_j}^2 \leq c |u|_{q,\emptyset}^2;$$

since $\lim_{h \to 0} \mu(\emptyset) = 0$, one also has $\lim_{h \to 0} |\Pi u - \tilde{\Pi} u|_{q,\emptyset} = 0.$

REFERENCES


