JEAN DESCLOUX
NABIL NASSIF
JACQUES RAPPAZ

On spectral approximation. Part 1. The problem of convergence


<http://www.numdam.org/item?id=M2AN_1978__12_2_97_0>

© AFCET, 1978, tous droits réservés.
L’accès aux archives de la revue « RAIRO – Analyse numérique » implique l’accord avec les conditions générales d’utilisation (http://www.numdam.org/legal.php). Toute utilisation commerciale ou impression systématique est constitutive d’une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM
Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
http://www.numdam.org/
Abstract. — One studies the problem of the numerical approximation of the spectrum of non-compact operators in Banach spaces. Special results are derived for the selfadjoint case. An example is presented.

1. INTRODUCTION

Let \( X \) be a complex Banach space with norm \( \| \cdot \| \), \( A \) be a bounded linear operator in \( X \) with spectrum \( \sigma (A) \) and resolvent set \( \rho (A) \). The problem is the numerical computation of \( \sigma (A) \). To this end we introduce a sequence \( \{ X_h \} \) of finite dimensional subspaces of \( X \) and the linear operators \( A_h : X_h \rightarrow X_h \); \( \sigma (A) \) is then approximated by the spectrum \( \sigma (A_h) \) of \( A_h \).

In many practical methods (Galerkin for example), \( A_h \) is the restriction to \( X_h \) of an operator \( B_h : X \rightarrow X \) such that \( B_h (X) \subseteq X_h \); then, except for the eigenvalues \( 0 \), \( A_h \) and \( B_h \) have the same eigenvalues and corresponding invariant subspaces.

Let us introduce some notations. For any complex number \( z \in \rho (A) \) [resp. \( z \in \rho (A_h) \)], \( R_z (A) = (z - A)^{-1} : X \rightarrow X \) [resp. \( R_z (A_h) = (z - A_h)^{-1} : X_h \rightarrow X_h \)] is the resolvent operator. For \( z_0 \in \mathbb{C} \) and \( \Lambda \subset \mathbb{C} \), \( \delta (z_0, \Lambda) = \inf_{z \in \Lambda} \| z - z_0 \| \) is the distance from \( z_0 \) to \( \Lambda \). For \( x \in X \), \( Y \) and \( Z \) closed subspaces of \( X \), we set:

\[
\delta (x, Y) = \inf_{y \in Y} \| x - y \|, \quad \delta (Y, Z) = \sup_{\| y \| = 1} \delta (y, Z),
\]

\[
\hat{\delta} (Y, Z) = \max (\delta (Y, Z), \delta (Z, Y)),
\]

where \( \hat{\delta} (Y, Z) \) is the gap between \( Y \) and \( Z \). For an operator \( C \), we set

\[
\| C \|_h = \sup_{\| x \| = 1} \| C x \|.
\]

(*) Manuscrit reçu le 10 juin 1977.
(1) Département de Mathématiques, École Polytechnique fédérale de Lausanne, Suisse.
(2) Department of Mathematics, American University of Beirut, Liban.
Finally, let $\Gamma$ be a Jordan curve in the revolvent set $\rho (A)$ and $\Lambda \subset \mathbb{C}$ be the domain limited by $\Gamma$; we define the spectral projectors $E : X \rightarrow X$ and $E_h : X_h \rightarrow X_h$ by

$$E = (2 \pi i)^{-1} \int_\Gamma R_z(A) \, dz, \quad E_h = (2 \pi i)^{-1} \int_\Gamma R_z(A_h) \, dz;$$

$E_h$ is defined only if $\Gamma$ is contained in the resolvent set $\rho (A_h)$.

We now list some desirable properties of spectral approximation of $A$ by $A_h$:

a) for any $K \subset \rho (A)$ compact, there exists $h_0$ such that $K \subset \rho (A_h)$, $\forall h \leq h_0$;

b) $\forall z \in \sigma (A)$, $\lim_{h \to 0} \delta (z, \sigma (A_h)) = 0$;

g) $\forall u \in E(X)$, $\lim_{h \to 0} \delta (u, E_h(X_h)) = 0$; in particular if $\Lambda \cap \sigma (A) \neq \emptyset$,

then for $h$ small enough $\Lambda \cap \sigma (A_h) \neq \emptyset$;

$\delta$) $\lim_{h \to 0} \delta (E_h(X_h), E(X)) = 0$;

e) If $E(X)$ is finite dimensional, then $\lim_{h \to 0} \delta (E_h(X_h), E(X)) = 0$; in particular for $h$ small enough, the sums of the algebraic multiplicities of the eigenvalues of $A$ and $A_h$ contained in $\Lambda$ are equal.

If $X$ is a Hilbert space and if $A$ and $A_h$ are selfadjoint, the condition $\delta$ can be refined; for an interval $I$, $E_I$ and $E_{hI}$ denote the spectral projectors of $A$ and $A_h$ relative to $I$; we introduce the condition:

0) for the intervals $J \subset I$, $J$ closed bounded, $I$ open, $\lim_{h \to 0} \delta (E_{hJ}(X_h), E_I(X)) = 0$.

Let us consider some conditions which could possibly ensure the preceding properties:

a) $\lim_{h \to 0} \| A - B_h \| = 0$;

b) $\lim_{h \to 0} B_h = A$ strongly;

c) $\bigcup_{h \in \mathbb{H}} \{ B_h x \mid \| x \| \leq 1 \}$ is relatively compact;

d) for any sequence $x_h \in X_h$, $\| x_h \| \leq 1$, the sequence $(A - A_h)x_h$ is relatively compact;

e) for any sequence $x_h \in X_h$, $\lim_{h \to 0} x_h = x$, one has $\lim_{h \to 0} A_h x_h = A x$.

P1) $\lim_{h \to 0} \| A - A_h \|_h = 0$;

P2) $\forall x \in X$, $\lim_{h \to 0} \delta (x, X_h) = 0$. 

R.A.I.R.O. Analyse numérique/Numerical Analysis
First we remark that none of these conditions, separately or together can ensure property $\beta$; however if $A$ and $B_h$ are selfadjoint operators in the Hilbert space $X$, then $b$ implies $\beta$ (see [5], p. 210 and 431).

If $d$ and $e$ are satisfied, then the sequence $A_h$ compactly approximates $A$ in the sense of Vainikko [9] (in fact Vainikko defines this notion in a slightly more general context); then properties $\alpha$, $\gamma$ and $\varepsilon$ are satisfied; in particular, to an isolated eigenvalue $\lambda$ of $A$ of finite algebraic multiplicity $m$ correspond the eigenvalues $\mu_{1h}, \mu_{2h}, \ldots$, of $A_h$ converging to $\lambda$ with total algebraic multiplicities $m$.

If $c$ is satisfied, then the set of $B_h$'s are collectively compact in the sense of Anselone [1]; together with $b$, $A$ is necessarily a compact operator and the sequence $B_h$ compactly approximates $A$ in the sense of Vainikko; one can deduce properties $\alpha$, $\gamma$, $\varepsilon$.

In this paper we shall study conditions P1, P2. P1 is clearly inspired by $a$ but less restrictive; indeed, since $B_h$ is compact, $a$ can be satisfied only if $A$ is compact. P1, P2 imply that the sequence $A_h$ compactly approximates $A$ in the sense of Vainikko. In section 2, we shall prove not only properties $\alpha$, $\gamma$ and $\varepsilon$ but also $\delta$; the proofs are simple and all the arguments can be found in [5]. (Of course, $a$ will also imply $\alpha$, $\gamma$, $\delta$, $\varepsilon$.)

In section 3, we consider the particular case where $X$ is a Hilbert space, $A$ and $A_h$ are selfadjoint. From what precedes, it follows that P1, P2 imply $\alpha$, $\beta$, $\gamma$, $\delta$, $\varepsilon$. In fact one has more; we shall prove: P1 $\iff$ $\theta$ (for all $I, J$); at the light of this result, P1 appears as a natural condition.

It should not be necessary emphasize that the interest of the different conditions, $a$, $b$, $\ldots$, consists not only in the results they imply but also in the possibility to realize them in practical situations. If $A$ is compact and $B_h$ is obtained by a Galerkin method using $X_h$, then $a$ will follow automatically from P2. Condition $c$ has been used successfully in connection with integral operators (see [1]). Curiously enough, to our knowledge, the concept of compact approximation of Vainikko has been applied so far only for finite differences methods approximating two points boundary value problems (i.e. compact operators) (see [10]).

As far as we are concerned, our goal was to compute the spectrum of some differential operators with non compact inverse arising from plasma physics by the Galerkin method. The situation can be formalized in the following way. $a$ and $b$ are given continuous sesquilinear forms on $X$; furthermore, one supposes $a$ coercive; $A$ and $A_h$ are defined by the relations: $a (A u, v) = b (u, v)$, $\forall u, v \in X$, $a (A_h u, v) = b (u, v)$, $\forall u, v \in X_h$. In this
case, $P_1$ is equivalent to the pure approximation property:

$$\lim_{h \to 0} \sup_{x \in X_h, \|x\|=1} \delta(Ax, X_h) = 0.$$ 

This condition can be considered for itself, i.e., for a general bounded operator $A$ in the Banach space $X$; at the present time we know two fundamental cases for which it is satisfied: 1) $A$ is compact (one supposes $P_2$ fulfilled); 2) $X = H^m(\Omega)$, $\Omega \subset \mathbb{R}^n$, $\{X_h\}$ is a family of finite element subspaces, $Au = \omega \cdot u$ (multiplication operator) where $\omega$ is a fixed sufficiently regular function. With the help of these two examples, we analyze briefly in section 4 a partial differential operator suggested by the physics of plasma. Note that this operator can also be treated by a different method developed in [7] by J. Rappaz. For an one-dimensional example, see also [3].

There exist many relations between the conditions $a$, $b$, ..., $P_2$; some of them are analyzed in [4]. Let us quote one of them: if $A_h$ is obtained by the Galerkin method (situation described above) then $P_1$, $P_2 \Leftrightarrow d$, $P_2$; in other words, if $P_2$ is satisfied, $A_h$ is a compact approximation of $A$ in the sense of Vainikko if and only if $P_1$ is satisfied.

Finally let us mention a generalization for closed operators which is developed in [4]. Suppose that $A$ is not a bounded operator in $X$, but simply a closed operator. Set $\delta(A_h, A) = \delta(G_h, G)$ where $G$ and $G_h$ are the graphs of $A$ and $A_h$ considered as subspaces in $X \times X$. $P_1$ is replaced by $P_1'$:

$$\lim_{h \to 0} \delta(A_h, A) = 0.$$ 

Then, as in the bounded case, one has: $P_1', P_2 \Leftrightarrow \alpha$, $\gamma$, $\delta$, $\varepsilon$; in the Hilbert case where $A$ and $A_h$ are selfadjoint: $P_1' \Leftrightarrow \theta$ (for all $I$, $J$). However we shall not present here the proof of these results since we have no specific example to exhibit.

2. THE BANACH CASE

We consider the situation and notations defined in the beginning of the introduction. In particular, $X$ is a Banach space of norm $\|\cdot\|$, $\{X_h\}$ is a sequence of finite dimensional subspaces of $X$, $A$ and $A_h$ are linear bounded operators in $X$ and $X_h$ respectively; for an operator $C$, $\|C\|_h = \sup_{x \in X_h, \|x\|=1} \|Cx\|$.

We also recall the definitions of properties $P_1$, $P_2$:

$$P_1: \lim_{h \to 0} \|A - A_h\|_h = 0; \quad P_2: \forall x \in X, \lim_{h \to 0} \delta(x, X_h) = 0.$$

R.A.I.R.O. Analyse numérique/Numerical Analysis
**Lemma 1:** One supposes \( P_1 \) and let \( F \subset \rho(A) \) be closed. Then there exists a constant \( C \) independent of \( h \) such that for \( h \) small enough we have:

\[
\| R_z(A_h) \|_h \leq C, \quad \forall z \in F.
\]

**Proof:** There exists \( C > 0 \) such that

\[
\| (z-A)u \| \geq 2C\| u \|, \quad \forall u \in X, \quad z \in F.
\]

By \( P_1 \) we have for \( h \) small enough \( \| (A-A_h)u \| \leq C\| u \|, \quad \forall u \in X_h. \) Then we obtain for \( u \in X_h, \ z \in F: \)

\[
\| (z-A_h)u \| \geq \| (z-A)u \| - \| (A-A_h)u \| \geq C\| u \|,
\]

since \( X_h \) is finite dimensional, this proves in particular the existence of \( R_z(A_h) \).

As a direct consequence this property of stability, we have:

**Theorem 1:** One supposes \( P_1 \) and let \( \Omega \subset C \) be an open set containing \( \sigma(A) \). Then there exists \( h_0 > 0 \) such that \( \sigma(A_h) \subset \Omega, \quad \forall h < h_0. \)

Let now \( \Gamma \subset \rho(A) \) be a smooth Jordan curve. We introduce (see for example [5], p. 178) the continuous spectral projectors \( E: X \to X \) and \( E_h: X_h \to X_h \) defined by

\[
E = (2\pi i)^{-1} \int_{\Gamma} R_z(A) \, dz \quad \text{and} \quad E_h = (2\pi i)^{-1} \int_{\Gamma} R_z(A_h) \, dz.
\]

By theorem 1, \( E_h \) is well defined for \( h \) sufficiently small.

**Lemma 2:** One supposes \( P_1 \). Then

\[
\lim_{h \to 0} \| E - E_h \|_h = 0.
\]

**Proof:** For \( h \) small enough we have

\[
\| E - E_h \|_h \leq (2\pi)^{-1} \int_{\Gamma} \| R_z(A) - R_z(A_h) \|_h \, dz
\]

\[
= (2\pi)^{-1} \int_{\Gamma} \| R_z(A)(A-A_h) R_z(A_h) \|_h \, dz
\]

\[
\leq (2\pi)^{-1} \int_{\Gamma} \| R_z(A) \| \cdot \| A-A_h \|_h \cdot \| R_z(A_h) \|_h \, dz.
\]

Taking in account \( P_1 \) and lemma 1 one gets the result. ■
One deduces immediately from lemma 2:

**Theorem 2:** One supposes P1. Then

\[ \lim_{h \to 0} \delta(E_h(X_h), E(X)) = 0. \]

**Theorem 3:** One supposes P1 and P2. Then for all \( x \in E(X) \):

\[ \lim_{h \to 0} \delta(x, E_h(X_h)) = 0. \]

**Proof:** Let \( x \in E(X) \). By P2 there exists \( x_h \in X_h \) with \( \lim_{h \to 0} \| x - x_h \| = 0. \) Then

\[ \| x - E_h x_h \| = \| E x - E_h x_h \| \leq \| E(x - x_h) \| + \| (E - E_h) x_h \| \leq \| E \| \| x - x_h \| + \| E - E_h \|_h \| x_h \|. \]

One uses lemma 2 and the continuity of \( E \).

Let \( n \) and \( n_h \) be the dimensions of \( E(X) \) and of \( E_h(X_h) \). Theorem 3 shows that if \( n = \infty \) then \( n_h = \infty \). If \( n < \infty \) then theorem 3 shows that

\[ \lim_{h \to 0} \delta(E(X), E_h(X_h)) = 0 \]

and with theorem 2 we shall have

\[ \lim_{h \to 0} \delta(E(X), E_h(X_h)) = 0. \]

Consequently we shall have \( n = n_h \) when \( h \) is small enough (see [5], p. 200). In particular if \( \Lambda \) is the domain of \( C \) limited by \( \Gamma \) and if \( \Lambda \cap \sigma(A) \neq \emptyset \) then \( \Lambda \cap \sigma(A_h) \neq \emptyset \) for \( h \) small enough.

**Remark:** In this section, we have verified the properties \( \alpha, \gamma, \delta, \varepsilon \) stated in the introduction. That property \( \beta \) cannot be obtained from P1 and P2 is shown by an example in [9], p. 12.

### 3. The Selfadjoint Case

In this section \( X \) is a complex Hilbert space with scalar product \( (.,.) \) and norm \( \| . \| \), \( \{ X_h \} \) is a sequence of finite dimensional subspaces; \( A \) and \( A_h \) are selfadjoint operators in \( X \) and \( X_h \) respectively. We recall some notations already defined in the introduction. For an operator \( C \), \( \| C \|_h = \sup_{\| x \|_h = 1} \| C x \|. \)

For an interval \( I \) (non necessarily finite) \( \overset{\circ}{I} \) is the interior of \( I \), \( \overline{I} \) is the closure of \( I \), \( E_I : X \to X \) and \( E_{h,I} : X_h \to X_h \) are the spectral projectors of \( A \) and \( A_h \).
associated to \( I \) (we shall use without explicit reference the spectral theory contained in [8], pp. 259-274). Besides P1 and P2, we introduce for convenience other properties:

\[
P1: \lim_{h \to 0} \| A - A_h \|_h = 0; \quad P2: \forall x \in X, \lim_{h \to 0} \delta(x, X_h) = 0;
\]

P1 a: \( \forall \) intervals \( I, J \) with \( \bar{J} \subset \bar{I} \), one has

\[
\lim_{h \to 0} \delta(E_{h,J}(X_h), E_I(X)) = 0;
\]

P1 b: \( \forall \) intervals \( I, J \) with \( \bar{J} \subset \bar{I} \), one has

\[
\lim_{h \to 0} \| E_{h,J} - E_I E_{h,J} \|_h = 0.
\]

P1 c: \( \forall \) intervals \( I, J \) with \( I \cap \bar{J} = \emptyset \), one has

\[
\lim_{h \to 0} \| E_I E_{h,J} \|_h = 0.
\]

The results of this section are contained in the following three theorems.

**Theorem 4:** The properties P1, P1 a, P1 b and P1 c are equivalent.

**Theorem 5:** One supposes P1 and P2. Let \( I \) and \( J \) be intervals with \( \bar{I} \subset J \); then \( \forall x \in E_I(X), \lim_{h \to 0} \delta(x, E_{h,J}(X_h)) = 0. \)

**Theorem 6:** One supposes P1 and P2; then \( \forall \lambda \in \sigma(A), \lim_{h \to 0} \delta(\lambda, \sigma(A_h)) = 0. \)

Since \( E_I \) and \( E_{h,J} \) are orthogonal projectors, P1 a is clearly equivalent to P1 b. Then theorem 4 follows from lemmas 3, 5, 6. Theorem 6 which corresponds to property \( \beta \) of the introduction is an almost obvious consequence of theorem 5. We now prove the remainder results.

**Lemma 3:** P1 b and P1 c are equivalent.

**Proof:** Let \( I \cap \bar{J} = \emptyset \) and suppose P1 b. There exists an interval \( P \) such that \( I \cap P = \emptyset \) and \( \bar{J} \subset \bar{P} \); then

\[
\lim_{h \to 0} \| E_I E_{h,J} \|_h \leq \lim_{h \to 0} \| (I - E_P) E_{h,J} \|_h = 0;
\]

consequently P1 c is verified. The converse implication follows from similar arguments.

For convenience we introduce the orthogonal projector \( \Pi_h \) of \( X \) on \( X_h \), i.e. \( (x - \Pi_h x, y) = 0, \forall y \in X_h, x \in X \), and \( B_h : X \to X \) defined by \( B_h = A_h \Pi_h. \)
Clearly $\Pi_h$ and $B_h$ are selfadjoint, $\sigma(B_h) = \sigma(A_h) \cup \{0\}$. If $J$ is an interval of $\mathbb{R}$ we define $F_{h,J}$ as the spectral projector relative to $B_h$ and $J$; we have $F_{h,J}x = E_{h,J}x$ for all $x$ in $X_h$.

**Lemma 4:** One supposes $P_1$. Let $f: \mathbb{R} \to \mathbb{R}$ be continuous. Then

$$\lim_{h \to 0} \| f(A) - f(B_h) \|_h = 0.$$  

**Proof:** $P_1$ implies the existence of $h_0 > 0$ and $M$ such that $\|A\| < M$, $\|B_h\| < M$, $\forall h < h_0$. We first prove lemma 4 for polynomials. It suffices to consider $F(\lambda) = \lambda^k$ with $k = 0, 1, 2, \ldots$. The case $k = 0$ is trivial and the case $k = 1$ is a consequence of $P_1$. Suppose the relation correct for $k = N$ and let us prove it for $k = N + 1$.

We have

$$A^{N+1} - B_h^{N+1} = A^N(A - B_h) + (A^N - B_h^N)B_h$$

and thus

$$\| A^{N+1} - B_h^{N+1} \|_h \leq \| A \|_{N+1} \| A - B_h \|_h + \| A^N - B_h^N \|_h \| B_h \|_h.$$  

Consequently we obtain $\| A^{N+1} - B_h^{N+1} \|_h \to 0$ as $h \to 0$. Consider now the general case.

Let $\varepsilon > 0$ fixed. There exists a polynomial $p$ such that $|f(\lambda) - p(\lambda)| < \varepsilon/3$, $\forall \lambda$ with $|\lambda| \leq M$. One has for $h < h_0$

$$\| f(A) - f(B_h) \|_h \leq \| f(A) - p(A) \|_h + \| p(A) - p(B_h) \|_h + \| p(B_h) - f(B_h) \|_h.$$  

Thus

$$\| f(A) - f(B_h) \|_h < \frac{2\varepsilon}{3} + \| p(A) - p(B_h) \|_h.$$  

Lemma 4 then follows from the result for polynomials. $\blacksquare$

**Lemma 5:** $P_1 \Rightarrow P_1 c.$

**Proof:** Let $I$ and $J$ be intervals of $\mathbb{R}$ such that $I \cap J = \emptyset$. Let $\varphi, \psi: \mathbb{R} \to [0, 1]$ be continuous functions such that $\varphi(x) = 1$ if $x \in I$, $\psi(x) = 1$ if $x \in J$ and $\varphi(x) \psi(x) = 0$, $\forall x \in \mathbb{R}$. Then $\varphi(A) \psi(A) = 0$, $\| \varphi(A) \| \leq 1$ and one has:

$$\| \varphi(A) \psi(B_h) \|_h \leq \| \varphi(A) \psi(A) \|_h + \| \varphi(A)(\psi(A) - \psi(B_h)) \|_h \leq \| \psi(A) - \psi(B_h) \|_h.$$  

By lemma 4 we obtain $\lim_{h \to 0} \| \varphi(A) \psi(B_h) \|_h = 0.$
But $E_I \varphi(A) = E_I$ and $\psi(B_h) F_{h,J} = F_{h,J}$ so that
\[
\lim_{h \to 0} \left\| E_I E_{h,J} \right\|_h = \lim_{h \to 0} \left\| E_I \varphi(A) \psi(B_h) F_{h,J} \right\|_h = 0. \]

**Lemma 6:** $P_1 \Rightarrow P_1$

**Proof:** We first remark that the sequence of operators $B_h$ is uniformly bounded; indeed let $I = (-||A|| - 1, ||A|| + 1)$, $J = (||A|| + 2, \infty)$ or $J = (-\infty, -||A|| - 2)$; then $E_I$ is the identity, $P_1 \Rightarrow$ implies $\lim_{h \to 0} \left\| E_{h,J} \right\|_h = 0$

and for $h$ small enough, $\sigma(A_h) \subseteq (-||A|| - 2, ||A|| + 2)$. Consequently let $M$ be such that $||A|| \leq M$, $||B_h|| \leq M, \forall h$.

Let $\varepsilon > 0$ be fixed. We prove that $\lim \sup_{h \to 0} \left\| A - B_h \right\|_h \leq \varepsilon$. Let $\lambda_k = k(\varepsilon/3), k = 0, \pm 1, \pm 2, \ldots$, $I_k$ be the open interval $(\lambda_k - (\varepsilon/3), \lambda_k + (\varepsilon/3))$, $J_k$ be the semi-closed interval $[\lambda_k - (\varepsilon/6), \lambda_k + (\varepsilon/6))$. In order to simplify the notations we set $G_k = E_{I_k}$ and $G_{h,k} = F_{h,J_k}$.

By $P_1 \Rightarrow$ and Lemma 3 we have:
\[
\lim_{h \to 0} \left\| G_{h,k} - G_k G_{h,k} \right\|_h = 0, \quad \forall k \tag{1}
\]
and
\[
\lim_{h \to 0} \left\| G_k G_{h,k} \right\|_h = 0 \quad \text{if} \quad |k - l| \geq 2.
\]

Let for $x \in X_h$,
\[
W_h(x) = \left\| (A - B_h)x \right\|^2 = \left\| (A - B_h)(\sum_k G_{h,k} x) \right\|^2.
\]

We can write
\[
W_h(x) = W_{1,h}(x) + \sum_{|k-l| \geq 2} W_{h,k,l}(x),
\]
where
\[
W_{h,k,l}(x) = ((A - B_h) G_{h,k} x, (A - B_h) G_{h,l} x)
\]
and
\[
W_{1,h}(x) = \sum_{|k-l| \leq 1} W_{h,k,l}(x).
\]

The indices $k$ and $l$ vary between $-N$ and $N$ where $N$ is a number independent of $h$, larger then $3M/\varepsilon$. It suffices to show:
\[
\lim \sup_{h \to 0} \left( \sup_{x \in X_h} \left\| W_{1,h}(x) \right\|_h \right) \leq \varepsilon^2 \tag{2}
\]
vol. 12, n° 2, 1978
and
\[
\lim_{h \to 0} \left( \sup_{x \in X_h \atop ||x|| \leq 1} W_{h,k,l}(x) \right) = 0 \quad \text{for} \quad |k-l| \geq 2. \tag{3}
\]

By Schwarz inequality, one gets
\[
W_{1,h}(x) \leq \frac{1}{2} \sum_{|k-l| \leq 1} \left( ||(A-B_h) G_{h,k} x ||^2 + ||(A-B_h) G_{h,k} x ||^2 \right)
\leq 3 \sum_k ||(A-B_h) G_{h,k} x ||^2. \tag{4}
\]

But
\[
(A-B_h) G_{h,k} x = (A-\lambda_k) G_{h,k} x - (B_h-\lambda_k) G_{h,k} x. \tag{5}
\]

We have
\[
||(B_h-\lambda_k) G_{h,k} x || \leq ||(B_h-\lambda_k) G_{h,k} || \cdot || G_{h,k} x ||
\leq \frac{\epsilon}{6} || G_{h,k} x ||. \tag{6}
\]

\[
(A-\lambda_k) G_{h,k} x = (A-\lambda_k) G_k G_{h,k} x + (A-\lambda_k)(G_{h,k} x - G_k G_{h,k} x)
\]

and thus
\[
||(A-\lambda_k) G_{h,k} x || \leq ||(A-\lambda_k) G_k || \cdot || G_{h,k} x ||
+ || A-\lambda_k || \cdot || G_{h,k} - G_k G_{h,k} ||_h \cdot || G_{h,k} x ||
\leq \left( \frac{\epsilon}{3} + || A-\lambda_k || \cdot || G_{h,k} - G_k G_{h,k} ||_h \right) || G_{h,k} x ||. \tag{7}
\]

By replacing (6) and (7) in (5), and (5) in (4) one gets
\[
W_{1,h}(x) \leq 3 \left\{ \frac{\epsilon}{2} + \max_k ( || A-\lambda_k || \cdot || G_{h,k} - G_k G_{h,k} ||_h ) \right\}^2 \sum_k || G_{h,k} x ||^2.
\]

But \( \sum_k || G_{h,k} x ||^2 = || x ||^2 \); then using (1) one gets (2).

It remains to verify (3):
\[
W_{h,k,l}(x) = (A G_{h,k} x, A G_{h,l} x) - (B_h G_{h,k} x, A G_{h,l} x)
- (A G_{h,k} x, B_h G_{h,l} x) + (B_h G_{h,k} x, B_h G_{h,l} x)
= (x, \sum_k A^2 G_{h,k} x) - (B_h x, \Pi_h G_{h,k} A G_{h,l} x)
- (x, \Pi_h G_{h,k} A G_{h,l} B_h x) + (B_h x, \Pi_h G_{h,k} G_{h,l} B_h x).
\]

In order to establish (3), it suffices to show that
\[
\lim_{h \to 0} || \Pi_h G_{h,k} A^j G_{h,l} ||_h = 0 \quad \text{if} \quad |k-l| \geq 2, \quad j = 0, 1, 2.
\]
Suppose \(|k-l| \geq 2\) and \(j \geq 0\). One has

\[\Pi_h G_{h,k} A^j G_{h,l} = \Pi_h G_{h,k} G_k A^j A^l G_{h,l} + \Pi_h G_{h,k} (I - G_k) A^j G_{h,l}\]

and thus

\[||\Pi_h G_{h,k} A^j G_{h,l}||_h \leq ||\Pi_h G_{h,k} A^j|| ||G_k G_{h,l}||_h + ||\Pi_h G_{h,k} (I - G_k)|| \cdot ||A^j G_{h,l}||.\]

But \(||\Pi_h G_{h,k} A^j|| \leq ||A||^j\), \(||A^j G_{h,l}|| \leq ||A||^j\) and by (1)

\[\lim_{h \to 0} ||G_k G_{h,l}||_h = 0;\]

furthermore since \(\Pi_h, G_{h,k}\) and \(G_k\) are selfadjoint,

\[\lim_{h \to 0} ||\Pi_h G_{h,k} (I - G_k)|| = \lim_{h \to 0} ||(I - G_k) G_{h,k}||_h = 0.\]

**Proof of theorem 5:** Let \(P\) and \(Q\) be intervals such that \(P \cup Q = \mathbb{R} - J\). Then \(\overline{P} \cap \tilde{I} = \overline{Q} \cap \tilde{I} = \emptyset\) and by P1:

\[\lim_{h \to 0} ||E_I - E_{h,J}||_h = \lim_{h \to 0} ||E_I E_{h,P} + E_I E_{h,Q}||_h = 0.\]

We remark that \(\Pi_h F_{h,J} = F_{h,J} \Pi_h = E_{h,J} \Pi_h\) so that

\[||\Pi_h E_I - E_{h,J} \Pi_h E_I|| = ||\Pi_h (E_I - F_{h,J} E_I)|| = \frac{1}{||E_{h,J} E_I||_h} \cdot \frac{1}{||E_I - E_{h,J} E_I||_h} = \frac{1}{||E_{h,J} E_I||_h} \cdot \frac{1}{||E_I - E_{h,J} E_I||_h} = \frac{1}{||E_I - E_{h,J} E_I||_h}.\]

Consequently

\[\lim_{h \to 0} ||\Pi_h E_I - E_{h,J} \Pi_h E_I|| = 0.\]

Let \(x \in E_I (X)\). Then by P2 \(\lim_{h \to 0} \frac{1}{||x - \Pi_h x||} = 0\) and by the preceding relation one get

\[||x - E_{h,J} \Pi_h x|| \leq ||x - \Pi_h x|| + ||\Pi_h E_I x - E_{h,J} \Pi_h E_I x|| \to 0.\]

**4. EXAMPLES**

Let \(X\) be a Banach space of norm \(|\cdot|\), \(\{X_h\}\) be a sequence of finite dimensional subspaces of \(X\), \(A\) be a linear bounded operator in \(X\); in this section we are concerned with the concrete verification of the two conditions:

\[P2: \forall x \in X, \lim_{h \to 0} \delta(x, X_h) = 0; \quad P3: \lim_{h \to 0} \sup_{x \in X_h, ||x|| = 1} \delta(A x, X_h) = 0.\]

vol. 12, n° 2, 1978
As mentioned in the introduction, P3 is equivalent to P1 for appropriate Galerkin methods.

When \( A \) is compact, \( P_2 \Rightarrow P_3 \); in fact, one has even more:

**Theorem 7:** Suppose \( A \) compact and \( P_2 \) verified. Then

\[
\lim_{h\to 0} \sup_{x \in X, \|x\| = 1} \delta(Ax, X_h) = 0.
\]

**Proof:** We briefly recall the classical argument. Let \( \varepsilon > 0 \) be given; one chooses a finite covering of \( \{ A x \mid \|x\| \leq 1 \} \) by balls of radius \( \varepsilon/2 \) and centers \( y_1, \ldots, y_N \). By P2, there exists \( h_0 > 0 \) such that \( \delta(y_k, X_h) < \varepsilon/2 \) for \( h < h_0, \ k = 1, 2, \ldots, N \). Then \( \delta(Ax, X_h) < \varepsilon \) for \( x \in X, \|x\| = 1, h < h_0 \).

Consider the following simple example. Let \( X = L^2(0, 1) \) and \( X_h \) be the space of piecewise constant functions on the intervals \([k-1,h, kh]\), \( k = 1, 2, \ldots, 1/h \), where \( 1/h \) is an integer; \( A \) is the multiplication operator, \((Af)(t) = (\omega f)(t)\) where \( \omega \in C^0[0,1] \); by using the uniform continuity of \( \omega \), one easily verifies P3. In fact, this is particular case of a general property of finite elements which has been first used by Nitsche and Schatz [6] and which can be stated in the following way; let \( \omega \) be a smooth function on a domain \( \Omega \subset \mathbb{R}^n \), \( \{ S_h \} \) be a family of finite element subspaces of \( H^m(\Omega) \); then for \( u \in S_h, \inf_{v \in S_h} \| \omega u - v \|_{H^m} \leq c \| u \|_{H^m} \), where \( c \) depends on \( \omega \) but not on \( u \). Of course, this property has to be verified in each specific case; for triangular polynomial elements, see for example [2].

The multiplication operator, in connection with compact operators, is a basic tool for the treatment of more complicated situations. In [3], we have analyzed a one-dimensional problem with two components from plasma physics; note that the method has been applied in a very successful code used in several laboratories. In the rest of this section, we shall be concerned with a similar two-dimensional problem with three components which presents new difficulties.

Let \( \Omega = (0, 1) \times (0, 1) \subset \mathbb{R}^2 \), \( X = H^1_0(\Omega) \times L^2(\Omega) \times L^2(\Omega) \) (in the following, we shall write simply \( H^1_0, L^2 \), \( \| . \| \) be the natural norm in \( X \); \( X \) is a subset of the Hilbert space \( (L^2)^3 \) of scalar product \( (., .)_{(L^2)^3} \); for an element of \( X \), we use the notation \( \mathbf{u} = (u_1, u_2) \) where \( u_1 \in H^1_0, u_2 \in (L^2)^2 \). We introduce the following sesquilinear form on \( X \):

\[
a(\mathbf{u}, \mathbf{v}) = \int_{\Omega} \left\{ \alpha \text{grad} u_1 \cdot \text{grad} \tilde{v}_1 + \beta \text{grad} u_1 \cdot \tilde{v}_2 + \gamma u_2 \cdot \text{grad} \tilde{v}_1 + \delta u_1 \tilde{v}_1 + \theta u_2 \cdot \tilde{v}_2 + u_1 \xi \cdot \tilde{v}_2 + \eta u_2 \tilde{v}_1 \right\}; \quad (1)
\]
\( a, P, y, 5, 6, 2 \)- and \( T \) are given complex continuous functions on \( \Omega \); one supposes \( \text{Re}(\alpha) > 0 \), \( \text{Re}(\alpha - (\beta y/\theta)) > 0 \) and also a coercive on \( X \). We define \( A : X \to X \) by the relation \( a(Af, v) = (f, v)_{(L^2)} \), \( \forall f, v \in X \).

In order to get some intuitive feeling about this problem, we consider the particular case where

\[
a(u, v) = \int_\Omega \left\{ \text{grad} u_1 \cdot \text{grad} \bar{v}_1 + \text{grad} u_1 \cdot \bar{v}_2 + u_1 \cdot \text{grad} \bar{v}_1 + u_1 \bar{v}_1 + 2u_2 \cdot \bar{v}_2 \right\}; \tag{2}
\]

if \( \mu^{-1} \) is an eigenvalue of \( A \), the corresponding eigenfunction \( u \), if it is sufficiently smooth, will satisfy the system of partial differential equations:

\[
-\Delta u_1 - \text{div} u_2 = (\mu - 1) u_1; \quad \text{grad} u_1 + u_2 = (\mu - 1) u_2; \quad u_1 = 0 \text{ on } \partial \Omega;
\]
on remarks that the left member of the first equation is obtained by taking the divergence of the left member of the second equation. Let \( \alpha_1, \alpha_2, \ldots \) be the eigenvalues of the Laplacian operator, \( \phi_1, \phi_2, \ldots \in H^1_0(\Omega) \) be a corresponding total orthogonal set of eigenfunctions, i.e. \( -\Delta \phi_k = \alpha_k \phi_k \).

One easily verifies that \( A \) has a pure point spectrum composed of the eigenvalues \( \lambda = 1, \lambda = 0.5 \) and \( \lambda_k = 1/(2 + \alpha_k), \quad k = 1, 2, \ldots \); \( \{(\phi, -\partial_x \phi, -\partial_y \phi) \mid \phi \in H^1_0\} \) and \( \{(0, \partial_x \psi, -\partial_y \psi) \mid \psi \in H^1\} \) are the invariant subspaces corresponding to \( \lambda = 1 \) and \( \lambda = 0.5 \) whereas \( (\phi_k, \partial_x \phi_k, \partial_y \phi_k) \) is an eigenvector corresponding to \( \lambda_k \).

We come back to the sesquilinear form (1); in general the spectrum of \( A \) will be much more complicated than in case (2); in particular for the selfadjoint case, \( A \) will not have a pure point spectrum.

Let us now define a sequence of finite element subspaces \( X_h \). We set \( h = 1/N \), \( N \) integer; \( \Omega \) is divided in \( N^2 \) equal squares and each of these squares is subdivided in two triangles by the diagonal of positive slope; \( K_h \subset H^1 \) is the set of piecewise linear functions corresponding to this triangularization and \( S_h = K_h \cap H^1_0 \);

\[
T_{1h} = \{ (\partial_x \phi, \partial_y \phi) \mid \phi \in S_h \}, \quad T_{2h} = \{ (\partial_y \psi, -\partial_x \psi) \mid \psi \in K_h \},
\]

\( T_h = T_{1h} \oplus T_{2h} \) and finally we set \( X_h = S_h \times T_h \).

**Theorem 8:** *The conditions P2 and P3 are satisfied for the example described above.*

For the sake of briefness, but without changing the main arguments, we shall give the proof of theorem for the simplified form

\[
a(u, v) = \int_\Omega \left\{ \text{grad} u_1 \cdot \text{grad} \bar{v}_1 + \text{grad} u_1 \cdot \bar{v}_2 + \gamma u_1 \cdot \bar{v}_1 + \theta u_2 \cdot \bar{v}_2 \right\},
\]

vol. 12, no 2, 1978
where, by an argument of regularization, we can suppose, without loss of generality, that $\gamma$ and $\theta \in C^\infty(\Omega)$.

We first note that the subspaces

\[
\{(\partial_x \varphi, \partial_y \varphi) \mid \varphi \in H_0^1\} \quad \text{and} \quad \{(\partial_y \psi, -\partial_x \psi) \mid \psi \in H^1\}
\]

are orthogonal in $(L^2)^2$ and that their direct sum is precisely $(L^2)^2$ (one uses Fourier series); one easily deduces from these facts that

\[
\lim_{h \to 0} \inf_{g \in \mathcal{T}_h} ||f-g||_{(L^2)^2} = 0, \quad \forall f \in (L^2)^2
\]

and finally that the property $P_2$ is satisfied. It remains to verify $P_3$.

**Lemma 7:** Let $\omega \in C^\infty(\overline{\Omega})$. There exist $\varepsilon_h$, $\lim \varepsilon_h = 0$, such that

\[
\inf_{v \in \mathcal{T}_h} \|\omega f-v\|_{(L^2)^2} \leq \varepsilon_h \|f\|_{(L^2)^2}, \quad \forall f \in \mathcal{T}_h.
\]

**Proof:** Let

\[
G : (L^2)^2 \to H_0^1 \times H^1, \quad f \to (\varphi, \psi) \quad \text{so that} \quad f = (\partial_x \varphi + \partial_y \psi, \partial_y \varphi - \partial_x \psi)
\]

and the $L^2$-norm of $\psi$ is minimum; then $G$ is continuous and

\[
G(\mathcal{T}_h) \subset S_h \times K_h.
\]

$S_h$ and $K_h$ satisfy the Nitsche-Schatz property mentioned above and for $f \in \mathcal{T}_h$ there will exist $\xi \in S_h$ and $\eta \in K_h$ with

\[
\|\omega \varphi - \xi\|_{H^1} + \|\omega \psi - \eta\|_{H^1} \leq ch \|f\|_{(L^2)^2}.
\]

One has

\[
\omega f = (\partial_x (\omega \varphi), \partial_y (\omega \varphi)) \n + (\partial_y (\omega \psi), -\partial_x (\omega \psi)) - (\varphi \partial_x \omega, \varphi \partial_y \omega) - (\psi \partial_y \omega, -\psi \partial_x \omega);
\]

The first and second terms of the right member are approximated by $(\partial_x \xi, \partial_y \xi)$ and $(\partial_y \eta, -\partial_x \eta)$ with an error $\leq ch \|f\|_{(L^2)^2}$; for the third term, one remarks that the mapping $(L^2)^2 \to (L^2)^2, \quad f \to (\varphi \partial_x \omega, \varphi \partial_y \omega)$ is compact; by theorem 7, there exists $w \in \mathcal{T}_h$ with

\[
\|(\varphi \partial_x \omega, \varphi \partial_y \omega) - w\|_{(L^2)^2} \leq \delta_h \|f\|_{(L^2)^2},
\]

where $\lim_{h \to 0} \delta_h = 0$; the last term can be treated in a similar way. $\blacksquare$

**Proof of property $P_3$ in theorem 8:** Let $f \in X_h$, $u = Af$, i.e.: 

\[
a(u, v) = (f, v)_{(L^2)^2}, \quad \forall v \in X.
\]
Setting \( v_1 = 0 \) in (3), one obtains:

\[
\mathbf{u}_2 = \frac{1}{\theta} (\mathbf{f}_2 - \nabla u_1).
\]  

(4)

Replacing \( \mathbf{u}_2 \) in (3) by (4) and setting \( v_2 = 0 \), one gets

\[
b(u_1, v_1) = \int_{\Omega} f_1 \bar{v}_1 - \int_{\Omega} \frac{\gamma}{\theta} f_2 \cdot \nabla \bar{v}_1, \quad \forall v_1 \in H^1_0, \]

(5)

where \( b(\varphi, \psi) = \int_{\Omega} (1 - \gamma/\theta) \nabla \varphi \nabla \psi \) is a continuous and coercive [one has supposed \( \text{Re} \left( 1 - \gamma/\theta \right) > 0 \)] sesquilinear form on \( H^1_0 \). Let \( (\varphi, \psi) = G f_2 \) (defined in lemma 7) and set \( u_1 = w + \gamma/(\gamma - \theta) \varphi \); replacing in (5) \( u_1 \) by this last expression and \( f_2 \) in function of \( \varphi \) and \( \psi \), one gets after some calculations an equation for \( w \) of the form \( b(w, v) = \ldots \), \( \forall v \in H^1_0 \), where the right member depends on \( \varphi \) and \( \psi \), but not on the derivatives of \( \varphi \) and \( \psi \); one deduces that the mapping \( X \rightarrow H^1_0, f \rightarrow w \) is compact so that by theorem 7 there exists \( p \in S_h \) with \( \| p - w \|_{H^1} \leq \varepsilon_h \| f \| \), where \( \varepsilon_h \) will denote here and in the following a generic sequence converging to zero. Since \( \varphi \in S_h \), there exists \( q \in S_h \) with \( \| \gamma/\gamma - \theta \|_{H^1} \leq c h \| \varphi \|_{H^1} \); setting \( r = p + q \), one has \( \| u_1 - r \|_{H^1} \leq \varepsilon_h \| f \| \). In order to approximate \( u_2 \), one first approximates \( u_1 \) in (4) by \( r \) and apply lemma 7: there exists \( s \in T_h \) such that \( \| u_2 - s \|_{L^2} \leq \varepsilon_h \| f \| \); finally, setting \( g = (r, s) \in X_h \) one has \( \| \mathbf{u} - g \| \leq \varepsilon_h \| f \| \), which proves property P3.

**Remarks:**

1) In the proof of theorem 8 we have used several times the compactness argument of theorem 7. Supposing the coefficients \( \alpha, \beta, \ldots \) sufficiently smooth, we can avoid it and obtain, instead of P3, the estimate

\[
\sup_{f \in X_h, \| f \| = 1} \delta(A f, X_h) = O(h).
\]

2) Some elements of this example are essential; adding in the form (1) the term \( \int \partial_x u_1 \partial_y \bar{v}_1 \) changes completely the structure of the problem; on the other hand the shape of \( \Omega \) (in as much it remains simply connected), the choices of \( S_h \) and \( K_h \) play no important roles.

3) Property P2 can be strengthened by the estimate (that we shall use in part 2 of this paper):

\[
\inf_{v \in X_h} \| \mathbf{u} - v \| \leq c h \| \mathbf{u} \|_{H^2 \times (H^1)^3}, \quad \forall \mathbf{u} \in (H^1_0 \cap H^2) \times (H^1)^2.
\]
4) A priori, it would seem more natural to use, instead $X_h$, the subspaces $\tilde{X}_h = S_h \times (C_h)^2$ where $C_h$ is the set of piecewise constant functions on the triangularization. Clearly $X_h \subset \tilde{X}_h$ so that P2 is satisfied for $\tilde{X}_h$; however, in general, P1 will not be verified for $\tilde{X}_h$. More precisely, we prove in [4] the following results; let $a_1 (u, v) = \int \nabla u_1 \cdot \nabla \bar{v}_1 + \nabla u_2 \cdot \bar{v}_2 + u_1 \bar{v}_2 + u_2 \bar{v}_1$, $a_2$ be the adjoint form $a_2 (u, v) = \tilde{a}_1 (v, u)$; then P1 is verified for $a_1$ and $\tilde{X}_h$ but is not verified for $a_2$ and $\tilde{X}_h$.

REFERENCES