Jaroslav Milota

Universal almost optimal variational methods

Czechoslovak Mathematical Journal, Vol. 21 (1971), No. 2, 291–310

Czechoslovak Mathematical Journal, 21 (96) 1971, Praha

UNIVERSAL ALMOST OPTIMAL VARIATIONAL METHODS

JAROSLAV MILOTA, Praha

(Received January 8, 1970)

1. INTRODUCTION

I. BABUŠKA investigated optimal quadrature formulae for periodic functions that were universal for a large class of spaces (see [3]). In this paper I want to describe how some of his ideas may be used in the examination of variational methods for boundary-value problems. Optimal and almost optimal formulae (irrespectively of universality) were examined partly in [1], [2]. Some of the following results were published in [8].

We consider the equation

$$(1,1) \qquad Ax = f$$

where A is a positive definite linear operator which is defined on a dense set $D(A)$ in a separable Hilbert space H. The scalar product in H is denoted by $(,)$. Let H_A be the completion of $D(A)$ with respect to the scalar product

$$(1,2) \qquad (x, y)_A = (Ax, y) .$$

Under the above properties the problem (1,1) has the uniquely determined weak solution for all $f \in H$. Thus we obtain an operator that is the inverse operator to the Friedrichs self-adjoint extension $\tilde{A}$ of A (see e.g. [6]). $\tilde{A}^{-1}$ is a bounded operator from H into H_A. It is also bounded as an operator $H \to H$ (A is a positive definite operator). By the definition of a weak solution, we have

$$(1,3) \qquad (\tilde{A}^{-1}f, x)_A = (f, x)$$

for all $x \in H_A$.

In this paper I investigate two finite dimensional numerical methods – the Ritz method and the least square method. These methods can be examined with respect to the asymptotic behaviour of errors of solutions for such class of right-hand sides that the corresponding solutions belong to a compact set. In [9] I investigated almost optimal finite dimensional approximations of compact sets. The results of the above paper will be used in §§ 3–5.

The second aim of this paper is the examination of errors of the Ritz method and the least square method with respect to minimal errors for various classes of right-hand sides (which leads to the concept of universality of numerical methods). In §§ 6, 7 we shall define suitable classes of such spaces and prove theorems on universality.

2. FINITE DIMENSIONAL APPROXIMATIONS

We shall begin this section with the study of the Ritz method. It is well known that this method is a projection method in H_A. Let $L(\varphi_1, \ldots, \varphi_n)$ denote the subspace generated by $\varphi_1, \ldots, \varphi_n$ and let $P_n^{A,\Phi}$ ($\Phi = (\varphi_n)$) be the projection in H_A onto $L(\varphi_1, \ldots \ldots, \varphi_n)$. Then the n-th Ritz approximation of a weak solution $\tilde{A}^{-1}f$ of (1,1) equals to $P_n^{A,\Phi}\tilde{A}^{-1}f$ and $P_n^{A,\Phi}\tilde{A}^{-1}f \overset{H_A}{\rightarrow} \tilde{A}^{-1}f$ iff $\tilde{A}^{-1}f \in \overline{L(\Phi)}$ (we use iff instead of if and only if). The convergence theorem is valid for all $f \in H$ iff Φ is a complete sequence in H_A. If $(\varphi_n) \subset D(\tilde{A})$ and $(\tilde{A}\varphi_n)$ is a complete sequence in H then (φ_n) is a complete sequence in H_A. These statements are evident from (1,3).

Proposition 1. $H_A = H_{\tilde{A}}$.

Proof. As $\tilde{A}$ is the extension of A we have only to prove $H_{\tilde{A}} \subset H_A$. Let $x \in H_{\tilde{A}}$ and $(x_n) \subset D(\tilde{A})$ such that $x_n \overset{H_{\tilde{A}}}{\rightarrow} x$. By the definition of $\tilde{A}$ there exists $(f_n) \subset H$ such that

$$(x_n, g)_A = (f_n, g) = (x_n, g)_{\tilde{A}}$$

holds for all $g \in H_A$. This means that (x_n) is a weak fundamental sequence in H_A. Using the weak completeness of the Hilbert space now, we obtain $y \in H_A$ as the weak limit in H_A of (x_n). So $(y, g)_A = (x, g)_{\tilde{A}}$ for all $g \in H_A$. Therefore the element y is uniquely determined by x and this mapping is a one-to-one correspondence between H_A and $H_{\tilde{A}}$.

The following theorem plays an important role in our investigations. Its part can be found in [6].

Theorem 1. *Let A be a positive definite operator and $\tilde{A}$ be its self-adjoint extension. Then the following statements are equivalent.*

(i) $\tilde{A}^{-1} : H \rightarrow H_A$ *is a completely continuous operator.*

(ii) $\tilde{A}^{-1} : H \rightarrow H$ *is a completely continuous operator.*

(iii) $\tilde{A}$ *has a non-decreasing sequence* (λ_n) *of positive eigenvalues such that* $\lim_{n \to +\infty} \lambda_n = +\infty$. *The corresponding sequence* (e_n) *of orthonormal (in H) eigenfunctions is a base in H.*

(iv) *The embedding of H_A into H is completely continuous.*

Proof. 1. Owing to the continuity of the embedding of H_A into H, (i) implies (ii).

2. (ii) $\Rightarrow$ (iii). $\tilde{A}^{-1}$ is a completely continuous self-adjoint operator and therefore it has a sequence (e_n) of eigenfunctions that forms a base in $\overline{R(\tilde{A}^{-1})}$. But $D(A) \subset \subset R(\tilde{A}^{-1})$ and hence (e_n) is a base in H.

3. (iii) $\Rightarrow$ (iv). If $f \in H_A$ then $f \in H$ and, by the assumption, $f = \sum_{n=1}^{+\infty} (f, e_n) e_n$ and $\|f\|_A^2 = \sum_{n=1}^{+\infty} |(f, e_n)|^2 \lambda_n$. Let $\|f\|_A \leqq 1$ and $\varepsilon > 0$. Then

$$\sum_{i=n_0}^{+\infty} |(f, e_i)|^2 \leqq \frac{1}{\lambda_{n_0}} \sum_{i=n_0}^{+\infty} |(f, e_i)|^2 \lambda_i \leqq \frac{1}{\lambda_{n_0}} < \varepsilon$$

for sufficiently large natural n_0. So the H_A-unit sphere is a compact set in H.

4. (iv) $\Rightarrow$ (ii). This implication follows immediately from the continuity of $\tilde{A}^{-1} : H \to H_A$.

5. (iii) $\Rightarrow$ (i). Let $f \in H$, $f = \sum_{n=1}^{+\infty} (f, e_n) e_n$. Due to the continuity of $\tilde{A}^{-1}$, $\tilde{A}^{-1}f = \sum_{n=1}^{+\infty} ((f, e_n)/\lambda_n) e_n$. Similarly as in the third part of this proof, we obtain for $\|f\| \leqq 1$

$$\sum_{i=n_0}^{+\infty} \frac{|(f, e_i)|^2}{\lambda_i} \leqq \frac{1}{\lambda_{n_0}} < \varepsilon$$

for sufficiently large n_0. So the $\tilde{A}^{-1}$-image of the H-unit sphere is a compact set in H_A.

Whenever we shall deal with the Ritz method we shall suppose that some condition of Theorem 1 is satisfied. Hence $M = \tilde{A}^{-1}(S(1))$ where $S(1)$ is the unit sphere in H is a compact set in H_A. For $(\varphi_n) \subset H_A$ we can define

$$(2,1) \qquad \varrho_n^{(A)}(M; \varphi_1, \ldots, \varphi_n) = \sup_{g \in M} \|g - P_n^{\Phi,A} g\|_A$$

where $P_n^{\Phi,A}$ is the projection in H_A onto the linear subspace $L(\varphi_1, \ldots, \varphi_n)$ generated by $\varphi_1, \ldots, \varphi_n$. The best n-dimensional approximation can be determined by

$$(2,2) \qquad \varrho_n^{(A)}(M) = \inf_{\varphi_1, \ldots, \varphi_n} \varrho_n^{(A)}(M; \varphi_1, \ldots, \varphi_n).$$

The asymptotic behaviour of $\varrho_n^{(A)}(M; \varphi_1, \ldots, \varphi_n)/\varrho_n^{(A)}(M)$ gives the possibility to speak about the suitability of an approximating sequence.

Definition 1. A sequence $(\varphi_n) \subset H_A$ is said to be an *almost optimal Ritz approximation* of (1,1) if

$$(2,3) \qquad \varrho_n^{(A)}(M; \varphi_1, \ldots, \varphi_n) = O(\varrho_n^{(A)}(M)), \quad n \to +\infty.$$

In § 4 we shall show some sufficient conditions for the almost optimality.

We turn now to the least square method. Let $(\varphi_n) \subset D(A)$ be a linear independent sequence. The n-th least square approximation $y_n = \sum_{i=1}^{n} a_i \varphi_i$ of $\tilde{A}^{-1}f$ is usually determined by

$$\|Ay_n - f\| = \min_{y \in L(\varphi_1, \ldots, \varphi_n)} \|Ay - f\| ,$$

i.e. the unknown coefficients $a_1, \ldots, a_n$ are the solutions of

$$\sum_{i=1}^{n} a_i(A\varphi_i, A\varphi_j) = (f, A\varphi_j) , \quad j = 1, \ldots, n .$$

From the form of this system we can see that the least square method can be described as a projection method in the space H_{A^2} where H_{A^2} is the completion of $D(A)$ with respect to the scalar product

$$(2,4) \qquad (x, y)_{A^2} = (Ax, Ay) .$$

(See [2].) We need $\tilde{A}^{-1}f \in H_{A^2}$ only.

Proposition 2. *There exists a continuous mapping of H_{A^2} into H.*

Proof. Let $x \in D(A)$ and $Ax = f$ (i.e. $x = \tilde{A}^{-1}f$). Then

$$(2,5) \qquad \|x\|_{A^2} = \|f\|_H \geqq \frac{1}{\|\tilde{A}^{-1}\|_{H \to H}} \|x\|_H .$$

For $x \in H_{A^2}$ there exists a sequence $(x_n) \subset D(A)$ such that $x_n \xrightarrow{H_{A^2}} x$. By (2,5), (x_n) forms a fundamental sequence in H, i.e. $x_n \xrightarrow{H} y$ and

$$\|x\|_{A^2} \geqq \frac{1}{\|\tilde{A}^{-1}\|_{H \to H}} \|y\|_H .$$

Now, we have only to show that the mapping $x \to y$ is an injection. Let $y = 0$. If $x_n \xrightarrow{H_{A^2}} x$ then (Ax_n) is a fundamental sequence in H. Let $Ax_n \xrightarrow{H} z$. As $y = 0$ then $z = 0$. By the definition of the scalar product in H_{A^2}, we have

$$(x_n, f)_{A^2} = (Ax_n, Af) \to 0$$

for all $f \in D(A)$. This means that $(x, f)_{A^2} = 0$ for all $f \in D(A)$ and so $x = 0$.

Proposition 3. *$H_{A^2} = D(\bar{A})$ where $\bar{A}$ is the smallest closed extension of A and $\|x\|_{A^2} = \|\bar{A}x\|$.*

Proof. Let $x \in H_{A^2}$ and $(x_n) \subset D(A)$ such that $x_n \xrightarrow{H_{A^2}} x$. Then $x_n \xrightarrow{H} x$ and $Ax_n \xrightarrow{H} y$. This means that $x \in D(\bar{A})$ and $\bar{A}x = y$. Further

$$\|x\|_{A^2} = \|y\| = \|\bar{A}x\| .$$

On the other hand, for $x \in D(\bar{A}) \setminus D(A)$ there exists a sequence $(x_n) \subset D(A)$ such that $x_n \xrightarrow{H} x$ and $Ax_n \xrightarrow{H} \bar{A}x$. So (x_n) forms a fundamental sequence in H_{A^2} and, by Proposition 2, $x_n \xrightarrow{H_{A^2}} x$.

According to the last proposition we shall suppose A to be a self-adjoint operator (then $A = \bar{A}$) and we shall interpret the least square method as a projection method in H_{A^2}.

We remark that the convergence theorem for this method is valid for all right-hand sides iff (φ_n) is a complete sequence in H_{A^2} and this is true iff $(A\varphi_n)$ is a complete sequence in H $(R(A) = H)$.

As $A^{-1}(S(1))$ is not a compact set in H_{A^2} the situation is different from that of the Ritz method. Let $N \subset H$, $(\varphi_n) \subset H_{A^2}$ and

$$\varrho_n^{(A^2)}(A^{-1}(N); \varphi_1, \ldots, \varphi_n) = \sup_{g \in A^{-1}(N)} \|g - P_n^{\Phi,A^2} g\|_{A^2} \tag{2,6}$$

where P_n^{Φ,A^2} is the projection in H_{A^2} onto $L(\varphi_1, \ldots, \varphi_n)$. We need to suppose that N is a compact set in H to be able to speak about the convergence of (2,6) to zero (see § 5). We remark that for $\psi_n = A\varphi_n$ we have

$$\varrho_n^{(A^2)}(A^{-1}(N); \varphi_1, \ldots, \varphi_n) = \sup_{f \in N} \|f - P_n^{\Psi} f\| = \varrho_n(N; \psi_1, \ldots, \psi_n) . \tag{2,7}$$

Now we put

$$\varrho_n(N) = \varrho_n^{(A^2)}(A^{-1}(N)) = \inf_{\psi_1, \ldots, \psi_n} \varrho_n(N; \psi_1, \ldots, \psi_n)$$

and define

Definition 2. A sequence $(\varphi_n) \subset D(A)$ is called *an almost optimal least square approximation of* (1,1) for $N \subset H$ if

$$\varrho_n^{(A^2)}(A^{-1}(N); \varphi_1, \ldots, \varphi_n) = O(\varrho_n^{(A^2)}(A^{-1}(N))) , \quad n \to +\infty . \tag{2,8}$$

It is evident that (φ_n) is an almost optimal least square approximation of (1,1) for N iff (ψ_n) $(A\varphi_n = \psi_n)$ is an almost optimal approximation of N, i.e. iff

$$\varrho_n(N; \psi_1, \ldots, \psi_n) = O(\varrho_n(N)) , \quad n \to +\infty .$$

It is quite clear now that we need some knowledge of approximations of compact sets in Hilbert spaces. In the next section we shall present them using [9].

3. SOME FACTS ABOUT APPROXIMATIONS OF COMPACT SETS IN HILBERT SPACES

Let $T : H_1 \to H$ be a completely continuous operator. Then there exist orthonormal sequences $(e_n) \subset H_1, (h_n) \subset H$ and a non-increasing sequence (μ_n) of positive numbers such that

$$Tf = \sum \mu_n (f, e_n)_1 \, h_n \tag{3,1}$$

for all $f \in H_1$. If $N = T(S_1(1))$ ($S_1(1)$ is the unit sphere in H_1) then

$$\varrho_n(N) = \varrho_n(N; h_1, \ldots, h_n) = \mu_{u+1} \tag{3,2}$$

(see Theorem 1 in [9]). If $(\varphi_n) \subset H$ then we have the following statement for the convergence to zero of $\varrho_n(N; \varphi_1, \ldots, \varphi_n)$: If (φ_n) is a complete sequence in H then

$$\lim_{n \to +\infty} \varrho_n(N; \varphi_1, \ldots, \varphi_n) = 0 \tag{3,3}$$

(see the first part of Theorem 2 in [9]).

To be able to speak about an almost optimal approximation of N we need some further concepts.

A sequence (φ_n) is called strongly minimal (or strongly maximal) if there exists a positive constant c_1 (or c_2) such that for the eigenvalues $(\mu_k^{(n)})_{k=1,\ldots,n;n=1,\ldots}$ of the Gramm matrices $((\varphi_i, \varphi_j))_{i,j=1,\ldots,n;n=1,\ldots}$ the inequality $c_1 \leqq \mu_k^{(n)}$ (or $\mu_k^{(n)} \leqq c_2$) holds. If (φ_n) is strongly minimal then there exists a uniquely determined biorthogonal sequence $(\omega_n) \subset \overline{L(\Phi)}$ where $L(\Phi)$ is the linear hull of (φ_n).

The following theorem will be useful (see Theorem 3 in [9]).

Theorem 2. *Let a sequence $(\varphi_n) \subset H$ have the biorthogonal sequence (ω_n) and let (ε_n) be an orthonormal base in H. Then the following statements are equivalent.*

(A) *(φ_n) is strongly minimal.*

(B) *(ω_n) is strongly maximal.*

(C) *The operator $U_1 : H \to l^2$ which is defined by $U_1 f = ((f, \omega_n))$ is linear and bounded.*

(D) *The linear operator U_2 which is defined on $L(\Phi)$ by $U_2 \varphi_n = \varepsilon_n$ has a bounded extension on H.*

(E) *The operator U_3 which is defined on l^2 by $U_3((\alpha_n)) = \sum_{n=1}^{+\infty} \alpha_n \omega_n$ is linear and bounded.*

(F) *The operator U_4 which is defined on $L(\varepsilon)$ by $U_4 \varepsilon_n = \omega_n$ has a bounded extension on H.*

We denote by H_Φ the completion of $L(\Phi)$ with respect to the scalar product

$$(\varphi_i, \varphi_j) = \delta_{ij}\,.$$

Theorem 3. (i) *If (φ_n) is strongly minimal and its biorthogonal (ω_n) is complete then there exists a continuous embedding of H into H_Φ.*

(ii) *If (φ_n) is strongly maximal and there exists a biorthogonal sequence (ω_n) to (φ_n) then there exists a continuous embedding of H_Φ into H.*

(iii) *If (φ_n) is strongly minimal and strongly maximal and complete in H then (φ_n) and its biorthogonal (ω_n) are bases in H and H_Φ and H are topologically equivalent.*

(See Corollary 1 – 3 of Theorem 3 in [9].) If a sequence (φ_n) satisfies the assumptions of the part (iii) of Theorem 3 then it is said to be a Riesz base in H (see [4]).

The first sufficient condition for the almost optimality has the following form (see Theorem 4 in [9]):

Theorem 4. *Let T be a completely continuous operator in the form* (3,1). *Let $(\varphi_n) \subset R(T)$ constitute a Riesz base in H and let $((1/\mu_n)\, T^*\omega_n)$ be strongly maximal in H where (ω_n) is the biorthogonal base to (φ_n). Then (φ_n) is an almost optimal approximation of $N = T(S(1))$.*

Further we have (see Theorem 6 in [9]):

Theorem 5. *Let T, U be completely continuous operators such that $R(T) = R(U)$. Then (φ_n) is an almost optimal approximation for N_T iff it has the same property for N_U.*

We remark that this theorem was proved in [9] by using Lemma 2, which can be a little strenghtened in the following way.

Proposition 4. *Let T, U be linear bounded operators. Then $R(T) \subset R(U)$ iff there exists a linear bounded operator A such that $T = UA$.*

The proof is the same as that in [9]. We shall use this proposition in § 6. We shall need the following theorem in § 5 (see Theorem 7 in [9]).

Theorem 6. *Let (φ_n) be a Riesz base in H and an almost optimal approximation for N_T where T is a completely continuous operator. Let C be a linear bounded operator such that C^{-1} exists and it is also bounded and $C(R(T)) = R(T)$. Then (ψ_n) $(C\varphi_n = \psi_n)$ is an almost optimal approximation for $C^{-1}(N_T)$.*

4. ALMOST OPTIMAL RITZ APPROXIMATIONS

Let (e_n) be the sequence of eigenfunctions of $\tilde{A}$ orthonormal in H. Then for $x \in D(A)$ we have $\tilde{A}x = \sum_{n=1}^{+\infty} \lambda_n (x, e_n)\, e_n$ and $\tilde{A}^{-1} f = \sum_{n=1}^{+\infty} ((f, e_n)/\lambda_n)\, e_n$ for $f \in H$. Using (3,1), (3,2) we obtain

Theorem 7. (i) $\varrho_n^{(A)}(M) = \varrho_n^{(A)}(M; e_1, \ldots, e_n) = 1/\sqrt{\lambda_{n+1}}$.

(ii) *Let (φ_n) be a complete sequence in H_A. Then*

$$\lim_{n \to +\infty} \varrho_n^{(A)}(M; \varphi_1, \ldots, \varphi_n) = 0 . \tag{4,1}$$

On the other hand, if (4,1) *holds then (φ_n) is a complete sequence in H_A.*

But the convergence theorem does not say too much about the suitability of an approximating sequence (φ_n). Therefore we study almost optimal approximations (see Definition 1). Using Theorem 4 we have

Theorem 8. *Let* (φ_n) *constitute a Riesz base in* H_A *and let* (ω_n) *be the biorthogonal sequence to* (φ_n) *in* H_A. *Let* $[\sqrt{(\lambda_n)}\,\tilde{A}^{-1}\omega_n]$ *be strongly maximal in* H_A. *Then* (φ_n) *is an almost optimal Ritz approximation of* (1,1).

Remark. Let $(\varphi_n) \subset D(A)$. Then $[(1/\sqrt{\lambda_n})\,A\varphi_n]$ is the biorthogonal sequence in H_A to the sequence $[\sqrt{(\lambda_n)}\,\tilde{A}^{-1}\omega_n]$ and $[\sqrt{(\lambda_n)}\,\tilde{A}^{-1}\omega_n]$ is strongly maximal in H_A iff $[(1/\sqrt{\lambda_n}]\,A\varphi_n)$ is strongly minimal in H_A.

Proof. We have

$$\left[\sqrt{(\lambda_n)}\,\tilde{A}^{-1}\omega_n,\ \frac{1}{\sqrt{\lambda_k}}\,A\varphi_k\right]_A = \sqrt{\left(\frac{\lambda_n}{\lambda_k}\right)}\,(\omega_n, A\varphi_k) = \sqrt{\left(\frac{\lambda_n}{\lambda_k}\right)}\,(\omega_n, \varphi_k)_A = \delta_{n,k}\,.$$

Now, the last statement of this remark follows from the parts (A), (B) of Theorem 2.

We further note that the strong minimality of an approximating sequence (φ_n) is a necessary and sufficient condition for the stability in the sense of Mihlin [7]. The strong minimality and strong maximality of (φ_n) is a sufficient condition for the stability of Gaussian elimination for the solution of equations giving unknown coefficients (see [2]).

By Theorem 5, we immediately have

Theorem 9. *Let* A, B *be self-adjoint positive definite operators such that* $D(A) = D(B)$ *and let* $A^{-1} : H \to H_A$ *be completely continuous. Then* (φ_n) *is an almost optimal Ritz approximation of* (1,1) *iff it has the same property for the equation* $Bx = f$.

The last theorem is an extension of a theorem from [2] for similar boundary-value problems for ordinary differential equations and it provides the possibility to use eigenfunctions of more simple boundary-value problems for a good approximation.

Example. Let A be given an the differential operator

$$-\frac{\mathrm{d}}{\mathrm{d}t}\left[p(t)\frac{\mathrm{d}}{\mathrm{d}t}\,x(t)\right] + q(t)\,x(t) \tag{4,2}$$

with the boundary conditions

$$x(-1) = x(1) = 0\,. \tag{4,3}$$

We suppose that $(\mathrm{d}/\mathrm{d}t)\,p(t)$, $q(t)$ are continuous on $\langle -1, 1\rangle$ and $p(t) \geqq p_0 > 0$, $q(t) \geqq 0$ for all $t \in \langle -1, 1\rangle$. Putting $p(t) \equiv 1$, $q(t) \equiv 0$ and $p(t) = t^2 - 1$, $q(t) \equiv 1$

we get operators B, C. Operators A, B, C are self-adjoint positive definite operators in $L^2(-1, 1)$ and $D(A) = D(B) = D(C)$. Hence we can use eigenfunctions of B, C, i.e. $(\sin n\pi t/n\pi)$, $(\int_{-1}^{t} P_n(\tau)\,d\tau)$ where P_n is the n-th Legendre polynomial, as suitable approximations of (1,1).

5. ALMOST OPTIMAL LEAST SQUARE APPROXIMATIONS

The situation for the least square method is different from that for the Ritz method as $M = A^{-1}(S(1))$ is not a compact set in H_{A^2}. We shall suppose A to be a self-adjoint positive definite operator. If we define $\varrho_n^{(A^2)}(M)$ as in (2,6) then the convergence theorem does not hold. Hence we suppose $M = A^{-1}(N)$ where N is a compact set in H. So we obtain the following theorem, which is analogous to Theorem 7.

Theorem 10. *Let N be a compact set in H, A be a self-adjoint positive definite operator. Let $(\varphi_n) \subset D(A)$ such that $(A\varphi_n)$ is a complete sequence in H. Then*

$$\lim_{n \to +\infty} \varrho_n^{(A^2)}(A^{-1}(N); \varphi_1, \ldots, \varphi_n) = 0 . \tag{5,1}$$

On the other hand, if (5,1) holds and $\bigcup_{n=1}^{+\infty} nN$ is a dense set in H then $(A\varphi_n)$ is a complete sequence in H.

In the following we suppose $N = T(S_1(1))$ where $T: H_1 \to H$ is a completely continuous operator and $S_1(1)$ is the unit sphere in H_1. Let T have the form

$$Tf = \sum_{n=1}^{+\infty} \mu_n(f, g_n)\, h_n \tag{5,2}$$

for all $f \in H_1$ where (μ_n) is a non-increasing sequence of s-numbers of T, $\lim_{n \to +\infty} \mu_n = 0$ and (g_n), (h_n) are orthonormal sequences. The following theorem is a modification of Theorem 7, (i) for this case.

Theorem 11. *Under the above assumptions*

$$\varrho_n^{(A^2)}(A^{-1}(N)) = \varrho_n^{(A^2)}(A^{-1}(N), A^{-1}h_1, \ldots, A^{-1}h_n) = \mu_{n+1} \,.$$

By the definition of H_{A^2}, a sequence $(\varphi_n) \subset H_{A^2}$ is strongly minimal (strongly maximal, Riesz base) in H_{A^2} iff (ψ_n), $A\varphi_n = \psi_n$, is strongly minimal (strongly maximal, Riesz base) in H. So we have

Theorem 12. *Let N be a compact set in H. Then $(\varphi_n) \subset H_{A^2}$ is an almost optimal least square approximation of (1,1) for N iff (ψ_n) is an almost optimal approximation of N in H.*

We shall now give an analogy of Theorem 9 for the least square method. If A, B are self-adjoint positive definite operators such that $D(A) = D(B)$ then there exists a linear bounded operator C with bounded inverse such that $A^{-1} = B^{-1}C$, i.e.

$$B = CA \tag{5,3}$$

which follows from Proposition 4. The following theorem is a generalization of a theorem from [2].

Theorem 13. *Let $T: H_1 \to H$, $U: H_2 \to H$ be completely continuous operators such that $R(T) = R(U)$. Let A, B be self-adjoint positive definite operators such that $D(A) = D(B)$. Let the operator C from (5,3) map $R(T)$ onto $R(T)$. Then (φ_n) is an almost optimal least square approximation of (1,1) for N_T iff it is an almost optimal least square approximation of $Bx = f$ for N_U.*

Proof. We put $A\varphi_n = \psi_n$, $B\varphi_n = \eta_n$. According to (5,3), $\eta_n = C\psi_n$. Let us suppose (φ_n) to be an almost optimal least square approximation of (1,1) for N_T. By Theorem 12, (ψ_n) is an almost optimal approximation of N_T in H. As $R(T) = R(U)$ we can use Theorem 5 to obtain that (ψ_n) is an almost optimal approximation of N_U in H. The operator C satisfies the assumptions of Theorem 6 and therefore (η_n) is an almost optimal approximation of N_U in H. Using again Theorem 12 we get that (φ_n) is an almost optimal least square approximation of $Bx = f$ for N_U.

Remark. T can be taken as an operator of embedding W_2^k into L_2. In W_2^k, (e_n) can be chosen as an orthonormal base that is orthogonal in L_2. If B is simpler than A and such that $\varphi_n = B^{-1}e_n$ can be found then we can make use of the fact that (φ_n) is an almost optimal least square approximation of (1,1) for $A^{-1}(S_{W_2^k}(1))$.

6. UNIVERSALITY OF THE RITZ METHOD

We suppose during this section that (φ_n) satisfies all the assumptions of Theorem 8 and $(\varphi_n) \subset D(A)$. We note that eigenfunctions of a self-adjoint positive definite B for which $D(A) = D(B)$ have all necessary properties. We denote $v_n = (1/\sqrt{\lambda_n})\, A\varphi_n$. If Σ is the class of non-decreasing positive sequences $\sigma = (\sigma_n)$ we are able to define the spaces

$$H_\sigma^v = \{f = \sum_{n=1}^{+\infty} a_n v_n;\ \|f\|_{\sigma,v}^2 = \sum_{n=1}^{+\infty} |a_n|^2 \sigma_n^2 < +\infty\}\,, \tag{6,1}$$

$$H_\sigma^\Phi = \{f = \sum_{n=1}^{+\infty} b_n \varphi_n;\ \|f\|_{\sigma,\Phi}^2 = \sum_{n=1}^{+\infty} |b_n|^2 \sigma_n^2 < +\infty\}\,. \tag{6,2}$$

By Theorem 3, (iii), the space H_1^Φ where $1 = (1, \ldots)$ is topologically equivalent to H_A, i.e. there exist positive constants $\mathscr{K}_1$, $\mathscr{K}_2$ such that

$$\mathscr{K}_2 \|f\|_A \leqq \|f\|_{1,\Phi} \leqq \mathscr{K}_1 \|f\|_A \tag{6,3}$$

holds for all $f \in H_A$.

Proposition 5. *There exists a continuous embedding of H_A into H_1^v and if (v_n) is moreover strongly maximal in H then there exists a continuous embedding of H_1^v into H.*

Proof. 1. Let $f \in H_A$. Then $\sum_{n=1}^{+\infty} (f, A^{-1}\omega_n) \sqrt{(\lambda_n)}\, v_n = \sum_{n=1}^{+\infty} (f, \omega_n) \sqrt{(\lambda_n)}\, v_n$ is convergent in H_1^v because (v_n) is strongly minimal in H_A. As $(A^{-1}\omega_n)$ is complete in H_A this series is convergent to f. Hence the embedding of H_A into H_1^v is a one-to-one correspondence. By part (C) of Theorem 2, there exists a constant $\mathscr{K} > 0$ such that

$$\|f\|_{1,v} = \left[\sum_{n=1}^{+\infty} |(f, \omega_n)|^2 \lambda_n\right]^{1/2} \leqq \mathscr{K} \|f\|_A .$$

Thus the embedding is continuous.

2. Let $f = \sum_{i=1}^{n} a_i v_i$. Then $f \in H$ and, by the strong maximality of (v_n) in H, there exists a positive constant $\mathscr{C}$ such that

$$\|f\|^2 \leqq \mathscr{C} \sum_{i=1}^{n} |a_i|^2 = \mathscr{C} \|f\|_{1,v}^2 .$$

Thus the linear hull $L(v)$ of the sequence (v_n) with the norm $\| \ \|_{1,v}$ can be continuously mapped into H. As H_1^v is the closure of $L(v)$ with respect to the norm $\| \ \|_{1,v}$ we obtain the result.

The assumption of the continuity of the embedding of H_1^v into H will play an important role in what follows. Hence, this assumption will be made. It is almost obvious from part 2 of the proof of Proposition 5 that this assumption is equivalent to the strong maximality of (v_n) in H.

Remark. *Let (φ_n) be the eigenfunctions of a self-adjoint positive definite operator B for which $D(A) = D(B)$ and let (φ_n) be normed in H_B. Then (v_n) is strongly maximal in H.*

Proof. By Proposition 4, there exists a linear bounded operator $C : H \to H$ such that $B^{-1} = A^{-1} . C$ and, by the minimaximum principle of eigenvalues (see [5]), there exist positive constants c_1, c_2 such that

$$c_1 \leqq \frac{\mu_n}{\lambda_n} \leqq c_2$$

where (μ_n) denotes the sequence of eigenvalues of B. If we put $\phi_n = \varphi_n \sqrt{\mu_n}$ then (ϕ_n) is an orthonormal base in H and

$$v_n = \frac{A\varphi_n}{\sqrt{\lambda_n}} = \frac{\mu_n}{\sqrt{\lambda_n}} AB^{-1}\varphi_n = \frac{\mu_n}{\sqrt{\lambda_n}} C\varphi_n = \sqrt{\left(\frac{\mu_n}{\lambda_n}\right)}\, C\phi_n .$$

Using now part (F) of Theorem 2 we complete the proof.

Proposition 6. Let $\sigma, \tilde{\sigma} \in \Sigma$. If moreover $\sigma_n = O(\tilde{\sigma}_n)$ $(\sigma_n = o(\tilde{\sigma}_n))$ then the embedding of $H^\nu_{\tilde{\sigma}}$ into H^ν_σ is continuous (completely continuous).

Proof. Let $f \in H^\nu_{\tilde{\sigma}}$ and $f = \sum_{n=1}^{+\infty} a_n v_n$. In the case of $\sigma_n = O(\tilde{\sigma}_n)$, it is

$$\|f\|^2_{\sigma,\nu} = \sum_{n=1}^{+\infty} |a_n|^2 \sigma_n^2 \leqq \mathscr{K}^2 \sum_{n=1}^{+\infty} |a_n|^2 \tilde{\sigma}_n^2 = \mathscr{K}^2 \|f\|^2_{\tilde{\sigma},\nu}.$$

If $\sigma_n = o(\tilde{\sigma}_n)$ then for any $\varepsilon > 0$ there exists n_0 such that

$$\sum_{i=n_0+1}^{+\infty} |a_i|^2 \sigma_i^2 \leqq \varepsilon^2 \sum_{i=n_0+1}^{+\infty} |a_i|^2 \tilde{\sigma}_i^2 \leqq \varepsilon^2 \|f\|^2_{\tilde{\sigma},\nu}.$$

Hence $S^\nu_{\tilde{\sigma}}(1)$ (the unit sphere in $H^\nu_{\tilde{\sigma}}$) is a compact set in H^ν_σ.

Theorem 14. *The operator $\tilde{A}^{-1}$ is a continuous mapping of H^ν_σ $(\sigma \in \Sigma)$ onto $H^\Phi_{\sigma\sqrt{\lambda}}$ where $\sqrt{(\lambda)} = (\sqrt{(\lambda_1)}, \ldots)$.*

Proof. According to Propositions 5,6, H^ν_σ can be continuously mapped into H. Hence for $f \in H^\nu_\sigma$, $\tilde{A}^{-1} f \in H^\Phi_1$ and $\tilde{A}^{-1}$ is continuous as an operator which maps H^ν_σ into H^Φ_1. For $f = \sum_{n=1}^{+\infty} a_n v_n$ we have $\tilde{A}^{-1} f = \sum_{n=1}^{+\infty} (a_n/\sqrt{\lambda_n})\, \varphi_n$. Now it is quite clear that

$$\|f\|_{\sigma,\nu} = \|\tilde{A}^{-1} f\|_{\sigma\sqrt{\lambda},\Phi}.$$

We also obtain directly that $\tilde{A}^{-1}$ is an operator onto $H^\Phi_{\sigma\sqrt{\lambda}}$.

We denote $M_\sigma = \tilde{A}^{-1}(S^\nu_\sigma(1))$ now. By Theorem 14 and Proposition 6, M_σ is a compact set in H_A for any $\sigma \in \Sigma$. Our next aim will be to determine in H_A the error of an almost optimal approximation $P^{\Phi,A}_n$ for a compact set M_σ. To be able to compare we denote

$$\varrho^{(A)}_n(M_\sigma) = \inf_{L(\eta_1,\ldots,\eta_n)} \sup_{f \in M_\sigma} \inf_{g \in L(\eta_1,\ldots,\eta_n)} \|f - g\|_A. \tag{6,4}$$

The following proposition is a special case of more general theorem of Tihomirov (see [10]).

Proposition 7. *Let E_{n+1} be $n+1$-dimensional subspace of H_A with the norm $\|f\|_{n+1} \leqq \|f\|_A$. Let $S_{n+1}(\alpha)$ be the sphere of radius α in E_{n+1}. If*

$$S_{n+1}(\alpha) \subset M \tag{6,5}$$

then

$$\varrho^{(A)}_n(M) \geqq \alpha.$$

Proof. By (6,4), it can be found immediately that

$$\varrho^{(A)}_n(M) \geqq \varrho^{(A)}_n(M \cap E_{n+1}) \geqq \varrho^{(A)}_n(S_{n+1}(\alpha)) \geqq \varrho^{(E_{n+1})}_n(S_{n+1}(\alpha)) = \alpha \varrho^{(E_{n+1})}_n(S_{n+1}(1)) = \alpha$$

because for any $\eta_1, \ldots, \eta_n$ there exists $f \in S_{n+1}(1)$ such that $\|f\|_{n+1} = 1$ and $f \perp L(\eta_1, \ldots, \eta_n)$.

We use the symbol $\asymp$ to denote the weak equivalence, i.e. $a_n \asymp b_n$ iff there exist positive constants $\mathscr{K}_1, \mathscr{K}_2$ such that

$$\mathscr{K}_1|b_n| \leqq |a_n| \leqq \mathscr{K}_2|b_n|$$

for all natural n.

Theorem 15. *The following asymptotic formula is valid*:

$$\varrho_n^{(A)}(M_\sigma) \asymp \sup_{f \in M_\sigma} \|f - P_n^{\Phi,A} f\|_A \asymp \frac{1}{\sigma_{n+1}\sqrt{\lambda_{n+1}}}\,. \tag{6,6}$$

Proof. 1. We use Proposition 7 to obtain a lower bound of $\varrho_n^{(A)}(M_\sigma)$. Let $E_{n+1} = L(\varphi_1, \ldots, \varphi_{n+1})$, $\|f\|_{n+1} = \|f\|_A$. Then $E_{n+1} = \tilde{A}^{-1}(L(v_1, \ldots, v_n))$. Putting $\alpha = 1/(\mathscr{K}_1 \sqrt{(\lambda_{n+1})}\, \sigma_{n+1})$ where $\mathscr{K}_1$ is a constant from (6,3), we have for

$$f = \sum_{i=1}^{n+1} a_i \varphi_i \in S_{n+1}(\alpha),\; g = Af = \sum_{i=1}^{n+1} a_i \sqrt{(\lambda_i)}\, v_i$$

and

$$\|g\|_{\sigma,v}^2 = \sum_{i=1}^{n+1} |a_i|^2 \lambda_i \sigma_i^2 \leqq \lambda_{n+1}\sigma_{n+1}^2 \sum_{i=1}^{n+1} |a_i|^2 =$$
$$= \lambda_{n+1}\sigma_{n+1}^2 \|f\|_{1,\Phi}^2 \leqq \mathscr{K}_1^2 \lambda_{n+1}\sigma_{n+1}^2 \|f\|_A^2 \leqq 1\,.$$

Hence (6,5) holds and

$$\varrho_n^{(A)}(M_\sigma) \geqq \frac{1}{\mathscr{K}_1 \sigma_{n+1}\sqrt{\lambda_{n+1}}}\,.$$

2. By (6,4), the inequality

$$\sup_{f \in M_\sigma} \|f - P_n^{\Phi,A} f\|_A \geqq \varrho_n^{(A)}(M_\sigma)$$

holds and we have to estimate the left-hand side in (6,6) to complete the proof. Let $g = \sum_{n=1}^{+\infty} a_n v_n \in S_\sigma^v(1)$. Then $f = \tilde{A}^{-1} g = \sum_{n=1}^{+\infty} (a_n/\sqrt{\lambda_n})\, \varphi_n$ and, by the definitions of projection and the strong maximality of (φ_n) in H_A, we have

$$\|f - P_n^{\Phi,A} f\|_A \leqq \left\| \sum_{i=n+1}^{+\infty} \frac{a_i}{\sqrt{\lambda_i}} \varphi_i \right\|_A \leqq \mathscr{C} \left[\sum_{i=n+1}^{+\infty} \frac{|a_i|^2}{\lambda_i} \right]^{1/2} \leqq \tag{6,7}$$
$$\leqq \frac{\mathscr{C}}{\sqrt{(\lambda_{n+1})}\, \sigma_{n+1}} \left[\sum_{k=1}^{+\infty} |a_k|^2 \sigma_k^2\right]^{1/2} \leqq \frac{\mathscr{C}}{\sqrt{(\lambda_{n+1})}\, \sigma_{n+1}}\,.$$

As a simple application of the last theorem we have

Theorem 16. *Let A be in the form (4,2), (4,3) on the interval (0, 1) and let the right-hand side f of (1,1) have all derivatives and let its compact support lie on (0,1) and $\|f^{(n+1)}\|_{L^2}/\|f^{(n)}\|_{L^2} = o(n)$. Then*

$$\|A^{-1}f - P_n^{\Phi,A}A^{-1}f\|_A = O\left(\frac{\|f^{(n+1)}\|_{L^2}}{n!}\right).$$

Proof. We put $H = L^2(0,1)$ and, according to the assumptions of f, f belongs to H_A and, by Proposition 5, f is also an element of H_1^{ν}. Therefore $f = \sum_{n=1}^{+\infty} a_n v_n$ where $a_n = \sqrt{(\lambda_n)}\,(f, A^{-1}\omega_n)_A = \sqrt{(\lambda_n)}\,(f, \omega_n)$. We denote

$$\omega_n^{(-k)}(t) = \int_0^t \left[\int_0^{\tau_k} \dots \int_0^{\tau_2} \omega_n(\tau_1)\,d\tau_1 \dots\right] d\tau_k\,.$$

Using the Hölder inequality we obtain

$$|\omega_n^{(-n)}(t)| \leqq \mathscr{K}\,\frac{2^{n-1}}{(2n-1)!!}\,t^{n-1/2}$$

where $(2n-1)!! = 1.3.5\dots(2n-1)$ and (ω_n) is bounded in H by $\mathscr{K}$. We remark that $\mathscr{K}$ does not depend on n. Now

$$(f, \omega_n) = \int_0^1 f(t)\,\omega_n(t)\,dt = (-1)^n \int_0^1 f^{(n)}(t)\,\omega_n^{(-n)}(t)\,dt$$

and, again applying the Hölder inequality, we have

$$|(f, \omega_n)| \leqq \mathscr{K}\|f^{(n)}\|\,\frac{2^{n-1}}{(2n-1)!!\,\sqrt{2n}}\,.$$

If we put

$$\sigma_n = \frac{(2n)!}{2^{2n}n!\,\sqrt{(n)}\,\|f^{(n)}\|\,\sqrt{\lambda_n}}$$

then $\sum_{n=1}^{+\infty} |(f, \omega_n)|^2\,\lambda_n\sigma_n^2$ is convergent and hence $f \in H_\sigma^{\nu}$. We use now (6,7) to obtain

$$\|A^{-1}f - P_n^{\Phi,A}A^{-1}f\|_A \leqq \mathscr{C}\,\frac{2^{2n+2}(n+1)!\,\sqrt{(n+1)}\,\|f^{(n+1)}\|}{(2n+2)!}\,.$$

By the Stirling formulae $n! \asymp (n/e)^n\sqrt{n}$ we have

$$\frac{2^{2n+2}\sqrt{(n+1)}\,(n+1)!}{(2n+2)!} \asymp \frac{1}{n!}\,.$$

The method of the proof of Theorem 15 gives the following theorem:

Theorem 17. *Let σ, $\tilde{\sigma}$ be such elements of Σ that $1 \leqq \tilde{\sigma}_1/\sigma_1 \lesseqq \ldots$ and $\lim_{n\to+\infty} \tilde{\sigma}_n/\sigma_n = +\infty$. Then*

$$\varrho_n^{(H^{\Phi}{}_{\sigma\sqrt{\lambda}})}(M_{\tilde{\sigma}}) = \varrho_n^{(H^{\Phi}{}_{\sigma\sqrt{\lambda}})}(M_{\tilde{\sigma}}; \varphi_1, \ldots, \varphi_n) = \frac{\sigma_{n+1}}{\tilde{\sigma}_{n+1}} . \tag{6,8}$$

Proof. 1. We put $E_{n+1} = L(\varphi_1, \ldots, \varphi_{n+1})$, $\|f\|_{n+1} = \|f\|_{\sigma\sqrt{\lambda},\Phi}$ and $\alpha = \sigma_{n+1}/\tilde{\sigma}_{n+1}$. Then for $f = \sum_{k=1}^{n+1} a_k\varphi_k \in S_{n+1}(\alpha)$ we have $g = Af = \sum_{k=1}^{n+1} a_k \sqrt{(\lambda_k)}\, v_k$ and

$$\|g\|_{\tilde{\sigma},v}^2 = \sum_{k=1}^{n+1} |a_k|^2 \lambda_k \tilde{\sigma}_k^2 \leqq \frac{\tilde{\sigma}_{n+1}^2}{\sigma_{n+1}^2} \|f\|_{\sigma\sqrt{\lambda},\Phi}^2 \leqq 1 .$$

So $S_{n+1}(\alpha) \subset M_{\tilde{\sigma}}$ and, by Proposition 7,

$$\varrho_n^{(H^{\Phi}{}_{\sigma\sqrt{\lambda}})}(M_{\tilde{\sigma}}) \geqq \frac{\sigma_{n+1}}{\tilde{\sigma}_{n+1}} .$$

2. Let $S_{\tilde{\sigma}}^{v}(1)$ denote the unit sphere in $H_{\tilde{\sigma}}^{v}$. If $f = \sum_{k=1}^{+\infty} a_k v_k \in S_{\tilde{\sigma}}^{v}(1)$ then $\tilde{A}^{-1}f = \sum_{k=1}^{+\infty} (a_k/\sqrt{\lambda_k})\, \varphi_k$ and thus $P_n^{\Phi,\sigma\sqrt{\lambda}}\tilde{A}^{-1}f = \sum_{k=1}^{n} (a_k/\sqrt{\lambda_k})\, \varphi_k$ where $P_n^{\Phi,\sigma\sqrt{\lambda}}$ is the projection in $H_{\sigma\sqrt{\lambda}}^{\Phi}$ onto $L(\varphi_1, \ldots, \varphi_n)$. As

$$\|\tilde{A}^{-1}f - P_n^{\Phi,\sigma\sqrt{\lambda}}\tilde{A}^{-1}f\|_{\sigma\sqrt{\lambda},\Phi}^2 = \sum_{k=n+1}^{+\infty} |a_k|^2 \sigma_k^2 \leqq \frac{\sigma_{n+1}^2}{\tilde{\sigma}_{n+1}^2} \|f\|_{\tilde{\sigma},v}^2 \leqq \frac{\sigma_{n+1}^2}{\tilde{\sigma}_{n+1}^2} ,$$

the estimate

$$\varrho_n^{(H^{\Phi}{}_{\sigma\sqrt{\lambda}})}(M_{\tilde{\sigma}}; \varphi_1, \ldots, \varphi_n) \leqq \frac{\sigma_{n+1}}{\tilde{\sigma}_{n+1}}$$

is valid. Arguments 1 and 2 prove (6,8).

For the definition of a universal approximation we consider a system $(H_\beta)_{\beta\in B}$ of Hilbert spaces H_β which are partially ordered by the relation $\prec$ where $H_\beta \prec H_{\bar{\beta}}$ iff there exists a completely continuous embedding of $H_{\bar{\beta}}$ into H_β.

Definition 3. Let $(H_\beta^1)_{\beta\in B}$, $(H_\gamma^2)_{\gamma\in\Gamma}$ be systems of Hilbert spaces. Let T be a linear continuous operator which maps H_β^1 into H_γ^2. Let (L_n) be a sequence of linear operators: $H_\beta^1 \to L(\psi_1, \ldots, \psi_n) \subset H_\gamma^2$. Then (L_n) is said to be a *universal approximation* of T with respect to $(H_\beta^1)_{\beta\in B}$, $(H_\gamma^2)_{\gamma\in\Gamma}$, (ψ_n) if there exists a constant $\mathscr{K}$ such that

$$\frac{\sup_{f\in S_{\bar{\beta}}^1(1)} \|Tf - L_n f\|_{H_\gamma^2}}{\varrho_n^{(H_\gamma^2)}(T(S_{\bar{\beta}}^1(1)); \psi_1, \ldots, \psi_n)} \leqq \mathscr{K} \tag{6,9}$$

holds for all natural n and $H_\beta^1 \prec H_{\bar{\beta}}^1$.

In (6,9) $S_{\beta}^{1}(1)$ stands for the unit sphere in H_{β}^{1}. It follows from the above assumptions that $T(S_{\beta}^{1}(1))$ is a compact set in H_{γ}^{2}.

Theorem 18. *The sequence* $(P_n^{\Phi,A}\tilde{A}^{-1})$ *is a universal approximation of* $\tilde{A}^{-1}$ *with respect to* $(H_{\sigma}^{\nu})_{\sigma\in\Sigma}$, $(H_{\sigma\sqrt{\lambda}}^{\Phi})_{\sigma\in\Sigma}$, (φ_n).

Proof. For $f = \sum_{n=1}^{+\infty} a_n v_n \in H_{\sigma}^{\nu}$ we have $\tilde{A}^{-1}f \in H_{\sigma\sqrt{\lambda}}^{\Phi}$ (see Theorem 14) and

$$\tilde{A}^{-1}f = P_n^{\Phi,A}\tilde{A}^{-1}f + P_n^{\Phi,\sigma\sqrt{\lambda}}(I - P_n^{\Phi,A})\,\tilde{A}^{-1}f + (I - P_n^{\Phi,\sigma\sqrt{\lambda}})\,\tilde{A}^{-1}f$$

where I denotes the unit operator. Therefore (we can suppose without a loss of generality that $\tilde{A}^{-1}f \notin L(\varphi_1, \ldots, \varphi_n)$)

$$\frac{\|(I - P_n^{\Phi,A})\,\tilde{A}^{-1}f\|_{\sigma\sqrt{\lambda},\Phi}^2}{[\varrho_n^{(H^{\Phi}{}_{\sigma\sqrt{\lambda}})}(\tilde{A}^{-1}f;\varphi_1,\ldots,\varphi_n)]^2} = 1 + \frac{\|P_n^{\Phi,\sigma\sqrt{\lambda}}(I - P_n^{\Phi,A})\,\tilde{A}^{-1}f\|_{\sigma\sqrt{\lambda},\Phi}^2}{[\varrho_n^{(H^{\Phi}{}_{\sigma\sqrt{\lambda}})}(\tilde{A}^{-1}f;\varphi_1,\ldots,\varphi_n)]^2}\,.$$

Now

$$\varrho_n^{(H^{\Phi}{}_{\sigma\sqrt{\lambda}})}(\tilde{A}^{-1}f;\varphi_1,\ldots,\varphi_n) = \Big[\sum_{k=n+1}^{+\infty} |a_k|^2\,\sigma_k^2\Big]^{1/2} \geqq$$
$$\geqq \sigma_{n+1}\sqrt{(\lambda_{n+1})}\,\varrho_n^{(H_1{}^{\Phi})}(\tilde{A}^{-1}f;\varphi_1,\ldots,\varphi_n)$$

and for all $g = \sum_{k=1}^{+\infty} b_k\varphi_k \in H_{\sigma\sqrt{\lambda}}^{\Phi}$ it is

$$\|P_n^{\Phi,\sigma\sqrt{\lambda}}g\|_{\sigma\sqrt{\lambda},\Phi}^2 = \sum_{k=1}^{n} |b_k|^2\,\lambda_k\sigma_k^2 \leqq \sigma_{n+1}^2\lambda_{n+1}\|P_n^{\Phi,1}g\|_{1,\Phi}^2\,.$$

Hence we have obtained

$$(6,10)\qquad \frac{\|(I - P_n^{\Phi,1})\,\tilde{A}^{-1}f\|_{\sigma\sqrt{\lambda},\Phi}}{\varrho_n^{(H^{\Phi}{}_{\sigma\sqrt{\lambda}})}(\tilde{A}^{-1}f;\varphi_1,\ldots,\varphi_n)} \leqq \frac{\|(I - P_n^{\Phi,A})\,\tilde{A}^{-1}f\|_{1,\Phi}}{\varrho_n^{(H_1{}^{\Phi})}(\tilde{A}^{-1}f;\varphi_1,\ldots,\varphi_n)}\,.$$

Using now (6,3) and the definition of projection we have

$$\|(I - P_n^{\Phi,A})\,\tilde{A}^{-1}f\|_{1,\Phi} \leqq \mathscr{K}_1\|(I - P_n^{\Phi,A})\,\tilde{A}^{-1}f\|_A$$

and

$$\varrho_n^{(H_1{}^{\Phi})}(\tilde{A}^{-1}f;\varphi_1,\ldots,\varphi_n) = \|(I - P_n^{\Phi,1})\,\tilde{A}^{-1}f\|_{1,\Phi} \geqq \mathscr{K}_2\|(I - P_n^{\Phi,1})\,\tilde{A}^{-1}f\|_A \geqq$$
$$\geqq \mathscr{K}_2\|(I - P_n^{\Phi,A})\,\tilde{A}^{-1}f\|_A\,.$$

From this and (6,10) we can find

$$\frac{\|(I - P_n^{\Phi,A})\,\tilde{A}^{-1}f\|_{\sigma\sqrt{\lambda},\Phi}}{\varrho_n^{(H^{\Phi}{}_{\sigma\sqrt{\lambda}})}(\tilde{A}^{-1}f;\varphi_1,\ldots,\varphi_n)} \leqq \frac{\mathscr{K}_1}{\mathscr{K}_2}$$

which gives (6,9).

Remark. Theorems 17, 18 show that $P_n^{\Phi,A}\tilde{A}^{-1}$ yields also an almost optimal approximation of $\tilde{A}^{-1}$ in $H^{\Phi}_{\sigma\sqrt{\lambda}}$, $\sigma \in \Sigma$.

7. UNIVERSALITY OF THE LEAST SQUARE METHOD

We shall suppose (ψ_n) to be a Riesz base in H. In this case (φ_n) $(A\varphi_n = \psi_n)$ will be a Riesz base in H_{A^2}. We define the system of spaces $(H^{\Phi}_{\sigma})_{\sigma\in\Sigma}$ as in (6,2) and

$$(7,1) \qquad H^{\Psi}_{\sigma} = \{f = \sum_{n=1}^{+\infty} a_n\psi_n;\ \|f\|^2_{\sigma,\Psi} = \sum_{n=1}^{+\infty} |a_n|^2\,\sigma_n^2 < +\infty\}\,.$$

By Theorem 3, (iii), H^{Φ}_1 and H_{A^2} are topologically equivalent, i.e. there exist positive constants $\mathscr{K}^{(1)}$, $\mathscr{K}^{(2)}$ such that

$$(7,2) \qquad \mathscr{K}^{(2)}\|f\|_{A^2} \leqq \|f\|_{1,\Phi} \leqq \mathscr{K}^{(1)}\|f\|_{A^2}$$

holds for all $f \in H_{A^2}$. The same statement holds for H^{Ψ}_1 and H. Proposition 6 is also true for spaces H^{Ψ}_{σ}. Instead of Theorem 14 we have the following theorem on the regularity of solutions.

Theorem 19. *The operator A^{-1} is a continuous mapping of $H^{\Psi}_{\sigma}(\sigma \in \Sigma)$ onto H^{Φ}_{σ}.*

Proof. By Proposition 6, $H^{\Psi}_{\sigma}(\sigma \in \Sigma)$ is continuously mapped into H and $A^{-1}f \in H_{A^2}$ for $f = \sum_{n=1}^{+\infty} a_n\psi_n$ and $A^{-1}f = \sum_{n=1}^{+\infty} a_n\varphi_n$. The rest of the statement is obvious from the definitions of the norms.

Using this theorem we can prove theorems analogous to those in § 6.

Theorem 20. *Let $\sigma \in \Sigma$ such that $\lim_{n\to+\infty} \sigma_n = +\infty$. Then*

$$(7,3) \qquad \varrho_n^{(A^2)}(A^{-1}(S^{\Psi}_{\sigma}(1))) \asymp \varrho_n^{(A^2)}(A^{-1}(S^{\Psi}_{\sigma}(1));\ \varphi_1, \ldots, \varphi_n) \asymp \frac{1}{\sigma_{n+1}}\,.$$

Proof. The proof is almost the same as that of Theorem 15.

1. Put $E_{n+1} = L(\varphi_1, \ldots, \varphi_{n+1})$ and $\alpha = 1/\mathscr{K}^{(1)}\sigma_{n+1}$ where $\mathscr{K}^{(1)}$ is the constant from (7,2). Then

$$\|f\|^2_{\sigma,\Phi} = \sum_{i=1}^{n+1} |a_i|^2\,\sigma_i^2 \leqq \sigma_{n+1}^2\|f\|^2_{1,\Phi} \leqq \mathscr{K}^{(1)^2}\sigma_{n+1}^2\|f\|^2_A \leqq 1$$

for $f = \sum_{i=1}^{n+1} a_i\varphi_i \in S^{(A^2)}_{n+1}(\alpha)$.

By Proposition 7, we have

$$\varrho_n^{(A^2)}(S_\sigma^\Phi(1)) \geqq \frac{1}{\mathcal{K}^{(1)}\sigma_{n+1}}\,.$$

2. Let $f = \sum_{i=1}^{+\infty} a_i\varphi_i \in S_\sigma^\Phi(1)$. By the definition of projection and (7,2), it is

$$(7,4) \qquad \|f - P_n^{\Phi,A^2}f\|_{A^2} \leqq \|f - P_n^{\Phi,1}f\|_{A^2} \leqq \frac{1}{\mathcal{K}^{(2)}}\,\|f - P_n^{\Phi,1}f\|_{1,\Phi} =$$

$$= \frac{1}{\mathcal{K}^{(2)}}\,[\sum_{i=n+1}^{+\infty} |a_i|^2]^{1/2} \leqq \frac{1}{\mathcal{K}^{(2)}\sigma_{n+1}}\,[\sum_{i=n+1}^{+\infty} |a_i|^2\,\sigma_i^2]^{1/2} \leqq$$

$$\leqq \frac{1}{\mathcal{K}^{(2)}\sigma_{n+1}}\,\|f\|_{\sigma,\Phi} \leqq \frac{1}{\mathcal{K}^{(2)}\sigma_{n+1}}\,.$$

The proof is completed.

Corollary. *Under the above assumptions of Theorem* 16

$$\|A^{-1}f - P_n^{\Phi,A^2}A^{-1}f\|_{A^2} = O\left(\frac{\|f^{(n+1)}\|_{L^2}}{n!}\right), \quad n \to +\infty$$

is true.

Proof. We mention the necessary changes in the proof of Theorem 16. In the case of the least square method we have

$$\|A^{-1}f - P_n^{\Phi,A^2}A^{-1}f\|_{A^2} = \|f - P_n^\Psi f\|\,.$$

Being a Riesz base in H, (ψ_n) has the biorthogonal sequence (η_n) that is bounded in H. As $f \in H$ we can write $f = \sum_{n=1}^{+\infty} (f, \eta_n)\,\psi_n$. The estimate of (f, η_n) is the same. For sufficiently large natural n we can put

$$\sigma_n = \frac{(2n)!}{2^{2n} n!\,\sqrt{(n)}\,\|f^{(n)}\|}$$

again. The rest of the proof follows the same lines except of using (7,4) instead of (6,7).

Theorem 21. *Let σ, $\tilde\sigma$ be such elements of Σ that $1 \leqq \tilde\sigma_1/\sigma_1 \leqq \ldots$ and $\lim_{n\to+\infty} \tilde\sigma_n/\sigma_n = +\infty$. Then*

$$(7,5) \qquad \varrho_n^{(H_\sigma\Phi)}(A^{-1}(S_{\tilde\sigma}^\Psi(1))) = \varrho_n^{(H_\sigma\Phi)}(A^{-1}(S_{\tilde\sigma}^\Psi(1)); \varphi_1, \ldots, \varphi_n) = \frac{\sigma_{n+1}}{\tilde\sigma_{n+1}}\,.$$

Proof. 1. Put $E_{n+1} = L(\varphi_1, \ldots, \varphi_{n+1})$ and $\alpha = \sigma_{n+1}/\tilde{\sigma}_{n+1}$. For

$$f = \sum_{k=1}^{n+1} a_k \varphi_k \in S_{n+1}^{\sigma,\Phi}(\alpha) \quad \text{we have} \quad g = Af = \sum_{k=1}^{n+1} a_k \psi_k$$

and

$$\|g\|_{\tilde{\sigma},\Psi}^2 = \sum_{k=1}^{n+1} |a_k|^2 \tilde{\sigma}_k^2 \leqq \frac{\tilde{\sigma}_{n+1}^2}{\sigma_{n+1}^2} \sum_{k=1}^{n+1} |a_k|^2 \sigma_k^2 = \frac{\tilde{\sigma}_{n+1}^2}{\sigma_{n+1}^2} \|f\|_{\sigma,\Phi}^2 \leqq 1,$$

which means that $S_{n+1}^{\sigma,\Phi}(\alpha) \subset A^{-1}(S_{\tilde{\sigma}}^{\Psi}(1))$ and, by Proposition 7,

$$\varrho_n^{(H_\sigma \Phi)}(A^{-1}(S_{\tilde{\sigma}}^{\Psi}(1))) \geqq \frac{\sigma_{n+1}}{\tilde{\sigma}_{n+1}}.$$

2. If $f = \sum_{k=1}^{+\infty} a_k \psi_k \in S_{\tilde{\sigma}}^{\Psi}(1)$ then $A^{-1}f = \sum_{k=1}^{+\infty} a_k \varphi_k$ and $P_n^{\Phi,\sigma} A^{-1} f = \sum_{k=1}^{n} a_k \varphi_k$, i.e.

$$\|A^{-1}f - P_n^{\Phi,\sigma} A^{-1} f\|_{\sigma,\Phi}^2 = \sum_{k=n+1}^{+\infty} |a_k|^2 \sigma_k^2 \leqq \frac{\sigma_{n+1}^2}{\tilde{\sigma}_{n+1}^2} \sum_{k=n+1}^{+\infty} |a_k|^2 \tilde{\sigma}_k^2 \leqq \frac{\sigma_{n+1}^2}{\tilde{\sigma}_{n+1}^2}.$$

Hence

$$\varrho_n^{(H_\sigma \Phi)}(A^{-1}(S_{\tilde{\sigma}}^{\Psi}(1)); \varphi_1, \ldots, \varphi_n) \leqq \frac{\sigma_{n+1}}{\tilde{\sigma}_{n+1}}.$$

The proof is complete.

Theorem 22. *The sequence $(P_n^{\Phi,A^2} A^{-1} f)$ is a universal approximation of A^{-1} with respect to $(H_\sigma^{\Psi})_{\sigma\in\Sigma}$, $(H_\sigma^{\Phi})_{\sigma\in\Sigma}$, (φ_n).*

Proof. The proof is again analogous to that of Theorem 18. By Theorem 19, A^{-1} is again a linear bounded operator: $H_\sigma^{\Psi} \to H_\sigma^{\Phi}$ for all $\sigma \in \Sigma$. Let $H_{\tilde{\sigma}}^{\psi}$ be continuously mapped into H_σ. Then

$$A^{-1}f = P_n^{\Phi,A^2} A^{-1} f + P_n^{\Phi,\sigma}(I - P_n^{\Phi,A^2}) A^{-1} f + (I - P_n^{\Phi,\sigma}) A^{-1} f$$

for $f = \sum_{k=1}^{+\infty} a_k \psi_k \in H_{\tilde{\sigma}}^{\Psi} \setminus L(\psi_1, \ldots, \psi_n)$ and (see the proof of Theorem 18)

$$\frac{\|(I - P_n^{\Phi,A^2}) A^{-1} f\|_{\sigma,\Phi}^2}{[\varrho_n^{(H_\sigma \Phi)}(A^{-1}f; \varphi_1, \ldots, \varphi_n)]^2} = 1 + \frac{\|P_n^{\Phi,\sigma}(I - P_n^{\Phi,A^2}) A^{-1} f\|_{\sigma,\Phi}^2}{[\varrho_n^{(H_\sigma \Phi)}(A^{-1}f; \varphi_1, \ldots, \varphi_n)]^2} \leqq$$

$$\leqq \frac{\|(I - P_n^{\Phi,A^2}) A^{-1} f\|_{1,\Phi}}{[\varrho_n^{(H_1 \Phi)}(A^{-1}f; \varphi_1, \ldots, \varphi_n)]^2}.$$

With respect to the topological equivalence of H_{A^2} and H_1^Φ we obtain

$$(7,6) \qquad \frac{\|(I - P_n^{\Phi,A^2}) A^{-1}f\|_{1,\Phi}}{\varrho_n^{(H_1{}^\Phi)}(A^{-1}f; \varphi_1, \ldots, \varphi_n)} \leqq \frac{\mathscr{K}^{(1)}}{\mathscr{K}^{(2)}}$$

where $\mathscr{K}^{(1)}$, $\mathscr{K}^{(2)}$ are the constants from (7,2).

Corollary. $(P_n^{\Phi,A^2}A^{-1}f)$ *is an almost optimal least square approximation of* (1,1) *even in* H_σ^Φ, $\sigma \in \Sigma$ *and*

$$\sup_{\|f\|_{\tilde{\sigma},\Psi} \leqq 1} \|A^{-1}f - P_n^{\Phi,A^2}A^{-1}f\|_{\sigma,\Phi} \leqq \frac{\mathscr{K}^{(1)}}{\mathscr{K}^{(2)}} \frac{\sigma_{n+1}}{\tilde{\sigma}_{n+1}}$$

if $H_{\tilde{\sigma}}^\Psi$ *can be mapped into* H_σ^Ψ *completely continuously.*

Proof. We obtain the result if we use Theorems 21, 22 and (7,5), (7,6).

References

[1] *Babuška I., Sobolev S. L.:* Optimization of numerical methods (in Russian), Apl. matematiky *10* (1965), 2, 96–129.

[2] *Babuška I., Práger M., Vitásek E.:* Numerical processes in differential equations, Praha 1966.

[3] *Babuška I.:* Über universal optimale Quadraturformeln, Apl. matematiky *13* (1968), 304—338, 388–404.

[4] *Bari N. K.:* Biorthogonal systems and bases in Hilbert spaces (in Russian), Utschennye zapiski MGU *4* (1951), 69–107.

[5] *Dunford N., Schwartz T. J.:* Linear operators, Part II, New York 1963.

[6] *Mihlin S. G.:* Problem of the minimum of a quadratic functional (in Russian), Moskva 1952.

[7] *Mihlin S. G.:* Concrete realization of variational methods (in Russian), Moskva 1966.

[8] *Milota J.:* Universal almost optimal formulae for solutions of boundary-value problems for ordinary differential equations, Apl. matematiky 13 (1968), *2*, 163–167.

[9] *Milota J.:* Almost optimal approximations of compact sets in Hilbert spaces, Comm. Math. Univ. Carolinae *10* (1969), 1, 121–140.

[10] *Tihomirov V. M.:* The best approximations of sets in functional spaces and the theory of the best approximation (in Russian), Uspehi matem. nauk *15* (1960), 3, 81–120.

Author's address: Praha 8 - Karlín, Sokolovská 83, ČSSR (Matematicko-fyzikální fakulta KU).