MATHEMATICS OF COMPUTATION, VOLUME 32, NUMBER 144
OCTOBER 1978, PAGES 947–954

Simultaneous Approximation in Scales of Banach Spaces*

By James H. Bramble and Ridgway Scott

Abstract. The problem of verifying optimal approximation simultaneously in different norms in a Banach scale is reduced to verification of optimal approximation in the highest order norm. The basic tool used is the Banach space interpolation method developed by Lions and Peetre. Applications are given to several problems arising in the theory of finite element methods.

1. Introduction. In many papers concerning the mathematical analysis of finite element methods, certain approximation properties are assumed. In particular, it is often supposed that a given function may be approximated by a function in another space and that this approximation is "optimal" simultaneously in different norms. More precisely, let Ω be a bounded domain in $\mathbf{R}^N$ and $H^s = W_2^s(\Omega)$ the Sobolev space of order s with norm $\|\cdot\|_s$ (cf. Lions and Magenes [9]). Let k and r be positive integers with $k < r$, and let $\{S_h: 0 < h < 1\}$ be a family of subspaces of H^k. The following hypothesis is often made (cf. Babuška [1], Baker [3], Bramble and Thomée [6], Douglas and Dupont [8], Natterer [10], Nitsche and Schatz [11], Raviart [12], and Schultz [13]):

A. There exists a constant C_A such that, for $u \in H^r$ and $h \in]0, 1[$,

$$\inf_{\chi \in S_h} \left\{ \sum_{j=s}^{k} h^j \|u - \chi\|_j \right\} \leqslant C_A h^r \|u\|_r.$$

Here, s is some integer (positive or negative) less than k. We show as a particular case of our main result that, under mild restrictions on the boundary of Ω, the above statement is equivalent to the following:

B. There exists a constant C_B such that, for $u \in H^r$ and $h \in]0, 1[$,

$$\inf_{\chi \in S_h} h^k \|u - \chi\|_k \leqslant C_B h^r \|u\|_r.$$

More precisely, we show that assumption B implies A with s allowed to be arbitrary ($s < k$) and C_A depending only on C_B, r, and s; the implication A $\Rightarrow$ B is obvious.

Another consequence of our theorem in the next section is a simultaneous approximation version of the results of Babuška and Kellogg [2]: If B holds and if $u \in H^m$ for some m satisfying $k \leqslant m < r$, then

$$\inf_{\chi \in S_h} \sum_{j=s}^{k} h^j \|u - \chi\|_j = o(h^m) \tag{1.1}$$

Received March 10, 1977; revised January 6, 1978.
AMS (MOS) subject classifications (1970). Primary 65N30, 41A65.
* Work performed under the auspices of the Energy Research and Development Administration.

rather than just $O(h^m)$. These results have obvious generalizations to noninteger Sobolev spaces as well, and the more general case is treated in Section 4. The integer case was discussed here only for the purpose of exposition.

The following example indicates the nontriviality of the implication B $\Rightarrow$ A. Let u_h denote the projection of u onto S_h with respect to the H^k inner product:

$$\|u - u_h\|_k = \inf_{\chi \in S_h} \|u - \chi\|_k. \tag{1.2}$$

Then $u - u_h$ is orthogonal to S_h, and we have

$$\|u - u_h\|_k^2 = (u - u_h, u - u_h)_k = (u - u_h, u)_k,$$

where $(\ ,\ )_k$ denotes the inner product in H^k. Now suppose that the support of u is contained in a compact subset of Ω. Integrating by parts and using the generalized Schwarz inequality, we obtain

$$\|u - u_h\|_k^2 \leqslant c_k \|u - u_h\|_s \|u\|_{2k-s}, \tag{1.3}$$

where s is any real number not greater than k and c_k depends on k. Thus

$$\|u - u_h\|_s \geqslant \|u - u_h\|_k^2 / c_k \|u\|_{2k-s}, \tag{1.4}$$

and we conclude that the approximation rate for $u - u_h$ in any Sobolev norm cannot exceed double the original rate in H^k for arbitrary u. Therefore, we see that the element of S_h for which the infimum is attained in B is in general not the same as that in A. Note that this phenomena is not restricted to convergence in negative norms; the above example shows in particular that the H^2 projection onto quadratic splines does not have the optimal rate of approximation in $L^2 = H^0$.

Using a duality argument, one *can* show that B $\Rightarrow$ A provided that an *inverse assumption* holds for the family S_h. However, our proof below requires no inverse assumption. Finally, we remark that Bramble and Schatz in [4] and [5] assume only B in their treatment of the least squares approximation. Their proofs involve an infinite interation technique that is avoided here. Thus, the results of Baker [3] concerning least squares, combined with the result here that B $\Rightarrow$ A, lead to simplified proofs for this method under the original assumption B.

2. Preliminaries: Banach Space Interpolation. Recall the real method of interpolation by Lions and Peetre (cf. Butzer and Berens [7]). Let B_0 and B_1 be Banach spaces such that $B_1 \subset B_0$ with the inclusion map continuous, and let $|\cdot|_i$ denote the norm in B_i, $i = 0, 1$. Let $u \in B_0$ and $t > 0$, and define

$$K(u, t) = \inf\{ |u - v|_0 + t|v|_1 : v \in B_1 \}. \tag{2.1}$$

Then $K(u, t)$ is a continuous, increasing function of t for fixed u. (Here and below, the words increasing and decreasing are used in the nonstrict sense, i.e., one means not the other.) For real numbers θ and p in the ranges $0 < \theta < 1$ and $1 \leqslant p < \infty$, define

$$|u|_{\theta,p} \equiv \left(\int_0^\infty K(u, t)^p t^{-\theta p - 1}\, dt \right)^{1/p}. \tag{2.2}$$

With θ as before (and $p = \infty$), define

$$|u|_{\theta,\infty} \equiv \sup\{K(u, t)t^{-\theta} : t > 0\}. \tag{2.3}$$

Then $|\cdot|_{\theta,p}$ is a norm, and the associated Banach space is denoted by either $B_{\theta,p}$ or $[B_0, B_1]_{\theta,p}$. If λ and q are other indices such that $0 < \lambda < 1$ and $1 \leqslant q \leqslant \infty$, then we say

$$(\theta, p) < (\lambda, q) \Leftrightarrow \begin{cases} \theta < \lambda & \text{or} \\ \theta = \lambda & \text{and} \quad p > q. \end{cases} \tag{2.4}$$

If $(\theta, p) \leqslant (\lambda, q)$, the inclusions

$$B_1 \subset B_{\lambda,q} \subset B_{\theta,p} \subset B_0 \tag{2.5}$$

are valid together with the corresponding norm inequalities (continuous inclusions). In particular,

$$|u|_{\theta,p} \leqslant 2|u|_{\theta,q}, \qquad u \in B_{\theta,q}, \tag{2.6}$$

for all $\theta \in]0, 1[$ and all $p, q \in [1, \infty]$ (this is not the optimal constant, but it will suffice). Further, interpolation inequalities hold: If $(\theta, p) < (\lambda, q)$, there is a constant c depending only on θ and λ such that

$$|u|_{\theta,p} \leqslant c|u|_0^{1-\theta/\lambda}|u|_{\lambda,q}^{\theta/\lambda} \tag{2.7}$$

for all $u \in B_{\lambda,q}$. Finally, we make the convention that $B_{\theta,p} \equiv B_\theta$ for all p when $\theta = 0$ or 1, and we extend the relation (2.4) by defining $(0, p) < (\theta, q) < (1, r)$ for any $p, q, r \in [1, \infty]$ and any $\theta \in]0, 1[$. The above standard results can be found, e.g., in Butzer and Berens [7]. We now prove a result to be used to derive our main theorem.

LEMMA. *Let $u \in B_0$ and $t > 0$, and suppose $v \in B_1$ is such that*

$$|u - v|_0 + t|v|_1 \leqslant 2K(u, t).$$

Then, for any (θ, p) such that $u \in B_{\theta,p}$, we have

$$|u - v|_{\theta,p} \leqslant 3|u|_{\theta,p}. \tag{2.8}$$

If $\theta \in]0, 1[$ and $1 \leqslant p < \infty$, $(\lambda, q) \leqslant (\theta, p)$, and $u \in B_{\theta,p}$, then

$$|u - v|_{\lambda,q} \leqslant c\left[t^{\theta-\lambda}\left(\int_0^t K(u, s)^p s^{-\theta p-1}ds\right)^{1/p} + t^{-\lambda}K(u, t)\right], \tag{2.9}$$

where c depends only on θ and λ.

Proof. Our first claim is that, for all $s > 0$,

$$K(u - v, s) \leqslant 3K(u, \min\{s, t\}). \tag{2.10}$$

We have

$$K(u - v, s) \leqslant |u - v|_0 \leqslant 2K(u, t)$$

by the definition of K and our choice of v, so (2.10) is proved for $s \geqslant t$. For $s \leqslant t$, we argue as follows: For all $w \in B_1$,

$$K(u - v, s) \leqslant |u - v - (w - v)|_0 + s|w - v|_1$$

$$\leqslant |u - w|_0 + s|w|_1 + s|v|_1.$$

Taking the infimum over $w \in B_1$ and recalling the assumption on v, we have

$$K(u - v, s) \leqslant K(u, s) + 2sK(u, t)/t.$$

Because $t^{-1}K(u, t)$ is a decreasing function of t and $s \leqslant t$, we find

$$K(u - v, s) \leqslant 3K(u, s),$$

completing the proof of (2.10). We now prove the lemma.

First, since $K(u, t)$ is an increasing function of t, (2.10) implies that $K(u - v, s) \leqslant 3K(u, s)$ for all $s > 0$, and hence (2.8) is immediate for $\theta \in]0, 1[$. For the remaining cases, simply observe that, for $\theta = 0$ or 1, $K(u, t) \leqslant t^\theta |u|_\theta$ for all $t > 0$ and use the assumption on v.

To prove (2.9), first assume that $\lambda < \theta$. Then

$$\frac{1}{6}|u - v|_{\lambda,q} \leqslant \frac{1}{3}|u - v|_{\lambda,1} = \frac{1}{3}\int_0^\infty K(u - v, s)s^{-\lambda-1}\,ds$$

$$\leqslant \int_0^t K(u, s)s^{-\lambda-1}\,ds + K(u, t)\int_t^\infty s^{-\lambda-1}\,ds$$

$$\leqslant \left(\int_0^t K(u, s)^p s^{-\theta p-1}\,ds\right)^{1/p}\left(\int_0^t s^{(\theta-\lambda)p'-1}\,ds\right)^{1/p'} + \frac{1}{\lambda}K(u, t)t^{-\lambda}$$

$$= \left(\int_0^t K(u, s)^p s^{-\theta p-1}\,ds\right)^{1/p}((\theta - \lambda)p')^{-1/p'}t^{\theta-\lambda} + \frac{1}{\lambda}K(u, t)t^{-\lambda},$$

where we used (2.6), (2.10), and Hölder's inequality. This proves (2.9) in this case. When $\lambda = \theta$, the argument is similar:

$$\frac{1}{6}|u - v|_{\theta,q} \leqslant \frac{1}{3}|u - v|_{\theta,p}$$

$$\leqslant \left(\int_0^t K(u, s)^p s^{-\theta p-1}\,ds + K(u, t)^p\int_t^\infty s^{-\theta p-1}\,ds\right)^{1/p}$$

$$\leqslant \left(\int_0^t K(u, s)^p s^{-\theta p-1}\,ds\right)^{1/p} + (\theta p)^{-1/p}t^{-\theta}K(u, t).$$

This completes the proof of the lemma.

3. Main Results.

THEOREM. *Let θ and p be fixed, $0 < \theta < 1$ and $1 \leqslant p \leqslant \infty$. Let $\epsilon > 0$ and let S_ϵ be a subspace of $B_{\theta,p}$ such that*

$$\inf\{|u - \chi|_{\theta,p}: \chi \in S_\epsilon\} \leqslant \epsilon^{1-\theta}|u|_1 \quad \text{for all } u \in B_1. \tag{3.1}$$

Then for any $(\lambda, q) \geqslant (\theta, p)$ (with $\lambda = 1$ being allowed),

$$E(u) \equiv \inf\{|u - \chi|_0 + \epsilon^\theta |u - \chi|_{\theta,p}: \chi \in S_\epsilon\} \leqslant c\epsilon^\lambda |u|_{\lambda,q} \tag{3.2}$$

for all $u \in B_{\lambda,q}$, where c depends only on θ and λ. Furthermore, if $\lambda < 1$ and $q <$

∞, *and* (3.1) *holds for a family of subspaces* $\{S_\epsilon\colon 0 < \epsilon \leqslant \epsilon_0\}$, *then*

$$\lim_{\epsilon \downarrow 0} (E(u)/\epsilon^\lambda) = 0. \tag{3.3}$$

Proof. Define a parameter $\delta = \sup\{E(u)\colon |u|_1 = 1\}$. We will show that $\delta \leqslant c\epsilon$, and this then proves (3.2) in the case $\lambda = 1$. To begin with, note that $\delta < \infty$ in view of the containment relations (2.5). For any $v \in B_1$,

$$E(u) \leqslant E(u - v) + E(v) \tag{3.4}$$

because of the triangle inequality. Thus

$$E(u) \leqslant (|u - v|_0 + \epsilon^\theta |u - v|_{\theta,p}) + \delta |v|_1, \tag{3.5}$$

where we have used the fact that $0 \equiv \chi \in S_\epsilon$ and the definition of δ. Choose $v \in B_1$ such that

$$|u - v|_0 + \delta |v|_1 \leqslant 2K(u, \delta). \tag{3.6}$$

Then the lemma from the previous section implies that

$$E(u) \leqslant 2K(u, \delta) + 3\epsilon^\theta |u|_{\theta,p}. \tag{3.7}$$

Recalling the definition of the interpolation norm for $p = \infty$ and the norm inequality (2.6), we have

$$E(u) \leqslant 2\delta^\theta |u|_{\theta,\infty} + 3\epsilon^\theta |u|_{\theta,p} \leqslant 4(\delta^\theta + \epsilon^\theta)|u|_{\theta,p}. \tag{3.8}$$

This holds for any $u \in B_{\theta,p}$, and hence

$$\begin{aligned} E(u) &= \inf_{\chi \in S_\epsilon} E(u - \chi) \\ &\leqslant 4(\delta^\theta + \epsilon^\theta) \inf_{\chi \in S_\epsilon} |u - \chi|_{\theta,p}. \end{aligned}$$

Finally, invoking the assumption (3.1), we have

$$E(u) \leqslant 4(\delta^\theta + \epsilon^\theta)\epsilon^{1-\theta}|u|_1.$$

This implies that $\delta \leqslant 4(\delta^\theta + \epsilon^\theta)\epsilon^{1-\theta}$; and thus, the arithmetic-geometric mean inequality implies that $\delta \leqslant c\epsilon$ as claimed (here c depends only on θ).

As stated above, the fact that $\delta \leqslant c\epsilon$ proves (3.2) in the case $\lambda = 1$, and using this estimate for δ in (3.8) yields the case $(\lambda, q) = (\theta, p)$ as well. The general case when $\lambda = \theta$ then follows from the norm inequality (2.6). For the case $\theta < \lambda < 1$, we use (3.4) and the cases already derived: For any $v \in B_1$,

$$E(u) \leqslant E(u - v) + E(v) \leqslant c(\epsilon^\theta |u - v|_{\theta,p} + \epsilon |v|_1). \tag{3.9}$$

Using the norm interpolation inequality (2.7) and the arithmetic-geometric mean inequality, we have

$$\begin{aligned} E(u) &\leqslant c'(\epsilon^\theta |u - v|_0^{1-\theta/\lambda} |u - v|_{\lambda,q}^{\theta/\lambda} + \epsilon |v|_1) \\ &\leqslant c'(|u - v|_0 + \epsilon^\lambda |u - v|_{\lambda,q} + \epsilon |v|_1), \end{aligned} \tag{3.10}$$

where c' is a constant depending only on θ and λ. Now choose $v \in B_1$ such that

$|u - v|_0 + \epsilon |v|_1 \leqslant 2K(u, \epsilon)$. As above, we have

$$E(u) \leqslant c'(2K(u, \epsilon) + 3\epsilon^\lambda |u|_{\lambda,q})$$

$$\leqslant c'(2\epsilon^\lambda |u|_{\lambda,\infty} + 3\epsilon^\lambda |u|_{\lambda,q}) \leqslant 7c'\epsilon^\lambda |u|_{\lambda,q}.$$

This completes the proof of (3.2).

We now prove (3.3). Note that (3.10) holds also in the case $\lambda = \theta$, since in this case it follows directly from (3.9). Then with v chosen in (3.10) as in the lemma, (2.9) implies that

$$E(u) \leqslant c''\left[\left(\int_0^\epsilon K(u, t)^q t^{-\lambda q-1}\, dt\right)^{1/q} \epsilon^\lambda + K(u, \epsilon)\right],$$

where c'' depends only on θ and λ. Since $t^{-1}K(u, t)$ is a decreasing function of t, it follows that

$$\epsilon^{-\lambda}K(u, \epsilon) = \epsilon^{-1}K(u, \epsilon)\left(\frac{1}{(1-\lambda)q}\int_0^\epsilon t^{(1-\lambda)q-1}\, dt\right)^{1/q}$$

$$\leqslant \left(\frac{1}{(1-\lambda)q}\int_0^\epsilon K(u, t)^q t^{-\lambda q-1}\, dt\right)^{1/q}.$$

Therefore, for some constant c depending only on θ and λ,

$$E(u) \leqslant c\epsilon^\lambda \left(\int_0^\epsilon K(u, t)^q t^{-\lambda q-1}\, dt\right)^{1/q}.$$

Thus (3.3) follows from the fact that this integral tends to zero as $\epsilon \downarrow 0$.

COROLLARY. *Suppose $\theta_i \in [0, 1[$ and $p_i \in [1, \infty]$ are such that $(\theta_i, p_i) < (\theta, p)$ for $i = 1, \ldots, I-1$. Define $(\theta_I, p_I) = (\theta, p)$. Then, under the conditions of the theorem, the conclusions remain valid for $E(u)$ defined by*

$$E(u) \equiv \inf\left\{\sum_{i=1}^I \epsilon^{\theta_i}|u - \chi|_{\theta_i,p_i}:\ \chi \in S_\epsilon\right\},$$

except that the constant c depends on all the θ_i's as well.

Proof. From the norm interpolation inequality (2.7) and the arithmetic-geometric mean inequality, it follows that

$$\sum_{i=1}^I \epsilon^{\theta_i}|u - \chi|_{\theta_i,p_i} \leqslant c(|u - \chi|_0 + \epsilon^\theta |u - \chi|_{\theta,p}),$$

where c depends only on $\theta_1, \ldots, \theta_I$. Hence the result is a direct consequence of our estimate for the original $E(u)$.

4. Applications. We now give the proof of the equivalence of A and B in the introduction. To do so, we simply observe the well-known interpolation properties of the Sobolev spaces (cf. Lions and Magenes [9, pp. 98–99]):

$$[H^k, H^{k+m}]_{\theta,2} \simeq H^{k+\theta m}. \tag{4.1}$$

Of course, (4.1) is not known for general Ω, so we make (possibly) a restriction on

Ω by assuming (4.1) to hold. It is valid for $\Omega = \mathbf{R}^N$ and any Ω for which there exists an extension map $\Omega \to \mathbf{R}^N$ preserving Sobolev classes, e.g., if Ω satisfies a cone property. Given a family $\{S_h\colon 0 < h < 1\}$ satisfying condition **B** with constant C_B, let $S_\epsilon = S_h$ with $\epsilon = (C_B)^{(r-s)/(r-k)}h^{r-s}$. Define $B_0 = H^s$ and $B_1 = H^r$, $\lambda = 1$, and $\theta_i = i/(r-s)$ and $p_i = 2$ for $i = 0, 1, \ldots, k-s$. Applying the corollary in the last section, condition A follows (note that C_A depends on the constants coming from the isomorphisms (4.1) as well as C_B and the constant in the corollary). In fact, we are not restricted to integer Sobolev spaces. The same arguments yield

$$\inf_{\chi\in S_h}\left\{\sum_{i=0}^{I} h^{s_i}\|u-\chi\|_{s_i}\right\} \leqslant Ch^m\|u\|_m \tag{4.2}$$

for any real numbers $s_0, \ldots, s_I \in [-\infty, r[$ provided **B** holds for r and k real numbers with $r > k \equiv \max\{s_i\colon i = 0, \ldots, I\}$ and m is any real number such that $k \leqslant m \leqslant r$. Further, if $u \in H^m$ for some such m that is *less* than r, then (cf. Babuška-Kellogg [2])

$$\lim_{h\downarrow 0} h^{-m} \inf_{\chi\in S_h}\left\{\sum_{i=0}^{I} h^{s_i}\|u-\chi\|_{s_i}\right\} = 0. \tag{4.3}$$

In applying the corollary so far, we have always chosen the second interpolation index equal to 2. It is frequently useful to have $p = \infty$ in the corollary, with $p_i = 2$ for $i = 0, \ldots, I$ as before. In this case, the approximation result (4.3) is lost, but an analog of (4.2) remains valid:

$$\inf_{\chi\in S_h}\sum_{i=0}^{I} h^{s_i}\|u-\chi\|_{s_i} \leqslant Ch^{\theta r+(1-\theta)k}|u|_{[H^k,H^r]_{\theta,\infty}}. \tag{4.4}$$

For example, if u solves a second order elliptic boundary value problem with a right-hand side that is discontinuous but piecewise smooth, then $u \in [H^2, H^3]_{1/2,\infty}$ but no better [14]. For such a function u, (4.4) implies

$$\inf_{\chi\in S_h}\sum_{i=0}^{I} h^{s_i}\|u-\chi\|_{s_i} = O(h^{5/2});$$

and this is the best possible approximation rate.

Since (4.2)–(4.4) hold for s_i negative, we have a rather surprising result (cf. Bramble and Schatz [4] and [5]), namely that an *arbitrary* order of approximation is possible from S_h in an appropriate negative norm. To see how this can be, consider approximation by piecewise constant functions S_h on a uniform mesh on (0, 1) of size $h = 1/(2M)$. Define $\chi(u) \in S_h$ by the requirement, for all $a, b \in \mathbf{R}$,

$$\int_{jh}^{(j+2)h}(\chi(u)-u)(a+bx)\,dx = 0, \qquad j = 2i,\ i = 0, \ldots, M-1.$$

Note that $\chi(u)$ has (possibly) different values in $(jh, (j+1)h)$ and $((j+1)h, (j+2)h)$. The existence and uniqueness of this projection follows from the invertibility of the 2×2 matrix whose (m, n)th entry is $\int_{(j+m-1)h}^{(j+m)h} x^{n-1}\,dx$. We now show that $\chi(u)$ is simultaneously an optimal approximation in H^0 and H^{-2}. For any $\varphi \in H^2$, we have

$$\int_0^1 (u - \chi(u))\varphi\, dx = \int_0^1 (u - \chi(u))(\varphi - \hat{\varphi})\, dx$$

$$\leqslant \|u - \chi(u)\|_0 \|\varphi - \hat{\varphi}\|_0 \leqslant ch^2 \|u - \chi(u)\|_0 \|\varphi\|_2,$$

where $\hat{\varphi}$ is a piecewise linear function approximating φ to second order on the coarser mesh $\{0, 2/M, 4/M, \ldots, 1\}$ of size $2h$. Thus $\|u - \chi(u)\|_{-2} \leqslant ch^2 \|u - \chi(u)\|_0$, and the simultaneous optimality of $\chi(u)$ follows from its optimality in H^0, which can be proved by standard techniques. An optimal approximation in the $-m$ norm can be constructed by choosing

$$\int_{jh}^{(j+m)h} (\chi_m(u) - u) \sum_{i=0}^{m-1} a_i x^i\, dx = 0$$

for all a_i.

5. Acknowledgement. We would like to thank the referee for suggestions leading to a simplified proof of our main theorem.

Department of Mathematics
Cornell University
Ithaca, New York 14853

Applied Mathematics Department
Brookhaven National Laboratory
Upton, New York 11973

1. I. BABUŠKA, "The finite element method with penalty," *Math. Comp.*, v. 27, 1973, pp. 221–228.

2. I. BABUŠKA & B. KELLOGG, "Non uniform estimates for the finite element method," *Siam J. Numer. Anal.*, v. 12, 1975, pp. 868–875.

3. G. A. BAKER, "Simplified proofs of error estimates for the least squares method for Dirichlet's problem," *Math. Comp.*, v. 27, 1973, pp. 229–235.

4. J. H. BRAMBLE & A. H. SCHATZ, "Rayleigh-Ritz-Galerkin methods for Dirichlet's problem using subspaces without boundary conditions," *Comm. Pure Appl. Math.*, v. 23, 1970, pp. 653–674.

5. J. H. BRAMBLE & A. H. SCHATZ, "Least squares methods for $2m$th order elliptic boundary-value problems," *Math. Comp.*, v. 25, 1971, pp. 1–32.

6. J. H. BRAMBLE & V. THOMÉE, "Discrete time Galerkin methods for a parabolic boundary value problem," *Ann. Mat. Pura Appl.*, v. 101, 1974, pp. 115–152.

7. P. L. BUTZER & H. BERENS, *Semi-Groups of Operators and Approximation*, Springer-Verlag, Berlin and New York, 1967.

8. J. DOUGLAS, Jr. & T. DUPONT, "Galerkin methods for parabolic equations with nonlinear boundary conditions," *Numer. Math.*, v. 20, 1973, pp. 213–237.

9. J. L. LIONS & E. MAGENES, *Non-homogeneous Boundary Value Problems and Applications*, Vol. 1, Springer-Verlag, Berlin and New York, 1972.

10. F. NATTERER, "The finite element method for ill-posed problems," *RAIRO*, v. 11, 1977, pp. 271–278.

11. J. A. NITSCHE & A. H. SCHATZ, "Interior estimates for Ritz-Galerkin methods," *Math. Comp.*, v. 28, 1974, pp. 937–958.

12. P. A. RAVIART, "The use of numerical integration in finite element methods for solving parabolic equations," *Topics in Numerical Analysis*, J. J. H. Miller, ed., Academic Press, New York, 1973, pp. 233–264.

13. M. H. SCHULTZ, "L^2 error bounds for the Rayleigh-Ritz-Galerkin method," *SIAM J. Numer. Anal.*, v. 8, 1971, pp. 737–748.

14. R. SCOTT, "Applications of Banach space interpolation to finite element theory," *Functional Analysis Methods in Numerical Analysis* (M. Z. Nashed, Editor), Springer-Verlag, Berlin and New York.