SIAM J. NUMER. ANAL.
Vol. 10, No. 1, March 1973

A GENERALIZED RITZ-LEAST-SQUARES METHOD FOR DIRICHLET PROBLEMS*

JAMES H. BRAMBLE† AND JOACHIM A. NITSCHE‡

Abstract. A new projection method for the approximate solution of Dirichlet's problem is formulated and error estimates are given. The resulting method does not require subspaces satisfying any boundary conditions or inverse assumptions. The corresponding linear system is shown to have conditioning properties similar to those of the classical Ritz method.

1. Introduction. In a recent paper [10], Bramble and Schatz presented a method for the approximation of solutions of $2m$th order elliptic boundary value problems with essential boundary conditions. The importance of this work rests in the fact that the approximating functions need not satisfy any boundary conditions. The classical principles for self-adjoint problems always assume that some auxiliary conditions will be satisfied by the approximating functions. This requirement makes the classical Rayleigh–Ritz method difficult to apply in practice. The method of [10] is a least-squares method and essentially embeds the given problem in a higher order one for which the boundary conditions are natural. This leads to the solution of a system whose behavior is like that of a $4m$th order equation and hence could lead to difficulties with regard to conditioning of the linear systems involved. On the other hand, Nitsche [15] gave a method for the Dirichlet problem for Poisson's equation which leads to a system which does not have the aforementioned conditioning problems, but the quadratic form is in general not positive definite. Although the elements of the subspaces in his principle are not required to satisfy the usual boundary conditions, some side conditions "near the boundary" are required.

In the present paper we present a method which avoids the difficulties in both the abovementioned methods yet retains the good properties of each. That is, the present method requires only approximability properties of the subspaces yet has the "right" properties with regard to conditioning of the resulting system. The error estimates obtained are in a sense "optimal" with minimal assumptions on the approximating subspaces.

For other work in this direction we mention the papers of Aubin [2] and Babuška [5] who present and analyze a "penalty" method. The results in each case do not yield "optimal" approximations; i.e., the solution of the finite-dimensional problem does not reproduce the properties of "best approximation" possessed by the subspaces. Note also the work of Schultz [16] and Zlámal [18] in constructing subspaces whose elements vanish on the boundary for the application of the classical Rayleigh–Ritz method.

For the sake of clarity we present first detailed proofs of our results in the case of Dirichlet's problem for Poisson's equation. This is then followed by the formulation of the general case of the Dirichlet problem for a $2m$th order elliptic

* Received by the editors January 6, 1972. This work was supported in part by the National Science Foundation.

† Department of Mathematics, Cornell University, Ithaca, New York 14850.

‡ Institut fur Angew Math., Universitaet Freiburg, 78-Freiburg, Germany.

differential equation. This equation may be non-self-adjoint; it is sufficient that a coerciveness condition similar to Gårding's inequality hold true. Only a detailed formulation of the problem and the results are given since the difficulties in going from the case of Laplace's operator to the $2m$th order case are purely technical. The last section is concerned with estimates for the condition number of the linear systems in case the subspaces possess a basis with some typical properties.

2. Notations and preliminaries. Let Ω be a bounded domain in R^N with boundary $\partial\Omega$, which is, for convenience, of class C^∞. For nonnegative integers p, the Sobolev space $H^p(\Omega) = W_2^p(\Omega)$ is the completion of the space $C^\infty(\Omega)$ of infinitely differentiable real-valued functions in the norm

$$\|u\|_p = \left\{ \sum_{|\alpha| \leqq p} \|D^\alpha u\|^2 \right\}^{1/2},$$

where

$$\|u\| = \left\{ \int_\Omega u^2 \, dx \right\}^{1/2}.$$

$H^p(\Omega)$ is a Hilbert space with the inner product

$$(u, v)_p = \sum_{|\alpha| \leqq p} (D^\alpha u, D^\alpha v),$$

where

$$(u, v) = \int_\Omega uv \, dx.$$

If p is any positive number which is not an integer, then the Hilbert space $H^p(\Omega)$ is defined by interpolation between the Hilbert spaces with successive integers.

The Hilbert spaces $H^p(\partial\Omega)$ are defined as follows: Let $\Delta_{\partial\Omega}$ denote the Laplace–Beltrami operator on $\partial\Omega$. There exists a positive sequence $\{\lambda_j\}$ of eigenvalues and a corresponding sequence $\{w_j\}$ of eigenfunctions satisfying

$$-\Delta_{\partial\Omega} w_j = \lambda_j w_j \quad \text{on } \partial\Omega, \qquad j = 1, 2, \cdots,$$

where the w_j are orthonormal in $H^0(\partial\Omega) = L_2(\partial\Omega)$ with respect to the inner product

$$\langle u, v \rangle = \int_{\partial\Omega} uv \, d\sigma,$$

σ being the measure on $\partial\Omega$ induced by the Lebesgue measure in R^{N-1}. For any $s \geqq 0$ we define the inner product

$$\langle u, v \rangle_s = \sum_{j=1}^{\infty} \lambda_j^s \langle u, w_j \rangle \langle v, w_j \rangle.$$

$H^s(\partial\Omega)$ is then the completion of $C^\infty(\partial\Omega)$ in the norm

$$|u|_s = \langle u, u \rangle_s^{1/2}.$$

In the later sections we shall need some known a priori estimates. These we state now as lemmas. For proofs see [1] and [14].

LEMMA 1. *For* $u \in H^1(\Omega)$ *we have* $u|_{\partial\Omega} \in H^{1/2}(\partial\Omega)$ *and there is a constant* c_1 *independent of* u *such that*

$$\|u\|_1^2 \leqq c_1\{D(u, u) + |u|^2\}.$$

Here $D(u, v)$ *stands for the Dirichlet integral*

$$D(u, v) = \int_\Omega \sum_{i=1}^N \frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_i}\,dx.$$

LEMMA 2. *For any* $u \in H^1(\Omega)$ *and* $\varepsilon > 0$ *there is a constant* c_2 *independent of* u *and* ε *such that*

$$|u|^2 \leqq c_2\{\varepsilon^{-1}\|u\|^2 + \varepsilon\|u\|_1^2\}.$$

Denote by ∇_S the surface gradient vector of a function $u \in H^1(\partial\Omega)$. Hence $|\nabla_S u|$ is the $L_2(\partial\Omega)$-norm of the absolute value of the surface gradient.

Remark. If $\bar{u}, \bar{v}$ are any sufficiently smooth extensions of u, v to a neighborhood of $\partial\Omega$, then we may define

$$\langle \nabla_S u, \nabla_S v\rangle = \int_{\partial\Omega}\left\{\sum_{i=1}^N \frac{\partial \bar{u}}{\partial x_i}\frac{\partial \bar{v}}{\partial x_i} - \frac{\partial \bar{u}}{\partial n}\frac{\partial \bar{v}}{\partial n}\right\}d\sigma.$$

In case $N = 2$ we have $|\nabla_S u|^2 = \int_{\partial\Omega}|du/ds|^2\,ds$, where d/ds denotes differentiation with respect to the arc length of $\partial\Omega$.

LEMMA 3. *For any* $u \in H^1(\partial\Omega)$ *there is a constant* c_3 *independent of* u *such that*

$$|u|_1^2 \leqq c_3\{|\nabla_S u|^2 + |u|^2\}.$$

LEMMA 4. *Let* $0 \leqq t \leqq s$ *and* $\varepsilon > 0$ *be given. Then there are constants* c_4 *and* c_5 *depending only on* s *and* Ω *such that for* $u \in H^{s+t}(\Omega)$,

$$\|u\|_s^2 \leqq c_4\|u\|_{s-t}\|u\|_{s+t} \leqq (c_4/2)\{\varepsilon^{-1}\|u\|_{s-t}^2 + \varepsilon\|u\|_{s+t}^2\},$$

and for $u \in H^{s+t}(\partial\Omega)$,

$$|u|_s^2 \leqq c_5|u|_{s-t}|u|_{s+t} \leqq (c_5/2)\{\varepsilon^{-1}|u|_{s-t}^2 + \varepsilon|u|_{s+t}^2\}.$$

We define as usual the Laplace operator

$$\Delta u = \sum_{i=1}^N \frac{\partial^2 u}{\partial x_i^2}.$$

LEMMA 5. *For* $u \in H^s(\Omega)$ *with* $s \geqq 2$ *there is a constant* c_6 *independent of* u *such that*

$$\|u\|_s^2 \leqq c_6\{\|\Delta u\|_{s-2}^2 + |u|_{s-1/2}^2\}.$$

LEMMA 6. *Let* u *be a harmonic function* ($\Delta u = 0$ *in* Ω) *with* $u|_{\partial\Omega} \in H^{s-1/2}(\partial\Omega)$ *for* $s \geqq 1/2$. *Then* $u \in H^s(\Omega)$ *and there is a constant* c_7 *independent of* u *such that*

$$\|u\|_s^2 \leqq c_7|u|_{s-1/2}^2.$$

3. A quadratic functional and its positivity. In this section we shall consider the functional

$$K(u, v) = -(u, \Delta v) - (\Delta u, v) - D(u, v) + h^2(\Delta u, \Delta v) + \gamma\{h^{-1}\langle u, v\rangle + h\langle \nabla_S u, \nabla_S v\rangle\} \tag{3.1}$$

for a suitably chosen positive number γ. It is in terms of this functional that our approximation method is defined and hence its positivity is of importance. Specifically, we have the following theorem.

THEOREM 1. *There exist a positive constant* γ_0 *and a positive constant* $c = c(\gamma_0)$ *such that for* $0 < h \leqq 1$ *and* $u \in H^2(\Omega)$,

$$\|u\|_1^2 + h^2\|\Delta u\|^2 + h^{-1}|u|^2 + h|\nabla_S u|^2 \leqq cK(u, u). \tag{3.2}$$

Proof. Let $u = y + z$, where $\Delta y = 0$ in Ω and $z = 0$ on $\partial\Omega$. Then $D(y, z) = 0$ and hence $D(u, u) = D(y, y) + D(z, z)$. This gives

$$K(u, u) = D(z, z) - 2\langle u, z_n\rangle + h^2\|\Delta z\|^2 + D(y, y) - 2\langle y, y_n\rangle + \gamma\{h^{-1}|u|^2 + h|\nabla_S u|^2\},$$

where z_n, y_n are the outward normal derivatives of z and y on $\partial\Omega$. Now since $\Delta y = 0$,

$$D(y, y) - 2\langle y, y_n\rangle = -D(y, y).$$

Thus

$$K(u, u) = D(z, z) - 2\langle u, z_n\rangle + h^2\|\Delta z\|^2 - D(y, y) + \gamma\{h^{-1}|u|^2 + h|\nabla_S u|^2\}.$$

Now by Lemmas 1 and 5 with $s = 2$ since $z = 0$ on $\partial\Omega$ we have, for some $c_8 > 0$,

$$K(u, u) \geqq c_8\{\|z\|_1^2 + h^2\|z\|_2^2\} - 2\langle u, z_n\rangle - D(y, y) + \gamma\{h^{-1}|u|^2 + h|\nabla_S u|^2\}. \tag{3.3}$$

For any $\delta > 0$ we have

$$2\langle u, z_n\rangle \leqq \delta h^{-1}|u|^2 + \delta^{-1}h|z_n|^2, \tag{3.4}$$

and as an easy consequence of Lemma 2,

$$|z_n|^2 \leqq c_2'\{h^{-1}\|z\|_1^2 + h\|z\|_2^2\} \tag{3.5}$$

for a suitable c_2'. Hence combining (3.4) and (3.5), we have

$$2\langle u, z_n\rangle \leqq \delta h^{-1}|u|^2 + c_2'\delta^{-1}\{\|z\|_1^2 + h^2\|z\|_2^2\}. \tag{3.6}$$

Thus we obtain from (3.6) and (3.3),

$$\begin{aligned} K(u, u) \geqq [c_8 - c_2'\delta^{-1}]\{\|z\|_1^2 + h^2\|z\|_2^2\} - D(y, y) \\ + (\gamma - \delta)\{h^{-1}|u|^2 + h|\nabla_S u|^2\}. \end{aligned} \tag{3.7}$$

Now since $\Delta y = 0$ in Ω we obtain from Lemma 6,

$$D(y, y) \leqq \|y\|_1^2 \leqq c_7|u|_{1/2}^2,$$

and from Lemma 4 with $s = t = 1/2$ and $\varepsilon = h$,

$$D(y, y) \leqq \tfrac{1}{2}c_5c_7\{h^{-1}|u|^2 + h|u|_1^2\}. \tag{3.8}$$

Combining (3.8) and Lemma 3 since $h \leqq 1$, it follows that

$$D(y, y) \leqq c_9\{h^{-1}|u|^2 + h|\nabla_S u|^2\} \tag{3.9}$$

for some constant c_9. From (3.9) and (3.7), we obtain

$$K(u,u) \geqq [c_8 - c_2'\delta^{-1}]\{\|z\|_1^2 + h^2\|z\|_2^2\} + D(y,y) + (\gamma - \delta - 2c_9)\{h^{-1}|u|^2 + h|\nabla_S u|^2\}. \tag{3.10}$$

Now (3.10) holds for any γ and any $\delta > 0$. Choose δ so that $c_8 - c_2'\delta^{-1} \geqq c_8/2$ and then γ_0 so that for $\gamma \geqq \gamma_0$ the inequality $\gamma - \delta - 2c_9 \geqq c_8/2$ is valid. With this choice we have

$$K(u,u) \geqq \tfrac{1}{2}c_8\{\|z\|_1^2 + h^2\|z\|_2^2\} + D(y,y) + \tfrac{1}{2}c_8\{h^{-1}|u|^2 + h|\nabla_S u|^2\}. \tag{3.11}$$

From Lemma 1 it follows with $y = u$ on $\partial\Omega$ that

$$D(y,y) + h^{-1}|u|^2 \geqq c_1^{-1}\|y\|_1^2. \tag{3.12}$$

Combining (3.11) with (3.12) gives, with some $c_{10} > 0$,

$$K(u,u) \geqq c_{10}\{\|z\|_1^2 + \|y\|_1^2 + h^2\|z\|_2^2 + h^{-1}|u|^2 + h|\nabla_S u|^2\}. \tag{3.13}$$

Clearly $\|u\|_1^2 \leqq 2\|z\|_1^2 + 2\|y\|_1^2$ and $\|\Delta u\|^2 = \|\Delta z\|^2 \leqq 2\|z\|_2^2$. Hence from (3.13) we obtain, for some $c = c(\gamma_0) > 0$,

$$K(u,u) \geqq c^{-1}\{\|u\|_1^2 + h^2\|\Delta u\|^2 + h^{-1}|u|^2 + h|\nabla_S u|^2\},$$

which is just (3.2).

4. An approximation method and error estimates. We shall consider the Dirichlet problem

$$\begin{aligned} -\Delta u &= f \quad \text{in } \Omega, \\ u &= g \quad \text{on } \partial\Omega \end{aligned} \tag{4.1}$$

with $f \in H^s(\Omega)$ and $g \in H^p(\partial\Omega)$ for appropriate s and p. The bilinear functional $K(\psi, u)$ has the property that for u satisfying (4.1), $s \geqq 0$ and $p \geqq 1$,

$$K(\psi,u) = (\psi,f) - \langle\psi_n, g\rangle - h^2(\Delta\psi, f) + \gamma\{h^{-1}\langle\psi,g\rangle + h\langle\nabla_S\psi, \nabla_S g\rangle\},$$

and hence this functional can be computed in terms of f and g for each $\psi \in H^2(\Omega)$.

Now let us consider a one-parameter family of subspaces S^h, $0 < h \leqq 1$, of $H^2(\Omega)$ such that S^h, for fixed h, is closed with respect to the norm $K^{1/2}(\cdot,\cdot)$. That this is a norm follows from Theorem 1. We then have the following theorem.

THEOREM 2. *Let u be the solution of* (4.1) *with $s \geqq 0$ and $p \geqq 1$. There exists a unique $u_h \in S^h$ such that*

$$K(\chi, u - u_h) = 0$$

for all $\chi \in S^h$. Furthermore,

$$K(u - u_h, u - u_h) = \inf_{\chi\in S^h} K(u - \chi, u - \chi).$$

Proof. This follows immediately from the projection theorem by considering the closure of $H^2(\Omega)$ with respect to the norm $K^{1/2}(\cdot,\cdot)$ as a Hilbert space with the obvious inner product, and using the fact that S^h is a closed subspace of this space.

We shall consider u_h as an approximate solution of the problem (4.1) in the subspace S^h and define the error $e = e_h = u - u_h$. Note that $u_h = R_h u$, where R_h

is a linear projection operator. We obtain from a direct application of Theorem 1 the following error estimates.

THEOREM 3. *Let u be the solution of* (4.1) *with $s \geqq 0$ and $p \geqq 1$ and e defined as above. Then*

$$\|e\|_1^2 \leqq cK(e, e),$$

$$|e|^2 \leqq chK(e, e),$$

$$|e|_1^2 \leqq ch^{-1}K(e, e)$$

and

$$\|\Delta e\|^2 \leqq ch^{-2}K(e, e),$$

where c is a constant independent of e and h.

With the help of this theorem and our a priori estimates, we obtain the next result.

THEOREM 4. *Let e be as above. Then there is a constant c independent of h and e such that*

$$\|e\|_{3/2}^2 \leqq ch^{-1}K(e, e).$$

Proof. Let $e = y + z$ with $\Delta y = 0$ in Ω and $z = 0$ on $\partial\Omega$. Then from Lemma 5,

$$\|y\|_{3/2}^2 \leqq c_6|y|_1^2 = c_6|e|_1^2. \tag{4.2}$$

Combining (4.2) with Theorem 3, we have

$$\|y\|_{3/2}^2 \leqq ch^{-1}K(e, e) \tag{4.3}$$

with an appropriate c. Now from Lemma 4 with $s = 3/2$, $t = 1/2$ and $\varepsilon = h$ we have

$$\|z\|_{3/2}^2 \leqq \tfrac{1}{2}c_4\{h^{-1}\|z\|_1^2 + h\|z\|_2^2\}, \tag{4.4}$$

and from Lemma 5 since $z = 0$ on $\partial\Omega$,

$$\|z\|_2^2 \leqq c_6\|\Delta z\|^2 = c_6\|\Delta e\|^2. \tag{4.5}$$

From Lemma 1 and the fact that $D(y, z) = 0$ it follows that

$$\|z\|_1^2 \leqq c_1D(z, z) \leqq c_1D(e, e) \leqq c_1\|e\|_1^2. \tag{4.6}$$

Combining (4.4), (4.5) and (4.6) with Theorem 1, we obtain

$$\|z\|_{3/2}^2 \leqq ch^{-1}K(e, e). \tag{4.7}$$

The result now follows from (4.3) and (4.7).

We next consider the error in the L_2-norm. In order to do this we define $v \in H^2(\Omega)$ to be the solution of

$$\begin{aligned} -\Delta v &= e \quad \text{in } \Omega, \\ v &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4.8}$$

Then we have

$$K(v, e) = -(v, \Delta e) + (e, e) - D(v, e) - h^2(e, \Delta e). \tag{4.9}$$

But since $v = 0$ on $\partial\Omega$,

$$-(v, \Delta e) - D(v, e) = 0.$$

Thus it follows from (4.9) and the arithmetic-geometric mean inequality that

$$\|e\|^2 = K(v, e) + h^2(e, \Delta e) \leqq K(v, e) + \frac{1}{2}\|e\|^2 + \frac{h^4}{2}\|\Delta e\|^2. \tag{4.10}$$

Hence by Theorem 3,

$$\|e\|^2 \leqq 2K(v, e) + ch^2 K(e, e). \tag{4.11}$$

We now derive an estimate where for the first time we make an approximability assumption about S^h.

THEOREM 5. *Suppose that e is defined as above and that S^h is such that*

$$h^{-2} \sup_{\|w\|_2 \leqq 1} \inf_{\chi \in S^h} K(w - \chi, w - \chi) \leqq c_{11}$$

for some constant c_{11} independent of h. Then for a suitable constant c,

$$\|e\|^2 \leqq ch^2 K(e, e). \tag{4.12}$$

Proof. Note that for $\chi \in S^h$,

$$K(v, e) = K(v - \chi, e).$$

Hence from (4.11) we have with v defined by (4.8),

$$\|e\|^2 \leqq 2 \inf_{\chi \in S^h} K(v - \chi, e) + ch^2 K(e, e).$$

Since $K(\cdot, \cdot)$ is a positive definite symmetric form, it follows from Schwarz's inequality that

$$\|e\|^2 \leqq 2\sqrt{K(e, e)}\sqrt{\inf_{\chi \in S^h} K(v - \chi, v - \chi)} + ch^2 K(e, e).$$

From this it follows easily that

$$\|e\|^2 \leqq \delta h^{-2} \inf_{\chi \in S^h} K(v - \chi, v - \chi) + (c + \delta^{-1})h^2 K(e, e) \tag{4.13}$$

for any $\delta > 0$. From the hypothesis of Theorem 5, we have

$$h^{-2} \inf_{\chi \in S^h} K(v - \chi, v - \chi) \leqq c_{11}\|v\|_2^2, \tag{4.14}$$

and from Lemma 5,

$$\|v\|_2^2 \leqq c_6\|e\|^2. \tag{4.15}$$

Combining (4.13), (4.14) and (4.15) and choosing $\delta^{-1} = 2c_6 c_{11}$, we obtain (4.12) for a suitable constant c.

5. Approximating subspaces. Now we introduce some approximability assumptions on the subspaces S_h.

DEFINITION 1. The family of subspaces $\{S^h | 0 < h \leqq 1\}$ is of *approximation type* $(k.r)_\Omega$ if there is a constant $c_{k.r}$ such that for all $u \in H^r(\Omega)$,

$$\inf_{\chi \in S^h} \sum_{j=0}^{k} h^j \|u - \chi\|_j \leqq c_{k.r} h^r \|u\|_r. \tag{5.1}$$

In the normal case $k < r$. It can be shown that the following is true provided $\partial\Omega$ has the cone property.

LEMMA 7. *If $\{S^h\}$ is of approximation type $(k.r)_\Omega$, then it is also of type $(\kappa.\rho)$ for any κ, ρ with $\kappa \leqq k$, $\rho \leqq r$ and $\kappa \leqq \rho$.*

The construction of subspaces with this property is discussed, for example, in papers of Babuška [4], Bramble and Hilbert [6], Bramble and Zlámal [7], Di Guglielmo [11], [12], Fix and Strang [13], Schultz [16], Strang [17] and Zlámal [18].

With the help of Lemma 2 and the corresponding inequality for $\nabla_S u$:

$$|\nabla_S u|^2 \leqq c_{12}\{\varepsilon^{-1}\|u\|_1^2 + \varepsilon\|u\|_2^2\}, \tag{5.2}$$

where because of Lemma 3, $|\nabla_S u|$ could also be replaced by $|u|_1$, we find with $\varepsilon = h$ for any $\chi \in S^h$ which fulfills

$$\sum_{j=0}^{2} h^j\|u - \chi\|_j \leqq 2c_{2.r}h^r\|u\|_r, \qquad r \geqq 2,$$

also

$$h^{1/2}|u - \chi| + h^{3/2}|u - \chi|_1 \leqq c_{13}h^r\|u\|_r. \tag{5.3}$$

Now let $\{S^h\}$ be of approximation type $(k.r)_\Omega$ with $k \geqq 2$. Then as a consequence of (5.1) and (5.3), we get immediately

$$K(e, e) = \inf_{\chi \in S^h} K(u - \chi, u - \chi) \leqq c_{14}h^{2\rho - 2}\|u\|_\rho.$$

Here ρ may be any number with $2 \leqq \rho \leqq r$. This guarantees that the needed approximability assumption of Theorem 5 is true.

THEOREM 6. *Assume $\{S^h\}$ is of approximation type $(k.r)_\Omega$ with $k \geqq 2$. Then the following error estimates are valid with $2 \leqq \rho \leqq r$:*

$$\|e\|_\sigma \leqq ch^{\rho-\sigma}\|u\|_\rho \qquad \text{for } 0 \leqq \sigma \leqq 3/2, \tag{5.4a}$$

$$|e|_\sigma \leqq ch^{\rho-\sigma-1/2}\|u\|_\rho \quad \text{for } 0 \leqq \sigma \leqq 1, \tag{5.4b}$$

$$\|\Delta e\| \leqq ch^{\rho-2}\|u\|_\rho, \tag{5.4c}$$

where c is constant independent of e and h.

Remark. Inequality (5.4a) for $\sigma = 0$ and $\sigma = 3/2$ implies this relation for all σ between these limits.

The term $|\nabla_S u|^2$ in $K(\cdot\,,\,\cdot)$ could be cumbersome in practice. It is interesting that this term can sometimes be omitted. Obviously it will be sufficient if the corresponding quadratic functional

$$K_1(u, u) = K(u, u) - \gamma h|\nabla_S u|^2$$

is positive definite on the subspaces S^h. This will be true if

$$h|\nabla_S \chi|^2 \leqq c_{15}K_1(\chi, \chi) \quad \text{for all } \chi \in S^h \tag{5.5}$$

with a constant c_{15} independent of h. Then we have on S^h:

$$K_1(\chi, \chi) \leqq K(\chi, \chi) \leqq (1 + \gamma c_{15})K_1(\chi, \chi).$$

We shall discuss two different inverse assumptions.

DEFINITION 2. The sequence of subspaces $\{S^h|0 < h \leqq 1\}$ is of *inverse type* $(k.r)_\Omega$ (resp. $(k.r)_{\partial\Omega}$) if there is a constant $c_{k.r}$ (resp. $c'_{k.r}$) such that for all $0 < h \leqq 1$ and all $\chi \in S^h$,

$$\|\chi\|_r \leqq c_{k.r} h^{k-r} \|\chi\|_k$$

$$(\text{resp. } |\chi|_r \leqq c'_{k.r} h^{k-r} |\chi|_k).$$

Evidently (5.5) will be fulfilled if $\{S^h\}$ is of inverse type $(1.2)_{\partial\Omega}$. In order to see that also the property of being of inverse type $(1.2)_\Omega$ for the subspaces will be sufficient for (5.5), we look at (5.2) with $\varepsilon = h$:

$$\begin{aligned} h|\nabla_S \chi|^2 &\leqq c_{12}\{\|\chi\|_1^2 + h^2\|\chi\|_2^2\} \\ &\leqq c_{12}(1 + c_{1.2}^2)\|\chi\|_1^2 \end{aligned}$$

which gives us with Lemma 1,

$$h|\nabla_S \chi|^2 \leqq c_{16}\{D(\chi, \chi) + |\chi|^2\}.$$

Both terms $D(\chi, \chi)$ and $|\chi|^2$ (the last even with the factor h^{-1}) appear in $K_1(\chi, \chi)$.

COROLLARY 6.1. *If the subspaces are of inverse type* $(1.2)_\Omega$ *or* $(1.2)_{\partial\Omega}$, *then the error estimates of Theorem* 6 *are also valid if* $u_h \in S^h$ *is the element of* S^h *such that*

$$\inf_{\chi \in S^h} K_1(u - \chi, u - \chi) = K_1(u - u_h, u - u_h).$$

In case $\{S^h\}$ is of inverse type $(1.2)_\Omega$, we have a further error estimate.

COROLLARY 6.2. *If the subspaces are of inverse type* $(1.2)_\Omega$ *at least, then we have in addition to* (5.4a) *also*

$$\|e\|_2 \leqq ch^{\rho-2}\|u\|_\rho, \qquad 2 \leqq p \leqq r.$$

The proof is obvious.

Up to now the influence of f and g, the given data, is mixed in the error estimates. In order to get a separation, let u_1 and u_2 be the solutions of

$$\begin{aligned} -\Delta u_1 &= f \quad \text{in } \Omega, \\ u_1 &= 0 \quad \text{on } \partial\Omega \end{aligned}$$

and

$$\begin{aligned} -\Delta u_2 &= 0 \quad \text{in } \Omega, \\ u_2 &= g \quad \text{on } \partial\Omega \end{aligned}$$

and let $u_{1.h}$ and $u_{2.h}$ be the corresponding solutions of our variation principle. Because of the linearity we have $u_h = u_{1.h} + u_{2.h}$. Our error estimates together with Lemma 5 and Lemma 6 give, for example,

$$\begin{aligned} \|u_1 - u_{1.h}\| &\leqq ch^\rho\|u_1\|_\rho \leqq c_{17}h^\rho\|f\|_{\rho-2}, \\ \|u_2 - u_{2.h}\| &\leqq ch^\rho\|u_2\|_\rho \leqq c_{18}h^\rho|g|_{\rho-1/2}. \end{aligned}$$

By combining these inequalities but using different values for ρ in the two cases according to the possible different regularity of f and g, we get

$$\|e\| = \|u - u_h\| \leqq c_{19}\{h^{\lambda+2}\|f\|_\lambda + h^{\mu+1/2}|g|_\mu\}.$$

Because of the assumed property of the subspaces we must have $\lambda \leqq r - 2$ and $\mu \leqq r - 1/2$; otherwise, the exponents of h must be replaced by r.

6. Generalizations. Without going into details we discuss in this section the corresponding variation principle for solving the boundary value problem

$$
\begin{aligned}
Au &= \sum_{|\alpha|,|\beta| \leqq m} (-1)^{|\alpha|} D^\alpha (a_{\alpha\beta} D^\beta u) = f \quad \text{in } \Omega, \\
B_j u &= \sum_{|\alpha| \leqq j} b_{j\alpha} D^\alpha u = g_j \quad \text{on } \partial\Omega, \qquad j = 0, 1, \cdots, m-1,
\end{aligned} \tag{6.1}
$$

for a general not necessarily self-adjoint elliptic differential operator of order $2m$ with inhomogeneous Dirichlet boundary conditions. The B_j are assumed to be such that for any z with zero boundary values, i.e., with

$$B_j z = 0 \quad \text{on } \partial\Omega, \qquad j = 0, 1, \cdots, m-1, \tag{6.2}$$

and any w, we have

$$a(w, z) = (Aw, z) = \int_\Omega \left\{ \sum_{|\alpha|,|\beta| \leqq m} a_{\alpha\beta} D^\alpha z D^\beta w \right\} dx, \tag{6.3}$$

and that for all z with (6.2) we have the coercive condition

$$a(z.z) \geqq c_{20} \|z\|_m^2 \tag{6.4}$$

with a positive constant c_{20}. Corresponding to (3.1) we now define the bilinear functional

$$
\begin{aligned}
K(v, w) &= (Av, w) + (v, A^*w) - a(v, w) + h^{2m}(Av, A^*w) \\
&\quad + \gamma h^{1-2m} \sum_{j=0}^{m-1} \{h^{2j}\langle B_j v, B_j w\rangle + h^{2\mu}\langle \nabla_S^{\mu-j} B_j v, \nabla_S^{\mu-j} B_j w\rangle\}.
\end{aligned} \tag{6.5}
$$

Here A^* denotes the formal adjoint differential operator and μ may be any integer with $m \leqq \mu < 2m$.

For a sequence of subspaces $\{S^h\}$ of approximation type $(k.r)_\Omega$ (now we need $2m \leqq k$), the approximation $u_h \in S^h$ of the solution u of (6.1) is defined by

$$K(u - u_h, \chi) = 0 \quad \text{for all } \chi \in S^h. \tag{6.6}$$

It may be noted that u_h is computed by means of the given data f and g_j only. Parallel to the developments of § 2 one can prove that the functional $K(\cdot, \cdot)$ has a symmetric part which is positive definite for γ sufficiently large. Therefore the solution u_h is unique. In the same way quasi-optimal error bounds, with $\rho \leqq r$,

$$\|u - u_h\|_\sigma \leqq c_{22} h^{\rho-\sigma} \|u\|_\rho \tag{6.7}$$

are derived for all σ with $0 \leqq \sigma \leqq \mu + 1/2$, and moreover,

$$\|A(u - u_h)\| \leqq c_{23} h^{\rho-2m} \|u\|_\rho. \tag{6.8}$$

While (6.7) for $\sigma = \mu + 1/2$ as well as (6.8) is merely a direct consequence of (6.6), the estimate (6.7) in case $\sigma = 0$ is proved by using a "partial" duality relation combining the scalar product in $L_2(\Omega)$ with the bilinear form $K(\cdot, \cdot)$: For a given w, let the function z be the solution of

$$A^*z = w \quad \text{in } \Omega,$$
$$B_jz = 0 \quad \text{on } \partial\Omega.$$

Then we have

$$(v, w) = K(v, z) - h^{2m}(Av, w).$$

The choice $v = w = u - u_h$ leads by means of arguments similar to § 2 to the needed estimate. Inequality (6.7) for values σ with $0 < \sigma < \mu + 1/2$ then follows by interpolation.

Finally we remark that the terms $\langle \nabla_S^{\mu-j} B_j v, \nabla_S^{\mu-j} B_j w \rangle$ in the functional $K(\cdot, \cdot)$ may be omitted, if the subspaces are of inverse type $(k.\mu)_{\partial\Omega}$ or $(k.\mu + 1)_\Omega$ with some $k < \mu$ or $k < \mu + 1$.

7. Conditioning of the linear system. It is known that for systems coming from the Rayleigh–Ritz method applied to the Dirichlet problem for Poisson's equation the condition number is proportional to h^{-2} for a large class of approximating subspaces. Hereby the condition number of an invertible matrix M is defined as $\|M\| \times \|M^{-1}\|$ and is relevant to the practical problem of solving the linear system with matrix M. Conditions under which one can obtain the abovementioned estimate will be assumed in the present situation and we will then estimate the condition number coming from the functional $K(u, u)$ defined by (3.1).

Suppose $\{S^h\}$ is a one-parameter family of subspaces of finite dimension $n = n_h$ of $H^k(\Omega)$ with the following property: There exists a set of basis elements $\{\varphi_1, \cdots, \varphi_n\}$ such that for all real vectors $\{\alpha_1, \cdots, \alpha_n\}$ the two inequalities

$$\lambda h^N \sum_{j=1}^{n} a_j^2 \leqq \sum_{i,j=1}^{n} (\varphi_i, \varphi_j)\alpha_i\alpha_j,$$

$$\sum_{|\beta| \leqq l} \sum_{i,j=1}^{n} (D^\beta\varphi_j, D^\beta\varphi_j)\alpha_i\alpha_j \leqq \Lambda h^{N-2l} \sum_{j=1}^{n} \alpha_j^2$$

are valid for all $l \leqq k$ with positive constants λ, Λ independent of h. We consider for this choice of a basis the matrix $\mathfrak{R} = (K(\varphi_i, \varphi_j))$ of the linear system corresponding to (3.1). By definition, $\mathfrak{R}$ is symmetric and by Theorem 1 it is positive definite. We can now prove the following theorem.

THEOREM 7. *Under the above assumptions on $\{S^h\}$ with $k = 2$,*

$$\|\mathfrak{R}\| \times \|\mathfrak{R}^{-1}\| \leqq Ch^{-2},$$

where C is a constant independent of h.

Proof. Let w be any element of $H^2(\Omega)$. Then from Theorem 1 we have

$$\|w\|^2 \leqq c_{24}K(w, w).$$

From the definition of $K(u, u)$ and Lemma 2 applied to the boundary terms, it follows that

$$K(w, w) \leqq c_{25} \sum_{j=0}^{2} h^{2j-2}\|w\|_j^2.$$

Hence for a suitable c_{26} we have

$$c_{26}^{-1}\|w\|^2 \leqq K(w, w) \leqq c_{26} \sum_{j=0}^{2} h^{2j-2}\|w\|_j^2.$$

Now for arbitrary $w \in S^h$ we have $w = \sum_{j=1}^{n} \varphi_j\alpha_j$, and therefore,

$$\|w\|^2 = \sum_{i,j=1}^{n} (\varphi_i, \varphi_j)\alpha_i\alpha_j,$$

and because of the first assumption,

$$\lambda h^N \sum_{j=1}^{n} \alpha_j^2 \leqq \|w\|^2.$$

Clearly,

$$\sum_{j=0}^{2} h^{2j-2}\|w\|_j^2 = \sum_{k=0}^{2} h^{2k-2} \sum_{|\beta| \leqq k} \sum_{i,j=1}^{n} (D^\beta\varphi_i, D^\beta\varphi_j)\alpha_i\alpha_j$$

$$\leqq 3\Lambda h^{N-2} \sum_{j=1}^{n} \alpha_j^2$$

by the second assumption. For

$$K(w, w) = \sum_{i,j=1}^{n} K(\varphi_i, \varphi_j)\alpha_i\alpha_j,$$

we find by combining the above inequalities that

$$c_{27}^{-1}h^N \sum_{j=1}^{n} a_j^2 \leqq \sum_{i,j=1}^{n} K(\varphi_i, \varphi_j)\alpha_i\alpha_j \leqq c_{27}h^{N-2} \sum_{j=1}^{n} \alpha_j^2$$

with a constant c_{27} depending only on λ, Λ and not on h. Hence the largest eigenvalue λ_n of $\mathfrak{R}$ is bounded from above by $c_{27}h^{N-2}$ and the smallest λ_1 is bounded from below by $c_{27}^{-1}h^N$. Since $\mathfrak{R}$ is symmetric, we have

$$\|\mathfrak{R}\| \times \|\mathfrak{R}^{-1}\| = \lambda_n/\lambda_1 \leqq c_{27}^2 h^{-2}.$$

REFERENCES

[1] S. AGMON, *Lectures on Elliptic Boundary Value Problems*, Van Nostrand, Princeton, N.J., 1965.

[2] J. P. AUBIN, *Approximation des problèms aux limites non homogènes et régularité de la convergence*, Calcolo, 6 (1969), pp. 117–139.

[3] ———, *Variational methods for non-homogeneous boundary value problems*, SYNSPACE (1970), B. E. Hubbard, ed., Academic Press, New York, 1971.

[4] I. BABUŠKA, *Approximation by hill functions*, Tech. Note BN-648, Univ. of Maryland, College Park, 1970.

[5] ———, *Numerical solution of boundary value problems by the perturbed variational principle*, Tech. Note BN-624, Univ. of Maryland, College Park, 1969.

[6] J. H. BRAMBLE AND S. HILBERT, *Bounds for a class of linear functionals with application to Hermite interpolation*, Numer. Math., 16 (1971), pp. 362–369.

[7] J. H. BRAMBLE AND M. ZLÁMAL, *Triangular elements in the finite element method*, Math. Comp., 24 (1970), pp. 809–820.

[8] J. H. BRAMBLE AND A. H. SCHATZ, *Rayleigh–Ritz–Galerkin methods for Dirichlet's problem using subspaces without boundary conditions*, Comm. Pure Appl. Math., 23 (1970), pp. 653–675.
[9] ———, *On the numerical solution of elliptic boundary value problems by least squares approximation of the data*, SYNSPADE (1970), B. E. Hubbard, ed., Academic Press, New York, 1971.
[10] ———, *Least squares methods for $2m^{th}$ order elliptic boundary value problems*, Math. Comp., 25 (1970), pp. 1–33.
[11] F. DI GUGLIELMO, *Construction d'approximations des espaces de Sobolev sur des reseaux en simplexes*, Calcolo, 6 (1969), pp. 279–331.
[12] ———, *Méthode des éléments finis: Une famille d'approximations des espaces de Sobolev par les translates de p fonctions*, Ibid., 7 (1970), pp. 185–234.
[13] G. FIX AND G. STRANG, *Fourier analysis of the finite element method in Ritz–Galerkin theory*, Studies in Appl. Math., 48 (1969), pp. 265–273.
[14] J. L. LIONS AND E. MAGENES, *Problèmes aux limites non homogènes et applications*, vol. 1, Dunod, Paris, 1968.
[15] J. NITSCHE, *Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind*, Abh. Math. Sem. Univ. Hamburg, 36 (1970/71), pp. 9–15.
[16] M. H. SCHULTZ, *Rayleigh–Ritz–Galerkin methods for multidimensional problems*, this Journal, 6 (1969), pp. 523–538.
[17] G. STRANG, *The finite element method and approximation theory*, SYNSPADE (1970), B. E. Hubbard, ed., Academic Press, New York, 1971.
[18] M. ZLÁMAL, *On the finite element method*, Numer. Math., 12 (1968), pp. 394–409.