71-2

Research Report Number 71 - 2

Department of Computer Science

YALE UNIVERSITY

New Haven, Connecticut 06520

April 1971

L^2 -ERROR BOUNDS FOR THE RAYLEIGH-RITZ-GALERKIN METHOD

by

#3 Martin H. Schultz 71-2

This research supported in part by the National Science Foundation, GP 11326, and the Office of Naval Research (NR 044-401)

§0. Introduction.

In this paper we obtain new a priori bounds for the error involved in approximating the solutions to a wide variety of boundary value problems for linear and nonlinear elliptic partial differential equations by the Rayleigh-Ritz-Galerkin method.

Throughout this paper, let $\Omega \subset R^N$ be an open set, m be a positive integer, $W_0^{m,2}(\Omega)$ denote the closure of the real-valued functions $f \in C_0^\infty(\Omega)$ with respect to the norm $||f||_m \equiv (\int_\Omega \sum_{0 \leq |\alpha| \leq m} |D^\alpha f|^2 dx)^{1/2}$, where we have used the standard multi-index notation, cf. [21], $W^{m,2}(\Omega)$ denote the closure of the real-valued functions $f \in C^\infty(\Omega)$ with respect to $|| \circ ||_m$, and K denote a positive constant not necessarily the same at each occurrence. If we are considering a problem of order 2m, the "classical" a priori error bounds are in the Sobolev norm $|| \circ ||_m$, cf. [18], for these types of result and a list of references. Such error bounds trivially induce error bounds in the lower order Sobolev norms, $|| \circ ||_j$, $0 \leq j \leq m-1$. But these induced error bounds are <u>not</u> sharp.

Recently, J. Nitsche has shown how to directly obtain error bounds in the norm $|| \circ ||_0$ for the special case of $m = 1$, $N = 1$, and $\Omega \equiv (a,b)$, cf. [14]. S. Hilbert has adapted Nitsche's technique to the special case of $m = 1$ and $N = 2$, cf. [11]. In this paper we develop generalizations of Nitsche's technique and show how they can be applied to semilinear elliptic problems, linear and semilinear forced vibration problems, and eigenvalue problems as well as linear elliptic problems.

Roughly speaking, if $||e_h||_{m,2} = O(h^p)$ as $h \to 0$, where e_h denotes the error in approximating the solution to a problem by using the Rayleigh-Ritz-Galerkin method over a finite dimensional subspace parameterized by h, then we obtain $||e_h||_j = O(h^{p+m-j})$ as $h \to 0$, for all $0 \leq j \leq m-1$.

We remark that analogous results have been obtained by entirely different techniques for the case of semi-linear problems with $N = 1$ and special subspaces by Hulme, cf. [12], Perrin, Price, and Varga, cf. [15], and Schultz, cf. [20], and the case of linear problems with the $\Omega \equiv \mathbb{R}^N$ and special subspaces by Fix and Strang, cf. [10].

§1. Linear Elliptic Problems.

Throughout this paper, let H be any closed subspace of $W^{m,2}(\Omega)$ such that $W_0^{m,2}(\Omega) \subset H \subset W^{m,2}(\Omega)$, and $a(u,v)$ be a real-valued, bounded, bilinear form over H such that there exist constants $0 < \gamma \leq \mu$ such that

(1.1) $$\gamma||u||_m^2 \leq a(u,u) \leq \mu||u||_m^2, \quad \text{for all } u \in H.$$

Given a real-valued function $g \in W^{0,2}(\Omega)$, our problem is to find $u \in H$ such that

(1.2) $$a(u,v) = (g,v)_0 \equiv \int_\Omega g(x)v(x)dx, \quad \text{for all } v \in H.$$

If S is any finite dimensional subspace of H, then the Galerkin method is to find $u_S \in S$ such that

(1.3) $$a(u_S,w) = (g,w)_0, \quad \text{for all } w \in S.$$

The following Theorem is a combination of the results of [5] and [7].

Theorem 1.1. If (1.1) holds, (1.2) has a unique solution, u, (1.3) has a unique solution, u_S, for every finite dimensional subspace of H and

(1.4) $$||u - u_S||_m \leq \gamma^{-1}\mu \inf_{y \varepsilon S} ||u - y||_m.$$

Inequality (1.4) has been combined with results from approximation theory by a number of authors, cf. [18] for a list of references, to give a priori error bounds for the Galerkin method. In this paper we will show that under an additional hypothesis we can give a priori error bounds directly, in $W^{j,2}(\Omega)$, where $0 \leq j$ m-1.

We now make the additional assumption that a(u,v) is strongly coercive over H, i.e., that every solution, u, (1.2) is in $W^{2m,2}(\Omega)$ and that there exists a positive constant, ρ, such that

(1.5) $$||u||_{2m} \leq \rho||g||_0 \quad \text{for all } g \ \varepsilon \ W^{0,2}(\Omega).$$

Using a fundamental technique of Nitsche, cf. [14], we have the following generalization of analogous results of [5], [7], and [18].

Theorem 1.2. Let (1.1) hold, a(u,v) be strongly coercive over H, and C be a collection of finite dimensional subspaces of H such that if $g \ \varepsilon \ W^{2m,2}(\Omega) \cap H$, then there exists a positive, real-valued function E on C such that

(1.6) $$\inf_{y \varepsilon S} ||g - y||_m \leq E(S)||g||_p ,$$

for some $m \leq p \leq 2m$ and all $S \in C$. Then if u_S denotes the Galerkin approximation to u in S,

(1.7) $$||u - u_S||_0 \leq \rho\gamma^{-1}\mu^2 E(S) \inf_{y\varepsilon S} ||u - y||_m, \quad \text{for all } S \in C.$$

Proof. Given $S \in C$, define $\psi_S \equiv (u - u_S)/||u - u_S||_0$ and a new bilinear form $b(u,v) = a(v,u)$. We consider the problem

(1.8) $$b(u,v) = (\psi_S,v)_0 \quad \text{for all } v \in H,$$

which has a unique solution, ϕ, by Theorem 1.1.

Thus, $b(\phi,u - u_S) = a(u - u_S,\phi) = (\psi_S,u - u_S) = ||u - u_S||_0$. If $s \in S$, then by the definition of the Galerkin method $a(u - u_S, \phi - s) = (\psi_S,u - u_S)_0$ and hence

(1.9) $$||u - u_S||_0 = a(u - u_S, \phi - s) \leq \mu||u - u_S||_m||\phi - s||_m$$

$$\leq \mu E(S)||\phi||_p||u - u_S||_m \leq \rho\mu E(S)||\psi_S||_0||u - u_S||_m.$$

The result followsbby applying (1.4) to the right hand side of (1.9). QED.

As a corollary of the preceeding result, we have

Theorem 1.3. Let (1.1) hold, a(u,v) be strongly coercive over H, P be a collection of positive parameters, h, and $\{S_{m,r,h}\}_{h\varepsilon P}$, $m \leq r$, be a collection of finite dimensional subspaces of H such that there exists a positive constant M such that $\inf_{y\varepsilon S_{k,r,h}} ||\phi - y||_\ell < M h^{j-\ell}||\phi||_j$ for all

$0 \leq \ell \leq m \leq j \leq r$, $h \in P$, and $\phi \in W^{j,2}(\Omega) \cap H$. Then if u_h denotes the Galerkin approximation to $u \in W^{t,2}(\Omega)$, $2m \leq t$, in $S_{m,r,h}$,

$$(1.10) \qquad ||u - u_h||_0 \leq \rho\gamma^{-1}\mu^2 M^2 h^{s+p-2m}||u||_t, \text{ for all } h \in P,$$

where $s \equiv \min(r, 2m)$ and $p \equiv \min(r,t)$.

The construction of such spaces has attracted much attention recently. Examples include spaces of multivariate spline functions, cf. [18], spaces of multivariate L-spline functions, cf. [18], and finite element spaces, cf. [8] and [10].

Following Agmon, cf. [1], we say that a set $\Omega \subset R^n$ has the restricted cone condition if $\partial\Omega$ has a locally finite open covering $\{0_i\}$ and corresponding cones $\{C_i\}$ with vertices at the origin and the property that $x + C_i \subset \Omega$ for $x \in \Omega \cap 0_i$.

As a corollary of Theorem 1.3, we have the following result.

Theorem 1.4. Let the hypotheses of Theorem 1.3 hold and Ω have the restricted cone condition. Then there exists a positive constant, K, such that

$$(1.11) \qquad ||u - u_h||_j \leq K\, h^{s+p-2m-j}||u||_t, \text{ for all } h \in P \text{ such that}$$

$0 < h \leq 1$, $0 \leq j \leq m$, where $s \equiv \min(r,2m)$, $p \equiv \min(r,t)$, and u_h denotes the Galerkin approximation to $u \in W^{t,2}(\Omega)$, $2m \leq t$, in $S_{m,r,h}$.

Proof. For each $0 \leq j \leq m$ and $u \in W^{m,2}(\Omega)$, let

$|u|_j \equiv (\sum_{|\alpha|=j} \int_\Omega [D^\alpha u(x)]^2 dx)^{1/2}$. By Theorem 3.3, pg.24 of [1], there exists a positive constant, η, such that

$$(1.12) \qquad |u|_j^2 \le \eta(\varepsilon^{m-j}|u|_m^2 + \varepsilon^{-j}|u|_0^2),$$

for all $1 \le j \le m-1$, $u \in W^{m,2}(\Omega)$, and $0 < \varepsilon \le 1$. The result follows by combining (1.12) with Theorems 1.1 and 1.4 and setting $\varepsilon = h^2$. In fact, we obtain

$$|u|_j^2 \le \frac{K}{2}(h^{4m+2p-4m-2j} + h^{2s+2p-4m-2})\ ||u||^2_{W^{t,2}(\Omega)} \le K\ h^{2s+2p-4m-2j}||u||^2_{W^{t,2}(\Omega)}$$

since $s \equiv \min\ (r,2m)$. QED.

For completeness, we now give some examples of boundary value problems for linear elliptic partial differential equation which give rise to strongly coercive forms. But first we introduce some additional terminology. Following Agmon, cf. [1], again we define an open set $\Omega \subset R^n$ to be of class C^k, $k \ge 1$, if for every $x \in \partial\Omega$ there is an open neighborhood U such that for some $1 \le i \le n$, $U \cap \partial\Omega$ has the representation $x_i = g(x^i) \in C^k$, $x^i \equiv (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ in U^i, the projection of U onto the hyperplane $x_i = 0$, and $U \cap \Omega$ is contained in the half-cylinder $\{x | x_i > g(x^i) \text{ for } x^i \in U^i\}$.

We will discuss the Neumann problem for second order problems in C^2 domains, then we will discuss the Dirichlet problem for even order problems in smooth domains, and finally the Dirichlet problem for second order problems in convex polygons in the plane.

We consider the problem of finding the generalized solution of

(1.13) $$-\sum_{i,j=1}^{n} D_i(a_{ij}(x)D_j u(x)) + c(x)u(x) = f(x),\ x \in \Omega,$$

where Ω is bounded, open subset of R^n, $a_{ij}(x)$, $1 \leq i, j \leq n$, and $c(x)$ are real-valued, $C^1(\overline{\Omega})$ functions, $c(x) \geq k > 0$ for all $x \in \Omega$, $a_{ij}(x) = a_{ji}(x)$, $1 \leq i, j \leq n$, $x \in \Omega$, and $\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq k_1|\xi|^2$ for any $x \in \Omega$ and any $\xi \in R^n$ with $|\xi|^2 \equiv \xi_1^2 + \ldots + \xi_n^2$, subject to the Neumann boundary conditions

(1.14) $$\sum_{i,j=1}^{n} a_{ij}(x)\ D_j u(x)\ \cos(v,x_i) = 0, \text{ for } x \in \partial\Omega,$$

where v is the outer normal on $\partial\Omega$. We define a bilinear form on $W^{1,2}(\Omega)$ by

$$a(u,v) \equiv \int_\Omega \left[\sum_{i,j=1}^{n} a_{ij}(x)D_i u(x)D_j v(x) + c(x)u(x)v(x) \right]dx.$$

By our assumptions,

$$\min (k,k_1)||u||_1^2 \leq |a(u,u)| \leq \max\left(\max_{x\in\Omega} |c(x)|\ ,\ \max_{1\leq i,j\leq n} \max_{x\in\Omega}|a_{ij}(x)|\right)||u||_1^2,$$

so that (1.1) holds. By a generalized solution of (1.13) – (1.14) we mean a function $u \in W^{1,2}(\Omega)$ such that

(1.15) $$a(u,v) = (f,v)_0, \quad \text{for all } v \in W^{1,2}(\Omega).$$

Lions has shown, cf. [13, pg. 111], that if Ω is of class C^2 and the above hypotheses hold, then $a(u,v)$ is strongly coercive over $H \equiv W^{1,2}(\Omega)$.

Now we consider the problem of finding the generalized solution of

(1.16) $$\sum_{0\leq|\alpha|,|\beta|\leq m} (-1)^{|\alpha|} D^{\alpha}[a_{\alpha\beta}(x)D^{\beta}u(x)] = f(x), \quad x \in \Omega,$$

where Ω is a bounded, open subset of R^n, and $a_{\alpha\beta}(x)$, $0\leq|\alpha|,|\beta|\leq m$, are real-valued functions with $a_{\alpha\beta}(x) \in C^{|\alpha|}(\overline{\Omega})$ such that there exists a positive constant, γ, with

$$\int_\Omega \sum_{0\leq|\alpha|,|\beta|\leq m} a_{\alpha\beta}(x)D^{\alpha}wD^{\beta}w \, dx \geq \gamma||w||_m^2$$

for all $w \in W_0^{m,2}(\Omega)$, subject to the Dirichlet boundary conditions

(1.17) $$D^{\alpha}u(x) = 0, \text{ for all } x \in \partial\Omega, \quad 0 \leq |\alpha| \leq m-1.$$

We define a bilinear form on $W_0^{m,2}(\Omega) \times W_0^{m,2}(\Omega)$ by

$$a(u,v) \equiv \int_\Omega \sum_{0\leq|\alpha|,|\beta|\leq m} a_{\alpha\beta}(x)D^{\beta}u(x)D^{\alpha}v(x)dx.$$

By our assumptions,

$$\gamma||u||_m^2 \le |a\ (u,u)| \le (\max_{0\le|\alpha|,|\beta|\le m}\ \max_{x\in\Omega} |a_{\alpha\beta}(x)|)\ ||u||^2$$

so that (1.1) holds. By a generalized solution of (1.16) - (1.17) we mean a function $u \in W_0^{m,2}(\Omega)$ such that $a(u,v) = (f,v)_0$, for all $v \in W_0^{m,2}(\Omega)$. Agmon has shown, cf. [1, pg. 132], that if Ω is of class C^{2m} and the above hypotheses hold, then $a(u,v)$ is strongly coercive over $H \equiv W_0^{m,2}(\Omega)$. Moreover, if in addition, $n = 2$, $m = 2$, $a_{01}(x) = a_{10}(x)$ for all $x \in \Omega$, and Ω is a convex polygon, then Birman and Skvortsov have shown, cf., [3], that $a(u,v)$ is strongly coercive over $H \equiv W_0^{1,2}(\Omega)$.

§2. Linear Forced Vibration Problems.

In this section, we extend the methods and results of §1 to linear forced vibration problems. Finite difference methods for approximating the solutions of such problems have been discussed in [4]. In particular, we consider the problem of finding $u \in H$ such that

(2.1) $$a(u,v) + (pu,v)_0 = (g,v)_0,$$

for all $v \in H$, where we assume that (i) (1.1) holds, (ii) $a(u,v)$ is symmetric, i.e., $a(u,v) = a(v,u)$ for all $u,v \in H$, (iii) $p(x)$ is real-valued and belongs to $L^{\infty}(\Omega)$, (iv) $g(x)$ is real-valued and belongs to $W^{0,2}(\Omega)$, and (v) if $g(x) \equiv 0$, then $u \equiv 0$ is the only solution of (2.1).

Theorem 2.1. If hypotheses (i) - (v) hold, there exists a linear, compact mapping, T, of $W^{0,2}(\Omega)$ into H such that u satisfies (2.1) if and only if

(2.2) $$a(u,v) = -\,a(Tpu,v) + a(Tg,v), \text{ for all } v \in H,$$

or

(2.3) $$u = -\,Tpu + Tg.$$

Proof. $F[v] \equiv (g,v)_0 - (pu,v)_0$ is a bounded linear functional of v with respect to the inner product $a(u,v)$ in H. In fact,

$$\begin{aligned}|F[v]| &\le \left(\max_{x\in\Omega} |p(x)|\ ||u||_0 + ||g||_0\right)||v||_0 \\ &\le \left(\max_{x\in\Omega} |p(x)|\ ||u||_0 + ||g||_0\right)\gamma^{-1/2}[a(v,v)]^{1/2}.\end{aligned}$$

Thus, by the Riesz Representation Theorem, cf. [21], (2.2) is equivalent to (2.1). Moreover, T is linear.

To show that T is compact from $W^{0,2}(\Omega)$ to $H \subset W^{m,2}(\Omega)$, we first observe that T takes sets bounded in $W^{0,2}(\Omega)$ into sets bounded in $W^{m,2}(\Omega)$ and hence by Rellich's Theorem, cf. [21], that T is compact from $W^{0,2}(\Omega)$ to $W^{0,2}(\Omega)$. Finally, if S is any bounded set in $W^{0,2}(\Omega)$, there exists a sequence $\{u_n\}_{n=1}^{\infty} \subset S$ such that Tu_n is a Cauchy sequence in $W^{0,2}(\Omega)$. But,

$$a(Tu_n - Tu_k, Tu_n - Tu_k) = ([u_n - u_k], Tu_n - Tu_k)_0 + (g, Tu_n - Tu_k)_0$$

$$\leq (||u_n - u_k||_0 + ||g||_0)||Tu_n - Tu_k||_0$$

and hence $\{Tu_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $W^{m,2}(\Omega)$. QED.

By applying the well-known Fredholm Alternative Theorem, cf. [16], to (2.3), we have

Theorem 2.2. If hypotheses (i) - (v) hold, equation (2.1) has a unique solution.

Now let $\{S_n\}_{n=1}^{\infty}$ be a sequence of finite dimensional subspaces of H such that

(2.4) $$\lim_{n\to\infty} \inf_{y \varepsilon S_n} ||\phi - y||_m = 0, \text{ for all } \phi \in H \text{ and } \{P_n\}_{n=1}^{\infty}$$

be the sequence of linear, orthogonal, projection mappings of H onto S_n

with respect to the inner product $a(u,v)$. We remark that $||P_n|| \leq (\gamma^{-1}\mu)^{1/2}$.

The Galerkin method for (2.1) is to determine $u_n \in S_n$ such that

(2.5) $$a(u_n,w) + (pu_n,w)_0 = (g,w)_0, \quad \text{for all } w \in S_n,$$

or equivalently

(2.6) $$u_n = - P_nTpu_n + P_nTg.$$

Theorem 2.2. If hypotheses (i) - (v) hold, the approximate problems (2.5) or (2.6) have a unique solution, u_n, for all n sufficiently large and

(2.7) $$||u - u_n||_m \leq (1 + \varepsilon_n) \inf_{y \in S_n} ||u - y||_m,$$

where $\varepsilon_n \to 0$ as $n \to \infty$.

Proof. Let $||(I - Tp)^{-1}|| = R$, $||Tp - P_nT\ || = \alpha_n$, and $||P_n(TpP_n - Tp)|| = \beta_n$.

By the compactness of T, α_n and $\beta_n \to 0$ as $n \to \infty$. Since $I - P_n Tp =$

$I - Tp + (Tp - P_nTp)$, $I - P_nTp$ is nonsingular for sufficiently large n.

The error bound (2.7) follows by observing that

$$(I - Tp)\ (u - u_n) + (Tp - P_n Tp)\ (u - u_n) + P_n(TpP_n - T)\ (u - u_n) =$$

$(I - Tp)\ (u - P_n u) + (Tp - P_n Tp)\ (u - P_n u)$, multiplying by $(I - Tp)^{-1}$

taking norms, and setting $\varepsilon_n \equiv \dfrac{R(2\alpha_n + \beta_n)}{1 - R\alpha_n - R\beta_n}$. QED.

We now prove an analogue of Theorem 1.2.

Theorem 2.4. Let hypotheses (i) - (v) hold, $a(u,v)$ be strongly coercive over H, and $\{S_n\}_{n=1}^{\infty}$ be a sequence of finite dimensional subspaces of H such that (2.4) holds and if $g \in W^{2,m2}(\Omega) \cap H$ there exists a sequence, $\{E_n\}_{n=1}^{\infty}$, of positive numbers such that

(2.8) $$\inf_{y \in H_n} ||g - y||_m \leq E_n ||g||_p,$$

for some $m \leq p \leq 2m$ and all $n \geq 0$. Then there exists a positive constant, K, such that for all n sufficiently large

(2.9) $$||u - u_n||_0 \leq KE_n \inf_{y \in S_n} ||u - y||_m.$$

Proof. Let $\psi_n \equiv (u - u_n)/||u - u_n||_0$ and ϕ_n denote the unique solution of $a(u - u_n, \phi_n) + (p(u - u_n), \phi_n)_0 = (\psi_n, u - u_n)_0 = ||u - u_n||_0$.

By the definition of the Galerkin method,

$$a(u - u_n, \phi_n - s) + (p(u - u_n), \phi_n - s)_0 = ||u - u_n||_0 \quad \text{for all } s \in S_n$$

and hence

$$||u - u_n||_0 \leq \mu ||u - u_n||_m ||\phi_n - s||_m + \max_{x \in \Omega} |p(x)| \; ||u - u_n||_0 ||\phi_n - s||_0 .$$

The result now follows as in the proof of Theorem 1.2, since by the strong coerciveness of a(u,v),

$$||\phi_n||_{2m} \leq \rho(||p\phi_n||_0 + ||\psi_n||_0)$$

$$\leq \rho(\max_{x\in\Omega} |p(x)| \; ||(I - Tp)^{-1}|| \; ||T|| + 1)||\psi_n||_0$$

$$= \rho(\max_{x\in\Omega} |p(x)| \; ||(I - Tp)^{-1}|| \; ||T|| + 1). \qquad \text{QED.}$$

§3. Semilinear Elliptic Boundary Value Problems

In this section, we extend the methods and results of the previous sections to semilinear elliptic boundary value problems. In particular, we consider the problem of finding $u \in H$ such that

$$(3.1) \qquad a(u,v) + (f(u),v)_\theta = 0,$$

for all $v \in H$, where we assume that (i) (1.1) holds, (ii) a(u,v) is symmetric, (iii) $f(x,u)$ and $\frac{\partial f}{\partial u}(x,u)$ are real-valued and continuous on $\Omega \times (-\infty,\infty)$, (iv) $|\frac{\partial f}{\partial u}(x,u)| \leq B$ for all $(x,u) \in \Omega \times (-\infty,\infty)$, and

(v) $\frac{\partial f}{\partial u}(x,u) \geq \eta > -\Lambda \equiv -\inf_{w \in H} a(w,w)/(w,w)_0$.

If S is any finite dimensional subspace of H, the Galerkin method is to find $u_S \in S$ such that

(3.2) $$a(u_S,w) + (f(u_S),w)_0 = 0 \quad \text{for all } w \in S.$$

We recall the following result from [7].

Theorem 3.1. Under hypotheses (i) - (v), the problem, (3.1) has a unique solution, u,

(3.3) $||u||_m \leq \gamma^{-1} (1 + \frac{\eta}{\Lambda})^{-1} ||f(0)||_0$, the problem (3.2) has a unique, u_S, for all S, and

(3.4) $$||u - u_S||_m \leq \gamma^{-1}(1 + \frac{\eta}{\Lambda})^{-1} (\mu + B) \inf_{y \in S} ||u - y||_m, \quad \text{for all } S.$$

Following the results of the previous sections, we have the next Theorem.

Theorem 3.2. Let hypotheses (i) - (v) hold, a(u,v) be strongly coercive over H, and C be a collection of finite dimensional subspaces of H such that if $g \in W^{2m,2}(\Omega) \cap H$, then there exists a positive, real-valued function E on C such that

(3.5) $$\inf_{y \in S} ||g - y||_m \leq E(S)||g||_p$$

for some $m \leq p \leq 2m$ and all $S \in C$. Then, if u_S denotes the Galerkin approximation to u in S,

(3.6) $||u - u_S||_0 \leq (\mu + B)^2\gamma^{-1}\rho[B\gamma^{-1}(1 + \frac{\eta}{\Lambda})^{-1} + 1](1 + \frac{\eta}{\Lambda})^{-1} E(S) \inf_{y \in S} ||u - y||_m$

for all $S \in C$.

Proof. Let S be any subspace in $C, \psi_S \equiv (u - u_S)/||u - u_S||_0$, and ϕ_S be the unique solution of the linear problem

(3.7) $a(v,\phi_S) + (\frac{\partial f}{\partial u}(\theta u - (1 - \theta) u_S)v, \phi_S)_0 = (\psi_S,v)_0$, for all $v \in H$, where $0 < \theta < 1$.

From Theorem 3.1, we have that $||\phi_S||_m \leq \gamma^{-1}(1 + \frac{\eta}{\Lambda})^{-1}$ and from the strong coerciveness of $a(u,v)$ we have that $||\phi_S||_{2m} \leq \rho[B\gamma^{-1}(1 + \frac{\eta}{\Lambda})^{-1} + 1]$.

Moreover, from (3.7) we have that for all $y \in S$,

(3.8) $$||u - u_S||_0 = a(u - u_S, \phi_S - y) + (\frac{\partial f}{\partial u}[u - u_S], \phi_S - y)$$

and hence

(3.9) $$||u - u_S||_0 \leq (\mu + B)^2 \gamma^{-1} \rho[B\gamma^{-1}(1 + \frac{\eta}{\Lambda})^{-1} + 1] \cdot (1 + \frac{\eta}{\Lambda})^{-1} E(S) \inf_{y \in S} ||u - y||_m.$$

QED.

§4. Semilinear Forced Vibration Problems.

In this section, we study the Galerkin method for approximating the solutions of semilinear, forced vibration problems. In particular, we consider the problem of finding $u \in H$ such that

$$(4.1) \qquad a(u,v) + (f(u),v)_0 = 0$$

for all $v \in H$, where we assume that (i) holds, (ii) $a(u,v)$ is symmetric, (iii) $f(x,u)$ and $\frac{\partial f}{\partial u}(x,u)$ are real-valued and continuous on $\Omega \times (-\infty,\infty)$, (iv) $|f(x,u)| \le B$ for all $(x,u) \in \Omega \times(-\infty,\infty)$, and

(v) $-\lambda_{j+1} < \eta_{j+1} \le \frac{\partial f}{\partial u}(x,u) \le \eta_j < -\lambda_j$ for all $(x,u) \in \Omega \times (-\infty,\infty)$,

where $0 < \lambda_1 \le \lambda_2 \le \ldots$ are the eigenvalues associated with $a(u,v)$ over H.

Let S_n be any finite dimensional subspace of H. Then the Galerkin method is to find $u_n \in S_n$ such that

$$(4.2) \qquad a(u_n,w) + (f(u_n),w)_0 = 0,$$

for all $w \in S_n$. We recall the following result from [17].

Theorem 4.1. Under hypotheses (i) - (v), problem (4.1) has a unique solution, u, and $||u||_m \le \gamma^{-1}B\ (\text{meas}\ \Omega)^{1/2}$. If $\{S_n\}_{n=1}^{\infty}$ is a sequence of finite dimensional subspaces of H such that

(4.3) $$\lim_{n\to\infty} \inf_{y\in S_n} ||u - y||_m = 0$$

for all $u \in H$, then problem (4.2) has a solution, u_n, which is unique for all sufficiently large n and

(4.4) $$||u - u_n||_m \leq (1 + \varepsilon_n) \inf_{y\in S_n} ||u - y||_m,$$

for all sufficiently large n, where $\varepsilon_n \downarrow 0$ as $n \to \infty$.

Using the Fredholm alternative, cf. [16], and hypotheses (i) - (v), we can prove the following Theorem in essentially the same way that we proved Theorem 3.2. The exact details are left to the reader.

Theorem 4.2. Let $a(u,v)$ be strongly coercive over H, hypotheses (i) - (v) hold, and $\{S_n\}_{n=1}^{\infty}$ be a sequence of finite dimensional subspaces of H such that (4.3) holds and such that if $g \in W^{2m,2}(\Omega) \cap H$, there exists a sequence, $\{E_n\}_{n=1}^{\infty}$, of positive real-numbers such that

(4.5) $$\inf_{y\in S_n} ||g - y||_m \leq E_n ||g||_p$$

for some $m \leq p \leq 2m$ and all $n \geq 1$. Then there exists a positive constant, K, such that

(4.6) $$||u - u_n||_0 \leq KE_n \inf_{y\in S_n} ||u - y||_m,$$

for all sufficiently large n.

§5. Eigenvalue Problems.

In this section, we study the Rayleigh-Ritz-Galerkin method for approximating the eigenvalues and eigenvectors of symmetric, elliptic problems. In particular, we consider the problem of finding λ and $u \not\equiv 0$ in H such that

$$(5.1) \qquad a(u,v) - \lambda(pu,v)_0 = 0,$$

for all $v \in H$, where we assume that (i) (1.1) holds, (ii) $a(u,v)$ is symmetric, and (iii) $p(x)$ is real-valued and $0 < \delta \leq p(x) \leq \Delta$ for all $x \in \Omega$.

From Theorem 2.1, we have that under hypotheses (i) - (iii) there exists a linear, compact mapping, T, from $W^{0,2}(\Omega)$ to $W^{m,2}(\Omega)$ such that (5.1) holds for λ and $u \in H$ for all $v \in H$ if and only if

$$(5.2) \qquad u - \lambda \cdot Tu = 0.$$

From the standard theory of compact, self-adjoint operators in Hilbert space, cf. [9], we have the following result.

Theorem 5.1. Under hypotheses (i) - (iii), problem (5.1) or (5.2) has a countable number of normalized, solutions $\{\lambda_i, \phi_i\}_{i=1}^{\infty}$. Moreover,

$0 < \lambda_1 \le \lambda_2 \le \cdots$ and $\lambda_i \to \infty$ as $i \to \infty$, $H \equiv \text{span}\,\{\phi_i\}_{i=1}^{\infty}$, $a(\phi_i,\phi_j) = \delta_{ij}$,

$(p\phi_i, \phi_j)_0 = \lambda_i \delta_{ij}$, and $\lambda_k = \inf\,\{a(w,w)/(pw,w)_0 \mid w \in H,\ (pw,\phi_i)_0 = 0,$

for all $1 \le i \le k - 1\}$ for all $k \ge 1$.

If S_n is any finite dimensional subspace of H, the Rayleigh-Ritz method is to find the extremal points, $\{\phi_i(S_n)\}_{i=1}^{\dim S_n}$, of $R[w] \equiv a(w,w)/(pw,w)_0$ over S_n. We now make the additional hypothesis that (iv) $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_k$, and recall the following result from [2] and [6].

<u>Theorem 5.2</u>. Let hypotheses (i) - (iv) hold and $\{S_n\}_{n=1}^{\infty}$ be any sequence of finite dimensional subspaces of H such that

$$(5.3) \qquad \lim_{n\to\infty} \{ \inf_{y \in S_n} ||\phi_i - y||_m \} = 0$$

for all $1 \le i \le k$. Then $\lambda_i(S_n) \downarrow \lambda_i$ as $n \to \infty$ for all $1 \le i \le k$, where $\lambda_i(S_n) \equiv R[\phi_i(S_n)]$, $||\phi_i - \phi_i(S_n)||_m \to 0$ as $n \to \infty$ for all $1 \le i \le k$.

Moreover, there exists a positive constant, K, such that

$$(5.4) \qquad 0 \le \lambda_k(S_n) - \lambda_k \le K \sum_{i=1}^{k} \inf_{y \in S_n} ||\phi_i - y||_m^2,$$

for all sufficiently large n and

$$(5.5) \qquad ||\phi_k - \phi_k(S_n)||_m \le K \sum_{i=1}^{k} \inf_{y \in S_n} ||\phi_i - y||_m,$$

for all sufficiently large n.

We now prove a new result which gives sharp $L^2(\Omega)$ error bounds for the approximate eigenfunctions. The technique of proof is analogous to the techniques of the previous section.

Theorem 5.3. Let a(u,v) be strongly coercive over H, hypotheses (i) - (iii) hold, $0 < \lambda_1 < \lambda_2$, and $\{S_n\}_{n=1}^{\infty}$ be a sequence of finite dimensional subspaces of H such that (5.3) holds and such that if $g \in W^{2m,2}(\Omega)$ H, there exists a sequence, $\{E_n\}_{n=1}^{\infty}$ of positive, real-numbers such that

$$(5.6) \qquad \inf_{y \in S_n} ||g - y||_m \le E_n ||g||_p$$

for some $m \le p \le 2m$ and all $n \ge 1$. Then there exists a positive constant, K, such that

$$(5.7) \qquad ||\phi_1 - \phi_1(S_n)||_0 \le K\{E_n \inf_{y \in S_n} ||\phi_1 - y||_m + \inf_{y \in S_n} ||\phi_1 - y||_m^2\}$$

for all sufficiently large n.

Proof. Let H_1 denote the orthogonal complement of ϕ_1 in H with respect to the inner product $(pu,v)_0$ and P_1 denote the corresponding orthogonal projection mapping onto H_1. Let $e_n \equiv \phi_1 - \phi_1(S_n) \equiv \alpha_n\phi_1 + \beta_n v_n$, where $v_n \equiv P_1 e_n/(P_1 e_n, P_1 e_n)_0^{1/2}$.

By the Fredholm Alternative, cf. [16], and ellipticity, the problem

(5.8) $$u_n + \lambda_1 Tpu_n = Tv_n$$

has a unique solution, θ_n, in H_1. Moreover, $\gamma(1 - \frac{\lambda_1}{\lambda_2})||\theta_n||_m^2 \leq \mu||v_n||_0||\theta_n||_0$ and hence $||\theta_n||_m \leq \gamma^{-1}\mu(1 - \frac{\lambda_1}{\lambda_2})^{-1}$.

Since $T(\sum_{i=2}^{\infty} \alpha_i\phi_i) = \sum_{i=2}^{\infty} \alpha_i\lambda_i^{-1} \phi_i \in H_1$, we may use the fact that $a(u,v)$ is strongly coercive to obtain

(5.9) $$||\theta_n||_{2m} \leq \rho[\lambda_1\Delta\gamma^{-1} \mu(1 - \frac{\lambda_1}{\lambda_2})^{-1} + 1]$$

Rewriting problem (5.8) we have

(5.10) $$a(\theta_n, e_n) - \lambda_1(p\theta_n, e_n) = (v_n, e_n)_0 = ||P_1 e_n||_0.$$

Using the well-known equivalence of the Rayleigh-Ritz method and the Galerkin method for this problem, we have

(5.11) $$||P_1e_n||_0 = a(\theta_n - y,e_n) - \lambda_1(p(\theta_n - y),e_n)_0 + (\lambda_1(S_n) - \lambda_1)(py,\theta_1(S_n)),$$

for all $y \in S_n$, and hence there exists a positive constant, K, such that

(5.12) $$\beta_n = ||P_1e_n||_0 \leq K\{E_n||e_n||_m + |\lambda_1(S_n) - \lambda_1|\}.$$

Thus, $\beta_n \to 0$ as $n \to \infty$ and hence $\alpha_n \to 0$ as $n \to \infty$. Moreover,

(5.13) $$\phi_1(S_n) = (1 - \alpha_n)\phi_1 + \beta_n v_n$$

and hence

(5.14) $$1 = (p\phi_1(S_n), \phi_1(S_n))_0 = (1 - \alpha_n)^2 + \beta_n^2.$$

Thus,

(5.15) $$\alpha_n = 1 \pm (1 - \beta_n^2)^{1/2}.$$

In face, since $\alpha_n \to 0$ as $n \to \infty$, we must have

(5.16) $$\alpha_n = 1 - (1 - \beta_n^2)^{1/2}.$$

Finally, we have $||e_n||_0 \leq \delta^{-1/2}(pe_n,e_n)_0^{1/2} \leq \delta^{-1/2}(\alpha_n^2 + \beta_n^2)^{1/2}$

and the result follows from the above estimates of α_n and β_n and Theorem 5.2.

QED.

§6. An Example.

In this section we consider a very simple example to show how the results of this paper may be applied, although in the interest of simplicity we will not attempt to obtain the best possible constants. In particular, we consider the problem

(6.1) $$-D[p(x)Du] + q(x)u = g(x), \quad 0 < x < 1$$

subject to the boundary conditions

(6.2) $$u(0) = u(1) = 0$$

where $p(x)$ is a real-valued, continuously differentiable function on $[0,1]$ with $p(x) \geq w > 0$ for all $x \in [0,1]$, $q(x)$ is a continuous, real-valued function on $[0,1]$,

(6.3) $$\gamma||u||_1^2 \leq \int_0^1 \{p(x)(Du)^2 + q(x)u^2\}dx \leq \mu||u||_1^2$$

for all $u \in W_0^{1,2}(0,1)$, and $g(x)$ is a real-valued function in $W^{0,2}(0,1)$.

We define $a(u,v) \equiv \int_0^1 \{p(x)DuDv + q(x)uv\}dx$ and consider the problem of finding $u \in W_0^{1,2}(0,1)$ such that

$$a(u,v) = (g,v)_0 \quad \text{for all } v \in W_0^{1,2}(0,1). \tag{6.4}$$

<u>Theorem 6.1.</u> Under the above hypotheses, $a(u,v)$ is strongly coercive and if u is the solution to (6.4) then

$$||D^2u||_0 \le w^{-1}\, ||g||_0(P_1\gamma^{-1} + Q\gamma^{-1} + 1), \tag{6.5}$$

where $P_1 \equiv \max_{x\varepsilon[0,1]} |Dp(x)|$ and $Q \equiv \max_{x\varepsilon[0,1]} |q(x)|$.

<u>Proof</u>. The fact that $u(x) \in W^{2,2}(0,1)$ is a standard result in ordinary differential equations, cf. [13]. Moreover, by setting $v = u$ in (6.4) and using (6.3) we have

$$||u||_1 \le \gamma^{-1}||g||_0. \tag{6.6}$$

Differentiating out equation (6.1) we get

$$-D^2u = \frac{1}{p(x)}\,[Dp(x)Du - q(x)u + g] \tag{6.6}$$

and the (6.5) follows by applying the bound (6.6). QED.

Now let $\Delta : 0 = x_0 < x_1 < \cdots < x_N < x_{N+1} = 1$ be a partition of the interval [0,1] and S_Δ be the set of all continuous, piecewise linear polynomials with respect to Δ which satisfy the boundary conditions (6.2), i.e. $s(x) \in S_\Delta$ if and only if $s(x)$ is a linear polynomial on each subinterval defined by Δ, $s(x) \in C[0,1]$, and $s(0) = s(1) = 0$. Combining Theorems 1.2 and 6.1 with Theorem 3.5 of [19] we obtain the following explicit error bound.

Theorem 6.2. Under the above hypotheses, if u_Δ denotes the Rayleigh-Ritz-Galerkin approximation, then

$$(6.7)\quad ||u - u_\Delta||_0 \le \gamma^{-1}\mu^2 w^{-2}||g||_0 (P_1\gamma^{-1} + Q\gamma^{-1} + 1)\ (1 + \bar{h}^2\,\pi^{-2})\bar{h}^2 ,$$

where $\bar{h} \equiv \max_{0 \le i \le N} (x_{i+1} - x_i)$, for all partitions Δ.

We remark that this result shows that we have a second order scheme even for nonuniform meshes!

References

[1] Agmon, S., Lectures on Elliptic Boundary Value Problems. (291 pp.). Princeton: Van Nostrand (1965).

[2] Birkhoff, G., C. deBoor, B. Swartz, and B. Wendroff, Rayleigh-Ritz approximation by piecewise cubic polynomials. SIAM J. Numer. Anal. 3, 188-203 (1966).

[3] Birman, M.S. and G.E. Skvortsov, On the summability of the highest order derivatives of the solution of the Dirichlet problem in a domain with piecewise smooth boundary. Izv. Vyssh. Uchebn. Zaved Matematika 30, 12-21 (1962).

[4] Bramble, J.H., Error estimates for difference methods in forced vibration problems. J. SIAM Numer. Anal. 3, 1-12 (1966).

[5] Céa, J., Approximation variationnelle des problèmes aux limites. Ann. Inst. Fourier (Grenoble) 14, 345-444 (1964).

[6] Ciarlet, P.G., M.H. Schultz, and R.S. Varga, Numerical methods of high-order accuracy nonlinear boundary value problems. III. Eigenvalue problems. Numer. Math. 12, 120-133 (1968).

[7] Ciarlet, P.G., M.H. Schultz, and R.S. Varga, Numerical methods of high-order accuracy for nonlinear boundary value problems. V. Monotone operators. Numer. Math. 13, 51-77 (1969).

[8] di Guglielmo. F., Construction d'approximations des espaces de Sobolev sur des reseaux en simplexes. Estratto da Calcole 6, 279-331 (1969).

[9] Gould, S.H., Variational Methods for Eigenvalue Problems (275 pp.) Toronto: University of Toronto Press (1966).

[10] Fix,G. and G. Strang, Fourier analysis of the finite element method in Ritz-Galerkin theory. To appear.

[11] Hilbert, S., Numerical methods for elliptic boundary value problems. Doctoral Thesis. University of Maryland, 1969.

[12] Hulme, B.L., Interpolation by Ritz approximation. J. Math. Mech. 18, 337-342 (1968).

[13] Lions, J.L., Problems aux limites dans les equations aux derivees partielles. Lecture Notes Universite de Montrial (1962).

[14] Nitsche, J., Ein kriterium fur die quasi-optimalitat des Ritzschen verfohrens. Num. Math. 11, 346-348 (1968).

[15] Perrin, F.M., H.S.,Price, and R.S. Varga, On higher-order numerical methods for nonlinear two-point boundary value problems. Numer. Math. 13, 180-198 (1969).

[16] Riesz, F. and B. Sz-Nagy, Functional Analysis, (468 pp.). New York: F. Ungar Publishing Co. (1955).

[17] Schultz, M.H., Error bounds for the Rayleigh-Ritz-Galerkin method. Journal of Mathematical Analysis and Applications 27, 524-533 (1969).

[18] Schultz, M.H., Multivariate spline functions and elliptic problems. Approximations with Special Emphasis on Spline Functions, 279-347. New York: Academic Press 1969.

[19] Schultz, M.H., Error bounds for polynomial spline interpolation. Mathematics of Computation 24, 507-515 (1970).

[20] Schultz, M.H., Elliptic spline functions and the Rayleigh-Ritz-Galerkin method. Mathematics of Computation 24, 65-80 (1970).

[21] Yosida, K., Functional Analysis (458 pp.). New York: Academic Press (1965).