MATHEMATICS OF COMPUTATION, VOLUME 33, NUMBER 146
APRIL 1979, PAGES 493–519

Convergence of a Finite Element Method for the Approximation of Normal Modes of the Oceans

By Mitchell Luskin

Abstract. This paper gives optimal order error estimates for the approximation of the spectral properties of a variant of the shallow water equations by a finite element procedure recently proposed by Platzman. General results on the spectral approximation of unbounded, selfadjoint operators are also given in this paper.

I. Introduction. We analyze in this paper an approximation procedure to compute the normal modes of the oceans. The procedure that we analyze is a finite element method proposed by Platzman [16].

Our procedure gives a finite dimensional operator whose properties are shown to approximate those of the unbounded selfadjoint operator associated with the differential equations modeling the physical problem. The error estimates given here are optimal for the procedure considered. The theoretical results on spectral approximation given in this paper generalize the error estimates of Bramble and Osborn [1] and Osborn [14] for compact operators and of Descloux, Nassif, and Rappaz [5], [6] for bounded operators. Results of Descloux, Rappaz, and the author on the spectral convergence of unbounded, closed (not necessarily selfadjoint) operators will appear in a later paper [4].

We model the time dependent behavior of the oceans by Laplace's tidal equations, a variant of the shallow water equations, as discussed by Platzman in [15] where normal modes of the Atlantic and Indian Oceans calculated by a finite difference procedure are presented. A Lanczos method used to solve the resulting matrix eigenproblem is discussed in [3]. The frequencies of the normal modes have limit points at 0 and ∞. The spectral properties of the associated unbounded, selfadjoint operator have been studied by Veltkamp [18].

Platzman [16], Dupont [7] and Scott [17] have found that the standard Galerkin method for hyperbolic systems [7], [8] can cause modes with high wavenumber to have low or zero frequency in time even though corresponding eigenfunctions of the differential equations with high wavenumber have a high frequency in time. This behavior is clearly unacceptable in an approximate procedure for finding eigenvalues and eigenvectors.

In Section 2, we study a hyperbolic system for which both the differential and Galerkin spectral properties can be computed analytically. The above-mentioned phenomenon is clearly evident in this example. This example is due to Platzman [16].

Received February 6, 1978.
AMS (MOS) subject classifications (1970). Primary 65N25, 65N30.

0025-5718/79/0000-0052/$07.75

We also study the properties of Platzman's alternative Galerkin method when applied to the example of Section 2. We find that eigenfunctions of this Galerkin method with high wavenumber have a high frequency in time. This agrees with what is to be expected from the differential problem.

In Section 3, we construct the operator, T, associated with Laplace's tidal equations. We also define the finite dimensional operators, $\{T^h\}$, associated with our finite element method.

We prove in Section 5 two properties concerning the convergence of T^h to T. In Section 4, error estimates for the appproximation of the spectral properties of T by T^h are derived from these two properties.

In Section 6, we improve the eigenvalue estimate given in Section 4. It can be seen by inspecting the example in Section 2 that the results on the convergence of the eigenspaces in Theorem 1 and the result on the convergence of the eigenvalues in Theorem 4 are optimal.

The procedure discussed here for the computation of eigenvalues and eigenfunctions of hyperbolic systems has many applications to time dependent problems. This is discussed in a later paper by the author [12].

II. An Example. In this section we study the approximation properties of two Galerkin methods when applied to the solution of a simple hyperbolic system with two dependent variables defined in one space dimension. This system models a one-dimensional channel.

We denote by $C(I)$ the space of continuous complex-valued functions on I and by P_1, the space of complex-valued, linear functions. Also, for $f, g \in L^2(I)$, we set $(f, g) = \int_0^1 f\bar{g}\,dx$. We denote by $H^1(I)$ the Sobolev space of functions with one weak derivative in $L^2(I)$ and norm

$$\|f\|^2_{H^1(I)} = \|f\|^2_{L^2(I)} + \|f'\|^2_{L^2(I)}.$$

We also set

$$H^1_0(I) = \{f \in H^1(I) \mid f(0) = f(1) = 0\}.$$

We consider the hyperbolic system

$$\begin{aligned} &\frac{\partial u}{\partial t} + \frac{\partial v}{\partial x} = 0, \qquad (x, t) \in (0, 1) \times (0, T), \\ &\frac{\partial v}{\partial t} + \frac{\partial u}{\partial x} = 0, \qquad (x, t) \in (0, 1) \times (0, T), \\ &v(0, t) = v(1, t) = 0, \qquad t \in (0, T), \\ &u(x, 0) = u_0(x), \quad v(x, 0) = v_0(x), \qquad x \in (0, 1). \end{aligned} \tag{2.1}$$

The associated eigenvalue problem is to find $\lambda \in C$, $u \in H^1(I)$, and $v \in H^1_0(I)$ such that

$$\lambda u + v_x = 0, \quad \lambda v + u_x = 0. \tag{2.2}$$

If we eliminate u from (2.2), we find that λ, v must satisfy

$$\lambda^2 v = v_{xx}, \qquad v(0) = v(1) = 0. \tag{2.3}$$

The solutions to (2.3) are well known to be

$$\lambda_k = ik\pi, \quad v_k = \sin k\pi x, \qquad k = 0, \pm 1, \pm 2, \ldots . \tag{2.4}$$

Thus, a complete set of eigenvalues-eigenvectors for (2.2) is

$$\lambda_k = ik\pi, \quad u_k = i \cos k\pi x, \quad v_k = \sin k\pi x \tag{2.5}$$

for $k = 0, \pm 1, \pm 2, \ldots$.

We now consider two approximation procedures for (2.2). Let $N > 0$ be a positive integer, $h = 1/N$, $x_i = ih$, $I = [0, 1]$, and $I_i = [x_{i-1}, x_i]$. Then set

$$\begin{aligned} M_h &= \{v \in C(I) \mid v|_{I_i} \in P_1 \text{ for } i = 1, \ldots, N\}, \\ M_h^0 &= M_h \cap H_0^1(I). \end{aligned} \tag{2.6}$$

Let the interpolation operator P: $C(I) \longrightarrow M_h$ be defined by the relations

$$Pv(x_i) = v(x_i) \quad \text{for } i = 0, \ldots, N. \tag{2.7}$$

A standard Galerkin method to approximately solve (2.1) is to determine U: $[0, T] \longrightarrow M_h$, V: $[0, T] \longrightarrow M_h^0$ such that

$$\begin{aligned} (U_t, W) + (V_x, W) &= 0, \qquad W \in M_h, \\ (V_t, W) + (U_x, W) &= 0, \qquad W \in M_h^0. \end{aligned}$$

The associated eigenproblem is to find $(\Gamma, U, V) \in \mathbf{C} \times M_h \times M_h^0$ such that

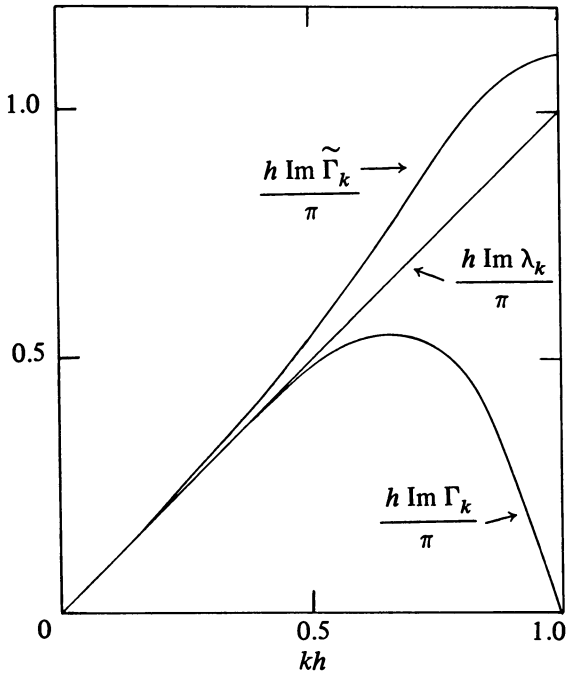
$$\begin{aligned} \Gamma(U, V) + (V_x, W) &= 0, \qquad W \in M_h, \\ \Gamma(V, W) + (U_x, W) &= 0, \qquad W \in M_h^0. \end{aligned} \tag{2.8}$$

A complete set of eigenvalues-eigenvectors for (2.8) is given by

$$\begin{aligned} &\Gamma_k = \frac{i3 \sin kh\pi}{h(2 + \cos kh\pi)}, \quad U_k = P(i \cos k\pi x), \\ &V_k = P(\sin k\pi x) \quad \text{for } k = 0, \pm 1, \ldots, \pm N. \end{aligned} \tag{2.9}$$

Note that $\Gamma_N = 0$, whereas $\lambda_N = iN\pi$. Clearly, this Galerkin method causes eigenfunctions with high wavenumber to have low or zero frequency in time, even though eigenfunctions of the differential problem with high wavenumber have a high frequency in time.

Suppose we examine the graphs of h Im Γ_K/π and h Im λ_k/π as functions of kh, $0 \leqslant kh \leqslant 1$. It is clear that if this procedure is used to compute modes whose frequencies are in a given interval of interest, one will also compute spurious modes due to the spectrum bending.



GRAPH 1 (from Platzman, [16])

In the simple context of (2.1), we can describe our proposed method as follows: Find U: $[0, T] \longrightarrow M_h$, φ: $[0, T] \longrightarrow M_h$ such that

$$(U_t, W) - (\varphi_x, W_x) = 0, \qquad W \in M_h,$$
$$(\varphi_{xt}, W_x) + (U_x, W_x) = 0, \qquad W \in M_h,$$
$$\int_I \varphi \, dx = 0.$$

Then (U, φ_x) is an approximation to (u, v). The associated eigenproblem is to determine $(\Gamma, U, \varphi) \in \mathbf{C} \times M_h \times M_h$ such that

$$\Gamma(U, W) - (\varphi_x, W_x) = 0, \qquad W \in M_h, \tag{2.10a}$$
$$\Gamma(\varphi_x, W_x) + (U_x, W_x) = 0, \qquad W \in M_h, \tag{2.10b}$$
$$\int_I \varphi \, dx = 0. \tag{2.10c}$$

Then we approximate an eigenvalue-eigenvector of (2.2) by (Γ, U, φ_x). Note that we have imposed the boundary values for v implicitly in the equations (2.10) rather than in the trial space as in (2.8). A complete set of eigenvalues-eigenvectors for (2.10) is

$$\tilde{\Gamma}_k = \frac{i}{h} 2 \sin(k\pi h/2) \left(1 - \frac{2}{3} \sin^2 \frac{k\pi h}{2}\right)^{-1/2}, \quad U_k = P(i \cos k\pi x),$$
$$\varphi_k = (\tilde{\Gamma}_k)^{-1} U_k = (\tilde{\Gamma}_k)^{-1} P(i \cos k\pi x) \quad \text{for } k = \pm 1, \ldots, \pm N, \tag{2.11}$$
$$\Gamma_0 = 0, \quad U_0 = 1, \quad \varphi_0 = 0.$$

Recall that v_k is approximated by φ_{k_x}.

We see from Graph 1 that the spectrum, $\{\text{Im } \tilde{\Gamma}_k\}$ does not "bend" as does the graph of $\{\text{Im } \Gamma_k\}$. In fact, it can be seen that $|\text{Im } \tilde{\Gamma}_k| \geqslant |\text{Im } \lambda_k|$ for all $k \geqslant 0$. If we add Γ times (2.10a) to (2.10b), we obtain

$$\Gamma^2(U, W) + (U_x, W_x) = 0, \qquad W \in M_h. \tag{2.12}$$

If we eliminate v from (2.2), we obtain

$$\lambda^2 u - u_{xx} = 0, \qquad x \in I, \qquad u_x(0) = u_x(1) = 0. \tag{2.13}$$

Since (2.12) is the standard Rayleigh-Ritz method for computing the eigenvalues λ^2 of (2.13), it follows that $|\text{Im } \tilde{\Gamma}_k| \geqslant |\text{Im } \lambda_k|$ for all k since it is well known that the approximate eigenvalues will be greater in magnitude than their corresponding differential eigenvalues.

III. The Problem for the Normal Modes of the Oceans. We shall introduce Platzman's procedure [16] to compute the normal modes of a homogeneous ocean in free motion without friction modeled by Laplace's "tidal" equations. The time dependent differential equations are:

$$\frac{\partial \zeta}{\partial t} = -\nabla \cdot \vec{u}, \qquad (x, t) \in \Omega \times (0, T), \tag{3.1a}$$

$$D^{-1}\frac{\partial \vec{u}}{\partial t} = -g\nabla\zeta - D^{-1} f \mathbf{k} \times \vec{u}, \qquad (x, t) \in \Omega \times (0, T), \tag{3.1b}$$

$$\int_\Omega \zeta \, dS = 0, \tag{3.1c}$$

$$\vec{u} \cdot \vec{n} = 0, \qquad (x, t) \in \partial\Omega \times (0, T), \tag{3.1d}$$

where Ω is an open set on the sphere in $\mathbf{R}^3$, $D = D(x)$ is ocean depth, g = gravity, f = Coriolis parameter, ζ = ocean surface elevation above mean level, and $\vec{u}$ = horizontal velocity. Also, ∇ is the horizontal gradient operator on the sphere, $\vec{\mathbf{k}}$ is the unit vertical vector, and $\vec{n}$ is the horizontal exterior normal to the ocean domain.

For the sake of simplicity in our analysis we shall always assume that $D = 1$. We shall also assume that our units are chosen so that $g = 1$. The Coriolis parameter, f, is given by $f = 2\omega \cos\theta$ where ω is the angular velocity of the earth and θ is the geographical co-latitude. So, f is a smooth function with bounded derivatives of all orders on Ω.

Although Ω, the ocean domain, is an open set on the sphere, we wish henceforth for simplicity to assume that Ω is a simply connected open set in $\mathbf{R}^2$ with a smooth boundary, $\partial\Omega$. In this case, we may represent

$$\vec{u}(x, t) = (u_1(x, t), u_2(x, t))$$

and

$$\mathbf{k} \times \vec{u} \equiv (-u_2, u_1).$$

Also, $\vec{n}$ is now the exterior normal to $\partial\Omega$. See Section 7 for a discussion of the case when Ω is not simply connected.

The eigenproblem for (3.1) is to find $(\lambda, \zeta, \vec{u})$ such that

$$\begin{aligned} &\lambda\zeta = -\nabla \cdot \vec{u}, && x \in \Omega, \\ &\lambda\vec{u} = -\nabla\zeta - f\mathbf{k} \times \vec{u}, && x \in \Omega, \\ &\int_\Omega \zeta \, dS = 0, \\ &\vec{u} \cdot \vec{n} = 0, && x \in \partial\Omega. \end{aligned} \tag{3.2}$$

Since $\vec{u} \cdot \vec{n} = 0$ on $\partial\Omega$ and Ω is simply connected, we can represent $\vec{u}$ as

$$\begin{aligned} &\vec{u} = -\nabla\varphi + \mathbf{k} \times \nabla\psi, && x \in \Omega, \\ &\frac{\partial\varphi}{\partial n} = 0, && x \in \partial\Omega, \\ &\int_\Omega \varphi \, dS = 0, \\ &\psi = 0, && x \in \partial\Omega. \end{aligned} \tag{3.3}$$

Here, φ represents an "irrotational" potential and ψ represents a "rotational" potential.

By applying the divergence and the operator $\nabla \cdot (\mathbf{k} \times \vec{u}) = -\partial u_2/\partial x_1 + \partial u_1/\partial x_2$ to (3.1b), we can derive equations for the time dependent behavior of the oceans in terms of the dependent variables (ζ, φ, ψ).

For complex-valued functions $y, z \in L^2(\Omega)$ we set $(y, z) = \int_\Omega y\bar{z} \, dS$ with the obvious modification for vector-valued functions. We also denote by $H^m(\Omega)$, for $m \geqslant 0$ an integer, the Sobolev space of functions with m weak derivatives in $L^2(\Omega)$ and norm

$$\|y\|^2_{H^m(\Omega)} = \sum_{|\alpha| \leqslant m} \|D^\alpha y\|^2_{L^2(\Omega)}.$$

We also set

$$H^1_0(\Omega) = \{y \in H^1(\Omega) | y = 0 \text{ on } \partial\Omega\}.$$

To construct equations for the time dependent behavior of (ζ, φ, ψ) in weak form, we first take the $L^2(\Omega)$ inner product of (3.1a) with a smooth function z. Since $\nabla \cdot (\mathbf{k} \times \nabla\psi) = 0$ and $\partial\varphi/\partial n = 0$ for $x \in \partial\Omega$, we obtain

$$\left(\frac{\partial\zeta}{\partial t}, z\right) = -(\nabla \cdot \vec{u}, z) = (\nabla \cdot \nabla\varphi, z) = -(\nabla\varphi, \nabla z). \tag{3.4a}$$

If we take the $L^2(\Omega)^2$ inner product of (3.1b) with $-\nabla z$ where z is a smooth function, we obtain

$$(\nabla\varphi_t - \mathbf{k} \times \nabla\psi_t, \nabla z) = (\nabla\varphi_t, \nabla z) = (\nabla\zeta + f\mathbf{k} \times \vec{u}, \nabla z). \tag{3.4b}$$

Similarly, if we take the $L^2(\Omega)^2$ inner product of (3.1b) with $\mathbf{k} \times \nabla z$ where z is a

smooth function such that $z = 0$ on $\partial\Omega$, we obtain

$$(3.4c)\qquad \begin{aligned}&(-\nabla\varphi_t + \mathbf{k} \times \nabla\psi_t, \mathbf{k} \times \nabla z)\\ &\quad = (\nabla\psi_t, \nabla z) = -(f\mathbf{k} \times \vec{u}, \mathbf{k} \times \nabla z) = -(f\vec{u}, \nabla z).\end{aligned}$$

Set

$$\begin{aligned}L^2_*(\Omega) &= \left\{z \in L^2(\Omega) \middle| \int_\Omega z\, dS = 0\right\},\\ H^1_*(\Omega) &= H^1(\Omega) \cap L^2_*(\Omega),\\ H^2_1(\Omega) &= \left\{z \in H^2(\Omega) \middle| \frac{\partial z}{\partial n} = 0 \text{ on } \partial\Omega\right\} \cap L^2_*(\Omega).\end{aligned}$$

We can formulate our eigenproblem as follows:

Find $(\lambda, \zeta, \varphi, \psi) = \mathbf{C} \times H^1_*(\Omega) \times H^2_1(\Omega) \times H^1_0(\Omega)$ such that

$$(3.5)\qquad \begin{aligned}&\lambda(\zeta, z) = -(\nabla\varphi, \nabla z), && z \in H^1(\Omega),\\ &\lambda(\nabla\varphi, \nabla z) = (\nabla\zeta - f\mathbf{k} \times \nabla\varphi - f\nabla\psi, \nabla z), && z \in H^1(\Omega),\\ &\lambda(\nabla\psi, \nabla z) = (f\nabla\varphi - f\mathbf{k} \times \nabla\psi, \nabla z), && z \in H^1_0(\Omega).\end{aligned}$$

We now define the linear operator, T, whose spectral properties we wish to approximate. We take H to be the complex Hilbert space defined by $H = L^2_*(\Omega) \times H^1_*(\Omega) \times H^1_0(\Omega)$ with inner product

$$\langle\varphi, \psi\rangle = \langle(\varphi_1, \varphi_2, \varphi_3), (\psi_1, \psi_2, \psi_3)\rangle = (\varphi_1, \psi_1) + (\nabla\varphi_2, \nabla\psi_2) + (\nabla\varphi_3, \nabla\psi_3)$$

and norm $\|\varphi\| = \langle\varphi, \varphi\rangle^{1/2}$.

We note that $H^1_*(\Omega)$ and $H^1_0(\Omega)$ are closed subspaces of $H^1(\Omega)$ and that $(\nabla z, \nabla z)^{1/2}$ defines a norm equivalent to the usual $H^1(\Omega)$ norm restricted to these subspaces. For a state $\varphi = (\varphi_1, \varphi_2, \varphi_3) \in H$, $\frac{1}{2}(\varphi_1, \varphi_1)$ is the potential energy of the state, $\frac{1}{2}(\nabla\varphi_2, \nabla\varphi_2)$ is the "irrotational" kinetic energy of the state, and $\frac{1}{2}(\nabla\varphi_3, \nabla\varphi_3)$ is the "rotational" kinetic energy of the state. Hence, $\|\varphi\|^2$ is twice the sum of the potential, rotational, and irrotational energies of the state of the system described by φ.

We define T to be the selfadjoint, unbounded operator with domain $D(T) = H^1_*(\Omega) \times H^2_1(\Omega) \times H^1_0(\Omega)$ such that for $(a, B, C) \in H^1_*(\Omega) \times H^2_1(\Omega) \times H^1_0(\Omega)$, $T(a, B, C) = (d, E, F) \in H$ satisfies the following equations:

$$(3.6a)\qquad (d, z) = i(\nabla B, \nabla z) = -i(\Delta B, z), \qquad z \in H^1(\Omega),$$

$$(3.6b)\qquad (\nabla E, \nabla z) = -i(\nabla a - f\mathbf{k} \times \nabla B - f\nabla C, \nabla z), \qquad z \in H^1(\Omega),$$

$$(3.6c)\qquad (\nabla F, \nabla z) = -i(f\nabla B - f\mathbf{k} \times \nabla C, \nabla z), \qquad z \in H^1_0(\Omega).$$

If λ is an eigenvalue of T with eigenvector (a, B, C), then $i\lambda$ is an eigenvalue of (3.2) (and (3.5)) with the corresponding eigenfunction $\zeta = a$ and $\vec{u} = -\nabla B + \mathbf{k} \times \nabla C$.

We may also define T in terms of the bilinear form

$$\begin{aligned}\mathcal{B}(\varphi, \psi) &= \mathcal{B}((\varphi_1, \varphi_2, \varphi_3), (\psi_1, \psi_2, \psi_3)) \\ &= i(\nabla\varphi_2, \nabla\psi_1) - i(\nabla\varphi_1 - f\mathbf{k} \times \nabla\varphi_2 - f\nabla\varphi_3, \nabla\psi_2) \\ &\quad - i(f\nabla\varphi_2 - f\mathbf{k} \times \nabla\varphi_3, \nabla\psi_3).\end{aligned}$$

Note that $\mathcal{B}(\cdot, \cdot)$ is not defined on $H \times H$. However, if $\varphi \in H^1_*(\Omega) \times H^2_1(\Omega) \times H^1_0(\Omega)$, then there exists a constant, $C(\varphi)$, depending on φ, such that

$$|\mathcal{B}(\varphi, \psi)| \leqslant C(\varphi)\|\psi\| \quad \text{for all} \quad \psi \in H.$$

In this case, we can define $T\varphi$ through the Riesz representation theorem so that

$$\langle T\varphi, \psi\rangle = \mathcal{B}(\varphi, \psi) \quad \text{for all} \quad \psi \in H.$$

Since there does not exist a constant $K > 0$ such that $C(\varphi) \leqslant K\|\varphi\|$ for $\varphi \in H^1_*(\Omega) \times H^2_1(\Omega) \times H^1_0(\Omega)$, $\mathcal{B}$ is not a continuous form on $H \times H$.

We shall now show that T is selfadjoint and unbounded. We define the operator T_1: $H^1_*(\Omega) \times H^2_1(\Omega) \times H^1_0(\Omega) \longrightarrow H$ by

$$T_1(\varphi_1, \varphi_2, \varphi_3) = (-i\Delta\varphi_2, i\varphi_1, 0). \tag{3.7}$$

It is easily checked that T_1 is an unbounded, selfadjoint operator.

We also define the bilinear form $\mathcal{B}_2(\cdot, \cdot)$ on $H \times H$ by

$$\mathcal{B}_2(\varphi, \psi) = i(f\mathbf{k} \times \nabla\varphi_2 + f\nabla\varphi_3, \nabla\psi_2) - i(f\nabla\varphi_2 - f\mathbf{k} \times \nabla\varphi_3, \nabla\psi_3).$$

Since there exists a constant $K > 0$ such that

$$\mathcal{B}_2(\varphi, \psi) \leqslant K\|\varphi\|\,\|\psi\|, \quad \forall\varphi, \psi \in H, \quad \text{and} \quad \mathcal{B}_2(\varphi, \psi) = \overline{\mathcal{B}_2(\psi, \varphi)},$$

there exists a bounded, selfadjoint operator T_2: $H \longrightarrow H$ defined by the relations $\langle T_2\varphi, w\rangle = \mathcal{B}_2(\varphi, w)$, $\forall w \in H$.

Now $T = T_1 + T_2$, so T is an unbounded, selfadjoint operator. Note that T_1 is the "zero-rotation" operator, i.e., if $\omega = 0$, then $T = T_1$.

We now consider how the spectral properties of T may be approximated by the spectral properties of operators, T^h, defined by means of a finite element approximation. We assume the existence of spaces of functions, M_h, parametrized by h, such that the following properties hold:

(1) $\dim M_h < \infty$.

(2) $M_h \subset H^1(\Omega)$.

(3) There exists a positive integer $r \geqslant 1$ and a constant K, independent of h, such that for $1 \leqslant s \leqslant r + 1$ and $z \in H^s(\Omega)$,

$$\inf_{\chi \in M_h} \{\|z - \chi\|_{L^2(\Omega)} + h\|z - \chi\|_{H^1(\Omega)}\} \leqslant Kh^s\|z\|_{H^s(\Omega)}. \tag{3.8a}$$

We define the spaces

$$M^0_h = M_h \cap H^1_0(\Omega), \quad M^*_h = M_h \cap H^1_*(\Omega).$$

We also assume that for $1 \leqslant s \leqslant r+1$, if $z \in H^s(\Omega) \cap H_0^1(\Omega)$, then

$$\inf_{\chi \in M_h^0} \{\|z-\chi\|_{L^2(\Omega)} + h\|z-\chi\|_{H^1(\Omega)}\} \leqslant Kh^s \|z\|_{H^s(\Omega)}, \tag{3.8b}$$

and if $z \in H^s(\Omega) \cap H_*^1(\Omega)$, then

$$\inf_{\chi \in M_h^*} \{\|z-\chi\|_{L^2(\Omega)} + h\|z-\chi\|_{H^1(\Omega)}\} \leqslant Kh^s \|z\|_{H^s(\Omega)}. \tag{3.8c}$$

(4) If $s \in C^\infty(\bar{\Omega})$, then there exists a constant $K = K(s)$ such that for all $\xi \in M_h$,

$$\inf_{\chi \in M_h} \|s\xi - \chi\|_{H^1(\Omega)} \leqslant Kh\|\xi\|_{H^1(\Omega)}. \tag{3.8d}$$

This is a version of the Nitsche-Schatz property [13].

(5) There exists a constant K, independent of h, such that the inverse property

$$\|\chi\|_{H^1(\Omega)} \leqslant Kh^{-1}\|\chi\|_{L^2(\Omega)}, \quad \forall \chi \in M_h, \tag{3.8e}$$

holds.

The above properties for M_h, with the exception of (3.8b), are satisfied by spaces of continuous, piecewise polynomials of degree r defined over a regular, quasi-uniform triangulation of Ω with the diameter of the largest triangle bounded by h [2]. However, if the spaces M_h are spaces of piecewise polynomials, then (3.8b) is impossible to verify for a smooth boundary, $\partial\Omega$. If the boundary, $\partial\Omega$, is a polygon, then (3.8b) can be achieved by spaces of piecewise polynomials. However, in the case of a polygonal domain, Ω, the components of the eigenfunctions of T will not even be in $H^2(\Omega)$ unless Ω is convex. Isoparametric elements [2] are often used to obtain zero values on the boundaries of approximating domains for finite element spaces with a high order of approximation. However, we feel that the analysis of these elements in the present context would obscure the main ideas of this paper.

We define the operator T^h: $M_h^* \times M_h^* \times M_h^0 \longrightarrow M_h^* \times M_h^* \times M_h^0$ by $T^h(a_h, B_h, C_h) = (d_h, E_h, F_h)$ if

$$(d_h, z) = i(\nabla B_h, \nabla z), \quad \forall z \in M_h^*, \tag{3.9a}$$

$$(\nabla E_h, \nabla z) = i(-\nabla a_h + f\mathbf{k} \times \nabla B_h + f\nabla C_h, \nabla z), \quad \forall z \in M_h^*. \tag{3.9b}$$

$$(\nabla F_h, \nabla z) = i(-f\nabla B_h + f\mathbf{k} \times \nabla C_h, \nabla z), \quad \forall z \in M_h^0. \tag{3.9c}$$

We note that by (3.8e) there exists a constant K such that

$$\|\nabla\chi\|_{L^2(\Omega)} \leqslant Kh^{-1}\|\chi\|_{L^2(\Omega)}, \quad \forall \chi \in M_h.$$

Hence, there exists $d_h \in M_h^*$ such that

$$(d_h, z) = i(\nabla B_h, \nabla z), \quad \forall z \in M_h^*.$$

We can also define T^h in terms of our bilinear form $\mathcal{B}$. We note that $\mathcal{B}(\cdot\,, \cdot)$

is defined on $(M_h^* \times M_h^* \times M_h^0) \times (M_h^* \times M_h^* \times M_h^0)$ and that there exists a constant $K > 0$ such that

$$|\mathcal{B}(\varphi, \psi)| \leqslant Kh^{-1}\|\varphi\| \, \|\psi\|$$

for $\varphi, \psi \in M_h^* \times M_h^* \times M_h^0$. Hence, for $\varphi \in M_h^* \times M_h^* \times M_h^0$, there exists $T^h\varphi \in M_h^* \times M_h^* \times M_h^0$ such that

$$\langle T^h\varphi, \psi\rangle = \mathcal{B}(\varphi, \psi), \quad \forall \psi \in M_h^* \times M_h^* \times M_h^0.$$

Note that $D(T^h) \not\subset D(T)$ and $D(T) \not\subset D(T^h)$. Also, if $\pi_{D(T^h)}$ is the orthogonal projection of H onto $M_h^* \times M_h^* \times M_h^0$, then $T^h \neq \pi_{D(T^h)}T$. However, it is easily checked that T^h is a selfadjoint linear operator when $M_h^* \times M_h^* \times M_h^0$ is given the inner product of H.

In [18], Veltkamp has analyzed the spectral properties of T by studying T as a perturbation of T_1. The spectral properties of T_1 are easily discovered. The eigenvalue $\lambda = 0$ is an eigenvalue of infinite multiplicity for T_1. All elements $(0, 0, \varphi_3)$, $\varphi_3 \in H_0^1(\Omega)$, are eigenfunctions. Thus, the modes with zero frequency have only rotational energy. These are the vorticity modes.

It can be seen by eliminating φ_2 from the eigenvalue equations associated with (3.7) that $(\varphi, -i\varphi/\lambda, 0)$ is an eigenvector with eigenvalue λ if

$$\begin{aligned} &\Delta\varphi = -\lambda^2\varphi, \qquad x \in \Omega, \\ &\int_\Omega \varphi \, dS = 0, \\ &\frac{\partial\varphi}{\partial n} = 0, \qquad x \in \partial\Omega. \end{aligned} \tag{3.10}$$

Thus, the eigenfunctions of T_1 corresponding to nonzero frequency have only potential and irrotational energy. These are the gravity modes.

So, the spectrum of T_1 consists of isolated eigenvalues with a limit point at $-\infty$ and ∞. All the nonzero eigenvalues have finite multiplicity and zero is an eigenvalue of infinite multiplicity. We assume that the addition of T_2 to T_1 perturbs the spectrum of T_1 so that the spectrum of T consists of isolated eigenvalues of finite multiplicity with limit points at 0, ∞, and $-\infty$; This is proven in a special case and conjectured in general in [18]. All modes now have both rotational and irrotational energy. We also note that even if T is one-to-one, T^{-1} will not be a bounded operator.

IV. The Approximation of Unbounded Operators. We now study how two properties concerning the convergence of T^h to T imply the convergence of the spectral properties of T^h to T. These properties for the ocean equations will be verified in Section V.

We let H be a separable Hilbert space and T be a selfadjoint operator in H with a pure point spectrum of eigenvalues of finite multiplicity. Let $\{\lambda_k\}_{k=-\infty}^{\infty}$ be the eigenvalues of T listed according to multiplicity with corresponding orthonormal eigenvectors $\{v_k\}$. We also suppose T^h to be a finite dimensional operator on a subspace $D(T^h) \subset H$ such that $T^h: D(T^h) \longrightarrow D(T^h)$, dim $D(T^h) < \infty$, and T^h is selfadjoint in $D(T^h)$ when

$D(T^h)$ is given the inner product of H. Let $\{\lambda_k^h\}_{k=1}^{N_h}$ be the eigenvalues of T^h with corresponding orthonormal eigenvectors, $\{v_k^h\}_{k=1}^{N_h}$.

The basic properties of spectral approximation of T by T^h will be shown to follow from the following two properties:

Property A. If v_k is one of the above eigenvectors of T and $\lambda_k \neq 0$, then $\exists\, v^h \in D(T^h)$ such that

$$\|v_k - v^h\| + \|Tv_k - T^h v^h\| \leqslant Kh^r, \tag{4.1}$$

where $K = K(\lambda_k)$.

Property B. There exists K_1, independent of h and λ_k^h, such that if v_k^h is one of the above eigenvectors of T^h with eigenvalue λ_k^h, then there exists $v \in D(T)$ such that

$$\|v - v_k^h\| \leqslant K_1|\lambda_k^h|h, \quad \|Tv - T^h v_k^h\| \leqslant K_1(|\lambda_k^h| + 1)h. \tag{4.2}$$

If $x \in H$ and N is a subspace of H, define

$$\operatorname{dist}(x, N) = \inf_{y \in N} \|x - y\|.$$

As in [10], given two closed subspaces M and N of H we define

$$\delta(M, N) = \sup_{x \in M, \|x\|=1} \operatorname{dist}(x, N), \qquad \hat{\delta}(M, N) = \max[\delta(M, N), \delta(N, M)].$$

We also wish to define two quantities measuring the separation of eigenvalues. We set

$$G_k^+ = \inf\{\lambda_j - \lambda_k| \text{ all } j \text{ such that } \lambda_j - \lambda_k > 0\},$$
$$G_k^- = \inf\{\lambda_k - \lambda_j| \text{ all } j \text{ such that } \lambda_k - \lambda_j > 0\},$$
$$G_k = \min\{G_k^+, G_k^-\}.$$

We assume that $G_k \neq 0$ for $\lambda_k \neq 0$.

THEOREM 1. *Suppose T and T^h are as above and satisfy Property* A *and Property* B. *Let $\lambda \neq 0$ be an eigenvalue of T of multiplicity n. We may assume without loss of generality that $\lambda = \lambda_1 = \cdots = \lambda_n$ (reorder the eigenvalues, if necessary). Let*

$$A_h = \{j|\lambda_j^h \in [\lambda - G_1^-/2, \lambda + G_1^+/2]\}.$$

Also set

$$V = \operatorname{span}\{v_1, \ldots, v_n\}, \quad V^h = \operatorname{span}\{v_j^h | j \in A_h\}.$$

Then there exists $K = K(\lambda) > 0$ such that

(i) $\max\{|\lambda - \lambda_j^h| | j \in A_h\} \leqslant Kh^r$,

(ii) $\hat{\delta}(V, V^h) \leqslant Kh^r$.

Remarks. It follows from (ii) of Theorem 1 that for h sufficiently small $\dim V = \dim V^h$. Hence, the cardinality of the set $A_h = n$ for h sufficiently small.

In what follows, K will denote a positive constant which may depend on λ, but which is always independent of h. We allow K to vary from equation to equation.

Note that it follows from Property A that to $\{v_1, \ldots, v_n\}$ we can associate $\{\tilde{v}_1^h, \ldots, \tilde{v}_n^h\} \subseteq D(T^h)$ such that

$$\|v_m - \tilde{v}_m^h\| + \|Tv_m - T^h\tilde{v}_m^h\| \leqslant Kh^r \quad \text{for } m = 1, \ldots, n. \tag{4.3}$$

Hence, we can define a linear map, L_h: $V \longrightarrow D(T^h)$, such that

$$L_h v_m = \tilde{v}_m^h \quad \text{for } m = 1, \ldots, n.$$

We can then deduce from (4.3) and the finite dimensionality of V that

$$\|v - L_h v\| + \|Tv - T^h L_h v\| \leqslant Kh^r \|v\| \quad \text{for } v \in V. \tag{4.4}$$

If M is a closed subspace of H, we shall denote the orthogonal projection onto M by π_M. We shall divide the proof of Theorem 1 into a series of lemmas.

LEMMA 1. *Let λ_k^h be an eigenvalue of T^h and let $\epsilon_k^h = \min_j |\lambda_j - \lambda_k^h|$. Then there exists K_1 (independent of h and λ_k^h) such that*

$$\epsilon_k^h \leqslant K_1 h \frac{((\lambda_k^h)^4 + (|\lambda_k^h| + 1)^2)^{1/2}}{(1 - K_1 |\lambda_k^h| h)}. \tag{4.5}$$

Furthermore, if $k \in A_h$, then

$$\operatorname{dist}(v_k^h, V) \leqslant Kh. \tag{4.6}$$

Remark. Note that it does not follow from (4.6) that $\delta(V^h, V) \leqslant Kh$ until we verify that $\dim V^h = \dim V$.

Proof. By Property B, there exists $v \in D(T)$ such that

$$\|v - v_k^h\| \leqslant K_1 |\lambda_k^h| h, \qquad \|Tv - T^h v_k^h\| \leqslant K_1 (1 + |\lambda_k^h|) h. \tag{4.7}$$

Let $v = \sum_{j=-\infty}^{\infty} \gamma_j \gamma_j$. Then

$$Tv = \sum_j \gamma_j \lambda_j v_j \tag{4.8}$$

and

$$\begin{aligned} Tv &= (Tv - T^h v_k^h) + T^h v_k^h \\ &= (Tv - T^h v_k^h) + \lambda_k^h (v_k^h - v) + \lambda_k^h v \\ &= (Tv - T^h v_k^h) + \lambda_k^h (v_k^h - v) + \sum_j \lambda_k^h \gamma_j v_j. \end{aligned} \tag{4.9}$$

After subtracting (4.8) from (4.9), we obtain

$$\sum_j (\lambda_j - \lambda_k^h) \gamma_j v_j = (Tv - T^h v_k^h) + \lambda_k^h (v_k^h - v). \tag{4.10}$$

Hence, it follows from the orthonormality of the family $\{v_k\}$ and (4.7) that

$$\sum_j (\lambda_j - \lambda_k^h)^2 \gamma_j^2 \leqslant K_1^2 h^2 [(|\lambda_k^h| + 1)^2 + (\lambda_k^h)^4]. \tag{4.11}$$

Thus,

$$(\epsilon_k^h)^2 \|v\|^2 = (\epsilon_k^h)^2 \sum_j \gamma_j^2 \leqslant K_1^2 h^2 [(|\lambda_k^h| + 1)^2 + (\lambda_k^h)^4]. \tag{4.12}$$

The result (4.5) follows from (4.12) and the estimate

$$\|v\| \geqslant \|v_k^h\| - K_1 |\lambda_k^h| h = 1 - K_1 |\lambda_k^h| h. \tag{4.13}$$

We now assume that $k \in A_h$, i.e., $\lambda_k^h \in [\lambda - G_1^-/2, \lambda + G_1^+/2]$. It follows from (4.5) that for h sufficiently small

$$|\lambda_j - \lambda_k^h| \geqslant G_1/2 \quad \text{for } j \neq 1, \ldots, n. \tag{4.14}$$

Hence, we can obtain from (4.11) that

$$(G_1/2)^2 \sum_{j \neq 1,\ldots,n} \gamma_j^2 \leqslant \sum_j (\lambda_j - \lambda_k^h)^2 \gamma_j^2 \leqslant Kh^2. \tag{4.15}$$

So, by (4.7) and (4.15)

$$\left\| v_k^h - \sum_{j=1,\ldots,n} \gamma_j v_j \right\| \leqslant \|v_k^h - v\| + \left\| \sum_{j \neq 1,\ldots,n} \gamma_j v_j \right\| \leqslant Kh. \quad \text{Q.E.D.} \tag{4.16}$$

LEMMA 2. $\delta(V, V^h) \leqslant Kh^r$.

Proof. Let $m \in \{1, \ldots, n\}$. By (4.4)

$$\|v_m - L_h v_m\| + \|Tv_m - T^h L_h v_m\| \leqslant Kh^r \|v_m\| = Kh^r. \tag{4.17}$$

Since $L_h v_m \in D(T^h)$, we can expand $L_h v_m = \sum_{j=1}^{N_h} \beta_j v_j^h$. Then

$$T^h L_h v_m = \sum_j \beta_j \lambda_j^h v_j^h \tag{4.18}$$

and

$$\begin{aligned} T^h L_h v_m &= (T^h L_h v_m - Tv_m) + Tv_m \\ &= (T^h L_h v_m - Tv_m) + \lambda(v_m - L_h v_m) + \lambda L_h v_m \\ &= (T^h L_h v_m - Tv_m) + \lambda(v_m - L_h v_m) + \sum_j \beta_j \lambda v_j^h. \end{aligned} \tag{4.19}$$

So, after subtracting (4.19) from (4.18) we obtain

$$\sum_j \beta_j (\lambda_j^h - \lambda) v_j^h = (T^h L_h v_m - Tv_m) + \lambda(L_h v_m - v_m). \tag{4.20}$$

From (4.17), we obtain from (4.20) and the orthonormality of $\{v_j^h\}$,

$$\sum \beta_j^2 (\lambda_j^h - \lambda)^2 \leqslant Kh^{2r}. \tag{4.21}$$

It follows from Lemma 1 that for h sufficiently small,

$$|\lambda_j^h - \lambda| \geqslant G_1/2 \quad \text{for } j \notin A_h.$$

Hence,

$$(G_1/2)^2 \sum_{j \notin A_h} \beta_j^2 \leqslant \sum_j \beta_j^2 (\lambda_j^h - \lambda)^2 \leqslant Kh^{2r}. \tag{4.22}$$

So, as in Lemma 1, we find from (4.17) and (4.22) that

$$\left\| v_m - \sum_{j \in A_h} \beta_j v_j^h \right\| \leqslant \|v_m - L_h v_m\| + \left\| \sum_{j \notin A_h} \beta_j v_j^h \right\| \leqslant Kh^r. \tag{4.23}$$

Note that $\Sigma_{j \in A_h} \beta_j v_j^h = \pi_{V^h} L_h v_m$. Thus, we have proved that for $m = 1, \ldots, n$, $\|v_m - \pi_{V^h} L_h v_m\| \leqslant Kh^r$. Hence, by the finite dimensionality of V,

$$\|v - \pi_{V^h} L_h v\| \leqslant Kh^r \|v\| \quad \text{for } v \in V. \quad \text{Q.E.D.} \tag{4.24}$$

LEMMA 3. $\hat{\delta}(V, V^h) \leqslant Kh^r$.

Proof. If we can show that dim V^h = dim V for h sufficiently small, then (4.6) implies that $\delta(V^h, V) \leqslant Kh$ for h sufficiently small. The above result and the bound $\delta(V, V^h) \leqslant Kh^r$ imply that $\hat{\delta}(V, V^h) < 1$ for h sufficiently small.

However, if M and N are two closed subspaces of a Hilbert space H and $\hat{\delta}(M, N) < 1$, then $\delta(M, N) = \delta(N, M) = \hat{\delta}(M, N)$ (see [10, p. 200]). Hence, for h sufficiently small $\hat{\delta}(V, V^h) = \delta(V, V^h) \leqslant Kh^r$.

Now it follows from Lemma 2 that $\delta(V, V^h) < 1$ for h sufficiently small. This implies that dim $V \leqslant$ dim V^h for h sufficiently small. It remains to show that dim $V^h \leqslant$ dim V.

We let n_h = dim V^h and suppose that $n_h > n$ = dim V. We may assume without loss of generality that $A_h = \{1, \ldots, n_h\}$ (renumber eigenvalues, if necessary).

By Lemma 1, for each $j \in A_h$ there exists $w_j \in V$ such that

$$\|v_j^h - w_j\| \leqslant Kh. \tag{4.25}$$

Since $n_h > n$, there exists $\{\alpha_s\}_{s=1}^n$ such that

$$w_{n+1} = \sum_{s=1}^{n} \alpha_s w_s. \tag{4.26}$$

So,

$$v_{n+1}^h = \sum_1^n \alpha_s v_s^h + (v_{n+1}^h - w_{n+1}) + \sum_1^n \alpha_s (w_s - v_s^h). \tag{4.27}$$

Let r be such that $|\alpha_r| = \max\{|\alpha_s| \,|s = 1, \ldots, n\}$. We may assume without loss of generality that $\alpha_r > 0$.

If we take the inner product of (4.27) with v_r^h, we obtain from the orthonormality of $\{v_j^h\}$

$$\begin{aligned} 0 &= \alpha_r + O(h) + \sum_1^n |\alpha_s| O(h) \\ &= \alpha_r + O(h) + n\alpha_r O(h). \end{aligned} \tag{4.28}$$

So,

$$\alpha_r \leqslant Kh. \tag{4.29}$$

But since

$$1 - Kh \leqslant \|w_s\| \leqslant 1 + Kh \quad \text{for } s = 1, \ldots, n,$$

we see from (4.26) that

$$1 - Kh \leqslant \|w_n\| \leqslant \sum_1^n |\alpha_s| \, \|w_s\| \leqslant n\alpha_r(1 + Kh). \tag{4.30}$$

So for h sufficiently small

$$\alpha_r \geqslant \frac{(1 - Kh)}{n(1 + Kh)} \geqslant \frac{1}{2n}. \tag{4.31}$$

However, (4.31) contradicts (4.29). Q.E.D.

We now state the following easily proved lemma [6].

LEMMA 4. *Let Y and Z be two subspaces of H such that* dim Y = dim Z. *Let P: $Y \longrightarrow Z$ be a linear operator such that*

$$\|Py - y\| \leqslant 2^{-1}\|y\|, \quad \forall\, y \in Y. \tag{4.32}$$

Then P is bijective and

$$\|P^{-1}z\| \leqslant 2\|z\|, \quad \forall\, z \in Z. \tag{4.33}$$

We wish to apply Lemma 4 to the map $\pi_{Vh}L_h$: $V \longrightarrow V^h$. It follows from (4.24) that for h sufficiently small,

$$\|\pi_{Vh}L_h v - v\| \leqslant \tfrac{1}{2}\|v\|, \quad \forall\, v \in V. \tag{4.34}$$

Hence, it follows by (4.33) that by choosing a new orthonormal basis for V, we may assume that there exist scalars $\{\alpha_m\}_{m=1}^n$ such that

$$\pi_{Vh}L_h v_m = \alpha_m v_m^h, \quad |\alpha_m| \geqslant \tfrac{1}{2} \quad \text{for } m = 1, \ldots, n. \tag{4.35}$$

LEMMA 5. $\max\{|\lambda - \lambda_j^h| \mid j \in A_h\} \leqslant Kh^r$.

Proof. We follow the proof of Lemma 2. Let $m \in \{1, \ldots, n\}$ and represent $L_h v_m = \Sigma_{j=1}^{N_h} \beta_j v_j^h$. Then, as in Lemma 2, we shall obtain $\Sigma\beta_j^2(\lambda_j^h - \lambda)^2 \leqslant Kh^{2r}$. However, it follows from (4.35) that $|\beta_m| = |\alpha_m| \geqslant \tfrac{1}{2}$. Thus,

$$(\tfrac{1}{2})^2(\lambda_m^h - \lambda)^2 \leqslant \Sigma\beta_j^2(\lambda_j^h - \lambda)^2 \leqslant Kh^{2r}.$$

So, $|\lambda_m^h - \lambda| \leqslant Kh^r$ for $m \in \{1, \ldots, n\} = A_h$. Q.E.D.

Lemmas 4 and 5 complete the proof of the theorem.

V. Analysis of the Finite Element Procedure. In this section, we shall show that Property A and Property B hold for the operators T and T^h defined in Section 3. We first prove a lemma concerning the regularity of the eigenfunctions of T.

LEMMA 6. *If $v_k = (a, B, C)$ is an eigenvector of T with eigenvalue $\lambda_k \neq 0$, then $a, B, C \in C^\infty(\bar{\Omega})$.*

Proof. It suffices to show that $a, B, C \in \bigcap_{m=0}^{\infty} H^m(\Omega)$. We show that $B \in H^{k+1}(\Omega)$, $a, C \in H^k(\Omega)$ implies that $B \in H^{k+2}(\Omega)$, $a, C \in H^{k+1}(\Omega)$. Since $B \in H^2(\Omega)$, $a, C \in H^1(\Omega)$, this will prove the lemma by induction.

By the definition of T, $C \in H_0^1(\Omega)$ satisfies

$$\lambda_k(\nabla C, \nabla w) = (if\nabla B + if\mathbf{k} \times \nabla C, \nabla w) \tag{5.1}$$

for $w \in H_0^1(\Omega)$. Since $\lambda_k \neq 0$ is real and $B \in H^{k+1}(\Omega)$, it follows by elliptic regularity [11] that $C \in H^{k+1}(\Omega)$.

Now a satisfies

$$i(\nabla a, \nabla w) = (-\lambda_k \nabla B + if\mathbf{k} \times \nabla B + if\nabla C, \nabla w) \tag{5.2}$$

for $w \in H^1(\Omega)$. So, since $B, C \in H^{k+1}(\Omega)$, it follows by elliptic regularity that $a \in H^{k+1}(\Omega)$.

Also, B satisfies

$$i(\nabla B, \nabla z) = \lambda(a, z), \qquad z \in H^1(\Omega). \tag{5.3}$$

Since $a \in H^{k+1}(\Omega)$, it follows that $B \in H^{k+2}(\Omega)$. Q.E.D.

THEOREM 2. *Property* A *holds for the operators T, T^h of Section* 3.

Proof of Theorem 2. Let $v_k = (a, B, C)$. Then we shall show that (4.1) is valid for some constant $K(\lambda_k)$ if we take $v^h = \pi_{D(T^h)} v_k$ where $\pi_{D(T^h)}$ is the orthogonal projection of H onto $D(T^h) = M_h^* \times M_h^* \times M_h^0$. Now $\pi_{D(T^h)} v_k = (a_h, B_h, C_h)$ satisfies

$$(a - a_h, z) = 0, \quad \forall z \in M_h^*, \tag{5.4a}$$

$$(\nabla(B - B_h), \nabla z) = 0, \quad \forall z \in M_h^*, \tag{5.4b}$$

$$(\nabla(C - C_h), \nabla z) = 0, \quad \forall z \in M_h^0. \tag{5.4c}$$

Now by Lemma 6, $a \in H^{r+1}(\Omega) \cap L_*^2(\Omega)$. Since a_h is the $L^2(\Omega)$ projection of a onto M_h^*, we have for $0 \leqslant s \leqslant r + 1$ that

$$\|a - a_h\|_{L^2(\Omega)} = \inf_{\chi \in M_h^*} \|a - \chi\|_{L^2(\Omega)} \leqslant Kh^s \|a\|_{H^s(\Omega)}. \tag{5.5a}$$

It then follows from (3.8e) that for $\chi \in M_h^*$

$$\begin{aligned} \|a - a_h\|_{H^1(\Omega)} &\leqslant \|a - \chi\|_{H^1(\Omega)} + \|\chi - a_h\|_{H^1(\Omega)} \\ &\leqslant \|a - \chi\|_{H^1(\Omega)} + Kh^{-1}\|\chi - a_h\|_{L^2(\Omega)} \\ &\leqslant \|a - \chi\|_{H^1(\Omega)} + Kh^{-1}\|a - \chi\|_{L^2(\Omega)} + Kh^{-1}\|a - a_h\|_{L^2(\Omega)}. \end{aligned}$$

Thus, we can obtain from (3.8c) that for $1 \leqslant s \leqslant r + 1$,

$$\|a - a_h\|_{H^1(\Omega)} \leqslant Kh^{s-1}\|a\|_{H^s(\Omega)}. \tag{5.5b}$$

We also have from Lemma 6 that $B \in H^{r+1}(\Omega) \cap H_1^2(\Omega)$ and $C \in H^{r+1}(\Omega) \cap H_0^1(\Omega)$. We obtain from (5.4b) and (5.4c) the result that for $1 \leqslant s \leqslant r+1$

$$\|\nabla(B - B_h)\|_{L^2(\Omega)} = \inf_{\chi \in M_h^*} \|\nabla(B - \chi)\| \leqslant Kh^{s-1}\|B\|_{H^s(\Omega)}, \tag{5.5c}$$

$$\|\nabla(C - C_h)\|_{L^2(\Omega)} = \inf_{\chi \in M_h^0} \|\nabla(C - \chi)\| \leqslant Kh^{s-1}\|C\|_{H^s(\Omega)}. \tag{5.5d}$$

So, by (5.5a), (5.5c) and (5.5d) we see that

$$\|v_k - v_h\| \leqslant Kh^r. \tag{5.6}$$

Let $Tv_k = (d, E, F)$ and $T^h v^h = (d_h, E_h, F_h)$.

By (5.4b), it follows that

$$(d - d_h, z) = (\lambda_k a - d_h, z) = 0, \quad \forall z \in M_h^*.$$

Hence, by (5.4a), $d_h = \lambda_k a_h$. So, we obtain from (5.5a) that

$$\|d - d_h\|_{L^2(\Omega)} = \lambda_k \|a - a_h\|_{L^2(\Omega)} \leqslant Kh^r. \tag{5.7}$$

Now by the definition of T and T^h we see that

$$(\nabla(E - E_h), \nabla z) = (-i\nabla(a - a_h) + if\mathbf{k} \times \nabla(B - B_h) + if\nabla(C - C_h), \nabla z)$$

for $z \in M_h$.

Hence, if $\tilde{E} \in M_h$,

$$\begin{aligned}(\nabla(\tilde{E} - E_h), \nabla z) = (\nabla(\tilde{E} - E) - i\nabla(a - a_h) + if\mathbf{k} \times \nabla(B - B_h) \\ + if\nabla(C - C_h), \nabla z) \quad \text{for } z \in M_h.\end{aligned} \tag{5.8}$$

If we take $z = \tilde{E} - E_h \in M_h$ in the above and use the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned}\|\nabla(E - E_h)\|_{L^2(\Omega)} \leqslant \|\nabla(\tilde{E} - E)\|_{L^2(\Omega)} + \|\nabla(a - a_h)\|_{L^2(\Omega)} \\ + K\|\nabla(B - B_h)\|_{L^2(\Omega)} + K\|\nabla(C - C_h)\|_{L^2(\Omega)}.\end{aligned} \tag{5.9}$$

Hence, by the triangle equality and (5.5), if we let $\tilde{E}$ approximate E, we can obtain by (3.8a)

$$\begin{aligned}\|\nabla(E - E_h)\|_{L^2(\Omega)} \leqslant 2\|\nabla(E - \tilde{E})\|_{L^2(\Omega)} + \|\nabla(a - a_h)\|_{L^2(\Omega)} \\ + K\|\nabla(B - B_h)\|_{L^2(\Omega)} + K\|\nabla(C - C_h)\|_{L^2(\Omega)} \leqslant Kh^r.\end{aligned} \tag{5.10}$$

Note that since $E = \lambda_k B$, $E \in H^{r+1}(\Omega)$.

We also have that

$$(\nabla(F - F_h), \nabla z) = (-if\nabla(B - B_h) + if\mathbf{k} \times \nabla(C - C_h), \nabla z)$$

for $z \in M_h^0$. A similar argument to the previous one shows that

$$\|\nabla(F - F_h)\|_{L^2(\Omega)} \leqslant Kh^r. \tag{5.11}$$

Hence, the result (5.7), (5.10), and (5.11) shows that

$$\|Tv_k - T^h v^h\| \leqslant Kh^r. \tag{5.12}$$

THEOREM 3. *Property* B *holds for the operators* T, T^h *of Section* 3.

Proof. In the following proof K_1 will denote a constant which is independent of λ_k^h and h, but which may vary from equation to equation.

Let $v_k^h = (a_h, B_h, C_h)$ and $T^h v_k^h = (d_h, E_h, F_h)$. We set $B \in H_1^2(\Omega)$ to be the solution to

$$-\Delta B = id_h. \tag{5.13}$$

Note that $\int_\Omega d_h \, dS = 0$ since $d_h \in M_h^*$. By elliptic regularity,

$$\|B\|_{H^2(\Omega)} \leqslant K_1 \|d_h\|_{L^2(\Omega)}.$$

We set $v = (a_h, B, C_h)$. Then $v \in D(T)$ and

$$\|v - v_k^h\| = \|\nabla(B - B_h)\|_{L^2(\Omega)}. \tag{5.14}$$

However, by (5.13)

$$(\nabla(B - B_h), \nabla z) = 0 \quad \text{for } z \in M_h. \tag{5.15}$$

So, since $\|v_k^h\| = 1$,

$$\begin{aligned} \|\nabla(B - B_h)\|_{L^2(\Omega)} &\leqslant K_1 h \|B\|_{H^2(\Omega)} \leqslant K_1 h \|d_h\|_{L^2(\Omega)} \\ &\leqslant K_1 h |\lambda_k^h| \, \|a_h\|_{L^2(\Omega)} \leqslant K_1 h |\lambda_k^h|. \end{aligned} \tag{5.16}$$

Thus, we have verified that $\|v_k^h - v\| \leqslant K_1 h |\lambda_k^h|$.

Set $Tv = (d, E, F)$. By (5.13), $d = d_h$.

Now

$$(\nabla(E - E_h), \nabla z) = i(f\mathbf{k} \times \nabla(B - B_h), \nabla z) \quad \text{for } z \in M_h. \tag{5.17}$$

By the argument used to obtain (5.9) from (5.8) we see that for $\hat{E} \in M_h$

$$\begin{aligned} \|\nabla(E - E_h)\|_{L^2(\Omega)} &\leqslant \|\nabla(\hat{E} - E)\|_{L^2(\Omega)} + K_1 \|\nabla(B - B_h)\|_{L^2(\Omega)} \\ &\leqslant \|\nabla(\hat{E} - E)\|_{L^2(\Omega)} + K_1 h \|B\|_{H^2(\Omega)} \\ &\leqslant \|\nabla(\hat{E} - E)\|_{L^2(\Omega)} + K_1 h \|d_h\|_{L^2(\Omega)}. \end{aligned} \tag{5.18}$$

So, we must estimate how well we can approximate E by members of M_h in the $H^1(\Omega)$ norm.

Now E satisfies

$$(\nabla E, \nabla z) = (-i\nabla a_h + if\mathbf{k} \times \nabla B + if\nabla C_h, \nabla z) \tag{5.19}$$

for $z \in H^1(\Omega)$. Note that in general E is only in $H^1(\Omega)$.

Since $C_h \in M_h^0$, by (3.8d) there exists $\chi \in M_h$ such that

$$\|\chi - fC_h\|_{H^1(\Omega)} \leqslant K_1 h \|C_h\|_{H^1(\Omega)} \leqslant K_1 h \|\nabla C_h\|_{L^2(\Omega)}. \tag{5.20}$$

Let $\tilde{E} = E + ia_h - i\chi$. Then

$$(\nabla \tilde{E}, \nabla z) = (if\mathbf{k} \times \nabla B + i\nabla(fC_h - \chi) - i(\nabla f)C_h, \nabla z) \tag{5.21}$$

for $z \in H^1(\Omega)$.

Let $\tilde{\tilde{E}} \in H^1_*(\Omega)$ satisfy

$$(\nabla \tilde{\tilde{E}}, \nabla z) = (if\mathbf{k} \times \nabla B - i(\nabla f)C_h, \nabla z) \quad \text{for } z \in H^1(\Omega). \tag{5.22}$$

Then

$$(\nabla(\tilde{E} - \tilde{\tilde{E}}), \nabla z) = (i\nabla(fC_h - \chi), \nabla z) \quad \text{for } z \in H^1(\Omega), \tag{5.23}$$

so

$$\|\nabla(\tilde{E} - \tilde{\tilde{E}})\|_{L^2(\Omega)} \leqslant \|\nabla(fC_h - \chi)\|_{L^2(\Omega)} \leqslant K_1 h \|\nabla C_h\|_{L^2(\Omega)}. \tag{5.24}$$

From (5.22), we see that $\tilde{\tilde{E}} \in H^2(\Omega)$ and

$$\begin{aligned} \|\tilde{\tilde{E}}\|_{H^2(\Omega)} &\leqslant K_1(\|B\|_{H^2(\Omega)} + \|C_h\|_{H^1(\Omega)}) \\ &\leqslant K_1(\|d_h\|_{L^2(\Omega)} + \|\nabla C_h\|_{L^2(\Omega)}). \end{aligned} \tag{5.25}$$

Hence, there exists $\psi \in M_h$ such that

$$\|\tilde{\tilde{E}} - \psi\|_{H^1(\Omega)} \leqslant K_1 h \|\tilde{\tilde{E}}\|_{H^2(\Omega)} \leqslant K_1 h(\|d_h\|_{L^2(\Omega)} + \|\nabla C_h\|_{L^2(\Omega)}). \tag{5.26}$$

So, let $\hat{E} = -ia_h + i\chi + \psi$. Then

$$E - \hat{E} = (\tilde{E} - ia_h + i\chi) - (-ia_h + i\chi + \psi) = (\tilde{E} - \tilde{\tilde{E}}) + (\tilde{\tilde{E}} - \psi).$$

Thus,

$$\begin{aligned} \|E - \hat{E}\|_{H^1(\Omega)} &\leqslant \|\tilde{E} - \tilde{\tilde{E}}\|_{H^1(\Omega)} + \|\tilde{\tilde{E}} - \psi\|_{H^1(\Omega)} \\ &\leqslant K_1 h \|\nabla C_h\|_{L^2(\Omega)} + K_1 h(\|d_h\|_{L^2(\Omega)} + \|\nabla C_h\|_{L^2(\Omega)}). \end{aligned} \tag{5.27}$$

We shall now estimate $\|\nabla(F - F_h)\|_{L^2(\Omega)}$. Since $C = C_h$, $F - F_h$ satisfies

$$(\nabla(F - F_h), \nabla z) = (-if\nabla(B - B_h), \nabla z) \quad \text{for all } z \in M_h^0. \tag{5.28}$$

So, for $\tilde{F} \in M_h^0$,

$$\begin{aligned} \|\nabla(F - F_h)\|_{L^2(\Omega)} &\leqslant 2\|\nabla(F - \tilde{F})\|_{L^2(\Omega)} + K_1 \|\nabla(B - B_h)\|_{L^2(\Omega)} \\ &\leqslant 2\|\nabla(F - \tilde{F})\|_{L^2(\Omega)} + K_1 h \|d_h\|_{L^2(\Omega)}. \end{aligned} \tag{5.29}$$

Therefore, we must approximate F by an element of M_h^0 in the $H^1(\Omega)$ norm. Now F satisfies

$$(\nabla F, \nabla z) = (-if\nabla B + if\mathbf{k} \times \nabla C_h, \nabla z) \quad \text{for } z \in H_0^1(\Omega). \tag{5.30}$$

However, if $z \in C_0^\infty(\Omega)$,

$$(5.31)\qquad \begin{aligned}(f\mathbf{k} \times \nabla C_h, \nabla z) &= (C_h, \nabla \cdot (f\mathbf{k} \times \nabla z)) \\ &= (C_h \nabla f, \mathbf{k} \times \nabla z) = -(\mathbf{k} \times C_h \nabla f, \nabla z).\end{aligned}$$

So, (5.31) is satisfied for all $z \in H_0^1(\Omega)$. From (5.30) we obtain

$$(5.32)\qquad (\nabla F, \nabla z) = (-if\nabla B - iC_h\mathbf{k} \times \nabla f, \nabla z) \quad \text{for } z \in H_0^1(\Omega).$$

Therefore, $F \in H^2(\Omega)$ and

$$(5.33)\qquad \|F\|_{H^2(\Omega)} \leqslant K_1(\|B\|_{H^2(\Omega)} + \|C_h\|_{H^1(\Omega)}) \leqslant K_1(\|d_h\|_{L^2(\Omega)} + \|\nabla C_h\|_{L^2(\Omega)}).$$

Hence, there exists $\tilde{F} \in M_h^0$ such that

$$(5.34)\qquad \|\nabla(F - \tilde{F})\|_{L^2(\Omega)} \leqslant K_1 h(\|d_h\|_{L^2(\Omega)} + \|\nabla C_h\|_{L^2(\Omega)}).$$

It follows from (5.29) and (5.34) that

$$(5.35)\qquad \|\nabla(F - F_h)\|_{L^2(\Omega)} \leqslant K_1 h(\|d_h\|_{L^2(\Omega)} + \|\nabla C_h\|_{L^2(\Omega)}).$$

Therefore, since $d = d_h = \lambda_k^h a_h$, by (5.27) and (5.35), we obtain

$$(5.36)\qquad \|Tv - T^h v_j^h\| \leqslant K_1(|\lambda_k^h| + 1)h. \quad \text{Q.E.D.}$$

VI. Improved Eigenvalue Estimates. In this section we get improved eigenvalue estimates for the approximation of T by T^h. These estimates cannot be deduced from Property A and Property B, but require addition analysis of the operators T and T^h defined in Section III. Our technique is to represent the eigenvalues in terms of Rayleigh quotients.

THEOREM 4. *Let $\lambda \neq 0$ be an eigenvalue of T of multiplicity n. We may assume as in Theorem 1 that $\lambda = \lambda_1 = \cdots = \lambda_n$. Also, set*

$$A_h = \{j | \lambda_j^h \in [\lambda - G_1^-/2, \lambda + G_1^+/2]\}.$$

Then there exists $K > 0$, independent of h, such that

$$(6.1)\qquad |\lambda - \lambda_j^h| \leqslant Kh^{2r} \quad \textit{for all } j \in A_h.$$

Proof. As in the proof of Theorem 1, we may assume for h sufficiently small that

$$(6.2)\qquad A_h = \{1, \ldots, n\}$$

and

$$(6.3)\qquad \pi_{V^h} L_h v_m = \alpha_m v_m^h, \quad |\alpha_m| \geqslant \tfrac{1}{2} \quad \text{for } m = 1, \ldots, n.$$

Recall that in Section V we were able to verify Property A with $L_h = \pi_{D(T^h)}$. So, we then have that $\pi_{V^h} L_h = \pi_{V^h} \pi_{D(T^h)} = \pi_{V^h}$.

We shall estimate

$$\lambda - \lambda_m^h = \frac{\langle Tv_m, v_m\rangle}{\langle v_m, v_m\rangle} - \frac{\langle T^h \pi_{Vh} v_m, \pi_{Vh} v_m\rangle}{\langle \pi_{Vh} v_m, \pi_{Vh} v_m\rangle} \tag{6.4}$$

for $m \in A_h$ (recall from (6.3) that $\pi_{Vh} v_m = \alpha_m v_m^h$) by estimating

$$\frac{\langle Tv_m, v_m\rangle}{\langle v_m, v_m\rangle} - \frac{\langle T^h L_h v_m, L_h v_m\rangle}{\langle L_h v_m, L_h v_m\rangle} \tag{6.5}$$

and

$$\frac{\langle T^h L_h v_m, L_h v_m\rangle}{\langle L_h v_m, L_h v_m\rangle} - \frac{\langle T^h \pi_{Vh} v_m, \pi_{Vh} v_m\rangle}{\langle \pi_{Vh} v_m, \pi_{Vh} v_m\rangle}. \tag{6.6}$$

We first estimate the term (6.6). Represent as in Lemma 2,

$$L_h v_m = \sum_{j=1}^{N_h} \beta_j v_j^h. \tag{6.7}$$

Then

$$\pi_{Vh} L_h v_m = \sum_{j=1}^{n} \beta_j v_j^h. \tag{6.8}$$

It follows from (4.22) and the orthonormality of the family $\{v_j^h\}$ that

$$|\langle L_h v_m, L_h v_m\rangle - \langle \pi_{Vh} v_m, \pi_{Vh} v_m\rangle| = \sum_{j \notin A_h = \{1,\dots,n\}} \beta_j^2 \leqslant Kh^{2r}. \tag{6.9}$$

Similarly, it follows from (4.21), (4.22) and the Cauchy-Schwarz inequality that

$$\begin{aligned} &|\langle T^h L_h v_m, L_h v_m\rangle - \langle T^h \pi_{Vh} v_m, \pi_{Vh} v_m\rangle| \\ &= \sum_{j \notin A_h} \lambda_j^h \beta_j^2 \leqslant \left(\sum_{j \notin A_h} \beta_j^2\right)^{1/2} \left(\sum_{j \notin A_h} (\lambda_j^h)^2 \beta_j^2\right)^{1/2} \leqslant Kh^{2r}. \end{aligned} \tag{6.10}$$

Since $\|L_h v_m\| \geqslant \|v_m\| - \|v_m - L_h v_m\| \geqslant 1 - Kh^r$ by (4.4) and $\|\pi_{Vh} L_h v_m\| = \|\pi_{Vh} v_m\| \geqslant \frac{1}{2}$ by (6.3), it follows from (6.9) and (6.10) that

$$\left| \frac{\langle T^h L_h v_m, L_h v_m\rangle}{\langle L_h v_m, L_h v_m\rangle} - \frac{\langle T^h \pi_{Vh} v_m, \pi_{Vh} v_m\rangle}{\langle \pi_{Vh} v_m, \pi_{Vh} v_m\rangle} \right| \leqslant Kh^{2r}. \tag{6.11}$$

We now turn to the term (6.5) and again recall that $L_h = \pi_{D(T^h)}$. Let $v_m = (a, B, C)$, $Tv_m = \lambda v_m = (d, E, F)$, $L_h v_m = (a_h, B_h, C_h)$ and $T^h L_h v_m = (d_h, E_h, F_h)$. Then

$$\begin{aligned} \langle v_m, v_m\rangle - \langle L_h v_m, L_h v_m\rangle &= (a, a) + (\nabla B, \nabla B) + (\nabla C, \nabla C) \\ &\quad - (a_h, a_h) - (\nabla B_h, \nabla B_h) - (\nabla C_h, \nabla C_h). \end{aligned} \tag{6.12}$$

In order to show that the left-hand side of (6.12) is $O(h^{2r})$, we need the negative norm estimates for the terms estimated in (5.5).

For $\varphi \in L^2(\Omega)$ and $s \geqslant 0$, we define the norms

$$\|\varphi\|_{H_*^{-s}} = \sup_{\substack{\psi \in H^s(\Omega) \cap L^2_*(\Omega) \\ \|\psi\|_{H^s(\Omega)}=1}} (\varphi, \psi)$$

and

$$\|\varphi\|_{H^{-s}} = \sup_{\|\psi\|_{H^s(\Omega)}=1} (\varphi, \psi).$$

Then it follows by well-known negative norm estimates for the L^2 and H^1 projections [1, pp. 538–539] that

$$\|a - a_h\|_{H_*^{-s}} \leqslant Kh^{s+t}\|a\|_{H^t(\Omega)} \quad \text{for } 0 \leqslant s \leqslant r+1, 0 \leqslant t \leqslant r+1,$$

$$\|B - B_h\|_{H_*^{-s}} \leqslant Kh^{s+t}\|B\|_{H^t(\Omega)} \quad \text{for } 0 \leqslant s \leqslant r-1, 1 \leqslant t \leqslant r+1,$$

$$\|C - C_h\|_{H^{-s}} \leqslant Kh^{s+t}\|C\|_{H^t(\Omega)} \quad \text{for } 0 \leqslant s \leqslant r-1, 1 \leqslant t \leqslant r+1.$$

Note that if $\varphi \in H^s(\Omega) \cap L^2_*(\Omega)$, then by (5.4a), (5.5a), and (3.8c) for $0 \leqslant s \leqslant r+1$,

$$|(a - a_h, \varphi)| = \inf_{\chi \in M_h^*} |(a - a_h, \varphi - \chi)| \leqslant \|a - a_h\|_{L^2(\Omega)} Kh^s \|\varphi\|_{H^s(\Omega)}.$$

Hence, for $0 \leqslant s \leqslant r+1, 0 \leqslant t \leqslant r+1$,

$$\|a - a_h\|_{H_*^{-s}} \leqslant Kh^s\|a - a_h\|_{L^2(\Omega)} \leqslant Kh^{s+t}\|a\|_{H^t(\Omega)}.$$

By our definition of a_h as the $L^2(\Omega)$ projection of a on M_h^*,

$$(a, a) - (a_h, a_h) = (a, a - a_h).$$

Since $a \in H^r(\Omega)$ and $\|a - a_h\|_{H^{-r}(\Omega)} \leqslant Kh^{2r}\|a\|_{H^r(\Omega)}$, it follows that

$$|(a, a) - (a_h, a_h)| \leqslant Kh^{2r}\|a\|^2_{H^r(\Omega)} \leqslant Kh^{2r}. \tag{6.13}$$

Similarly, using negative norm estimates on the $H^1(\Omega)$ projections, we obtain

$$\begin{aligned} &|(\nabla B, \nabla B) - (\nabla B_h, \nabla B_h)| \\ &\quad = |(\nabla B, \nabla(B - B_h))| \leqslant \|B\|_{H^{r+1}(\Omega)}\|B - B_h\|_{H_*^{-r+1}(\Omega)} \\ &\quad \leqslant Kh^{2r}\|B\|^2_{H^{r+1}(\Omega)} \leqslant Kh^{2r} \end{aligned} \tag{6.14}$$

and

$$\begin{aligned} &|(\nabla C, \nabla C) - (\nabla C_h, \nabla C_h)| \\ &\quad = |(\nabla C, \nabla(C - C_h))| \leqslant Kh^{2r}\|C\|^2_{H^{r+1}(\Omega)} \leqslant Kh^{2r}. \end{aligned} \tag{6.15}$$

Hence, it follows from (6.13), (6.14), and (6.15) that

$$|\langle v_m, v_m\rangle - \langle L_h v_m, L_h v_m\rangle| \leqslant Kh^{2r}. \tag{6.16}$$

We now consider

$$\begin{aligned}(6.17)\qquad &\langle Tv_m, v_m\rangle - \langle T^h L_h v_m, L_h v_m\rangle \\ &= (d, a) - (d_h, a_h) + (\nabla E, \nabla B) - (\nabla E_h, \nabla B_h) + (\nabla F, \nabla C) - (\nabla F_h, \nabla C_h).\end{aligned}$$

It follows from (3.6a) and (3.9a) that

$$(d, a) - (d_h, a_h) = i(\nabla B, \nabla a) - i(\nabla B_h, \nabla a_h).$$

Now by (5.4b) and the negative norm estimate for $a - a_h$,

$$\begin{aligned}(6.18)\qquad (\nabla B, \nabla a) - (\nabla B_h, \nabla a_h) &= (\nabla B, \nabla(a - a_h)) \\ &= -(\Delta B, a - a_h) \leqslant Kh^{2r}\|B\|_{H^{r+1}(\Omega)}\|a\|_{H^{r+1}(\Omega)} \leqslant Kh^{2r}.\end{aligned}$$

So,

$$(6.19)\qquad |(d, a) - (d_h, a_h)| \leqslant Kh^{2r}.$$

Continuing to the next term in (6.17),

$$\begin{aligned}(6.20)\qquad (\nabla E, \nabla B) - (\nabla E_h, \nabla B_h) &= (\nabla(E - E_h), \nabla B) \\ &= (\nabla(E - E_h), \nabla(B - B_h)) + (\nabla(E - E_h), \nabla B_h).\end{aligned}$$

It follows from (5.5c) and (5.10) that

$$|(\nabla(E - E_h), \nabla(B - B_h))| \leqslant Kh^{2r}.$$

Now from (3.6b) and (3.9b)

$$\begin{aligned}(6.21)\qquad (\nabla(E - E_h), \nabla B_h) &= -i(\nabla(a - a_h), \nabla B_h) \\ &\quad + i(f\mathbf{k} \times \nabla(B - B_h), \nabla B_h) + i(f\nabla(C - C_h), \nabla B_h).\end{aligned}$$

However, by (5.5) and (6.18),

$$(\nabla(a - a_h), \nabla B_h) = (\nabla(a - a_h), \nabla(B_h - B)) + (\nabla(a - a_h), \nabla B),$$

$$|(\nabla(a - a_h), \nabla(B_h - B))| \leqslant Kh^{2r},$$

$$|(\nabla(a - a_h), \nabla B)| \leqslant Kh^{2r}.$$

Also

$$\begin{aligned}(6.22)\qquad &(f\mathbf{k} \times \nabla(B - B_h), \nabla B_h) \\ &\quad = (f\mathbf{k} \times \nabla(B - B_h), \nabla(B_h - B)) + (f\mathbf{k} \times \nabla(B - B_h), \nabla B), \\ &\quad |(f\mathbf{k} \times \nabla(B - B_h), \nabla(B_h - B))| \leqslant Kh^{2r}.\end{aligned}$$

Observe that

$$(6.23)\qquad (f\mathbf{k} \times \nabla(B - B_h), \nabla B) = -(\nabla(B - B_h), f\mathbf{k} \times \nabla B).$$

Now define $\varphi \in H^1_*(\Omega)$ to be the solution to

$$\begin{aligned}(6.24)\qquad \Delta\varphi &= -\nabla \cdot (f\mathbf{k} \times \nabla B), \qquad x \in \Omega, \\ \frac{\partial\varphi}{\partial n} &= (-f\mathbf{k} \times \nabla B) \cdot \vec{n}, \qquad x \in \partial\Omega,\end{aligned}$$

where $\vec{n}$ is the exterior normal to the boundary of Ω. Then by elliptic regularity, $\varphi \in H^{r+1}(\Omega)$ and

$$\|\varphi\|_{H^{r+1}(\Omega)} \leqslant K\|B\|_{H^{r+1}(\Omega)}. \tag{6.25}$$

Furthermore, for $z \in M_h^*$,

$$-(\nabla(B - B_h), f\mathbf{k} \times \nabla B) = (\nabla(B - B_h), \nabla\varphi) = (\nabla(B - B_h), \nabla(\varphi - z)).$$

Thus, by (3.8a)

$$\begin{aligned} |(f\mathbf{k} \times \nabla(B - B_h), \nabla B)| &\leqslant Kh^{2r}\|B\|_{H^{r+1}(\Omega)}\|\varphi\|_{H^{r+1}(\Omega)} \\ &\leqslant Kh^{2r}\|B\|^2_{H^{r+1}(\Omega)} \leqslant Kh^{2r}. \end{aligned} \tag{6.26}$$

We now bound the last term in (6.21). We have by (5.5) and the negative norm estimates for $C - C_h$

$$\begin{gathered} (f\nabla(C - C_h), \nabla B_h) = (f\nabla(C - C_h), \nabla(B_h - B)) + (f\nabla(C - C_h), \nabla B), \\ |(f\nabla(C - C_h), \nabla(B_h - B))| \leqslant Kh^{2r}, \\ |(f\nabla(C - C_h), \nabla B)| = |(C - C_h, \nabla \cdot f\nabla B)| \leqslant Kh^{2r}. \end{gathered} \tag{6.27}$$

Hence, from (6.20)–(6.27) we obtain

$$|(\nabla E, \nabla B) - (\nabla E_h, \nabla B_h)| \leqslant Kh^{2r}. \tag{6.28}$$

Now by (5.4c)

$$\begin{aligned} (\nabla F, \nabla C) - (\nabla F_h, \nabla C_h) &= (\nabla(F - F_h), \nabla C) \\ &= (\nabla(F - F_h), \nabla(C - C_h)) + (\nabla(F - F_h), \nabla C_h). \end{aligned} \tag{6.29}$$

It follows from (5.11) and (5.5d) that

$$|(\nabla(F - F_h), \nabla(C - C_h))| \leqslant Kh^{2r}. \tag{6.30}$$

By (3.6c) and (3.9c),

$$(\nabla(F - F_h), \nabla C_h) = -i(f\nabla(B - B_h), \nabla C_h) + i(f\mathbf{k} \times \nabla(C - C_h), \nabla C_h). \tag{6.31}$$

Now by (5.5)

$$\begin{gathered} (f\nabla(B - B_h), \nabla C_h) = (f\nabla(B - B_h), \nabla(C_h - C)) + (f\nabla(B - B_h), \nabla C), \\ |(f\nabla(B - B_h), \nabla(C_h - C))| \leqslant Kh^{2r}. \end{gathered} \tag{6.32}$$

Also, by an adjoint argument similar to that of (6.24)–(6.26) it follows that $|(f\nabla(B - B_h), \nabla C)| \leqslant Kh^{2r}$.

Turning to the final term in (6.31), we have

$$\begin{aligned} &(f\mathbf{k} \times \nabla(C - C_h), \nabla C_h) \\ &\quad = (f\mathbf{k} \times \nabla(C - C_h), \nabla(C_h - C)) + (f\mathbf{k} \times \nabla(C - C_h), \nabla C) \end{aligned} \tag{6.33}$$

and by (5.5)

$$|(f\mathbf{k} \times \nabla(C - C_h), \nabla(C_h - C))| \leqslant Kh^{2r}.$$

Also, since $C \in H_0^1(\Omega)$ and by the negative norm estimates for $C - C_h$

$$\begin{aligned} |(f\mathbf{k} \times \nabla(C - C_h), \nabla C)| &= |(\nabla(C - C_h), f\mathbf{k} \times \nabla C)| \\ &= |(C - C_h, \nabla \cdot (f\mathbf{k} \times \nabla C))| \leqslant Kh^{2r}. \end{aligned} \tag{6.34}$$

Hence, it follows from (6.29)–(6.34) that

$$|(\nabla F, \nabla C) - (\nabla F_h, \nabla C_h)| \leqslant Kh^{2r}. \tag{6.35}$$

Thus, from (6.19), (6.28) and (6.35) we obtain

$$|\langle Tv_m, v_m\rangle - \langle T^h L_h v_m, L_h v_m\rangle| \leqslant Kh^{2r}. \tag{6.36}$$

Hence, it follows from (6.16) and (6.36) that

$$\frac{\langle Tv_m, v_m\rangle}{\langle v_m, v_m\rangle} - \frac{\langle T^h L_h v_m, L_h v_m\rangle}{\langle L_h v_m, L_h v_m\rangle} \leqslant Kh^{2r}. \quad \text{Q.E.D.}$$

VII. Remarks. It is easily seen that the estimates given here are uniform for parts of the spectrum of T in finite intervals not containing 0. However, the bounds degenerate for eigenvalues whose absolute value approaches 0 and ∞. This can be understood by considering the operator T_1. The eigenspace of zero frequency, $\{(0, 0, \psi) | \psi \in H_0^1(\Omega)\}$, does not contain only smooth functions (in $C^\infty(\bar{\Omega})$) and by (3.10) it is seen that the eigenspaces for high frequencies contain spatially highly oscillatory functions. Since T is a bounded perturbation of T_1, we can thus expect "rough" eigenfunctions for low and high frequencies which can only be resolved for small h. Also, note that since 0 is a limit point of eigenvalues of T, the gap between eigenvalues is small near 0.

It is clear that the first Galerkin method described in Section II will not satisfy Property A and Property B due to the "spectrum bending". Thus, one could not use that method to approximate the spectrum in some finite interval.

That there is no spectrum bending for the proposed method can be seen as follows. We define the operator T_1^h: $D(T_h) \longrightarrow D(T_h)$ by $T_1^h(a_h, B_h, C_h) = (d_h, -ia_h, 0)$ where $(d_h, z) = i(\nabla B_h, \nabla z)$ for $z \in M_h^*$. Then, if λ_h is a nonzero eigenvalue of T_1^h, we see by eliminating B_h from the eigenvalue equations that

$$\lambda_h^2(a_h, z) = (\nabla a_h, \nabla z) \quad \text{for } z \in M_h^*. \tag{7.1}$$

Thus, (7.1) is a Rayleigh-Ritz approximation to (3.10). Recall that the Rayleigh-Ritz eigenvalues are larger in magnitude than their corresponding differential eigenvalues. So, we see in the absence of the Coriolis terms why our proposed method does not exhibit spectrum bending and why it yields good spectral approximation results.

Set $T_2^h = T^h - T_1^h$. Define

$$\|T_2^h\|_h = \sup_{\substack{v \neq 0 \\ v \in D(T^h)}} \frac{\|T_2^h v\|}{\|v\|}.$$

Then the above properties are shared by the approximation of T by T^h since $T = T_1 + T_2$, where T_2 is a bounded linear operator in H, and $T^h = T_1^h + T_2^h$ where $\|T_2^h\|_h$ is bounded independently of h.

We have seen that the use of the Stokes-Helmholtz potentials as dependent variables in place of the horizontal transport vector leads to an improved procedure for calculating normal modes. The Stokes-Helmholtz potentials also have the advantage of being "coordinate free". In addition, since by (3.3)

$$\int_\Omega |\vec{u}|^2 \, dS = \int_\Omega |\nabla\varphi|^2 \, dS + \int_\Omega |\nabla\psi|^2 \, dS,$$

by determining whether the greater part of the kinetic energy is "rotational" or "irrotational" one can classify normal modes as vorticity modes or gravity modes [16].

If Ω is not simply connected and $\partial\Omega$ has a finite number of connected components $\{\partial\Omega\}_{i=1}^s$, then we have to allow ψ in (3.3) to have arbitrary constant values $\{c_i\}_{i=1}^s$ on the components, i.e.,

$$\psi(x) = c_i \quad \text{for } x \in \partial\Omega_i.$$

We then have to modify the definition of T^h by replacing M_h^0 with the space

$$M_h^c = \{W \in M_h \mid \text{there exists constants } \{c_i\}_{i=1}^s \text{ such that } W(x) = c_i \text{ for } x \in \partial\Omega_i\}.$$

If we assume optimal order approximation properties for M_h^c then all of the previous results remain valid in this case.

Acknowledgements. I am very grateful to Professor George Platzman for introducing me to the topics discussed in this paper, and for his interest in this work. I would also like to thank Professors Todd Dupont, Dianne O'Leary, John Osborn and Jeff Rauch for valuable discussions on the subject of this paper.

Department of Mathematics
University of Michigan
Ann Arbor, Michigan 48109

1. J. H. BRAMBLE & J. E. OSBORN, "Rate of convergence estimates for nonselfadjoint eigenvalue approximations," *Math. Comp.*, v. 27, 1973, pp. 525–549.

2. P. G. CIARLET, *Numerical Analysis of the Finite Element Method*, Univeristy of Montreal Press, Montreal, 1976.

3. A. K. CLINE, G. H. GOLUB & G. W. PLATZMAN, "Calculation of normal modes of oceans using a Lanczos method," *Sparse Matrix Computations*, J. Bunch and D. Rose (eds.), Academic Press, New York, 1976, pp. 409–429.

4. J. DESCLOUX, M. LUSKIN & J. RAPPAZ, "Approximation of the spectrum of closed operators. The determination of normal modes of a rotating basin". (To appear.)

5. J. DESCLOUX, N. NASSIF & J. RAPPAZ, "On spectral approximation. Part 1. The problem of convergence," *R.A.I.R.O. Numerical Analysis*, v. 12, no. 2, 1978, pp. 97–112.

6. J. DESCLOUX, N. NASSIF & J. RAPPAZ, "On spectral approximation. Part 2. Error estimates for the Galerkin method," *R.A.I.R.O. Numerical Analysis*, v. 12, no. 2, 1978, pp. 113–119.

7. T. DUPONT, Personal communication.

8. T. DUPONT, "Galerkin methods for modeling gas pipelines," *Constructive and Computational Methods for Differential and Integral Equations*, Lecture Notes in Math., vol. 430, Springer-Verlag, Berlin and New York, 1974.

9. T. DUPONT & H. RACHFORD, JR., "A Galerkin method for liquid pipelines," *Computational Methods in Applied Sciences and Engineering*, Lecture Notes in Econ. and Math. Systems, vol. 134, Springer-Verlag, Berlin and New York, 1976.

10. T. KATO, *Perturbation Theory for Linear Operators,* Die Grundelehren der Math. Wissenschaften, Band 132, Springer-Verlag, New York, 1966.

11. J. L. LIONS & E. MAGENES, *Problèmes aux Limites Non Homogènes et Applications,* vol. 1, Dunod, Paris, 1968.

12. M. LUSKIN, "A finite element method for first order hyperbolic systems." (To appear.)

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

14. J. E. OSBORN, "Spectral approximation for compact operators," *Math. Comp.*, v. 29, 1975, pp. 712–725.

15. G. W. PLATZMAN, "Normal modes of the Atlantic and Indian Oceans," *J. Physical Oceanography,* v. 5, 1975, pp. 201–221.

16. G. W. PLATZMAN, "Normal modes of the world Ocean. Part 1. Design of a finite-element baratropic model," *J. Physical Oceanography*, v. 8, 1978, pp. 323–343.

17. R. SCOTT, Personal communication.

18. G. W. VELTKAMP, *Spectral Properties of Hilbert Space Operators Associated with Tidal Motions,* Drukkerij Wed. G. Van Soest, Amsterdam, 1960.