ORDERS OF CONVERGENCE OF THE RAYLEIGH-RITZ AND WEINSTEIN-BAZLEY METHODS*

BY GEORGE FIX

HARVARD UNIVERSITY

Communicated by Garrett Birkhoff, September 23, 1968

In this note, a "best possible" order-of-convergence theorem is proved for *eigenfunctions* computed by the approximate method of Rayleigh-Ritz (ref. 2, pp. 175-176). This result parallels the corresponding result for *eigenvalues* established by Birkhoff, de Boor, Swartz, and Wendroff.¹ A similar order-of-convergence theorem will be proved for the approximate eigenfunctions and eigenvalues computed by the Weinstein-Bazley method (ref. 3, chap. 12). This latter method yields *lower bounds* to the eigenvalues, while the Rayleigh-Ritz approximate eigenvalues are *upper bounds*. These theorems concern perturbations of self-adjoint operators of a kind arising in many concrete Sturm-Liouville systems, and characterized abstractly by the following definition, in which $D(B)$ denotes the range of the operator B .

Definition: Self-adjoint operators L_0 and $L = L_0 + B$ on a Hilbert space $\mathcal{H}$ will be said to have *compatibly compact inverses* when:

- (a) L_0 and L are positive definite and have compact inverses,
- (b) $B = L - L_0$ is self-adjoint and $D(B) \supset D(L_0)$.

Note that L and L_0 have orthonormal bases of eigenvectors, which we denote by $\{u_k\}$ and $\{v_k\}$, respectively, with eigenvalues which can be ordered as

$$0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \leq \dots, \quad 0 < \mu_1 \leq \mu_2 \leq \dots \leq \mu_n \leq \dots \quad (1)$$

We let u_k^n , $1 \leq k \leq n$, be the k th approximate eigenvector obtained by the Rayleigh-Ritz method, i.e. as the (projective) "point" where the "Rayleigh quotient" $(Lu, u)/(u, u)$ assumes its k th critical value on the subspace

$$S_n = \text{Span}[v_j]_{j=1}^n. \quad (2)$$

THEOREM 1. Let L_0 and $L_0 + B$ be compatibly compact, and let Q_n be the projection of $\mathcal{H}$ onto S_n with

$$\|B(I - Q_n)u_k\| = O(n^{-r'}), \quad \|(I - Q_n)u_k\| = O(n^{-r}), \quad (3)$$

as $n \rightarrow \infty$. If also

$$\lambda_{k-s-1} < \lambda_k < \lambda_{k+t+1}, \quad (4)$$

then

$$\text{dist}\{u_k, \text{Span}[u_j^n]_{j=k-s}^{k+t}\} = O(n^{-r}), \quad r = \min(r', r''). \quad (5)$$

COROLLARY. If λ_k is simple, then

$$\|u_k - u_k^n\| = O(n^{-r}). \quad (5')$$

Observe that if B is bounded, then $r' = r'' = r$, and (5') states that the Rayleigh-Ritz approximation u_k^n has the same order of accuracy as the best approximation $Q_n u_k$ to u_k in S_n .

THEOREM 2. *In addition to the hypotheses (a) and (b), assume (c) that B is bounded with $(Bu, u) \geq \lambda \|u\|^2$ for all $u \in \mathcal{H}$. Then B^{-1} exists; hence so does the orthogonal projection P_n on $\text{Span}[B^{-1}v_j]_{j=1}^n$ in the Hilbert space $\mathcal{H}'$ associated with the inner product*

$$[u, v] = (Bu, v).$$

Let (3) hold, and let

$$\|(I - Q_n)Bu_k\| = O(n^{-p}), \quad (L[I - Q_n]u_k, [I - Q_n]u_k) = O(n^{-t}), \quad (6)$$

as $n \rightarrow \infty$. Then, if v_k^n and μ_k^n are the k th orthonormal eigenvector and eigenvalue of the n th intermediate Weinstein-Bazley operator (ref. 3, chap. 12)

$$L_n = L_0 + BP_n, \quad (7)$$

when λ_k is simple, we have

$$\|u_k - v_k^n\| = O(n^{-\nu'}), \quad \lambda_k - \mu_k^n = O(n^{-2\nu}), \quad 1 \leq k \leq n, \quad (8)$$

where

$$\nu' = \min(r, p), \quad \nu = \min[r, p, t/2]. \quad (8')$$

Alternatively, if λ_k is a multiple eigenvalue as in (4), then

$$\text{dist}\{u_k, \text{Span}[v_j^n]_{j=k-s}^{k+t}\} = O(n^{-\nu'}), \quad \lambda_k - \mu_k^n = O(n^{-\nu}). \quad (9)$$

Observe that if $p = r$, then the eigenvector estimate (8) is best possible.

Applications: Typical applications of Theorems 1 and 2 for various special cases are as follows; a detailed discussion will be published elsewhere.⁵

For boundary value problems which have smooth periodic eigenfunctions, such as

$$Lu = -u'' + q(x)u = \lambda u, \quad q(x + \pi) = q(x), \quad q \in C^\infty(-\infty, \infty), \quad (10)$$

subject to periodic boundary conditions, the above (with $L_0 u = -u''$, $Bu = qu$, $q > 0$) implies that both the Rayleigh-Ritz and Weinstein-Bazley approximations have an *infinite order of convergence*. For regular Sturm-Liouville problems in normal form,

$$Lu = -u'' + q(x)u, \quad u(0) = u(1) = 0, \quad (11)$$

Theorem 1 implies that the Rayleigh-Ritz approximate eigenvalues and eigenfunctions have an error of $O(n^{-r})$ and $O(n^{-s/2})$, respectively; Theorem 2 implies that the Weinstein-Bazley approximations have errors of $O(n^{-s})$ and $O(n^{-s/2})$.

Proof of Theorem 1: The basic idea of the proof is to show that $Q_n u_k$ is "close" to the k th Rayleigh-Ritz approximate eigenvector. This will be accomplished by using the following theorem due to Wilkinson (ref. 4, pp. 172-173).

THEOREM W (Wilkinson). *Let A be an $n \times n$ real symmetric matrix with eigenvalues $\Lambda_1 \leq \dots \leq \Lambda_n$ and orthonormal eigenvectors $\{U_j\}_{j=1}^n$. Let V be a vector and μ a real number such that for some k , $1 \leq k \leq n$,*

$$(U_k, V) \geq 0, \quad \|V\| = 1, \quad \epsilon = AV - \mu V, \quad \nu = V^T AV. \quad (12)$$

Then if Λ_k is simple,

$$\|U_k - V\|^2 \leq (\|\epsilon\|/\alpha_k)^2 + (\|\epsilon\|/\alpha_k)^4, \quad |\nu - \Lambda_k| \leq (\|\epsilon\|/\alpha_k) \times [1 - (\|\epsilon\|/\alpha_k)^2]^{-1}, \quad (13)$$

where $\alpha_k = \min_{l \neq k} |\Lambda_l - \nu|$. If

$$\Lambda_{k-s-1} < \Lambda_{k-s} \leq \dots \leq \Lambda_{k+t} < \Lambda_{k+t+1},$$

then

$$\text{dist} \{V, \text{Span } [U_j]_{j=k-s}^{k+t}\} \leq \|\epsilon\|/d_k, \quad \min |\nu - \Lambda_j| \leq \|\epsilon\|, \quad (14)$$

where $d_k = \min \{|\Lambda_l - \nu|_l \neq m, k-s \leq m \leq k+t\}$.

We shall use Theorem W by applying it to $V = Q_n u_k / \|Q_n u_k\|$ and $\mu = \lambda_k$. We first recall that the Rayleigh-Ritz approximate eigenvalues and eigenvectors are those of the n -dimensional operator

$$Q_n L Q_n = Q_n L_0 Q_n + Q_n B Q_n. \quad (15)$$

LEMMA 1. For $1 \leq k \leq n$, let

$$\epsilon_k = Q_n L Q_n u_k - \lambda_k Q_n u_k. \quad (16)$$

Then

$$\|\epsilon_k\| \leq \|B(I - Q_n)u_k\|. \quad (17)$$

Proof: Since Q_n is the orthogonal projection on the subspace spanned by the first n eigenfunctions of L_0 , we have $Q_n L_0 = L_0 Q_n$, $Q_n^2 = Q_n$. Thus, applying Q_n twice to $L_0 u_k + B u_k = \lambda_k u_k$ gives

$$Q_n L_0 Q_n u_k + Q_n B u_k = \lambda_k Q_n u_k. \quad (18)$$

Hence,

$$\epsilon_k = Q_n B Q_n u_k - Q_n B u_k, \quad (19)$$

and so

$$\|\epsilon_k\| \leq \|Q_n\| \|B(I - Q_n)u_k\| = \|B(I - Q_n)u_k\|. \quad (20)$$

This proves (17).

Since

$$\|Q_n u_k\|^2 = \|u_k\|^2 + \|(I - Q_n)u_k\|^2 = 1 + \|(I - Q_n)u_k\|^2, \quad (21)$$

(5) follows from (3), (16), (17), (21), and Theorem W (with $V = Q_n u_k / \|Q_n u_k\|$, $\mu = \lambda_k$, and $A = Q_n L Q_n$).

Proof of Theorem 2: The proof of Theorem 2 is similar to the proof of Theorem 1 in that Theorem W is used with $V = Q_n u_k / \|Q_n u_k\|$ and $\mu = \lambda_k$. The operator A , however, will be

$$Q_n L_n Q_n = Q_n L_0 Q_n + Q_n B P_n Q_n, \quad (22)$$

as in reference 3, chapter 12, and not $Q_n L Q_n$.

LEMMA 2. For $1 \leq k \leq n$, let

$$e_k = Q_n L_n Q_n u_k - \lambda_k Q_n u_k. \quad (23)$$

Then

$$\|e_k\| \leq \|B\| \|(I - Q_n)u_k\| + (\|B\|/\gamma) \|(I - Q_n)Bu_k\|. \quad (24)$$

Proof: From (18), we have

$$e_k = Q_n B P_n Q_n u_k - Q_n B u_k = -Q_n B(I - P_n)u_k - Q_n B P_n(I - Q_n)u_k. \quad (25)$$

Thus,

$$\|e_k\| \leq \|B(I - P_n)u_k\| + \|B P_n\| \|(I - Q_n)u_k\|. \quad (26)$$

Observe that if

$$|||w||| = [w, w]^{1/2} = (Bw, w)^{1/2}, \quad (27)$$

then from condition (c) of Theorem 2

$$\gamma |||w|||^2 \leq |||w|||^2 \leq \|B\| |||w|||^2.$$

Therefore, since $B P_n$ is self-adjoint in $\mathcal{H}$,

$$\|B P_n\| = \sup_{w \in \mathcal{H}} \frac{(B P_n w, w)}{(w, w)} \leq \sup \frac{[P_n w, w]}{|||w|||^2} \leq \sup \frac{|||w|||^2}{|||w|||^2} \leq \|B\|. \quad (28)$$

Also,

$$\|B[I - P_n]u_k\| \leq |||B|||^{1/2} |||[I - P_n]u_k|||. \quad (29)$$

Since B is self-adjoint in $\mathcal{H}'$,

$$|||B||| = \sup \frac{[Bw, w]}{[w, w]} = \sup \frac{(B^2 w, w)}{|||w|||^2} \leq \|B\|^2 \sup \frac{|||w|||^2}{|||w|||^2} \leq \|B\|^2 / \gamma. \quad (30)$$

Now, P_n is the orthogonal projection of $\mathcal{H}'$ onto $\text{Span } [B^{-1}v_i]_{i=1}^n$, and

$$u_k = \sum_{i=1}^n (Bu_k, v_i) B^{-1}v_i. \quad (31)$$

Thus

$$|||[I - P_n]u_k||| \leq |||\sum_{i=n+1}^{\infty} (Bu_k, v_i) B^{-1}v_i||| \leq (1/\gamma)^{1/2} |||[I - Q_n]Bu_k|||. \quad (32)$$

Combining (26), (28)–(30), and (32), we obtain (24).

Theorem 2 now follows from Wilkinson's theorem (with $V = Q_n u_k / \|Q_n u_k\|$, $\mu = \lambda_k$, and $A = Q_n L_n Q_n$), Lemma 2, (21), and

$$(LQ_n u_k, Q_n u_k) = \lambda_k + (L(I - Q_n)u_k, (I - Q_n)u_k). \quad (33)$$

* Research supported by the Office of Naval Research.

¹ Birkhoff, G., C. de Boor, B. Swartz, and B. Wendroff, "Rayleigh-Ritz approximation by piecewise cubic polynomials," *SIAM J. Numer. Anal.*, **3**, 188–203 (1966).

² Courant, R., and D. Hilbert, *Methods of Mathematical Physics* (New York: Interscience, 1953), vol. 1.

³ Gould, S. H., *Variational Methods for Eigenvalue Problems* (University of Toronto Press, 1966).

⁴ Wilkinson, J. H., *The Matrix Eigenvalue Problem* (Oxford University Press, 1965).

⁵ Birkhoff, G., and G. Fix, "Accurate numerical solution of eigenvalue problems," to appear in the Proceedings of the Twentieth Symposium on Applied Mathematics, American Mathematical Society.