

# Convergence theorems for intermediate problems. II

**Christopher A. Beattie**

Department of Mathematics, Virginia Tech,  
Blacksburg, VA 24061, USA

**W. M. Greenlee**

Department of Mathematics, University of Arizona,  
Tucson, AZ 85721, USA

(MS received 4 January 2001; accepted 15 January 2002)

Convergence theorems for the practical eigenvector free methods of Gay and Goerisch are obtained under a variety of hypotheses, so that our theorems apply to both traditional boundary-value problems and atomic problems. In addition, we prove convergence of the  $T^*T$  method of Bazley and Fox without an alignment of projections hypothesis required in previous literature.

## 1. Introduction

The method of intermediate problems of Weinstein and Aronszajn (cf. [19, 20]) provides a means of calculation for bounds to eigenvalues complementary to the Rayleigh–Ritz bounds. Convergence theorems for the method of intermediate problems in a generality that includes problems with essential spectra can be found in [5, 7–9, 12]. This paper supplements the general convergence theorems of [7, 9] in two respects. In each of [7, 9], convergence theorems for the  $T^*T$  method of Bazley and Fox require an alignment of projections hypothesis that is both inconvenient and ignored in computational practice. By examining this method directly and not relating the Bazley–Fox projections, commonly called inner projections by physicists, to the earlier Aronszajn or outer projections, we are able to remove the alignment hypothesis. The basic tool, as in [5, 7, 9], is a monotone convergence theorem of Weidmann [18].

We then turn our attention to the so-called eigenvector-free (EVF) methods, which, like truncation methods (cf. [5, 12, 19, 20]), provide a practical means of dealing with the Weinstein–Aronszajn determinant without severely constraining the choice vectors. These methods can be viewed as extensions of both Bazley’s method of special choice (cf. [19, 20]) and Temple–Lehmann methods (cf. [16, 17, 19]). They originate in Gay [10], and like Rayleigh–Ritz calculations, the choice vectors are selected as approximations to the unknown eigenvectors rather than to add a perturbation to a base operator. But precise means of relating the approximate eigenvalues to those desired awaited the work of Goerisch [11], who coined the phrase EVF (cf. also [4]).

EVF calculations, like special choice, require an operator inversion. Beattie and Goerisch [6] introduce a parameter and use a second projection to reduce the operator inversion to an equivalent Schur complement of a larger matrix. Convergence of the method of [6] has been proved by Lee [15] (cf. [14] also) when  $T^*T$  is relatively bounded with respect to the base operator  $A_0$ . Lee's projection estimates lead to rate-of-convergence estimates, provided they are proved for the projections. In § 3 we obtain convergence theorems for the procedure of [6], both for inner second projections as in [6] and for outer second projections without a relative boundedness hypothesis. Convergence of Gay's method then follows provided the parameter of [6] is retained, or the operator to be inverted is positive-definite. But, in the atomic problems considered by Gay, the parameter is numerically inconvenient and the inversion is of a positive, but not positive-definite operator. Corresponding convergence criteria are obtained in § 4, where the hypotheses appear to be rather special. But our assumptions are satisfied in [10], and other atomic problems, since the associated operator sums are known to be self-adjoint.

## 2. Convergence of the $T^*T$ method

Let  $\mathcal{H}$  be a separable complex Hilbert space with norm  $\|\cdot\|$  and inner product  $\langle \cdot, \cdot \rangle$ . Let  $A$  be a self-adjoint operator in  $\mathcal{H}$  with domain  $\mathcal{D}(A)$ . We assume that  $A$  is bounded below with spectrum that begins with isolated eigenvalues of finite multiplicity,

$$\lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_\infty(A),$$

and corresponding orthonormal eigenvectors  $u_1, u_2, \dots$ . From here on in,  $\lambda_\infty(A)$  denotes the least point of the essential spectrum of  $A$ , with the usual convention that  $\lambda_\infty(A) = \infty$  if  $A$  has compact resolvent. Denote the quadratic form corresponding to  $A$ , i.e. the closure of  $\langle Au, u \rangle$ , by  $a(u)$ , with domain  $\mathcal{D}(a)$ .

To apply the method of intermediate problems to obtain lower bounds to the eigenvalues of  $A$  below  $\lambda_\infty(A)$ , we assume that  $a(u)$  can be written explicitly as a sum  $a_0(u) + b(u)$ , where  $a_0(u)$  and  $b(u)$  are densely defined closed real-valued quadratic forms in  $\mathcal{H}$ , with  $a_0(u)$  bounded from below and  $b(u)$  non-negative. In the general theory of intermediate problems, the requirement that  $b(u)$  be closed can be relaxed somewhat, as in [7], but it is needed for a convergence theory of the  $T^*T$  method in  $\mathcal{H}$  as presented here. By definition,  $\mathcal{D}(a) = \mathcal{D}(a_0) \cap \mathcal{D}(b)$ , so that  $a_0 \leq a$  in the usual sense (cf. [13]), i.e.  $\mathcal{D}(a_0) \supset \mathcal{D}(a)$  and  $a_0(u) \leq a(u)$  for all  $u \in \mathcal{D}(a)$ . It is assumed that the base problem, i.e. the spectral problem for the self-adjoint operator  $A_0$  corresponding to  $a_0(u)$ , is solved and that the spectrum of  $A_0$  also begins with discrete eigenvalues of finite multiplicity,

$$\lambda_1(A_0) \leq \lambda_2(A_0) \leq \cdots \leq \lambda_\infty(A_0),$$

with corresponding orthonormal eigenvectors  $u_1^0, u_2^0, \dots$ . The second monotonicity principle (cf. [19]) implies that  $\lambda_\infty(A_0) \leq \lambda_\infty(A)$ , and for each  $i$  such that  $\lambda_i(A) < \lambda_\infty(A_0)$ ,  $\lambda_i(A_0)$  exists with  $\lambda_i(A_0) \leq \lambda_i(A)$ .

We now consider a version of the method of intermediate problems to improve the base problem lower bounds that is due to Bazley and Fox [3]. Let  $B$  be the self-adjoint operator in  $\mathcal{H}$  corresponding to  $b(u)$ . Then  $B$  is non-negative and  $b(u)$  is

the closure of the quadratic form  $\langle Bu, u \rangle$ . We assume that we can find explicitly another separable complex Hilbert space  $\mathcal{H}_*$  and a closed densely defined operator  $T : \mathcal{H} \rightarrow \mathcal{H}_*$  such that  $B = T^*T$ , where  $T^*$  is the adjoint of  $T$ . Then  $\mathcal{D}(T) = \mathcal{D}(b)$  and  $b(u) = \|Tu\|_*^2$  on  $\mathcal{D}(b)$ , where  $\|\cdot\|_*$  denotes the norm in  $\mathcal{H}_*$ . So  $a(u) = a_0(u) + \|Tu\|_*^2$ .

Select vectors  $\{p_i\}_{i=1}^k \subset \mathcal{D}(T^*) \subset \mathcal{H}_*$  and define the associated  $H_*$ -orthogonal projection

$$P_k u = \sum_{i,j=1}^k \langle u, p_i \rangle_* \mathcal{B}_{ij} p_j,$$

where  $\langle \cdot, \cdot \rangle_*$  is the inner product in  $\mathcal{H}_*$  and  $[\mathcal{B}_{ij}]$  is the Moore–Penrose generalized matrix inverse to the Gram matrix  $[\langle p_i, p_j \rangle_*]$ . For  $u \in \mathcal{D}(a_0) \cap \mathcal{D}(T)$ , define  $a_k(u) = a_0(u) + \|P_k T u\|_*^2$ . Since  $\mathcal{R}(P_k)$ , the range of  $P_k$ , is contained in  $\mathcal{D}(T^*)$ , we may write

$$P_k T u = \sum_{i,j=1}^k \langle u, T^* p_i \rangle \mathcal{B}_{ij} p_j.$$

Thus  $a_k$  is closable, with associated self-adjoint operator

$$A_k = A_0 + \sum_{i,j=1}^k \langle \cdot, T^* p_i \rangle \mathcal{B}_{ij} T^* p_j \supset A_0 + T^* P_k T,$$

with  $\mathcal{D}(A_k) = \mathcal{D}(A_0)$ . Henceforth,  $a_k$  denotes the closed quadratic form associated with  $A_k$ ,  $\mathcal{D}(a_k) = \mathcal{D}(a_0)$ . Since  $A_k$  is a finite-rank perturbation of  $A_0$ ,  $\lambda_\infty(A_k) = \lambda_\infty(A_0)$ , and by the second monotonicity principle, for each  $i$  such that  $\lambda_i(A) < \lambda_\infty(A_0)$ , we have

$$\lambda_i(A_0) \leq \lambda_i(A_k) \leq \lambda_i(A_{k+1}) \leq \lambda_i(A).$$

Previous convergence results for this method [7, 9] have required alignment of the projections  $P_k$  with  $\mathcal{R}(T)$ , i.e.  $\mathcal{R}(P_k T) \subset \mathcal{D}(T^*) \cap \mathcal{R}(T)$ . This additional hypothesis is a nuisance in practice, and arises from relating convergence results for the  $T^*T$  method to those for the earlier Aronszajn method [1]. The following theorem removes the alignment requirement. Recall (cf. [13]) that a linear submanifold  $\mathcal{M}$  of the domain of a closed operator  $S$  is called a *core* of  $S$  if the set of elements  $\{u, Su\}$  with  $u$  in  $\mathcal{M}$  is dense in the graph of  $S$ .

**THEOREM 2.1.** *If  $\text{span}\{p_i\}$  is a core for  $T^*$ , then  $\lambda_i(A_k) \rightarrow \lambda_i(A)$  as  $k \rightarrow \infty$  for each  $i$  such that  $\lambda_i(A) < \lambda_\infty(A_0)$ .*

*Proof.* By a theorem of Weidmann [18], it suffices to prove that  $A_k \rightarrow A$  as  $k \rightarrow \infty$  in the strong resolvent sense, and then also one has uniform convergence of the corresponding eigenprojections. For this purpose, note that (cf. [13])  $\{A_k\}$  converges in the strong resolvent sense to a self-adjoint operator  $A_\infty$  satisfying  $A_k \leq A_\infty \leq A$  for each  $k$ . If  $a_\infty$  denotes the quadratic form corresponding to  $A_\infty$ , we have  $a_\infty \leq a$ , and so we need only show that  $a_\infty = a$ .

Assume, without loss of generality, that  $\{p_i\}_{i=1}^\infty$  is an orthonormal sequence in  $\mathcal{H}_*$  and define  $a_*(u) = \lim_{k \rightarrow \infty} a_k(u)$  on

$$\mathcal{D}(a_*) = \{u \in \mathcal{D}(a_0) : \sup_k a_k(u) < \infty\} = \left\{u \in \mathcal{D}(a_0) : \sum_{i=1}^\infty |\langle u, T^*p_i \rangle|^2 < \infty\right\}.$$

Polarization shows that  $\mathcal{D}(a_*)$  is a subspace of  $\mathcal{H}$ . The real-valued quadratic form  $a_*$  is a least upper bound to  $\{a_k\}$  in the sense that any closed real-valued quadratic form that dominates each  $a_k$  also dominates  $a_*$ . In particular,  $a_* \leq a_\infty \leq a$ . We now show that  $a_\infty = a$  by showing that  $a_* = a$ .

Since  $\text{span}\{p_i\}$  is, by the core hypothesis, dense in  $\mathcal{H}_*$ , we have that, for any  $u \in \mathcal{D}(a) = \mathcal{D}(a_0) \cap \mathcal{D}(T)$ ,

$$\sum_{i=1}^\infty |\langle u, T^*p_i \rangle|^2 = \sum_{i=1}^\infty |\langle Tu, p_i \rangle_*|^2 = \|Tu\|_*^2 < \infty.$$

Thus  $u \in \mathcal{D}(a_*)$  and  $a_*(u) = a(u)$ , which shows that  $a_* \supset a$ .

Now pick  $u \in \mathcal{D}(a_*)$  and let

$$M^2 = \sum_{i=1}^\infty |\langle u, T^*p_i \rangle|^2 < \infty.$$

For any  $p$  in  $\text{span}\{p_i\}$ , there exists a natural number  $N$  and complex scalars  $\{\alpha_i\}_{i=1}^N$  such that

$$p = \sum_{i=1}^N \alpha_i p_i \quad \text{and} \quad \|p\|_*^2 = \sum_{i=1}^N |\alpha_i|^2.$$

Thus

$$|\langle u, T^*p \rangle| = \left| \sum_{i=1}^N \bar{\alpha}_i \langle u, T^*p_i \rangle \right| \leq \left[ \sum_{i=1}^N |\alpha_i|^2 \right]^{1/2} \left[ \sum_{i=1}^N |\langle u, T^*p_i \rangle|^2 \right]^{1/2} \leq M \|p\|_*.$$

Since  $\text{span}\{p_i\}$  is a core for  $T^*$ ,  $u \in \mathcal{D}(T^{**}) = \mathcal{D}(T) = \mathcal{D}(b)$ , so that  $a_* = a$ . □

For later use, we offer the following corollary to the proof of theorem 2.1.

**COROLLARY 2.2.** *If  $\text{span}\{p_i\}$  is a core for  $T^*$ , then  $T^*P_kT$  converges to  $B = T^*T$  in the strong resolvent sense.*

*Proof.* The proof is that of theorem 2.1 if one substitutes  $b$  for  $a$  and  $B$  for  $A$  throughout that proof. □

### 3. Convergence of the eigenvector-free method of Beattie and Goerisch

Since  $A_k$  is a finite-rank perturbation of  $A_0$ , computation of  $\{\lambda_i(A_k)\}$  is accomplished through consideration of the rank and inertia of the Weinstein–Aronszajn (or WA) matrix,

$$\mathcal{W}_k(\lambda) = [\langle p_i, p_j \rangle_* + \langle R_\lambda^0 T^* p_i, T^* p_j \rangle],$$

considered as a function of the eigenvalue parameter  $\lambda$  (cf. [4]). Henceforth,  $R_\lambda^0$  denotes the resolvent operator  $(A_0 - \lambda)^{-1}$ . For general choices of the projecting

vectors  $\{p_i\}$ , evaluation of  $\mathcal{W}_k(\lambda)$  is usually very difficult, since the action of  $R_\lambda^0$  is rarely accessible in a computationally usable form. Bazley and Fox [3] proposed the use of a ‘special choice’, determining  $\{p_i\}_{i=1}^k$  so that  $T^*p_i = u_i^0$ . This choice for  $\{p_i\}$  makes the evaluation of  $\mathcal{W}_k(\lambda)$  comparatively straightforward, but one is left with a highly constrained choice of the projecting vectors  $\{p_i\}$  and the possibly difficult task of solving the operator equations  $T^*p_i = u_i^0$ . The latter problem can be circumvented by the Bazley–Fox method of second projection, as follows.

Following [2, 3], let us define, for some  $\theta > 0$ ,  $B_{k,\theta} = T^*P_kT + \theta$ . The operator  $B_{k,\theta}$  is bounded, self-adjoint, positive-definite and its inverse is given explicitly by

$$B_{k,\theta}^{-1}v = \frac{1}{\theta} \left[ v - \sum_{l,m=1}^k \langle v, T^*p_l \rangle \mathcal{C}_{l,m} T^*p_m \right], \quad v \in H,$$

where  $[\mathcal{C}_{l,m}]$  is the Moore–Penrose generalized matrix inverse to

$$[\theta \langle p_i, p_j \rangle_* + \langle T^*p_i, T^*p_j \rangle].$$

Now choose a set of vectors  $\{\hat{p}_i\}_{i=1}^n$  and note that the corresponding projection onto  $\text{span}\{\hat{p}_i\}_{i=1}^n$ , which is orthogonal with respect to the inner product induced by  $B_{k,\theta}$ , is given by

$$\hat{P}_n^{k,\theta} = \sum_{i,j=1}^n \langle \cdot, B_{k,\theta} \hat{p}_i \rangle \mathcal{D}_{ij} \hat{p}_j,$$

where  $[\mathcal{D}_{ij}]$  is the Moore–Penrose generalized matrix inverse to  $[\langle B_{k,\theta} \hat{p}_i, \hat{p}_j \rangle]$ . Then we have that  $B_{k,\theta} \hat{P}_n^{(k,\theta)}$  is a bounded symmetric positive-semidefinite operator, and

$$\langle B_{k,\theta} \hat{P}_n^{(k,\theta)} v, v \rangle \leq \langle B_{k,\theta} v, v \rangle$$

for all  $v \in \mathcal{H}$ . Thus, if we define

$$A_{k,n}^\theta = A_0 - \theta + B_{k,\theta} \hat{P}_n^{(k,\theta)}, \tag{3.1}$$

it is clear that the eigenvalues of  $A_{k,n}^\theta$  give lower bounds to the corresponding eigenvalues of  $A$ , and that they are monotone increasing in both  $k$  and  $n$ . The associated  $n \times n$  WA matrix is given by

$$\mathcal{W}_{k,n}(\lambda) = [(\hat{p}_i + R_{\lambda+\theta}^0 B_{k,\theta} \hat{p}_i, B_{k,\theta} \hat{p}_j)]. \tag{3.2}$$

If one views the intermediate operators (3.1) in the traditional WA sense with fixed  $\theta$ ,  $\{p_i\}$  and  $\{\hat{p}_i\}$ , evaluation of (3.2) for varying  $\lambda$  is, in general, difficult unless one chooses  $\hat{p}_i = B_{k,\theta}^{-1}u_i^0$ , as originally envisioned in the method of second projection. If, however, as in Beattie and Goerisch [6], we define  $\mu = \lambda + \theta$  and then consider only those  $\lambda$  for which  $\lambda < \mu$ , and introduce the change of variable  $q_i = R_\mu^0 B_{k,\theta} \hat{p}_i$  into (3.2), we get

$$\mathcal{W}_{k,n}(\lambda) = [\langle B_{k,\theta}^{-1}(A_0 - \mu)q_i, (A_0 - \mu)q_j \rangle + \langle q_i, (A_0 - \mu)q_j \rangle], \tag{3.3}$$

and things improve considerably. Fix  $\mu$ ,  $\{q_i\}_{i=1}^n \subset \mathcal{D}(A_0)$  and  $\{p_i\}_{i=1}^k \subset \mathcal{D}(T^*)$  and observe that (3.3) is the WA matrix for an intermediate problem for each (fixed)  $\lambda < \mu$ . While (3.3) is easy to evaluate and is free of dependence on the base problem

eigenvectors  $u_i^0$ , thus leading to the name *eigenvector free* (EVF), as  $\lambda$  changes, so do  $\theta$ ,  $\{\hat{p}_i\}$  and  $[C_{lm}]$ . Thus the related intermediate problem changes, and not at all in a monotonic fashion. In [6] it is shown how to use the spectrum-slicing technique of [4] to interpret each  $\lambda < \mu$  as a lower bound to an eigenvalue  $\lambda_p$  of  $A$ . The transformation from (3.2) to (3.3) is analogous to that introduced by Gay [10] for the original Aronszajn method [1]. The related eigenvalue-counting problem for Gay’s method was first solved by Goerisch [11]. A simpler proof relating Goerisch’s result to WA matrices was given by Beattie [4].

The intermediate operators (3.1) are basic to the parametrized intermediate problems implicit in the WA matrix (3.2) used to compute with the EVF method of [6]. We shall assume henceforth in this section that  $\mu$  is in the resolvent set for  $A_0 : \lambda_{r-1}^0 < \mu < \lambda_r^0$ . This may be done without loss of generality to our convergence result, since, by a monotonicity theorem of [6], allowing  $\mu = \lambda_r^0$  provides bounds no worse than for  $\lambda_{r-1}^0 < \mu < \lambda_r^0$ . Thus convergence for  $\lambda_{r-1}^0 < \mu < \lambda_r^0$  implies convergence for  $\mu = \lambda_r^0$ . Likewise, we assume in this section, without loss of generality, that  $\{q_i\}$  is a linearly independent set. We need two lemmas preceding our convergence theorem.

LEMMA 3.1.  $A_{n,k}^\theta$  is holomorphic in  $\theta$  for  $\text{Re}(\theta) > 0$ . For  $\theta \in (0, \infty)$ ,  $A_{n,k}^\theta$  is self-adjoint and monotone increasing in  $n$  and  $k$ .

Proof. Consider the closed quadratic form associated with  $A_{n,k}^\theta$ ,

$$a_{n,k}^\theta(u) = a_0(u) - \theta \|u\|^2 + \sum_{i,j=1}^n \langle u, (A_0 - \mu)q_i \rangle \mathcal{D}_{ij}^{(k,\theta)} \langle (A_0 - \mu)q_j, u \rangle,$$

where  $[\mathcal{D}_{ij}^{(k,\theta)}]$  is the  $n \times n$  matrix inverse to

$$[\langle (T^*P_kT + \theta)^{-1}(A_0 - \mu)q_i, (A_0 - \mu)q_j \rangle].$$

$[\mathcal{D}_{ij}^{(k,\theta)}]$  is obviously holomorphic in the right half-plane, hence so are  $A_{n,k}^\theta$  and  $A_{n,k}^{\theta}$ . To establish the monotonicity assertions, fix  $\theta \in (0, \infty)$ .  $A_{n,k}^\theta$  is self-adjoint on  $\mathcal{D}(A_{n,k}^\theta) = \mathcal{D}(A_0)$ , and monotonicity with respect to  $n$  for fixed  $k$  follows from Bessel’s inequality in the inner product associated with  $(T^*P_kT + \theta)$ . To determine monotonicity with respect to  $k$  for fixed  $n$ , observe that

$$0 < (T^*P_kT + \theta) \leq (T^*P_{k+1}T + \theta) \quad \text{implies} \quad (T^*P_{k+1}T + \theta)^{-1} \leq (T^*P_kT + \theta)^{-1}.$$

Thus

$$[\langle (T^*P_{k+1}T + \theta)^{-1}(A_0 - \mu)q_i, (A_0 - \mu)q_j \rangle] \leq [\langle (T^*P_kT + \theta)^{-1}(A_0 - \mu)q_i, (A_0 - \mu)q_j \rangle],$$

implying in turn  $[\mathcal{D}_{ij}^{(k,\theta)}] \leq [\mathcal{D}_{ij}^{(k+1,\theta)}]$ , so that  $a_{n,k}^\theta(u) \leq a_{n,k+1}^\theta(u)$ . □

Lemma 3.1 allows one to conclude that the eigenvalues of  $A_{n,k}^\theta$  lying below the essential spectrum of  $A_0$  and their corresponding normalized eigenvectors have a piecewise analytic dependence on  $\theta$  for  $\theta \in (0, \infty)$  and, in particular, are continuous functions of  $\theta$ . Loss of analyticity can occur only at the isolated points where the curves  $\lambda_i^{(n,k)}(\theta)$  cross for two or more indices  $i$ . This loss of analyticity is an artifact of the standard numbering conventions for eigenvalues, and even at such crossing points  $\lambda_i^{(n,k)}(\theta)$  may be continued analytically from an adjacent non-crossing point where  $\lambda_i^{(n,k)}(\theta)$  is analytic (cf. [13]).

LEMMA 3.2. At values of  $\theta$  where  $\lambda_i^{(n,k)}(\theta)$  crosses no other eigenvalue curve and

$$\lambda_i^{(n,k)}(\theta) < \lambda_\infty(A_0),$$

$\lambda_i^{(n,k)}(\theta)$  is differentiable and

$$\frac{d}{d\theta} \lambda_i^{(n,k)} \geq -1,$$

with equality only if  $\lambda_i^{(n,k)}(\theta) = \lambda_j - \theta$  for some  $j$ , i.e. only if  $\lambda_i^{(n,k)}(\theta)$  is a persistent eigenvalue of the base operator  $A_0 - \theta$  in the terminology of [20].

*Proof.* The above discussion indicates that isolated eigenvalues of finite multiplicity are analytic in  $\theta$ , so, in particular, they are differentiable with differentiable normalized eigenfunctions  $u_i^\theta$ , where the dependence on  $n$  and  $k$  has been suppressed. Differentiation of  $\lambda_i^{(n,k)}(\theta) = \langle A_{n,k}^\theta u_i^\theta, u_i^\theta \rangle$  reveals

$$\frac{d}{d\theta} \lambda_i^{(n,k)} = -1 + \sum_{l,m=1}^n \langle u_i^\theta, (A_0 - \mu)q_l \rangle \mathcal{E}_{lm} \langle (A_0 - \mu)q_m, u_i^\theta \rangle,$$

where

$$\mathcal{E}_{ij} = \sum_{l,m=1}^n \mathcal{D}_{il}^{(k,\theta)} \langle (T^* P_k T + \theta)^{-1} (A_0 - \mu)q_l, (T^* P_k T + \theta)^{-1} (A_0 - \mu)q_m \rangle \mathcal{D}_{mj}^{(k,\theta)}$$

is positive-definite. Thus

$$\frac{d}{d\theta} \lambda_i^{(n,k)} \geq -1,$$

with equality only if  $\langle u_i^\theta, (A_0 - \mu)q_l \rangle = 0$  for  $l = 1, \dots, n$ , which would, in turn, imply

$$\lambda_i^{n,k}(\theta) = A_{n,k}^\theta u_i^\theta = (A_0 - \theta)u_i^\theta,$$

so that  $\lambda_i^{(n,k)}(\theta) = \lambda_j^0 - \theta$  for some  $j$ . □

As  $n$  and  $k$  increase, the eigenvalue curves  $\{\lambda_i^{(n,k)}(\theta)\}_{n,k}$  form a monotone increasing family of continuous functions to the extent that

$$\lambda_i^{(n,k)}(\theta) \leq \lambda_i^{(n+1,k)}(\theta) \quad \text{and} \quad \lambda_i^{(n,k)}(\theta) \leq \lambda_i^{(u,k+1)}(\theta)$$

(see figure 1). The ordinates of the points of intersection of the eigenvalue curves  $\{\lambda_i^{(n,k)}(\theta)\}_{n,k}$  and the vertical line  $\theta = \hat{\theta}$  give lower bounds to  $\lambda_i$  obtained by resolving the standard intermediate problems  $A_{n,k}^\theta$ , which are of second projection type, as described in [2, 3, 6]. The EVF method of [6], in contrast, produces lower bounds to  $\lambda_i$ , as the ordinates of the points of intersection of the eigenvalue curves  $\{\lambda_i^{(n,k)}(\theta)\}_{n,k}$  and the line  $\lambda + \theta = \mu$ . Lemma 3.2 guarantees that for  $\lambda_i < \mu$  these intersection points are uniquely defined wherever they exist. From figure 1, it is evident that the EVF method of [6] produces convergent lower bounds to every eigenvalue  $\lambda_i$  of  $A$  such that  $\lambda_i < \mu$ , provided that the eigenvalue curves  $\{\lambda_i^{(n,k)}(\theta)\}_{n,k}$  converge uniformly to  $\lambda_i$ .

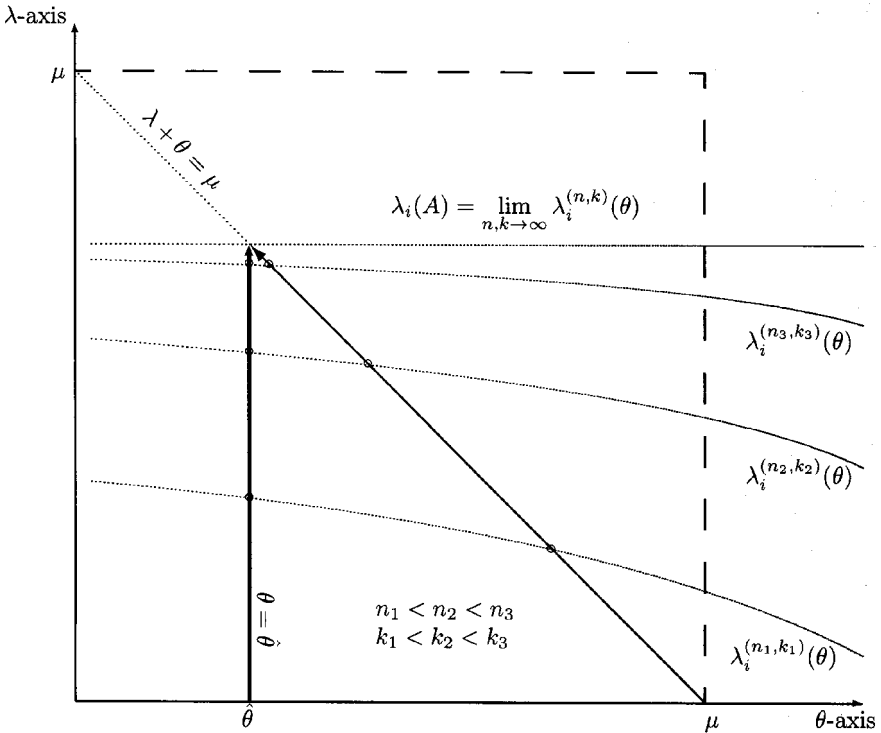


Figure 1. Pattern of convergence of standard intermediate problems  $A_{n,k}^{\hat{\theta}}$  versus the EVF method of [6]. The ordinates of circled points correspond to calculated quantities of three instances of each method.

**THEOREM 3.3.** *If  $\text{span}\{p_i\}$  is a core for  $T^*$  and  $\text{span}\{q_i\}$  is a core for  $A_0$ , then the EVF method of Beattie and Goerisch produces convergent eigenvalue lower bounds for each eigenvalue  $\lambda_i(A) < \mu < \lambda_\infty(A_0)$ .*

*Proof.* Suppose there are  $j$  eigenvalues of  $A$  strictly below  $\mu$ . Pick  $\delta$  such that  $0 < \delta < \mu - \lambda_j$ . We begin by showing pointwise convergence of the eigenvalue curves  $\{\lambda_i^{(n,k)}(\theta)\}_{n,k}$  for each  $\theta \in [\delta, \mu - \lambda_1(A_0)]$ . This will mean that

$$\lambda_i^{(n,k)}(\mu - \lambda_1(A_0)) > \lambda_1(A_0)$$

for  $n$  and  $k$  sufficiently large. This implies an intersection between the eigenvalue curve  $\lambda_i^{(n,k)}(\theta)$  and the line  $\lambda + \theta = \mu$ , i.e. that the EVF method of [6] produces a lower bound to  $\lambda_i$  for  $n$  and  $k$  sufficiently large.

Fix  $\theta \in [\delta, \mu - \lambda_1(A_0)]$  and pick  $\epsilon > 0$ . Define the intermediate operators

$$A_n^\theta = A_0 - \theta + [B + \theta]\hat{P}_n^{(\theta)},$$

where  $\hat{P}_n^{(\theta)}$  is a  $[B + \theta]$ -orthogonal projection onto  $\text{span}\{(B + \theta)^{-1}(A_0 - \mu)q_i\}_{i=1}^n$ . If  $\text{span}\{q_i\}$  is a core for  $A_0$ , then it is relatively straightforward to see that  $\text{span}\{(B + \theta)^{-1}(A_0 - \mu)q_i\}$  is a core for  $[B + \theta]$ . This, in turn, guarantees that



the lower  $j$  eigenvalues of  $A_n^\theta$ , listed as

$$\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \dots \leq \lambda_j^{(n)}$$

(with the fixed  $\theta$  dependence suppressed), converge to the associated eigenvalues of  $A$  as  $n \rightarrow \infty$  [7,9]. Pick  $N$  sufficiently large so that  $n \geq N$  implies  $0 \leq \lambda_i - \lambda_i^{(n)} \leq \frac{1}{2}\epsilon$  for each  $i \leq j$ .

Since  $(T^*P_kT + \theta)^{-1}$  converges strongly to  $(B + \theta)^{-1}$  by corollary 2.2, there exists  $K$  such that, for  $k \geq K$ ,

$$0 \leq [\mathcal{D}_{ij}^\infty] - [\mathcal{D}_{ij}^{(k,\theta)}] \leq \frac{1}{2}\epsilon [\mathcal{G}_{ij}] \in \mathbb{C}^{N \times N}, \tag{3.4}$$

where  $\mathcal{G}_{ij}$  is the matrix inverse to  $\{[(A_0 - \mu)q_i, (A_0 - \mu)q_j]\}$  and  $[\mathcal{D}_{ij}^\infty]$  is the matrix inverse to  $\{[(B + \theta)^{-1}(A_0 - \mu)q_i, (A_0 - \mu)q_j]\}$ . The inequalities of (3.4) may be used to assert

$$A_{N,K}^\theta \leq A_N^\theta \leq A_{N,K}^\theta + \frac{1}{2}\epsilon \sum_{i,j=1}^N \langle \cdot, (A_0 - \mu)q_i \rangle \mathcal{G}_{ij} (A_0 - \mu)q_j.$$

The final term may be recognized as  $\frac{1}{2}\epsilon$  times an  $\mathcal{H}$ -orthogonal projection onto  $\text{span}\{(A_0 - \mu)q_i\}_{i=1}^N$ , and hence is bounded in norm by  $\frac{1}{2}\epsilon$ . Thus, for  $n \geq N$  and  $k \geq K$ , we have

$$A_N^\theta - \frac{1}{2}\epsilon \leq A_{N,K}^\theta \leq A_{n,k}^\theta,$$

which implies that  $\lambda_i^{(N)} - \frac{1}{2}\epsilon \leq \lambda_i^{(n,k)}(\theta)$  for  $n \geq N$  and  $k \geq K$ . Thus we obtain

$$0 \leq \lambda_i - \lambda_i^{(n,k)}(\theta) \leq (\lambda_i - \lambda_i^{(N)}) + (\lambda_i^{(N)} - \lambda_i^{(n,k)}(\theta)) \leq \frac{1}{2}\epsilon + \frac{1}{2}\epsilon = \epsilon$$

for  $n \geq N$  and  $k \geq K$ .

The conclusion will be attained if, in particular, the eigenvalue curve  $\lambda_i^{(n,k)}(\theta)$  converges uniformly to the constant function  $\lambda_i$  on  $[\delta, \mu - \lambda_1(A_0)]$ . Since we have that  $\{\lambda_i^{(n,k)}(\theta)\}_{n,k}$  is a monotone increasing family of continuous functions that converges pointwise on the compact set  $[\delta, \mu - \lambda_1(A_0)]$ , Dini's theorem yields uniform convergence to  $\lambda_i$ . □

The EVF method of Beattie and Goerisch can also be implemented with Aronszajn (outer) projections rather than the inner or Bazley-Fox  $T^*T$  projections employed in [6]. In this case, the previous constructions are modified as follows. Select trial vectors  $\{r_i\}_{i=1}^k \subset \mathcal{D}(B)$  and let  $Q_k$  be the associated  $b$ -orthogonal projection onto  $\text{span}\{r_i\}_{i=1}^k$ , i.e.

$$Q_k u = \sum_{i,j=1}^k \langle u, B r_i \rangle \beta_{ij} r_j,$$

where  $[\beta_{ij}]$  is the Moore-Penrose generalized matrix inverse to  $[b(r_i, r_j)]$ . The corresponding intermediate quadratic form is

$$a_k(u) = a_0(u) + b(Q_k u)$$

for all  $u \in \mathcal{D}(a_k) = \mathcal{D}(a_0)$ , with corresponding self-adjoint operator

$$A_k = A_0 + BQ_k.$$

These intermediate operators yield non-decreasing lower bounds to the eigenvalues of  $A$  below  $\lambda_\infty(A_0)$ , and convergence theory for Aronszajn’s method has been developed in [7–9]. The corresponding WA matrix is

$$\mathcal{W}_k^a(\lambda) = [\langle r_i + R_\lambda^0 B r_i, B r_j \rangle],$$

where, as previously,  $R_\lambda^0$  is the resolvent operator  $(A_0 - \lambda)^{-1}$ . Evaluation is, in general, difficult unless one employs Bazley’s ‘special choice’,  $B r_i = u_i^0$ , a possibly onerous or impossible set of operator equations to solve. To avoid solution of these operator equations, one may again go to the Bazley–Fox method of second projection. For  $\theta > 0$ , the operator  $BQ_k + \theta$  is bounded, self-adjoint, positive-definite and, as a finite-rank perturbation of a multiple of the identity, has an explicitly computable inverse. So choose a set of vectors  $\{\hat{r}_i\}_{i=1}^n$  and let  $\hat{Q}_n^{(k,\theta)}$  be the corresponding projection onto  $\text{span}\{\hat{r}_i\}_{i=1}^n$ , which is orthogonal with respect to the inner product induced by  $BQ_k + \theta$ . Then  $(BQ_k + \theta)\hat{Q}_n^{(k,\theta)}$  is a bounded symmetric positive-semidefinite operator and

$$\langle (BQ_k + \theta)\hat{Q}_n^{(k,\theta)}v, v \rangle \leq \langle (BQ_k + \theta)v, v \rangle$$

for all  $v \in \mathcal{H}$ . Thus the eigenvalues of

$$A_{k,n}^{(a,\theta)} = A_0 - \theta + (BQ_k + \theta)\hat{Q}_n^{(k,\theta)} \tag{3.5}$$

give lower bounds to the corresponding eigenvalues of  $A$ , and they are monotone increasing in both  $k$  and  $n$ . The associated  $n \times n$  WA matrix is given by

$$\mathcal{W}_{k,n}^a(\lambda) = [\langle \hat{r}_i + R_{\lambda+\theta}^0 (BQ_k + \theta)\hat{r}_i, (BQ_k + \theta)\hat{r}_j \rangle]. \tag{3.6}$$

Viewing the intermediate operators (3.5) in the traditional WA sense via the matrix (3.6) is again usually difficult computationally, but one may invoke the EVF procedure of Beattie and Goerisch, which now proceeds as follows. Define  $\mu = \lambda + \theta$  and then consider only those  $\lambda$  for which  $\lambda < \mu$ . Further, introduce the change of variable  $s_i = R_\mu^0 (BQ_k + \theta)\hat{r}_i$  into (3.6) to get

$$\mathcal{W}_{k,n}^a(\lambda) = [\langle (BQ_k + \theta)^{-1}(A_0 - \mu)s_i, (A_0 - \mu)s_j \rangle + \langle s_i(A_0 - \mu)s_j \rangle]. \tag{3.7}$$

Fix  $\mu, \{s_i\}_{i=1}^n \subset \mathcal{D}(A_0)$  and  $\{r_i\}_{i=1}^k \subset \mathcal{D}(B)$ , and (3.7) is the WA matrix for an intermediate problem for each (fixed)  $\lambda < \mu$ . We will refer to this method as *the Beattie–Goerisch EVF method with outer* (i.e. Aronszajn) *second projection*. Convergence of this method follows as previously once the following lemma, whose proof is similar to a lemma of [7], is established.

LEMMA 3.4. *If  $\text{span}\{r_i\}$  is a core for  $B$ , then  $BQ_k$  converges to  $B$  in the strong resolvent sense.*

*Proof.* Pick  $\varepsilon$  with  $0 < \varepsilon < 1$  and define  $b_+(u) = b(u) + \varepsilon\|u\|^2$  for all  $u \in \mathcal{D}(b)$ . Then  $b_+(u)$  is a closed densely defined coercive quadratic form with corresponding self-adjoint operator  $B_+$  satisfying  $b_+(u) = \langle B_+u, u \rangle$  for all  $u \in \mathcal{D}(B_+) = \mathcal{D}(B)$ . Let  $Q_{k,+}$  be the  $b_+$ -orthogonal projection onto  $\text{span}\{r_i\}_{i=1}^k$ . Then, by lemma 3.5 of [7],

$$(1 - \varepsilon)\|u\|^2 + b_+(Q_{k,+}u) \leq \|u\|^2 + b(Q_k u) \leq \|u\|^2 + b(u).$$

So there are corresponding self-adjoint operators  $B_{k,+} = (1 - \varepsilon)I + B_+Q_{k,+}$  satisfying

$$(1 - \varepsilon)I \leq \dots \leq B_{k,+} \leq B_{k+1,+} \leq \dots \leq B + I, \\ I \leq \dots \leq I + BQ_k \leq I + BQ_{k+1} \leq \dots \leq B + I$$

and

$$(1 - \varepsilon)I + B_+Q_{k,+} = B_{k,+} \leq I + BQ_k, \tag{3.8}$$

where  $I$  is the identity operator on  $\mathcal{H}$ . Since  $\{B_{k,+}\}$  and  $\{I + BQ_k\}$  are monotone increasing families of positive-definite self-adjoint operators,  $B_{k,+}$  (respectively,  $I + BQ_k$ ) converges in the strong resolvent sense to a self-adjoint operator  $B_{\infty,+}$  (respectively,  $I + B_{\infty}$ ) satisfying  $B_{k,+} \leq B_{\infty,+} \leq I + B$ ,  $I + BQ_k \leq I + B_{\infty} \leq I + B$  and, from (3.8),  $B_{\infty,+} \leq I + B_{\infty} \leq I + B$ . If we can show that  $B_{\infty,+} = I + B$ , then necessarily  $B_{\infty} = B$ , and we are done. It is thus clearly sufficient to show that if  $b_{\infty,+}$  is the quadratic form corresponding to  $B_{\infty,+}$ , then  $b_{\infty,+} = b + 1$ .

Define now an auxiliary quadratic form

$$b_*(u) = (1 - \varepsilon)\|u\|^2 + \lim_k b_+(Q_{k,+}u)$$

for all  $u$  in

$$\mathcal{D}(b_*) = \left\{ u \in \bigcap_n \mathcal{D}(b_+(Q_{k,+}\cdot)) : \sup_k b_+(Q_{k,+}u) < \infty \right\} \\ = \left\{ u \in \mathcal{H} : \lim_k b_+(Q_{k,+}u) < \infty \right\}.$$

Then  $\mathcal{D}(b_*)$  is a linear subspace of  $\mathcal{H}$  and  $b_*(u)$  is a positive-definite quadratic form on  $\mathcal{D}(b_*)$ . In fact,  $b_*$  is a least upper bound for the family of forms  $\{(1 - \varepsilon)\|\cdot\|^2 + b_+(Q_{k,+}\cdot)\}$  in the sense that any closed real-valued quadratic form that dominates all of these forms also dominates  $b_*$ . Thus, in particular,  $\mathcal{D}(b_*) \supset \mathcal{D}(b)$  and  $b_* \leq b_{\infty,+} \leq b + 1$ . Hence we will show that  $b_{\infty,+} = b + 1$  by showing that  $b_* = b + 1$ .

Since  $\text{span}\{r_i\}$  is a core for  $B$ , it is also a core for  $B_+ = B + \varepsilon I$ , and hence for  $b_+$  as well. Thus  $Q_{k,+}$  converges strongly to the identity in the Hilbert space with norm  $b_+(\cdot)^{1/2}$ ,  $\mathcal{H}_{b_+}$ , which is continuously contained in  $\mathcal{H}$  and, for  $u \in \mathcal{D}(b_+)$ ,

$$b_+(u - Q_{k,+}u) = b_+(u) - b_+(Q_{k,+}u) \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

This implies that  $b_*(u) = b(u) + \|u\|^2$  for each  $u \in \mathcal{D}(b)$ , and we now need only show that  $\mathcal{D}(b_*) \subset \mathcal{D}(b_+) = \mathcal{D}(b)$ .

Take  $u \in \mathcal{D}(b_*)$ . Then  $\{b_+(Q_{k,+}u)\}$  is a Cauchy sequence of real numbers and the Pythagorean theorem implies

$$b_+(Q_{k,+}u - Q_{l,+}u) = |b_+(Q_{k,+}u) - b_+(Q_{l,+}u)| \rightarrow 0 \quad \text{as } k, l \rightarrow \infty.$$

Since  $b_+$  is coercive in each of  $\mathcal{H}_{b_+}$  and  $\mathcal{H}$ ,  $Q_{k,+}u$  converges in each of  $\mathcal{H}_{b_+}$  and  $\mathcal{H}$  to a (single) vector,  $w$  say. Additionally,

$$b_+(Q_{k,+}u - Q_{l,+}u) = b(Q_{k,+}u - Q_{l,+}u) + \varepsilon\|Q_{k,+}u - Q_{l,+}u\|^2,$$

and since  $b$  is a closed non-negative quadratic form in  $\mathcal{H}$ , we may conclude that  $u \in \mathcal{D}(b)$  if  $w = u$ .

For this purpose, let  $r \in \text{span}\{r_i\}_{i=1}^k$  for some index  $k$ . Then

$$\langle u - w, B_+r \rangle = \lim_{l \rightarrow \infty} \langle (I - Q_{l,+})u, B_+r \rangle = \lim_{l \rightarrow \infty} \langle u, B_+(I - Q_{l,+})r \rangle = 0.$$

Since  $\text{span}\{r_i\}$  is a core for  $B_+$ , we must conclude that  $w = u$ . □

**THEOREM 3.5.** *If  $\text{span}\{r_i\}$  is a core for  $B$  and  $\text{span}\{s_i\}$  is a core for  $A_0$ , the EVF method with outer second projection associated with the WA matrix (3.7) produces convergent eigenvalue lower bounds for each eigenvalue  $\lambda_i(A) < \mu < \lambda_\infty(A_0)$ .*

*Proof.* The analogues of lemmas 3.1 and 3.2 with outer second projection are obvious and, given lemma 3.4, the theorem is proved by copying the proof of theorem 3.3. □

We now consider convergence of EVF methods without second projection. Second projection has the distinct advantage that the solution of the operator equations associated with special choice or EVF methods reduces to the inversion of a positive-definite matrix. But when the operator equations are easy to solve, second projection, which lowers the approximating eigenvalues, may be unnecessary to obtain accurate approximations. In this section we retain the positive parameter  $\theta$  introduced for second projection, in order that  $B + \theta$  be boundedly invertible on  $\mathcal{H}$ . It is somewhat simpler here to proceed with outer, rather than inner, projections, so we will do so and note that in the present case the same convergence criteria are obtained.

So let  $\theta > 0$  and again select trial vectors  $\{r_i\}_{i=1}^n \subset \mathcal{D}(B) = \mathcal{D}(B + \theta)$ . Let  $Q_n^{(\theta)}$  be the associated  $(b + \theta)$ -orthogonal projection onto  $\text{span}\{r_i\}_{i=1}^n$ , i.e.

$$Q_n^{(\theta)} = \sum_{i,j=1}^n \langle u, (B + \theta)r_i \rangle \beta_{ij}^{(\theta)} r_j,$$

where  $[\beta_{ij}^{(\theta)}]$  is the matrix inverse to  $[b(r_i, r_j) + \theta \langle r_i, r_j \rangle]$ . The corresponding intermediate quadratic form is

$$a_n^{(\theta)}(u) = a_0(u) - \theta \|u\|^2 + b(Q_n^{(\theta)}u) + \theta \|Q_n^{(\theta)}u\|^2$$

for all  $u \in \mathcal{D}(a_n^{(\theta)}) = \mathcal{D}(a_0)$ , with associated self-adjoint operator

$$A_n^{(\theta)} = A_0 - \theta + (B + \theta)Q_n^{(\theta)}.$$

The corresponding WA matrix is

$$\mathcal{W}_n(\lambda) = [\langle r_i + R_{\lambda+\theta}^0(B + \theta)r_i, (B + \theta)r_j \rangle], \tag{3.9}$$

which is often difficult to compute with. The EVF procedure of Gay [10], for which the eigenvalue-counting problem has been resolved by Goerisch [11] (cf. also [4]), alleviates this as follows. Again, define  $\mu = \lambda + \theta$  and consider only those  $\lambda < \mu$ . Then introduce the change of variable  $s_l = R_\mu^0(B + \theta)r_i$  into (3.9) to get

$$\mathcal{W}_n(\lambda) = [((B + \theta)^{-1}(A_0 - \mu)s_i, (A_0 - \mu)s_j) + \langle s_i, (A_0 - u)s_j \rangle]. \tag{3.10}$$

Fix  $\mu$  and  $\{s_i\}_{i=1}^n \subset \mathcal{D}(A_0)$  and (3.10) is the WA matrix for an intermediate problem for each (fixed)  $\lambda < \mu$ . This is Gay's method with the parameter  $\theta$ , and convergence follows from the previous result.

**THEOREM 3.6.** *If  $\text{span}\{s_i\}$  is a core for  $A_0$ , Gay's EVF method with parameter  $\theta > 0$  produces convergent eigenvalue lower bounds for each eigenvalue  $\lambda_i(A) < \mu < \lambda_\infty(A_0)$ . If  $B$  is positive-definite, the conclusion also holds with  $\theta = 0$ .*

*Proof.* The analogues of lemmas 3.1 and 3.2 follow trivially, and then the proof for  $\theta > 0$  is contained in the proof of theorem 3.3. When  $B$  is positive-definite, it is apparent that this also follows with small negative values of  $\theta$ , which concludes the proof. □

An obvious consequence of theorem 3.6 is what appears to be a new convergence criterion for Temple–Lehmann methods (cf. [16, 17] and also [19]).

**COROLLARY 3.7.** *If  $\text{span}\{s_i\}$  is a core for  $A$ , then the right-definite Temple–Lehmann method produces convergent lower bounds to every eigenvalue  $\lambda_i(A) < \mu < \lambda_\infty(A)$ .*

#### 4. Convergence of eigenvector free methods without a parameter

As originally proposed by Gay [10], the WA matrix (3.9) was transformed into the WA matrix (3.10), with  $\theta = 0$  and  $B$  not necessarily positive-definite. This is clearly possible, provided that the vectors  $s_i$  and  $A_0 s_i$  are in  $\mathcal{R}(B)$ , the range of  $B$ , but convergence criteria are affected when  $B$  is not boundedly invertible. We now study this situation. Our approach requires that eigenvectors of  $A$  must be in  $\mathcal{D}(A_0 + B)$ , which is obviously the case if  $A = A_0 + B$ . When  $B$  is positive-definite and unbounded, the forthcoming convergence criteria may be weaker than that of theorem 3.6, due to the role of  $B^{-1}$  in (3.10).

We now set  $\theta = 0$  in (3.9) and let  $B^-$  be the generalized inverse of  $B$ , i.e.  $BB^- = I$  on  $\mathcal{R}(B)$ ,  $B^- = 0$  on  $\mathcal{R}(B)^\perp$ , and extension by linearity. Then, in (3.10),  $\lambda = \mu$ , and assuming that  $s_i$  and  $A_0 s_i$  are in  $\mathcal{R}(B)$ , equation (3.10) becomes

$$\mathcal{W}_n(\mu) = [\langle B^-(A_0 - \mu)s_i, (A_0 - \mu)s_j \rangle + \langle s_i, (A_0 - \mu)s_j \rangle]. \tag{4.1}$$

Lower bounds for the eigenvalues of  $A$  are then sought from the real zeros (if any) of the equation

$$\det \mathcal{W}_n(\mu) = \det[\langle B^-(A_0 - \mu)s_i, (A_0 - \mu)s_j \rangle + \langle s_i, (A_0 - \mu)s_j \rangle] = 0, \tag{4.2}$$

a so-called quadratic (in  $\lambda = \mu$ ) eigenvalue problem. In [10], Gay calculated accurate approximations to bound states (eigenvalues) of helium, but it remained until the work of Goerisch [11] to definitively relate zeros of (4.2) to eigenvalues of  $A$ —the eigenvalue-counting problem. The connection between zeros of (4.2) and the work of Goerisch is fully explained in §7 of [4]. Beattie writes  $B = T^*T$  in [4] and Goerisch used  $T = B^{1/2}$ , the positive square root of  $B$ , in [11]. The corresponding change in our conclusions is easily obtained and will be given at the end of this section. We note further that if one retains  $\theta > 0$ , i.e. employs the matrix (3.10) in the determinantal equation (4.2), one necessarily obtains real roots  $\mu$  for  $\theta$  large enough.

To obtain a convergence theorem, we first check for exactness of the method. This means setting  $s_i = u_i, i = 1, \dots, n$ , in the matrix (4.1) to see if the method produces  $\lambda_i = \lambda_i(A), i = 1, \dots, n$ , as values of  $\mu$  satisfying (4.2). In order to do this, we must assume that  $u_i$  and  $A_0u_i$  are in  $\mathcal{R}(B), i = 1, \dots, n$ , and in order to use the equations

$$Bu_i = \lambda_i u_i - A_0u_i, \quad u_i = B^-(\lambda_i u_i - A_0u_i), \quad i = 1, \dots, n, \tag{4.3}$$

we must further assume that  $u_i \in \mathcal{D}(A_0 + B), i = 1, \dots, n$ , so that

$$Au_i = (A_0 + B)u_i = \lambda_i u_i.$$

Now setting  $s_i = u_i, i = 1, \dots, n$ , in (4.1) gives the matrix

$$\begin{aligned} & [\{\{B^-(A_0 - \mu) + 1\}u_i, (A_0 - \mu)u_j\}] \\ &= [\{\{B^-(A_0 - \lambda_i + \lambda_i - \mu) + 1\}u_i, (A_0 - \lambda_j + \lambda_j - \mu)u_j\}] \\ &= [\langle(\lambda_i - \mu)B^-u_i, -Bu_j + (\lambda_j - \mu)u_j\rangle] \\ &= [(\lambda_i - \mu)\{\langle\lambda_j - \mu\rangle\langle B^-u_i, u_j\rangle - \langle u_i, u_j\rangle\}]. \end{aligned} \tag{4.4}$$

Since the  $i$ th row of this matrix vanishes for  $\mu = \lambda_i$ , the determinantal equation (4.2) is satisfied for  $\mu = \lambda_i, i = 1, \dots, n$ , and we have exactness. This immediately yields the following convergence criterion, but we reiterate the importance of the eigenvalue-counting theorems of [4, 11] for determining what eigenvalues of  $A$  may be approximated by a real root  $\mu$  of (4.2), since, in particular, there are  $2n$  (complex) roots of (4.2).

LEMMA 4.1. *If  $\{u_i\}_{i=1}^n \subset \mathcal{D}(A_0 + B), \{s_i\}_{i=1}^n$  and  $\{A_0s_i\}_{i=1}^n$  are in  $\mathcal{R}(B), \lambda_n(A) < \lambda_\infty(A_0)$ , and for each  $i = 1, \dots, n, s_i$  approximates  $u_i$  in the norm*

$$[\langle B^-A_0u, A_0u \rangle + \langle B^-u, u \rangle + \langle u, |A_0|u \rangle + \langle u, u \rangle]^{1/2}, \tag{4.5}$$

then there are  $n$  roots of (4.2) that approximate  $\lambda_i(A), i = 1, \dots, n$ . Here,  $|A_0|$  denotes the absolute value of  $A_0$  as defined by use of the spectral theorem.

Lemma 4.1 envisions finding approximations to the eigenvectors corresponding to each of the lower eigenvalues of  $A$  and obtaining lower bounds from (4.2). This does not envision taking  $n$  large in (4.2) to approximate the lowest  $n'$ , say, eigenvalues of  $A$ . The probable source of the approximate eigenvectors would be from a procedure of Rayleigh–Ritz type, e.g. finite-element approximations, for upper bounds. But the energy norm,  $[a(u)]^{1/2}$ , of Rayleigh–Ritz may or may not dominate the norm (4.5). When  $B^-$  is bounded, Rayleigh–Ritz eigenvectors for  $A^2$  approximate eigenvectors of  $A$  in a norm dominating (4.5), since then  $\|Au\|$  dominates  $\|A_0u\|$ , provided that  $A = A_0 + B$ .

To proceed from the approximability hypothesis of lemma 4.1 to a completeness assumption in the norm (4.5), we must supplement the computation (4.4). So let  $n' < n$ , and take  $s_i = u_i$  for  $i = 1, \dots, n'$  in (4.1). Then, if  $n' < m \leq n$ , the corresponding element of the matrix (4.1) is

$$\langle \{B^-(A_0 - \lambda_l + \lambda_l - \mu) + 1\}u_l, (A_0 - \mu)s_m \rangle = \langle (\lambda_l - \mu)B^-u_l, (A_0 - \mu)s_m \rangle, \tag{4.6}$$

and so by (4.4) and (4.6) it is still true that the  $l$ th row of the matrix (4.1) vanishes for  $\mu = \lambda_l$ . This leads to the following theorem.

**THEOREM 4.2.** *If  $\{u_i\}_{i=1}^{n'} \subset \mathcal{D}(A_0 + B)$ ,  $\{u_i\}_{i=1}^{n'}$ ,  $\{A_0 u_i\}_{i=1}^{n'}$ ,  $\{s_i\}$  and  $\{A_0 s_i\}$  are in  $\mathcal{R}(B)$ ,  $\lambda_{n'}(A) < \lambda_\infty(A_0)$  and  $\{s_i\}$  is complete in the norm (4.5), then there are  $n'$  roots of (4.2) that converge to  $\lambda_i(A)$ ,  $i = 1, \dots, n'$ , as  $n \rightarrow \infty$ .*

*Proof.* Since  $\{s_i\}$  is complete in the norm (4.5), for  $n$  sufficiently large, there are linear combinations  $t_i$ ,  $i = 1, \dots, n'$ , of  $\{s_i\}_{i=1}^n$  that approximate each  $u_i$ ,  $i = 1, \dots, n'$ , in the norm (4.5). The similarity transformation that maps  $s_i$  to  $t_i$  for each  $i = 1, \dots, n'$  and leaves  $s_i$  invariant for  $n' < l \leq n$  does not change the roots of the determinantal equation (4.2). The theorem thus follows from lemma 4.1.  $\square$

If  $B$  is written as  $T^*T$  and we define the generalized inverse,  $T^{*-}$ , of  $T^*$  by  $T^*T^{*-} = I$  on  $\mathcal{R}(T^*)$ ,  $T^{*-} = 0$  on  $\mathcal{R}(T^*)^\perp$ , and extension by linearity, it is natural to write (4.2) as

$$\det \mathcal{W}_n(\mu) = \det[\langle T^{*-}(A_0 - \mu)s_i, T^{*-}(A_0 - \mu)s_m \rangle_* + \langle s_i, (A_0 - \mu)s_m \rangle] = 0. \tag{4.7}$$

The analogues of equations (4.3) are

$$B u_i = \lambda_i u_i - A_0 u_i \quad \text{and} \quad T u_i = T^{*-}(\lambda_i u_i - A_0 u_i) \quad \text{for } i = 1, \dots, n'.$$

The following analogue of theorem 4.2 is now obvious.

**THEOREM 4.3.** *If  $\{u_i\}_{i=1}^{n'} \subset \mathcal{D}(A_0 + B)$ ,  $\{u_i\}_{i=1}^{n'}$ ,  $\{A_0 u_i\}_{i=1}^{n'}$ ,  $\{s_i\}$  and  $\{A_0 s_i\}$  are in  $\mathcal{R}(T^*)$ ,  $\lambda_{n'}(A) < \lambda_\infty(A_0)$  and  $\{s_i\}$  is complete in the norm*

$$[\langle T^{*-}A_0 u, T^{*-}A_0 u \rangle + \langle T^{*-}u, T^{*-}u \rangle + \langle u, |A_0|u \rangle + \langle u, u \rangle]^{1/2},$$

*then there are  $n'$  roots of (4.7) that converge to  $\lambda_i(A)$ ,  $i = 1, \dots, n'$ , as  $n \rightarrow \infty$ .*

### References

- 1 N. Aronszajn. Approximation methods for eigenvalues of completely continuous symmetric operators. In *Proc. of the Symp. on Spectral Theory and Differential Problems, Stillwater, OK*, pp. 179–202 (1951).
- 2 N. W. Bazley and D. W. Fox. A procedure for estimating eigenvalues. *J. Math. Phys.* **3** (1962), 469–471.
- 3 N. W. Bazley and D. W. Fox. Lower bounds to eigenvalues using operator decompositions of the form  $B^*B$ . *Arch. Ration. Mech. Analysis* **10** (1962), 352–360.
- 4 C. Beattie. An extension of Aronszajn’s rule: slicing the spectrum for intermediate problems. *SIAM J. Numer. Analysis* **24** (1987), 828–843.
- 5 C. Beattie. Some convergence results for intermediate problems that displace essential spectra. Applied Physics Laboratory Report No. 65, Johns Hopkins University, Laurel, MD (1982)
- 6 C. Beattie and F. Goerisch. Methods for computing lower bounds to eigenvalues of self-adjoint operators. *Numer. Math.* **72** (1995), 143–172.
- 7 C. Beattie and W. M. Greenlee. Convergence theorems for intermediate problems. *Proc. R. Soc. Edinb.* A **100** (1985), 107–122.
- 8 C. Beattie and W. M. Greenlee. Some remarks concerning closure rates for Aronszajn’s method. In *Numerische Behandlung von Eigenwertaufgaben* (ed. J. Albrecht, L. Collatz, W. Veltte and W. Wunderlich), vol. 5, International Series of Numerical Mathematics 96 (Basel: Birkhäuser, 1991).

- 9 R. D. Brown. Convergence criterion for Aronszajn's method and for the Bazley–Fox method. *Proc. R. Soc. Edinb.* A **108** (1988), 91–108.
- 10 J. G. Gay. A lower bound procedure for energy eigenvalues. *Phys. Rev. A* **135** (1964), A-1220–A-1226.
- 11 F. Goerisch. Eine von Eigenvektoren freie Fassung eines Verfahrens von Bazley. In *Numerische Behandlung von Eigenwertaufgaben* (ed. J. Albrecht and L. Collatz), vol. 2, International Series of Numerical Mathematics 43 (Basel: Birkhäuser, 1979).
- 12 W. M. Greenlee. A convergent variational method of eigenvalue approximation. *Arch. Ration. Mech. Analysis* **81** (1983), 279–287.
- 13 T. Kato. *Perturbation theory for linear operators* (Springer, 1976).
- 14 G. Lee. On lower bounds of eigenvalues for self-adjoint operators. *J. Korean Math. Soc.* **31** (1994), 477–492.
- 15 G. Lee. Convergence rate for lower bounds to self-adjoint operators. *J. Korean Math. Soc.* **33** (1996), 513–525.
- 16 N. J. Lehmann. Beiträge zur numerischem Lösung linearer Eigenwertprobleme. I. *Z. Angew. Math. Mech.* **29** (1949), 341–365.
- 17 N. J. Lehmann. Beiträge zur numerischem Lösung linearer Eigenwertprobleme. II. *Z. Angew. Math. Mech.* **30** (1950), 1–16.
- 18 J. Weidmann. Monotone continuity of the spectral resolution and the eigenvalues. *Proc. R. Soc. Edinb.* A **85** (1980), 131–136.
- 19 H. F. Weinberger. *Variational methods for eigenvalue problems* (Philadelphia, PA: SIAM, 1974).
- 20 A. Weinstein and W. Stenger. *Methods of intermediate problems for eigenvalues* (Academic, 1972).

(Issued 18 October 2002)