CENTER FOR PURE AND APPLIED MATHEMATICS
UNIVERSITY OF CALIFORNIA, BERKELEY

PAM- 517

REFINED INTERLACING PROPERTIES

R. O. Hill Jr. and B. N. Parlett

Dept. of Mathematics
Michigan State University
East Lansing, MI 48824

Dept. of Mathematics
and Computer Science Division
University of California
Berkeley, CA 94720

January 1991

This report was done with support from the Center for Pure and Applied Mathematics. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of the Center for Pure and Applied Mathematics or the Department of Mathematics.

REFINED INTERLACING PROPERTIES

R. O. Hill, Jr.* and B. N. Parlett†

Abstract: Interlacing results for eigenvalues due to Cauchy, Golub, and Kahan are extended and related to the last component of eigenvectors.

Key words: Eigenvalues, eigenvectors, interlacing.

AMS (MOS) subject classification:

15A18 primary,
15A42 secondary.

Abbreviated title: Refined Interlacing Properties

*Department of Mathematics, Michigan State University, East Lansing, MI 48824

†Mathematics Department and Computer Science Division, University of California at Berkeley, Berkeley, CA 94720.
This author gratefully acknowledges partial support from ONR contract N00014-90-J-1372.

Section 1, Background and Definitions As for back as 1821, in the Cours d'Analyze of the École Polytechnique, Augustin Cauchy published a proof of the following remarkable result. If any row, together with its matching column, is deleted from a real symmetric matrix, then the eigenvalues of the new matrix interlace the eigenvalues of the old one. In the presence of more information, much more can be said about the interlacing of eigenvalues and the relationship between the spacing and the corresponding eigenvectors.

An example of such results can be found in a 1972 paper by G. H. Golub [2] which discusses aspects of the Lanczos algorithm. Golub, seeking bounds for eigenvalues, constructs a special rank one update H to an $n \times n$ symmetric, tridiagonal matrix T_n. Letting $\tilde{T}_k$ be the leading principal $(k+1) \times (k+1)$ submatrix of H, Golub shows each interval determined by the eigenvalues of $\tilde{T}_k$ contains an eigenvalue of T_n. Our Theorem 1 extends this by replacing H with T_n itself.

In order to pursue these lines of investigation, some notational conventions will prove helpful. Let A_k denote the leading principal $k \times k$ submatrix of a real symmetric matrix $A_n = A$. The ordered eigenvalues of A_k are denoted by

$$\lambda_1^{(k)} \le \lambda_2^{(k)} \le \cdots \le \lambda_k^{(k)}$$

For each k we assume we have k orthonormal eigenvectors $z_i^{(k)}$ associated with $\lambda_i^{(k)}$ (where of course we are using the usual inner product on $\mathbb{R}^n$). When necessary, the jth entry of $z_i^{(k)}$ is denoted by $z_i^{(k)}(j)$. Frequently, we shall be concerned only with the magnitude of the last entry $z_i^{(k)}(k)$ and, for simplicity, we shall denote this by the last letter of the Greek alphabet

$$\omega_{ik} = |z_i^{(k)}(k)|$$

Denote by $p_k(x) = det(xI - A_k)$ the characteristic polynomial of A_k and let $p(x) = p_n(x)$.

Section 2 We now consider properties of the whole triangular set $\{\lambda_i^{(k)}\}$

$$\begin{array}{ccccccc}
 & & & \lambda_1^{(1)} & & & \\
 & & \lambda_1^{(2)} & & \lambda_2^{(2)} & & \\
 & \lambda_1^{(3)} & & \lambda_2^{(3)} & & \lambda_3^{(3)} & \\
\lambda_1^{(4)} & & \lambda_2^{(4)} & & \lambda_3^{(4)} & & \lambda_4^{(4)} \\
\cdot & & \cdot & & \cdot & & \cdot \quad \cdot
\end{array}$$

By Cauchy's Theorem, we know $\lambda_i^{(k} \leq \lambda_i^{(k-1)} \leq \lambda_{i+1}^{(k)}$, as indicated by the spacing in the table. The first question likely to arise after looking at this table is:

Does $[\lambda_1^{(2)}, \lambda_2^{(2)}]$ contain an eigenvalue of A_4?

An affirmative answer does not follow from Cauchy's theorem, and, in fact, the answer is *no* in general. However, the answer is *yes* when A is tridiagonal. Furthermore, the open interval $(\lambda_1^{(2)}, \lambda_2^{(2)})$ must contain an eigenvalue of A_4 if it is an unreduced tridiagonal matrix. (Recall, a tridiagonal matrix

$$T_k = \begin{bmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & & & \\ & & \ddots & & \\ & & & a_{k-1} & b_{k-1} \\ & & & b_{k-1} & a_k \end{bmatrix}$$

is called *unreduced* if all the b_i's are nonzero.) What is more this little known result can be proved by elementary means eminently suitable for introductory courses in linear algebra. We give brief, introductory proofs now, because we shall establish a more general result by other means later. Here is the general statement, which is simplified by letting $\lambda_0^{(k)} = -\infty$ and $\lambda_{k+1}^{(k)} = \infty$.

Theorem 1 Let T be an $n \times n$ real symmetric unreduced tridiagonal matrix with T_k its leading principal $k \times k$ submatrix. Then every interval $(\lambda_i^{(k)}, \lambda_{i+1}^{(k)})$ contains a distinct eigenvalue of T, $i = 0, 1, \ldots, k$.

Note Theorem 1 is true even if T is reduced provided that we replace the open intervals by closed intervals. This follows by using just a little care after breaking T up into tridiagonal blocks at each zero b_i. When $n = k + 1$ this is just Cauchy's Theorem; we shall need later the fact that Cauchy's Theorem yields strict inequalities for unreduced tridiagonal matrices.

Note Suppose A is a $n \times n$ symmetric matrix and T_k is a partial tridiagonalization of A (obtained perhaps by the Lanczos algorithm). Then each interval determined by the eigenvalues of T_k contains an eigenvalue of A, since if we complete the tridiagonalization we obtain a T_n similar to A.

Before proving Theorem 1 we consider the special case $n = k + 2$ given in Theorem 2, below. For this case, the proof is very apparent, it illustrates the general case, and more detailed information is obtained that not only has other interesting implications but also hints at the results to follow in the next section. In the statement of Theorem 2 we use the convention $(a, b) = (a)$ if $a = b$. Also, from before, a_{k+2} is the lowest diagonal entry of T_{k+2}, and it must lie in some interval $[\lambda_i^{(k)}, \lambda_{i+1}^{(k)}]$.

Theorem 2 Let T_{k+2} be unreduced and assume a_{k+2} lies in $[\lambda_i^{(k)}, \lambda_{i+1}^{(k)}]$.

(a) If $\lambda_i^{(k)} \le a_{k+2} \le \lambda_{i+1}^{(k+1)}$, then each of the $k + 2$ intervals

$$(-\infty,\ \lambda_1^{(k+1)}),\ (\lambda_1^{(k)},\ \lambda_2^{(k+1)}),\ \ldots,\ (\lambda_i^{(k)},\ a_{k+2}),\ (\lambda_{i+1}^{(k+1)},\ \lambda_{i+1}^{(k)}),\ \ldots$$
$$(\lambda_k^{(k+1)},\ \lambda_k^{(k)}),\ (\lambda_{k+1}^{(k+1)},\ \infty)$$

contains one of the $k + 2$ eigenvalues of T_{k+2}.

(b) If $\lambda_{i+1}^{(k+1)} \le a_{k+2} \le \lambda_{i+1}^{(k)}$ then (a) holds when the middle two intervals are replaced by $(\lambda_i^{(k)},\ \lambda_{i+1}^{(k+1)})$, $(a_{k+2},\ \lambda_{i+1}^{(k)})$.

These relationships are illustrated in Figure 1.

Figure 1. A possible graph of $y = p_{k+2}(x)$

Proof of Theorem 2 Assume that $a_{k+2} \le \lambda_{i-1}^{(k+1)}$ (and the proof in the case $a_{k+2} \ge \lambda_{i+1}^{(k+}$ is similar. By Cauchy's Interlace Theorem, we know p_{k+2} has zeros in the ntervals $(-\infty,\ \lambda_1^{(k+1)})$ and $(\lambda_{k+1}^{(k+1)},\ \infty)$. We shall prove one of the three remaining cases and leave the others as on exercise.

Pick an arbitrary j, $1 \le j < i$ and consider $(\lambda_j^{(k)},\ \lambda_{j+1}^{(k+1)})$. The usual expansion of $\det(xI - T_{k+2})$ along its bottom row gives the well-known recurrence relationship

$$(1) \qquad p_{k+2}(x) = (x - a_{k+2})p_{k+1}(x) - b_{k+1}^2 p_k(x)$$

We know $p_k(\lambda_j^{(k)}) = 0$ and $p_{k+1}(\lambda_{j+1}^{(k+1)}) = 0$. Using interlacing and the factorizations

$$p_k(x) = (x - \lambda_1^{(k)})\cdots(x - \lambda_j^{(k)})(x - \lambda_{j+1}^{(k)})\cdots(x - \lambda_k^{(k)})$$
$$p_{k+1}(x) = (x - \lambda_1^{(k+1)})\cdots(x - \lambda_j^{(k+1)})(x - \lambda_{j+1}^{(k+1)})\cdots(x - \lambda_{k+1}^{(k+1)})$$

we see $sgn\ p_k(\lambda_{j+1}^{(k+1)}) = (-1)^{k-j}$ and $sgn\ p_{k+1}(\lambda_j^{(k)}) = (-1)^{k-j+1}$. Since $\lambda_j^{(k)} < \lambda_i^{(k)} \le a_{k+2}$, $\lambda_j^{(k)} - a_{k+2} < 0$. Putting this altogether we have

$$sgn\ p_{k+2}(\lambda_{j+1}^{(k+1)}) = sgn[0 - b_{k+1}^2 p_k(\lambda_{j+1}^{(k+1)})] = -(-1)^{k-j} = (-1)^{k-j+1}$$
$$sgn\ p_{k+2}(\lambda_j^{(k)}) = sgn[(\lambda_j^{(k)} - a_{k+2})p_{k+1}(\lambda_j^{(k)}) - 0] = -(-1)^{k-j+1} = (-1)^{k-j}$$

Thus the polynomial p_{k+2} has a zero in $(\lambda_j^{(k)}, \lambda_{j+1}^{(k+1)})$. ∎

Before proving Theorem 1, there are two comments to make. First, Theorem 2 shows that though $\lambda_{j+1}^{(k+2)}$ is in the interval $(\lambda_j^{(k+1)}, \lambda_{j+1}^{(k+1)})$, it will tend to be closer to whichever endpoint is further from a_{k+2}; how close will be discussed further in Section 3. Second, Theorem 2 shows that each interval $(\lambda_j^{(k)}, \lambda_{j+1}^{(k)})$ contains a unique eigenvalue of T_{k+2} except for the interval that contains a_{k+2}. From this, if we knew bounds of all the a_j, $j \geq k + 2$, then we would know that intervals $(\lambda_j^{(k)}, \lambda_{j+1}^{(k)})$ outside those bounds still contained unique eigenvalues of T.

Proof of Theorem 1. Let $n \geq k + 3$. Let T_m^* be the trailing principal $(n-m+1) \times (n-m+1)$ submatrix of T, which is given by

$$T_m^* = \begin{bmatrix} a_m & b_m & & & \\ b_m & a_{m+1} & & & \\ & & \ddots & & \\ & & & a_{n-1} & b_{n-1} \\ & & & b_{n-1} & a_n \end{bmatrix}$$

and let p_m^* be its characteristic polynomial. Then

$$T = \left[\begin{array}{c|cccc|c} T_k & & & & & \\ \hline & & b_k & & & \\ & b_k & a_{k+1} & b_{k+1} & & \\ & & b_{k+1} & a_{k+2} & b_{k+2} & \\ & & & b_{k+2} & a_{k+3} & b_{k+3} \\ \hline & & & & b_{k+3} & T_{k+4}^* \end{array}\right]$$

Now expand $p(x) = det(xI - T)$ along the $(k + 2)$nd row to obtain (after a little algebra)

$$p(x) = p_{k+2}^*(x)p_{k+1}(x) - b_{k+1}^2 p_{k+3}^*(x)p_k(x)$$

The proof now proceeds in a manner very similar to the proof of Theorem 2. In each interval $(\lambda_j^{(k)}, \lambda_{j+1}^{(k)})$ there is at least one zero of p in either $(\lambda_j^{(k)}, \lambda_{j+1}^{(k+1)})$ or $(\lambda_{j+1}^{(k+1)}, \lambda_{j+1}^{(k)})$, depending on the relative position of $(\lambda_j^{(k)}, \lambda_{j+1}^{(k)})$ to the zeros of p_{k+2}^* and p_{k+3}^*. In this, you must use the fact that the zeros of p_{k+2}^* and p_{k+3}^* interlace each other, also (again by Cauchy's Interlace Theorem). The details are left as an exercise. ■

Section 3 If all that we wanted to show was that each closed interval $[\lambda_j^{(k)}, \lambda_{j+1}^{(k)}]$ contains a distinct eigenvalue of A, then the tridiagonal assumption in Theorem 1 is not necessary, as we shall show in Section 4. However it is necessary for our next result, which concerns what we call *crowded interlacing,* i. e., when the next eigenvalue along is very close to one of the eigenvalues between which it is interlaced,

$$\min(|\lambda_i^{(k)} - \lambda_i^{(k-1)}|, |\lambda_{i+1}^{(k)} - \lambda_i^{(k-1)}|) \ll |\lambda_{i+1}^{(k)} - \lambda_i^{(k)}|$$

In the context of the Lanczos algorithm it is important to understand when an eigenvalue of T_{k+1} is almost on top of an eigenvalue of T_k. This phenomenon is controlled by the numbers ω_{ik} introduced earlier. This is surprising but not without precedent. In studies of the inverse eigenvalue problem it was discovered independently by most researchers that an unreduced tridiagonal T_k is completely determined (to within ± signs) by the eigenvalues $(\lambda_i^{(n)})_1^n$ and the "weights" $(\omega_{in}^2)_1^n$. There are only $2n - 1$ free parameters here, since $\sum \omega_{in}^2 = 1$. Among physicists the parameters $(\lambda_i^{(n)}, \omega_{in})$ are called *action-angle variables*.

The following result implies, among other things, that

$$\omega_{1n}^2(\lambda_2^{(n)} - \lambda_1^{(n)}) < \lambda_1^{(n-1)} - \lambda_1^{(n)} < \omega_{1n}^2(\lambda_n^{(n)} - \lambda_1^{(n)})$$

Consequently the smaller is ω_{1n} the closer are $\lambda_1^{(n-1)}$ and $\lambda_1^{(n)}$.

Theorem 3 If T_n is an unreduced $n \times n$ tridiagonal matrix and

$1 \le k \le n$, then

$$\left.\begin{array}{c} \dfrac{\lambda_1^{(k-1)} - \lambda_1^{(k)}}{\lambda_k^{(k)} - \lambda_1^{(k)}} \\[2ex] \dfrac{\lambda_i^{(k)} - \lambda_{i-1}^{(k-1)}}{\lambda_i^{(k)} - \lambda_1^{(k)}} \dfrac{\lambda_i^{(k-1)} - \lambda_i^{(k)}}{\lambda_k^{(k)} - \lambda_i^{(k)}} \\[2ex] \dfrac{\lambda_k^{(k)} - \lambda_{k-1}^{(k-1)}}{\lambda_k^{(k)} - \lambda_1^{(k)}} \end{array}\right\} < \omega_{ik}^2$$

$$< \begin{cases} \dfrac{\lambda_1^{(k-1)} - \lambda_1^{(k)}}{\lambda_2^{(k)} - \lambda_1^{(k)}}, & i = 1 \\[2ex] \dfrac{\lambda_i^{(k)} - \lambda_{i-1}^{(k-1)}}{\lambda_i^{(k)} - \lambda_{i-1}^{(k)}} \dfrac{\lambda_i^{(k-1)} - \lambda_i^{(k)}}{\lambda_{i+1}^{(k)} - \lambda_i^{(k)}}, & i \ne 1, k \\[2ex] \dfrac{\lambda_k^{(k)} - \lambda_{k-1}^{(k-1)}}{\lambda_k^{(k)} - \lambda_{k-1}^{(k)}}, & i = k \end{cases}$$

Proof It is known that the tridiagonal T_k can be reconstructed from the $2k - 1$ values $(\lambda_i^{(k)})_1^k$, $(\lambda_i^{(k-1)})_1^{k-1}$. The expression for ω_{ik} is remarkably simple. As before, let $p_k(x) = \det(xI - T_k)$ be the characteristic polynomial of T_k. Then

$$\omega_{ik}^2 = -p_{k-1}(\lambda_i^{(k)})/p_k'(\lambda_i^{(k)})$$

A proof of this may be found in [4, p. 129]. It remains only to reorganize the expression above as a product of quotients:

$$
\omega_{ik}^{2} = \begin{cases} \displaystyle\prod_{j=1}^{k-1} \frac{\lambda_j^{(k-1)} - \lambda_1^{(k)}}{\lambda_{j+1}^{(k)} - \lambda_1^{(k)}}, & i = 1 \\ \displaystyle\prod_{j=1}^{i-1} \frac{\lambda_i^{(k)} - \lambda_j^{(k-1)}}{\lambda_i^{(k)} - \lambda_j^{(k)}} \prod_{j=i}^{k-1} \frac{\lambda_j^{(k-1)} - \lambda_i^{(k)}}{\lambda_{j+1}^{(k)} - \lambda_i^{(k)}}, & i \neq 1,\, k \\ \displaystyle\prod_{j=1}^{k-1} \frac{\lambda_k^{(k)} - \lambda_j^{(k-1)}}{\lambda_k^{(k)} - \lambda_j^{(k)}}, & i = k \end{cases}
$$

Since T_k is unreduced, Cauchy's Theorem yields $\lambda_j^{(k)} < \lambda_j^{(k-1)} < \lambda_{j+1}^{(k)}$, $j = 1, \ldots, k - 1$, so each factor in the above products is positive and less than one. To obtain the second inequality in the theorem, simply discard all factors in each $\prod$ term except for the smallest one.

In order to obtain the first inequality in the theorem, the products given above must be rearranged as follows:

$$
\prod_{j=1}^{k-1} \frac{\lambda_j^{(k-1)} - \lambda_1^{(k)}}{\lambda_{j+1}^{(k)} - \lambda_1^{(k)}} = \frac{\lambda_1^{(k-1)} - \lambda_1^{(k)}}{1} \left(\prod_{j=2}^{k-1} \frac{\lambda_j^{(k-1)} - \lambda_1^{(k)}}{\lambda_j^{(k)} - \lambda_1^{(k)}} \right) \frac{1}{\lambda_k^{(k)} - \lambda_1^{(k)}}
$$

$$
\prod_{j=1}^{i-1} \frac{\lambda_i^{(k)} - \lambda_j^{(k-1)}}{\lambda_i^{(k)} - \lambda_j^{(k)}} = \frac{1}{\lambda_i^{(k)} - \lambda_1^{(k)}} \left(\prod_{j=1}^{i-2} \frac{\lambda_i^{(k)} - \lambda_j^{(k-1)}}{\lambda_i^{(k)} - \lambda_{j+1}^{(k)}} \right) \frac{\lambda_i^{(k)} - \lambda_{i-1}^{(k)}}{1}
$$

$$
\prod_{j=i}^{k-1} \frac{\lambda_j^{(k-1)} - \lambda_i^{(k)}}{\lambda_{j+1}^{(k)} - \lambda_i^{(k)}} = \frac{\lambda_i^{(k-1)} - \lambda_i^{(k)}}{1} \left(\prod_{j=i+1}^{k-1} \frac{\lambda_j^{(k-1)} - \lambda_i^{(k)}}{\lambda_j^{(k)} - \lambda_i^{(k)}} \right) \frac{1}{\lambda_k^{(k)} - \lambda_i^{(k)}}
$$

$$
\prod_{j=1}^{k-1} \frac{\lambda_k^{(k)} - \lambda_j^{(k-1)}}{\lambda_k^{(k)} - \lambda_j^{(k)}} = \frac{1}{\lambda_k^{(k)} - \lambda_1^{(k)}} \left(\prod_{j=1}^{k-2} \frac{\lambda_k^{(k)} - \lambda_j^{(k-1)}}{\lambda_k^{(k)} - \lambda_{j+1}^{(k)}} \right) \frac{\lambda_k^{(k)} - \lambda_{k-1}^{(k+1)}}{1}
$$

Cauchy's Theorem shows that every quotient in the four products in parentheses shown above exceeds one. Deleting them yields the first inequality, as claimed. ■

Next we return to full symmetric matrices.

Section 4 We now extend the results of Theorem 1 beyond the tridiagonal case. The eigenvalues of the $k \times k$ matrix A_k define $k + 1$ intervals $(-\infty, \lambda_1^{(k)}]$, $[\lambda_1^{(k)}, \lambda_2^{(k)}]$, ..., $[\lambda_k^{(k)}, \infty)$, and we want to know when every principal supermatrix of A_k contains at least one eigenvalue in each of these intervals. We do not demand that these intervals be distinct. Recall our earlier convention that $\lambda_0^{(k)} = -\infty$ and $\lambda_{k+1}^{(k)} = \infty$.

Theorem 4 Let $A_n = \begin{bmatrix} A_k & C^t \\ C & U \end{bmatrix}$ be a symmetric partition of A_n. If $rank(C) = 1$, then each interval $[\lambda_i^{(k)}, \lambda_{i+1}^{(k)}]$, $i = 0, 1, \ldots, k$, contains at least one eigenvalue of A_n.

Proof Let $C = wv^t$, $w \in \mathbb{R}^{n-k}$, $v \in \mathbb{R}^k$. The case of the external intervals follows from Cauchy's Interlace Theorem, so pick an $i = 1, \ldots, k - 1$.

If $\lambda_i^{(k)} = \lambda_{i+1}^{(k)}$, find a unit eigenvector z of A_k associated with $\lambda_i^{(k)}$ and satisfying $v^t z = 0$. Now let $x = \begin{bmatrix} z \\ 0 \end{bmatrix}$ and it is trivial to verify that x is an eigenvector of A_n corresponding to $\lambda_i^{(k)}$.

Suppose $\lambda_i^{(k)} \neq \lambda_{i+1}^{(k)}$. Let $z = z_i^{(k)} \cos\psi + z_{i+1}^{(k)} \sin\psi$, where ψ will be determined later, let $x = \begin{bmatrix} z \\ 0 \end{bmatrix}$, and let $\gamma = \left(\lambda_i^{(k)} + \lambda_{i+1}^{(k)}\right)/2$. Then

$$(A - \gamma I)x = \begin{bmatrix} (\lambda_i^{(k)} - \gamma)\, z_i^{(k)} \cos\psi + (\lambda_{i+1}^{(k)} - \gamma)\, z_{i+1}^{(k)} \sin\psi \\ wv^t(z_i^{(k)} \cos\psi + z_{i+1}^{(k)} \sin\psi) \end{bmatrix}$$

Now take any ψ such that $v^t(z_i^{(k)} \cos\psi + z_{i+1}^{(k)} \sin\psi) = 0$. For such a ψ,

$$\|(A - \gamma I)x\|^2 = (\lambda_i^{(k)} - \gamma)^2 \cos^2\psi + (\lambda_{i+1}^{(k)} - \gamma)^2 \sin^2\psi$$
$$= \left(\frac{\lambda_{i+1}^{(k)} - \lambda_i^{(k)}}{2} \right)^2$$

It is a standard result (see [3, Th. 4-5-1]) that if $\|x\| = 1$ and $\|Ax - \gamma x\| = \delta$, then there is an eigenvalue of A in $[\gamma - \delta, \gamma + \delta]$. If we apply this to the above where $\delta = \left(\lambda_i^{(k)} - \lambda_{i+1}^{(k)}\right)/2$, we easily see that there is an eigenvalue of A in $[\gamma - \delta, \gamma + \delta] = [\lambda_i^{(k)}, \lambda_{i+1}^{(k)}]$, and we are done. ∎

In general it can be shown that if $rank(C) = r$, then each union of r abutting subintervals defined by the $\lambda_i^{(k)}$'s holds at least one eigenvalue of A_n. As soon as r exceeds k, the result becomes vacuous.

It is worth pointing out that Theorems 1 and 3 can both be obtained as special cases of Lehmann's optimal intervals. Lehmann's results [3] were published in the 1960's in German and are complicated by the use of an additional parameter. He assumes that A_k and C are known, but U is not, and then finds, for each $\xi \in \mathbb{R}$, the smallest interval centered at ξ that contains exactly j eigenvalues of A_n. The answer turns out to be that the radius [illegible]$_j$ of the smallest interval is the jth smallest singular value of $\begin{bmatrix} A_k - \xi I_k \\ C \end{bmatrix}$. It is not hard to see that if C has rank one and $\xi = \left(\lambda_i^{(k)} + \lambda_{i+1}^{(k)}\right)/2$, then $\sigma_{min}\begin{bmatrix} A_k - \xi I_k \\ C \end{bmatrix} = \left(\lambda_i^{(k)} - \lambda_{i+1}^{(k)}\right)/2$ as expected.

Section 5 An alternative approach to Lehmann's work was taken by W. Kahan. He proved the following refined interlace theorem. See [4, pp. 194-7].

Assume $A = A_n$ has the form

$$A = \begin{bmatrix} H & C^t & O^t \\ C & V & Z^t \\ O & Z & W \end{bmatrix}$$

where H in $m \times m$ and V is $k \times k$ (and O is the zero matrix). In applications, H and C are known but probably V, Z, and W are not. Being ignorant of V we replace it with a $k \times k$ X which we are free to choose and define an auxiliary matrix $M = M(X) = \begin{bmatrix} H & C^t \\ C & X \end{bmatrix}$. Denote the eigenvalues of M by $\mu_i = \mu_i(X)$ and assume $\mu_1 \le \cdots \le \mu_{m+k}$.

Theorem 5 (Kahan) Assume A, M, and X are given as above where X is any $k \times k$ matrix satisfying

$$V - X \text{ is invertible}$$

Then for each index $j = 1, \ldots, m$, the interval $[\mu_j, \mu_{j+k}]$ contains a different eigenvalue λ_i of A. In addition, for each index $i = 1, \ldots, k$, there is a different eigenvalue λ_i of A outside of the open interval (μ_i, μ_{i+m}).

The only blemish in this result is the unverifiable assumption that $V - X$ is invertible. Our final contribution is to remove this hypothesis by looking carefully at the general case when $V - X$ is singular.

Theorem 6 Kahan's Interlacing Theorem (Theorem 5) remains true if the hypothesis "$V - X$ is invertible" is removed.

Proof Let A, X, and M be as above, except assume that $V - X$ is singular. Let N = *Null Space* $(V - X)$, so that $N \oplus N^{\perp} = \mathbb{R}^k$. Picking orthonormal bases for N and $N^{\perp}$, we can change A and M so that $V - X = \begin{bmatrix} O & O \\ O & Y \end{bmatrix}$ where Y is invertible. Thus,

$$V = \begin{bmatrix} V_{11} & V_{21}^t \\ V_{21} & V_{22} \end{bmatrix} = X + \begin{bmatrix} O & O \\ O & Y \end{bmatrix}, \quad \text{so } X = \begin{bmatrix} V_{11} & V_{21}^t \\ V_{21} & X_{22} \end{bmatrix}, \; V_{22} = X_{22} + Y$$

Break up $Z = [\, Z_1 \,\vdots\, Z_2 \,]$, compatibly. Then for each $\varepsilon > 0$, let W_ε^* $= Z_2 Y^{-1} Z_2^t + \frac{1}{\varepsilon} Z_1 Z_1^t$ and obtain

$$A = \begin{bmatrix} H & C^t & & O \\ C & V_{11} & V_{21}^t & Z_1^t \\ & V_{21} & V_{22} & Z_2^t \\ O & Z_1 & Z_2 & W \end{bmatrix}$$

$$= \begin{bmatrix} H & C^t & & O \\ C & V_{11}-\varepsilon I & V_{21}^t & O \\ & V_{21} & X_{22} & O \\ O & O & & W-W_\varepsilon^* \end{bmatrix} + \begin{bmatrix} H & C^t & & O \\ C & \varepsilon I & O & Z_1^t \\ & O & Y & Z_2^t \\ O & Z_1 & Z_2 & W_\varepsilon^* \end{bmatrix}$$

$$= T + U, \; T = \begin{bmatrix} M_\varepsilon & \\ & W_\varepsilon \end{bmatrix}$$

Hence if we let $X_\varepsilon = \begin{bmatrix} V_{11}-\varepsilon I & V_{21}^t \\ V_{21} & X_{22} \end{bmatrix}$, we can now easily see that $V - X_\varepsilon$ is invertible so that Kahan's Interlacing Theorem applies to A, M_ε, and X_ε. If we carefully let $\varepsilon \longrightarrow 0$, then $M_\varepsilon \longrightarrow M$ (although $W_\varepsilon \not\rightarrow W$) so that the eigenvalues of M_ε go to the eigenvalues of M. Hence the conclusion follows, since A has only a finite number of eigenvalues. ∎

Bibliography

[1] G. H. Golub, *Some Uses of the Lanczos Algorithm in Numerical Linear Algebra* in *Topics in Numerical Analysis*, ed. J. J. H. Miller, 1973, pp. 173-184.

[2] Y. Ikebe, T. Inagake, and S. Miyamoto, The Monotonicity Theorem, Cauchy's Interlace Theorem, and the Courant-Fischer

Theorem, *The Am. Math. Monthly*, 94 (1987) pp. 352-354.

[3] N. J. Lehmann, *Optimale Eigenverteinschiessung*, Num. Math., 5 (1963) pp. 246-272.

[4] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, N.J., 1980.

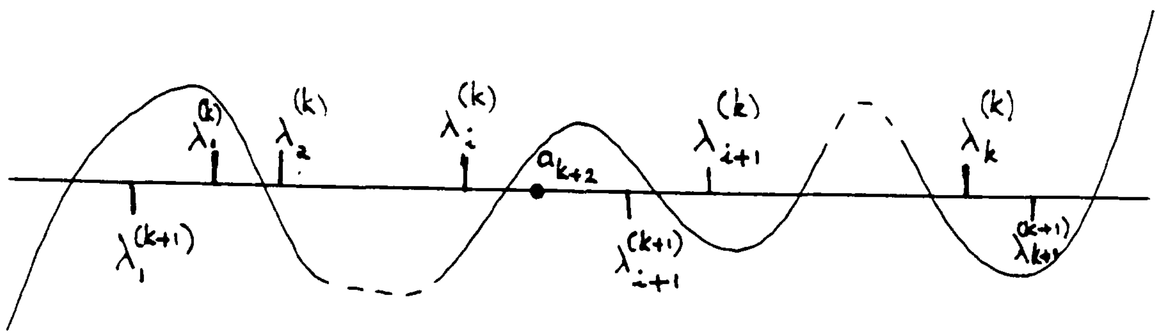

Figure 1. A possible graph of $y = p_{k+2}(x)$

Legal Notice