PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 48, Number 1, March 1975

# SPECTRA OF NEARLY HERMITIAN MATRICES

W. KAHAN[1]

ABSTRACT. When properly ordered, the respective eigenvalues of an $n \times n$ Hermitian matrix $A$ and of a nearby non-Hermitian matrix $A + B$ cannot differ by more than $(\log_2 n + 2.038)\|B\|$; moreover, for all $n \geq 4$, examples $A$ and $B$ exist for which this bound is in excess by at most about a factor 3. This bound is contrasted with other previously published over-estimates that appear to be independent of $n$. Further, a bound is found, for the sum of the squares of respective differences between the eigenvalues, that resembles the Hoffman-Wielandt bound which would be valid if $A + B$ were normal.

0. **Our problem.** How near are the eigenvalues of a nearly Hermitian matrix to those of a nearby Hermitian matrix? To be specific, let the $n \times n$ matrix $A$ be Hermitian ($A^* = A$) with eigenvalues $\alpha_j$ arranged in ascending order $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$, and let $B$ be an arbitrary $n \times n$ matrix, and index the eigenvalues $(\lambda_j + \iota\mu_j)$ of $A + B$ to have real parts $\lambda_j$ in ascending order $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. We seek bounds for differences like $|\lambda_j + \iota\mu_j - \alpha_j|$ or $\Sigma|\lambda_j + \iota\mu_j - \alpha_j|^2$ in terms of two norms of $B$, one of them

$$\|B\|_2 \equiv \sqrt{\operatorname{trace}(B^*B)} = \sqrt{\sum\sum|b_{ij}|^2},$$

and the other

$$\|B\| \equiv \max_{z \neq 0} \frac{\|Bz\|_2}{\|z\|_2}.$$

Slightly sharper bounds will be obtained by exploiting the decomposition of $B = X + \iota Y$ into its Hermitian and skew parts

$$X \equiv (B + B^*)/2 \quad \text{and} \quad \iota Y \equiv (B - B^*)/2$$

Received by the editors January 17, 1974.

*AMS (MOS) subject classifications* (1970). Primary 15A42, 15A60, 47A55, 65F15; Secondary 15A18, 15A57, 47A10, 47A30.

*Key words and phrases.* Non-Hermitian perturbation, eigenvalue error bounds, generalized Hoffman-Wielandt theorem, generalized Weyl inequality.

[1] This work was supported in part by a grant from the U. S. Office of Naval Research, contract no. ONR N00014-69-A-0200-1017.

whose norms are related to $B$'s via

$$\|B\|_2^2 = \|X\|_2^2 + \|Y\|_2^2, \qquad \|X\| \le \|B\|, \qquad \|Y\| \le \|B\| \le \|X\| + \|Y\|.$$

We shall prove that:

(i) *Every* $|\lambda_j - \alpha_j| \le \|X\| + \|Y\| \cdot (\log_2 n + 0.038)$ *and every* $|\mu_j| \le \|Y\|$, *and for every* $n \ge 4$ *there are matrices* $A$ *and* $B$ *for which* $X = 0$ *and the first inequality overestimates some* $|\lambda_j - \alpha_j|$ *by a factor less than* 3.

(ii) $\Sigma \mu_j^2 \le \|Y\|_2^2$ *and* $\sqrt{\Sigma(\lambda_j - \alpha_j)^2} \le \|X\|_2 + \sqrt{\|Y\|_2^2 - \Sigma\mu_j^2}$, *and nontrivial equality is possible.*

But before these claims are proved in §§ 2 and 3 of this paper, here is a survey of what has already been published about our problem.

**1. Survey.** This survey is drawn from texts like Wilkinson's (1965, pp. 93–109) and Householder's (1964, Chapter 3).

If $A + B$ were normal, the Hoffman-Wielandt theorem (1953) would imply, instead of (ii), $\Sigma(\lambda_j - \alpha_j)^2 + \Sigma\mu_j^2 \le \|B\|_2^2$; our weaker hypotheses lead to an inequality weaker by a factor of 2 at worst. If $B$ were Hermitian, the inequalities of H. Weyl (1911, Satz I) would imply (with $X = B$, $Y = 0$ and all $\mu_j = 0$) that $|\lambda_j - \alpha_j| \le \|B\|$, which is a special case of (i) that shows how much may be lost when $Y \ne 0$.

The best bounds pertinent to our problem which I have been able to draw directly from the earlier literature (especially from Wilkinson (1965, pp. 93–94)) involve the congruent truncated disks $D_j$ in the complex $(\lambda + \iota\mu)$-plane defined as follows (provided $B \ne 0$):

$$D_j = \{\lambda + \iota\mu\colon\ |\lambda + \iota\mu - \alpha_j| < \|B\| \text{ and } |\mu| \le \|Y\|\}.$$

Each eigenvalue $(\lambda_k + \iota\mu_k)$ of $A + B$ must lie in the closure of that connected component of the union $\bigcup D_j$ which includes $D_k$. For example, Figure 1 describes a situation with $n = 5$ which confines $(\lambda_1 + \iota\mu_1)$ to $\overline{D}_1$, $(\lambda_5 + \iota\mu_5)$ to $\overline{D}_5$, and the remaining three $(\lambda_j + \iota\mu_j)$'s to $\overline{D}_2 \cup \overline{D}_3 \cup \overline{D}_4$. For another example, consider a situation wherein the $D_j$'s form one long chain in which each $D_j$ slightly overlaps its neighbours as shown in Figure 2; without bounds like those proved in this paper there would be no way to explain why all eigenvalues $(\lambda_j + \iota\mu_j)$ do not flee like quicksilver to one end of the chain or the other. But (i) above prevents each $(\lambda_k + \iota\mu_k)$ from skipping past more than about $\frac{1}{2}\log_2 n$ of $D_k$'s immediate neighbours, and (ii) above restricts such long skips to at most a small fraction (about $2/(\log_2 n)^2$) of those $n$ eigenvalues.

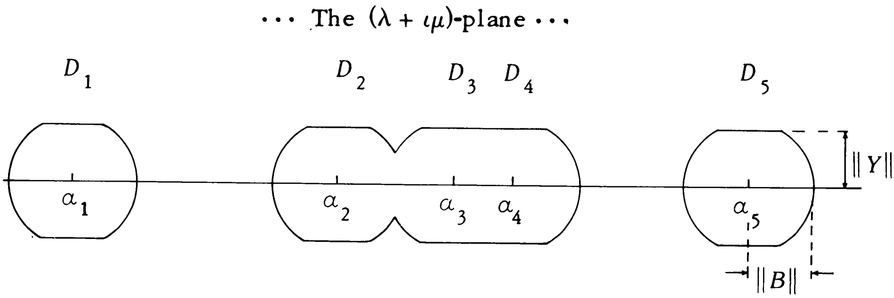


**Figure 1.** Five eigenvalue estimates

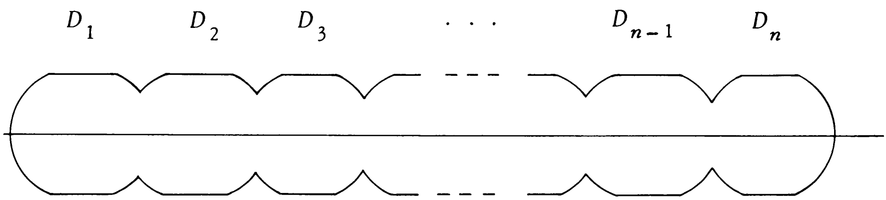


**Figure 2.** One long chain

2. **Proof of claim (i).** Our problem is invariant under unitary similarity, so Schur's theorem may be invoked to triangularize $A + B$ by a unitary similarity and then, without loss of generality, we may assume that $A + B$ was given as upper triangular at the outset. Say

$$A + B = \Lambda + \iota M + \iota U,$$

where $\Lambda \equiv \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$, $M \equiv \mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_n)$ and $U$ is upper triangular with zero for its diagonal. By taking Hermitian and skew parts we find

$$A + X = \Lambda + \iota(U - U^*)/2 \quad \text{and} \quad Y = M + (U + U^*)/2.$$

The last equation will be used below and in §3; for now the appropriate bound upon $M$ is found from Bendixson's inequality (cf. Householder (1964, p. 69)) which implies that every $|\mu_j| \le \|(A + B) - (A + B)^*\|/2 = \|Y\|$. The previous equation is ready for an application of Weyl's inequality; every

$$|\lambda_j - \alpha_j| \le \|\Lambda - A\| = \|X - \iota(U - U^*)/2\| \le \|X\| + \|U - U^*\|/2.$$

Next set $Z \equiv M + U$ and invoke a theorem published recently (1973) by the author; since $Z$'s eigenvalues $\mu_j$ are all real,

$$\|U - U^*\| = \|Z - Z^*\| \le \|Z + Z^*\| \cdot (\log_2 n + 0.038) = 2\|Y\| \cdot (\log_2 n + 0.038)$$

which, with the previous inequality, vindicates claim (i) but for the provision of an example.

Take the lower triangular $n \times n$ matrix

$$L \equiv \begin{pmatrix} 0 & & & & & & & \bigcirc \\ 1 & 0 & & & & & & \\ 1/2 & 1 & 0 & & & & & \\ 1/3 & 1/2 & 1 & 0 & & & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & & & \\ \frac{1}{n-2} & \frac{1}{n-3} & \cdots & 1/2 & 1 & 0 & & \\ \frac{1}{n-1} & \frac{1}{n-2} & \cdots & 1/3 & 1/2 & 1 & 0 \end{pmatrix}$$

and assemble $A = L + L^*$ and $B = L - L^*$. In the aforementioned paper (1973) the author demonstrated[2] that

$$2 \log n > \|A\| > 2 \log n - 2 \log 2 + 1/2 + 1/n$$

and that $\|B\| < \pi$. Consequently $A$ has at least one eigenvalue $\alpha_n \doteq 2 \log n$, but all the eigenvalues $(\lambda_j + \iota\mu_j)$ of $A + B = 2L$ are zero, so

$$|\lambda_n - \alpha_n| > \|B\| \cdot \frac{2}{\pi}\left(\log n - \log 2 + \frac{1}{4} + \frac{1}{2n}\right).$$

Compare this with assertion (i), noting that in this example $X = 0$ and $Y = B$ and that $(\log_2 n)/((2/\pi) \log n) \doteq 2.27$.

3. **Proof of claim (ii).** Continuing the analysis in § 2, observe first that

$$\begin{aligned} \|Y\|_2^2 &= \|M + (U + U^*)/2\|_2^2 \\ &= \|M\|_2^2 + \frac{1}{2} \|U\|_2^2 \quad \text{(because } M \text{ is diagonal and } U \text{ above the diagonal)} \\ &\geq \|M\|_2^2 = \sum \mu_j^2, \quad \text{as claimed in (ii).} \end{aligned}$$

[2] That paper contains an error due to faulty use of a slide rule. The assertion there that "$(2/\pi) \log n \doteq 0.92 \log_2 n$" should read "$\doteq (\log_2 n)/2.27$". Hence the phrase "about 8% when $n$ is large" should read "a factor less than 3 when $n \geq 4$" on pp. 235, 238 and 239.

Then we find

$$\sqrt{\sum(\lambda_j - \alpha_j)^2} \le \|\Lambda - A\|_2 \quad \text{(by the Hoffman-Wielandt theorem)}$$

$$= \|X - \iota(U - U^*)/2\|_2 \le \|X\|_2 + \|(U - U)^*/2\|_2$$

$$= \|X\|_2 + \|U\|_2/\sqrt{2} = \|X\|_2 + \sqrt{\|Y\|_2^2 - \sum \mu_j^2} \quad \text{as claimed in (ii).}$$

These inequalities reduce to something slightly stronger than the Hoffman-Wielandt theorem when $A + B$ is normal because then $U = 0$ so $\Sigma \mu_j^2 = \|Y\|_2^2$ and, hence, $\Sigma(\lambda_j - \alpha_j)^2 \le \|X\|_2^2$; whereas the Hoffman-Wielandt theorem in its raw form would imply only that the sum $\Sigma \mu_j^2 + \Sigma(\lambda_j - \alpha_j)^2 \le \|X\|_2^2 + \|Y\|_2^2 = \|B\|_2^2$. On the other hand, if we do not know separate bounds for $\|X\|_2$ and $\|Y\|_2$, but only one bound for $\|B\|_2$, we can still exploit (ii) as follows:

$$\sum(\lambda_j - \alpha_j)^2 + 2\sum \mu_j^2 \le \left(\|X\|_2 + \sqrt{\|Y\|_2^2 - \sum \mu_j^2}\right)^2 + 2\sum \mu_j^2$$

$$= 2\|X\|_2^2 + 2\|Y\|_2^2 - \left(\|X\|_2 - \sqrt{\|Y\|_2^2 - \sum \mu_j^2}\right)^2$$

$$\le 2(\|X\|_2^2 + \|Y\|_2^2) = 2\|B\|_2^2.$$

Although the last inequality is not as tight as that in (ii), both inequalities can be made nontrivial equalities by an example:

$$A \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \alpha_1 = -1, \qquad \alpha_2 = 1;$$

$$B \equiv \begin{pmatrix} \iota\mu_1 & 0 \\ -1 & \iota\mu_2 \end{pmatrix}, \qquad \|X\|_2^2 = \tfrac{1}{2}, \qquad \|Y\|_2^2 = \tfrac{1}{2} + \mu_1^2 + \mu_2^2, \qquad \|B\|_2^2 = 1 + \mu_1^2 + \mu_2^2;$$

$$A + B = \begin{pmatrix} \iota\mu_1 & 1 \\ 0 & \iota\mu_2 \end{pmatrix}, \qquad \lambda_1 = \lambda_2 = 0,$$

and

$$\sqrt{2} = \sqrt{\sum(\lambda_j - \alpha_j)^2} = \|X\|_2 + \sqrt{\|Y\|_2^2 - \sum \mu_j^2},$$

and finally

$$\sum(\lambda_j - \alpha_j)^2 + 2\sum \mu_j^2 = 2\|B\|_2^2.$$

4. **Caveat.** Sometimes our problem of §0 comes with the additional information that all of $(A + B)$'s eigenvalues are real. By itself this informa-

tion confers little advantage for the estimation of $\max_j |\lambda_j - \alpha_j|$ or $\Sigma(\lambda_j - \alpha_j)^2$ beyond what is already available from (i) and (ii), as we see from the two examples $A + B$ given above; both examples can have all eigenvalues real (i.e. zero).

But whence comes the knowledge that all of $(A + B)$'s eigenvalues are real? Frequently this is inferred from the existence of a positive definite Hermitian matrix $H$ for which $A + B = HA$. If bounds are known for $\|H\|$ and $\|H^{-1}\|$, then the eigenvalues $\lambda_j$ of $A + B$ compare with the eigenvalues $\alpha_j$ of $A$ as follows (cf. Weyl (1912, Satz IV)); for each $j$ either $1/\|H^{-1}\| \le \lambda_j/\alpha_j \le \|H\|$ or $\lambda_j = \alpha_j = 0$. These inequalities, if available, are generally sharper than the ones proved earlier in this paper.

Less often we may know that $V \equiv F(A + B)F^{-1}$ is Hermitian for some similarity $F$ whose condition number $\kappa \equiv \|F\| \cdot \|F^{-1}\|$ is known not to be large. In this case we may prove all $|\lambda_j - \alpha_j| \le \kappa \cdot \|B\|$, which is better than (i) whenever $\kappa < \log_2 n$, and also a sharper bound than Wilkinson's (1965, pp. 87–88) whenever the bounds for two different $\lambda_j$'s overlap. The following proof of the foregoing inequalities is adapted from an unpublished earlier report by the author (1967).

The polar factorization $F = QH$ provides a Hermitian positive definite $H \equiv (F^*F)^{1/2}$ with $\|H\| = \|F\|$ and $\|H^{-1}\| = \|F^{-1}\|$, as well as a unitary $Q$, and $Y \equiv Q^*VQ$ has the same eigenvalues $\lambda_j$ as have $V$ and $A + B$. Weyl's inequalities will yield the desired result $|\lambda_j - \alpha_j| \le \kappa \cdot \|B\|$ if we can prove $\|Y - A\| \le \kappa \cdot \|B\|$. But first let $x$ be a normalized $(x^*x = 1)$ eigenvector of $Y - A$ for which $(Y - A)x = \pm\|Y - A\|x$. Then

$$\begin{aligned}\|F\| \cdot \|B\| &= \|H\| \cdot \|F^{-1}VF - A\| = \|H\| \cdot \|H^{-1}(YH - HA)\| \ge \|YH - HA\| \\ &\ge |x^*(YH - HA)x| = |x^*(YH - HY)x + x^*H(Y - A)x| \\ &= |x^*(YH - HY)x \pm \|Y - A\|x^*Hx| = |\text{imaginary} \pm \text{real}| \\ &\ge \|Y - A\|x^*Hx \ge \|Y - A\|/\|H^{-1}\| = \|Y - A\|/\|F^{-1}\|.\end{aligned}$$

## REFERENCES

A. S. Householder (1964), *The theory of matrices in numerical analysis*, Blaisdell, New York. MR 30 #5475.

A. J. Hoffman and H. W. Wielandt (1953), *The variation of the spectrum of a normal matrix*, Duke Math. J. 20, 37–39. MR 14, 611.

W. Kahan (1967), *Inclusion theorems for clusters of eigenvalues of Hermitian matrices*, Computer Science Department, University of Toronto, Toronto, Ontario.

W. Kahan (1973), *Every $n \times n$ matrix $Z$ with real spectrum satisfies* $\|Z - Z^*\| \leq \|Z + Z^*\|(\log_2 n + 0.038)$, Proc. Amer. Math. Soc. 39, 235–241. MR 47 #1833.

H. Weyl (1911), *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichunger,..*, Math. Ann. 71, 441–479.

J. H. Wilkinson (1965), *The algebraic eigenvalue problem*, Clarendon Press, Oxford. MR 32 #1894.

ELECTRONICS RESEARCH LABORATORY, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720

*Current address*: Department of Mathematics, University of California, Berkeley, California 94720