PORTUGALIAE MATHEMATICA

ISSN 0032-5155

VOLUME 43

1985-1986

Edição da

SOCIEDADE PORTUGUESA DE MATEMÁTICA

PORTUGALIAE MATHEMATICA

Av. da República, 37-4.º

1000 LISBOA—PORTUGAL

PORTUGALIAE MATHEMATICA
Vol. 43 — Fasc. 1 — 1985-1986

PERTURBATION THEOREMS FOR THE MATRIX EIGENVALUE PROBLEM

L. ELSNER
Fakultät für Mathematik, Universität Bielefeld, Postfach 8640
4800 Bielefeld 1 — Federal Republic of Germany

ABSTRACT. An overview on some recent results concerning perturbations of eigenvalues of matrices is given.

1. It is well known that the eigenvalues of a matrix depend continuously on the entries of the matrix. For numerical considerations more quantitative statements are required. There are quite a few of them available in the literature, we refer to the books of Householder [13] and Marcus-Minc [14].

In this overview we restrict our attention to some aspects of this topic and present recent results on

- — comparisons between certain measures of the distance between spectra
- — global bounds for perturbations of spectra
- — inclusion theorems for the generalized eigenvalue problem.

2. Let A, B, C, ... denote complex $n \times n$ matrices.

For A, B with spectra $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ and $\sigma(B) = \{\mu_1, \dots, \mu_n\}$ we define as

$$S_A(B) = \max_i \min_j |\lambda_j - \mu_i|$$

Received December 4, 1984.
Subject classification: 15 A 18, 65 F 15.

the spectral variation of B with respect to A and

$$\nu(A, B) = \min_{\pi} \max_{i} | \lambda_i - \mu_{\pi(i)} |$$

the eigenvalue variation of A and B. Here π runs through all permutations of $\{1, \ldots, n\}$.

Besides this the following functions turn out to be useful

$$h_A(B) = \max \{S_A(tA + (1-t)B) : 0 \leqslant t \leqslant 1\}$$
$$g(A, B) = \max \{h_A(B), h_B(A)\}.$$

While $\nu(A, B)$ is the most natural measure of the distance between the spectra the other concepts are introduced because they can be bounded easily and can be related to $\nu(A, B)$.

We observe that the Hausdorff-distance between the sets $\sigma(A)$ and $\sigma(B)$, i. e. $\max(S_A(B), S_B(A))$ does not compare with $\nu(A, B)$, because for $n > 2$ it can be zero while $\nu(A, B) \neq 0$.

The following inequalities hold:

(2.1) $\nu(A, B) \leqslant (2n-1)\, h_A(B)$

(2.2) $\nu(A, B) \leqslant a_n\, g(A, B) \qquad a_n = \begin{cases} n & \text{if } n \text{ odd} \\ n-1 & \text{if } n \text{ even.} \end{cases}$

We note that (2.1) is essentially due to Ostrowski, but the formulation here has the advantage to be sharp, as can be seen from the example $A = \mathrm{diag}(0, 2, \ldots, 2n-2)$, $B = (2n-1) I_n$ (I_n = identity matrix), $h_A(B) = 1$, $\nu(A, B) = 2n - 1$. (see [6], [7]). The second inequality seems to be new. It is proved in [7], using the marriage-theorem. Also (2.2) is sharp.

This is shown by the examples $A = \mathrm{diag}(2, 4, \ldots, 2k, 0, \ldots, 0)$, $B = (2k+1) I_n - A$ $n = 2k+1$ or $n = 2k$ where $h_A(B) = h_B(A) = g(A, B) = 1$ and $\nu(A, B) = 2k + 1 = n$ in the case of n odd and $\nu(A, B) = 2k - 1 = n - 1$ in the even case.

The importance of (2.1) and (2.2) lies in the fact that most bounds for $S_A(B)$ available in the literature are also bounds on $h_A(B)$ and $g(A, B)$, hence provide bounds for $\nu(A, B)$. While those obtained by (2.1) are mostly in the literature (as (2.1) is «folklore»), the bounds via (2.2) are new and improve the known results by a factor of about 1/2. An example is given in the next chapter.

3. Let us call a bound of $S_A(B)$ or $\nu(A, B)$ global if it depends only on $\| A \|$, $\| B \|$ and $\| A\text{-}B \|$, where $\| \ \|$ is some matrix-norm.

Historically the first (though not completely fitting into this definition) is Ostrowski's result ([16])

$$(3.1) \quad S_A(B) \leqslant (n+2) \, [\max_{i,j} (| a_{ij} |, | b_{ij} |)]^{1-1/n} \, (1/n \sum_{i,j} | a_{ij} - b_{ij} |)^{1/n}$$

where a_{ij}, b_{ij} are the elements of A and B respectively. Obviously the righthand side is an upper bound for $g(A, B)$, too. The same holds for the bound given in [5]

$$(3.2) \quad S_A(B) \leqslant (1 + n^{-1/2}) \, n^{1/2n} \, M_E^{1-1/n} \| A - B \|_E^{1/n}$$

where $\| A \|_E^2 = \sum_{i,k=1}^{n} | a_{ik} |^2$ is the Euclidean matrix norm and $M_E = \max(\| A \|_E, \| B \|_E)$. (3.2) is the version of [6], in [5] the leading factor is slightly larger.

An analogous result for the spectral norm $\| \ \|_2$ can be found in [4], (see also [6]).

$$(3.3) \quad S_A(B) \leqslant n^{1/n} \, (2 M_2)^{1-1/n} \, \| A - B \|_2^{1/n}, \; M_2 = \text{Max}(\| A \|_2, \| B \|_2),$$

and in [10], S. Friedland showed that (3.3) holds for any operator norm.

The sharpest result, however, is the following ([9])

$$(3.4) \quad S_A(B) \leqslant (\| A \|_2 + \| B \|_2)^{1-1/n} \, \| A - B \|_2^{1/n}$$

which implies

$$(3.5) \quad g(A, B) \leqslant (2 M_2)^{1-1/n} \, \| A - B \|_2^{1/n}$$

Let us give a proof of (3.4):

If $\sigma_1 \leqslant \ldots \leqslant \sigma_n$ are the singular values of A, we know that $\| A - B \|_2 \geqslant \sigma_1$ for any singular B and $\sigma_n = \| A \|_2$. Hence

$$(3.6) \quad | \det A | = \sigma_1 \ldots \sigma_n \leqslant \| A - B \|_2 \, \| A \|_2^{n-1}$$

and upon replacing in (3.6) A by $A - \mu I$, B by $B - \mu I$, where μ is an eigenvalue of B, we get

$$| \det (A - \mu I) | \leqslant \| A - B \|_2 \, (\| A \|_2 + \| B \|_2)^{n-1}$$

and noting that $(S_A(B))^n \leqslant |\Pi(\lambda_i - \mu)| = |\det(A - \mu I)|$ for some eigenvalue μ of B, (3.4) follows.

This proof is much shorter and simpler than the proofs in [6] and has the additional advantage of being sharp. In fact, it is shown in [9], that equality holds in (3.4) iff $A = \varepsilon \| A \|_2 \cdot I$ and B has an eigenvalue $-\varepsilon \| B \|_2$ for some $\varepsilon \in C$, $|\varepsilon| = 1$. This is done there by giving a slightly longer proof of (3.4) using the Hadamard-inequality for det $(A - \mu I)$ and exploiting the additional information in the case of equality.

A similar sharpness result holds for (3.5).

From (3.5) and (2.2) we infer

$$(3.7) \quad \nu(A, B) \leqslant a_n (2 M_2)^{1-1/n} \| A - B \|_2^{1/n}.$$

It has been conjectured by S. Friedland that the factor $a_n \approx n$ may be replaced by 1 or at least a constant independent of n. This seems to be quite a hard problem.

Let us end this chapter by drawing attention to another conjecture, which was formulated by Mirsky 24 years ago:

If A, B both are normal then it is a consequence of the Bauer-Fike theorem ([1]) that

$$g(A, B) \leqslant \| A - B \|_2.$$

Mirsky conjectured [15]

$$(3.8) \quad \nu(A, B) \leqslant \| A - B \|_2.$$

To my knowledge the best result in this direction is given in [3] by Bhatia, Davis and Mc Intosh:

$\exists c$ independent of n such that

$$\nu(A, B) \leqslant c \| A - B \|_2,$$

for all A, B normal. Bhatia and Davis have shown in [2], that (3.8) holds for A, B both unitary.

4. In this chapter we want to report on some generalizations of classical perturbation theorems to the case of the generalized eigenvalue problem [8]. We prefer to write it in the form

$$(4.1) \quad \alpha B x = \beta A x$$

$(\alpha, \beta) \neq (0, 0) \in C^2$ is an eigenvalue of the matrix pair $Z = (A, B)$ if there exists $x \neq 0$ s. t. (4.1) holds. We consider (α, β) as a point in the projective complex plane with the chordal metric

$$(4.2)\quad \rho((\alpha, \beta), (\gamma, \delta)) = \frac{|\alpha\delta - \beta\gamma|}{\sqrt{|\alpha|^2 + |\beta|^2}\ \sqrt{|\gamma|^2 + |\delta|^2}}$$

It is known that if the matrix pair $Z = (A, B)$, which we view as a $n \times 2n$-complex matrix, is *regular*, i e.

$$(4.3)\quad \det(A - \lambda B) \not\equiv 0$$

then it has n eigenvalues (multiplicities counted in an appropriate way).

If $W = (C, D)$ is a regular pair with eigenvalues (γ_i, δ_i), $i = 1, \ldots, n$, we may define the spectral variation of W w.r.t. Z

$$S_Z(W) = \max_i \min_j \rho((\alpha_j, \beta_j), (\gamma_i, \delta_i))$$

and as above analogously $h_Z(W)$, $g(Z, W)$, $\nu(Z, W)$.

Having introduced distances between spectra we have to define distances of matrix pairs Z, W. It is not appropriate to use $\| Z - W \|$ for some matrix norm since, for T nonsingular, Z and $TZ = (TA, TB)$ have the same spectrum, hence the spectrum depends only on $\operatorname{Ker}(Z) \subset C^{2n}$ or on its orthogonal complement $L_Z = \{Z^T x : x \in C^n\}$.

With

$$(4.4)\quad P_Z = Z^H (ZZ^H)^{-1} Z$$

the orthogonal projector onto L_Z, we define the distances

$$(4.5)\quad d_2(Z, W) = \| P_Z - P_W \|_2$$

$$(4.6)\quad d_E(Z, W) = 1/\sqrt{2}\ \| P_Z - P_W \|_E$$

which are metrics on the Grassmann-manifolds $G_{n,2n}$ of the n-dimensional subspaces of C^{2n}.

We define a regular pair Z to be diagonalizable if there exists a basis of eigenvectors of Z. This is equivalent to the statement that there exist nonsingular S, T such that SAT and SBT are diagonal. Z is called normal if in addition the eigenvectors can be chosen orthonormal i. e. that there exist S nonsingular, T unitary such that SAT and SBT are diagonal. It can be proved

THEOREM 1. If $Z = (A, B)$ is a diagonalizable pair, W regular and SAT and SBT diagonal then

$$(4.7) \quad S_Z(W) \leqslant \| T \|_2 \, \| T^{-1} \|_2 \, d_2(Z, W)$$

This should be compared to the Bauer-Fike theorem [1]:

THEOREM 1'. If A is diagonalizable, $T^{-1} AT$ diagonal, then for any $n \times n$-matric C

$$(4.8) \quad S_A(C) \leqslant \| T \|_2 \, \| T^{-1} \|_2 \, \| A - C \|_2$$

The Hoffman-Wielandt theorem ([12]) is the following

THEOREM 2'. For A, C normal with eigenvalues $\{\lambda_i\}$, $\{\mu_i\}$

$$(4.9) \quad \nu(A, C) \leqslant \min_{\pi} \{ \sum_j |\lambda_j - \mu_{\pi(j)}|^2\}^{1/2} \leqslant \| A - C \|_E .$$

The generalized version is

THEOREM 2. Let Z, W be normal pairs with eigenvalues (α_i, β_i) and (γ_j, δ_j). Then

$$(4.10) \quad \nu(Z, W) = \min_{\pi} \max_{i} \rho((\alpha_i, \beta_i), (\gamma_{\pi(i)}, \delta_{\pi(i)}))$$

$$\leqslant \min [\sum_i \rho^2 ((\alpha_i, \beta_i), (\gamma_{\pi(i)}, \delta_{\pi(i)}))]^{1/2} \leqslant d_E(Z, W)$$

Denote the inverse function of $x \to x + x^2 + \ldots + x^n$ in $\mathbf{R}_+$ by g_n and define

$$(4.11) \quad S_n(d, r) = \begin{cases} r & d = 0 \\ d\,(g_n(d/r))^{-1} & d, r > 0 \\ 0 & r = 0 \end{cases}$$

The departure from normality of a matrix A is defined by $\Delta(A) = \{\min \| M \|_2 \mid M$ strictly upper triangular, $A = U(\Delta + M) U^H$, U unitary, Δ diagonal$\}$.

Then Henrici has shown ([11]):

THEOREM 3'. If $\Delta = \Delta(A)$ then

$$(4.12) \quad S_A(C) \leqslant S_n(\Delta, \| A - C \|_2)$$

Generalizing the departure from normality, it is possible to define a function m (Z) such that m (Z) $\geqslant 0$ for all Z and m (Z) $= 0$ iff Z is normal. Then we can prove

THEOREM 3. If Z, W are regular pairs then

(4.13) $S_Z(W) \leqslant S_n(m(Z), (1 + m(Z))\, d_2(Z, W))$.

In the case that Z is normal Theorems 1 and 3 both reduce to

(4.14) $S_Z(W) \leqslant d_2(Z, W)$.

We remark that the classical results (Theorems 1′, 2′, 3′) are not special cases of the general results. However they can be obtained via a limiting argument:

For given A, C consider the regular pairs $Z_\varepsilon = (I, \varepsilon A)$, $W_\varepsilon = (I, \varepsilon C)$. Applying Theorems 1, 2, 3 to Z_ε, W_ε and letting $\varepsilon \to 0$ results in Theorems 1′, 2′, 3′. This «derivation»-procedure was used previously by Stewart.

Finally let us mention some results on definite pairs.

Here a pair Z = (A, B) is called a definite pair, if A, B both are hermitian and

(4.15) $c(Z) = \min \{ | x^H (A + i B) x |,\ x^H x = 1 \} > 0$.

It is well known that definite pairs are diagonalizable.

The following result holds

THEOREM 4. If Z is a definite pair and W is regular then

(4.16) $S_Z(W) \leqslant \| Z \|_2 (c(Z))^{-1} d_2(Z, W)$
$S_Z(W) \leqslant (c(Z))^{-1} \| Z - W \|_2$

For the case of W definite Stewart obtained a similar result for ν (Z, W), ([17], Thm. 3.2). While Theorems 1, 2, 3 and the first inequality of Theorem 4 are proved in [8], the second inequality of Theorem 4 can be found in [19]. For further results see the overview by Sun [18].

REFERENCES

[1] F. L. Bauer and C. T. Fike — Norms and exclusion theorems, *Num. Math.*, 2; 137-141 (1960).

[2] R. Bhatia and C. Davis — A Bound for the Spectral Variation of a Unitary Operator, *Linear and Multilinear Algebra*, 15; 71-76 (1984).

[3] R. Bhatia, C. Davis and A. McIntosh — Perturbation of Spectral Subspaces and Solution of Linear Operator Equations, *Linear Algebra Appl.* 52-53; 45-67 (1983).

[4] R. Bhatia and S. Friedland — Variation of Grassmann Powers and Spectra, *Linear Algebra Appl.* 40; 1-18 (1981).

[5] R. Bhatia and K. Mukherjea — On the of Change of Spectra of Operators, *Linear Algebra Appl.* 27; 147-157 (1979).

[6] L. Elsner — On the Variation of Spectra of Matrices, *Linear Algebra Appl.*, 47; 127-138 (1982).

[7] L. Elsner, C. Johnson, J. Ross and J. Schönheim — A Generalized Matching Problem arising in estimating the Eigenvalue Variation of two Matrices, *Europ. J. Comb.*, 4; 133-136 (1983).

[8] L. Elsner and J. Sun — Perturbation Theorems for the Generalized Eigenvalue Problem, *Linear Algebra Appl.* 48; 341-357 (1982).

[9] L. Elsner — An optimal Bound for the Spectral Variation of two Matrices, *Linear Algebra Appl.* 71; 77-80 (1985).

[10] S. Friedland — Variation of Tensor Powers and Spectra, *Linear and Multilinear Algebra*, 12; 81-98 (1982).

[11] P. Henrici — Bounds for iterates, inverses, spectral variation and field of values of non-normal matrices, *Numer. Math.*, 4; 24-40 (1962).

[12] A. J. Hoffman and H. Wielandt — The variation of the spectrum of a normal matrix, *Duke Math. J.*, 20; 37-39 (1953).

[13] A. S. Householder — The Theory of Matrices in Numerical Analysis, Blaisdell, 1964.

[14] M. Marcus and H. Minc — A Survey of Matrix Theory and Matrix Inequalities, Allyn-Bacon, 1964.

[15] L. Mirsky — Symmetric Gauge Functions and unitarily invariant Norms, *Quarterly J. Math. Oxford Ser.*, 2, 11; 50-59 (1960).

[16] A. Ostrowski — Über die Stetigkeit von charakteristischen Wurzeln in Abhängigkeit von den Matrizenelementen, *J. Ber. DMV, Band* 60; 40-42 (1957).

[17] G. W. Stewart — Perturbation Bounds for the Definite Generalized Eigenvalue Problem, *Linear Algebra Appl.* 23; 69-85 (1979).

[18] Ji-guang Sun — Perturbation Analysis for the generalized Eigenvalue and the Generalized Singular Value Problem, p. 221-244 in Matrix pencils, ed. B. Kågström and A. Ruhe, *Lecture Notes in Mathematics*, 973, Springer-Verlag, 1983.

[19] L. Elsner and P. Lancaster — The Spectral Variation of Pencils of Matrices, *J. Comput. Math.* 3, n. 3 (1985).