Numerische Mathematik 4, 24—40 (1962)

Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices*

By

PETER HENRICI

Introduction

A *matrix* in this paper always means a square matrix of order n with complex elements; a *vector* means a column vector with n complex components. A *matrix norm* is a real-valued function ν defined on the space of matrices and satisfying the following relations for arbitrary matrices A and B and arbitrary complex scalars c:

(a) $\nu(A) \geqq 0$; $\nu(A) = 0$ if and only if $A = 0$.

(b) $\nu(cA) = |c|\,\nu(A)$.

(c) $\nu(A+B) \leqq \nu(A) + \nu(B)$.

(d) $\nu(AB) \leqq \nu(A)\,\nu(B)$.

A *vector norm* is a real-valued function defined on the space of vectors and satisfying relations analogous to (a), (b), and (c) above. By the *spectrum* of a matrix A we mean the totality of its eigenvalues, considered as a point set in the complex plane. The largest of the moduli of the eigenvalues of A is called the *spectral radius* of A and will be denoted by λ_A.

The following problems of computational linear algebra will be considered in this paper:

(i) To estimate the norms of the matrices A^n, $n = 1, 2, 3, \ldots$, in terms of λ_A.

(ii) To estimate the error $\tilde{x} - A^{-1}b$ of an approximate solution $\tilde{x}$ of the equation $Ax = b$ in terms of the residual $r = A\tilde{x} - b$ and the spectral radius $\lambda_{A^{-1}}$.

(iii) To estimate the distance of the spectrum of a matrix B from the spectrum of a matrix A in terms of a norm of $B - A$.

(iv) To find bounds for the field of values of a matrix A in terms of the spectrum of A. Here the field of values, $F(A)$, is defined as the set of complex numbers

$$\xi = \frac{x^* A x}{x^* x},$$

where x runs through all non-zero vectors. (The * denotes the conjugate transpose of a vector or matrix.)

Solutions to the above problems are classical (and indeed trivial in some cases) if A is *normal*, i.e., $AA^* = A^*A$. Solutions have also been constructed for non-normal A, but with less satisfactory results. Some of the bounds given

* The preparation of this paper was sponsored by the U.S. Army Research Office (Durham). Reproduction in whole or in part is permitted for any purpose of the United States Government. The paper was presented (in part) at the Invitational Symposium on Matrix Computation, Gatlinburg, Tenn., April 23—29, 1961.

depend on a knowledge of a matrix S in the representation $A = S J S^{-1}$, where J is the Jordan canonical form. Other bounds do not approach the classical bounds if A approaches a normal matrix. Contrary to this, the bounds given in the present paper depend — at most — on the eigenvalues of A. A knowledge of the Jordan canonical form is not required. Furthermore, our estimates approach the classical estimates for A normal. Our insistence on not using the Jordan form is motivated partly by reasons of computational convenience, and partly by the fact that the Jordan form is a discontinuous function on the space of matrices and is therefore ill suited for purposes of computation (see [6] for related remarks).

The principal tool in our investigation is a numerical-valued function on the space of matrices that serves as a measure of the departure from normality of the matrix. After some preliminary remarks on matrix norms (§ 1.1) this function will be defined in § 1.2. We shall then derive a bound for this function in terms of a rational function of the elements of the matrix (Theorem 1). In the subsequent sections the measure of non-normality will be applied to the solution of the four problems stated above.

The author is indebted to G. E. Forsythe, E. Sallin, O. Taussky Todd, H. Wielandt and, in particular, to B. J. Stone for a number of stimulating comments on the subject of this paper. It should be mentioned that Wielandt in [*28*] already has defined a measure of non-normality of a matrix. However, his measure is applicable only to matrices which are similar to a diagonal matrix, and to find an explicit bound for it again requires the knowledge of a matrix S effecting the diagonalization.

1. A measure of non-normality

1.1. *Some preliminaries on norms.* We shall frequently use the following special examples of matrix norms (see [*12*], [*20*]) and shall refer to them with special symbols $(A = (a_{ij}))$:

$$\sigma(A) = \max_{x \neq 0} \left[\frac{x^* A^* A x}{x^* x}\right]^{\frac{1}{2}} \quad \text{(spectral norm)}$$

$$\varrho(A) = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|$$

$$\gamma(A) = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|$$

$$\varepsilon(A) = \left[\sum_{1 \leq i, j \leq n} |a_{ij}|^2\right]^{\frac{1}{2}} \quad \text{(euclidean norm)}$$

$$\alpha(A) = \sum_{1 \leq i, j \leq n} |a_{ij}| .$$

Furthermore, if φ is a vector norm, then the function ν_φ defined by

$$\nu_\varphi(A) = \sup_{x \neq 0} \frac{\varphi(A x)}{\varphi(x)}$$

always defines a matrix norm. Matrix norms defined in this manner are called *lub* norms in [*4*]. The norms σ, ϱ, and γ defined above can be derived in this manner from suitable vector norms (see also [23], chapter 15). On the other hand, some matrix norms, such as the norms ε and α above, cannot be thus derived.

We shall use the following definitions:

A matrix norm ν is called *compatible* with a vector norm μ, if $\mu(A\,x) \leqq \nu(A)\,\mu(x)$ for all matrices A and vectors x. A lub norm is always compatible with the vector norm defining it.

A matrix norm ν will be called *unitarily invariant,* if $\nu(U A U^*) = \nu(A)$ for all A and all unitary U. The norms σ and ε are unitarily invariant, while the norms $\varrho, \gamma,$ and α are not.

A lub norm ν is called *axis-oriented* ([4], p. 138), if $\nu(D) = \max_{1 \leqq i \leqq n} |d_{ii}| = \lambda_D$ for any diagonal matrix $D = (d_{ij})$. Clearly, the lub norms σ, ϱ, and γ are axis-oriented.

A norm ν is said to *majorize* another norm μ if $\nu(A) \geqq \mu(A)$ for all matrices A. Clearly, α majorizes the four other norms considered above. The norm ε majorizes σ.

We shall require the following consequences of the defining properties (a), (b), (c), (d) of a norm (see [20] for proofs):

(I) *If λ_A denotes the spectral radius of A, then*

$$\nu(A) \geqq \lambda_A \tag{1.1}$$

for any matrix norm ν ([20], p. 3).

(II) *If μ and ν are any two matrix norms, then there exists a constant $p_{\mu\nu}$, depending only on these two norms, such that*

$$\mu(A) \leqq p_{\mu\nu}\,\nu(A) \tag{1.2}$$

for all matrices A (see [20], p. 4).

This last property is useful because it frequently permits one to reduce the investigation of a general property of norms to the study of a special norm. For the special norms introduced above and a few other norms, values of the constants $p_{\mu\nu}$ are given in the following article by STONE [25].

1.2. *The ν-departure from normality.* If A is any matrix, then a classical result due to SCHUR ([17], Theorem 10.4.1) states that there exist a unitary matrix U and a triangular matrix $T = (t_{ij})$ with $t_{ij} = 0$ for $j < i$ such that

$$A = U\,T\,U^*.$$

The matrix T is called a *Schur triangular form of A*. In general, T is not uniquely determined for a given A (see MCRAE [16]*). We put

$$T = D + M,$$

where D denotes the diagonal matrix whose main diagonal coincides with that of T (and thus is made up of the eigenvalues of A). It follows that $M = (m_{ij})$ has non-zero elements only to the right of the main diagonal and thus that

$$M^r = 0, \qquad r \geqq n. \tag{1.3}$$

* The author is indebted to G. E. FORSYTHE for this reference.

If ν is a norm, we define the *ν-departure from normality of A* by

$$\Delta_\nu(A) = \inf \nu(M), \tag{1.4}$$

where the inf is taken with respect to all M that can appear in a Schur triangular form. Since ε is unitarily invariant, $[\varepsilon(A)]^2 = [\varepsilon(T)]^2 = [\varepsilon(D)]^2 + [\varepsilon(M)]^2$. It follows that

$$\varepsilon(M) = \left\{[\varepsilon(A)]^2 - \sum_{i=1}^{n} |\lambda_i|^2\right\}^{\frac{1}{2}} \tag{1.5}$$

is independent of the special choice of the Schur triangular form. Since A is normal if and only if $\varepsilon(M) = 0$ (see [17], Theorem 10.3.8), it follows that $\Delta_\varepsilon(A)$ is zero if and only if A is normal. By property (II) above this statement is true for $\Delta_\nu(A)$ where ν is any norm.

1.3. *A bound for $\Delta_\varepsilon(A)$.* In this section we shall derive a bound for $\Delta_\varepsilon(A)$ which can be computed in an elementary way from the elements of A and which reduces to zero for A normal. By fact (II) above this makes it possible to construct bounds for $\Delta_\nu(A)$ for any norm ν.

Theorem 1. *For an arbitrary matrix A,*

$$\Delta_\varepsilon(A) \leqq \sqrt[4]{\frac{n^3 - n}{12}}\, \sqrt{\varepsilon(A^* A - A A^*)}\,. \tag{1.6}$$

Equality holds in (1.6) *if and only if A is unitarily similar to a matrix of the form*

$$\begin{pmatrix} \lambda & \alpha_1 & & & 0 \\ & \lambda & \alpha_2 & & \\ & & \lambda & \ddots & \\ & & & \ddots & \\ & & & \lambda & \alpha_{n-1} \\ 0 & & & & \lambda \end{pmatrix},$$

where λ is a complex constant, and where $|\alpha_k|^2 = a k(n-k)$, $k = 1, 2, \ldots, n-1$, for some $a \geqq 0$.

Proof. Let $T = D + M$ be a Schur triangular form of A, $A = U T U^*$, and define $\Gamma = (\gamma_{ij})$ by

$$\Gamma = T^* T - T T^*.$$

We first show that

$$[\varepsilon(M)]^2 \leqq \gamma_{22} + 2\gamma_{33} + \cdots + (n-1)\gamma_{nn}. \tag{1.7}$$

For the proof we observe that, setting $M = (m_{ij})$,

$$\gamma_{ii} = \sum_{k<i} |m_{ki}|^2 - \sum_{k>i} |m_{ik}|^2, \qquad i = 1, 2, \ldots, n \tag{1.8}$$

(empty sums are to be replaced by zero). We now proceed by induction with respect to n. Obviously the inequality (1.7) is true for matrices of order 1. We assume its validity for matrices of order n and use primes to refer to quantities corresponding to a matrix of order $n+1$ which has M as an upper lefthand submatrix of order n. In view of

$$\gamma'_{ii} = \gamma_{ii} - |m_{i,n+1}|^2, \qquad i = 1, 2, \ldots, n,$$

we have

$$\begin{aligned}[\varepsilon(M')]^2 &= [\varepsilon(M)]^2 + |m_{1,n+1}|^2 + |m_{2,n+1}|^2 + \cdots + |m_{n,n+1}|^2 \\ &\leqq \gamma_{22} + 2\gamma_{33} + \cdots + (n-1)\gamma_{nn} + \gamma'_{n+1,n+1} \\ &= \gamma'_{22} + |m_{2,n+1}|^2 + 2\gamma'_{33} + 2|m_{3,n+1}|^2 + \cdots \\ &\quad + (n-1)\gamma'_{nn} + (n-1)|m_{n,n+1}|^2 + \gamma'_{n+1,n+1} \\ &\leqq \gamma'_{22} + 2\gamma'_{33} + \cdots + n\gamma'_{n+1,n+1},\end{aligned}$$

by (1.8), proving (1.7) with n increased by one. We remark that equality holds in (1.7) if and only if $m_{ij}=0$ for $j>i+1$, i.e., if the triangular matrix T is "almost diagonal".

Subtracting from (1.7) $\frac{n-1}{2}$ times the identity

$$\gamma_{11} + \gamma_{22} + \cdots + \gamma_{nn} = 0$$

(which is a consequence of (1.8)), we find $[\varepsilon(M)]^2 \leqq \frac{1-n}{2}\gamma_{11} + \frac{3-n}{2}\gamma_{22} + \cdots + \frac{n-1}{2}\gamma_{nn}$, and, using the Cauchy inequality,

$$(1.9)\quad [\varepsilon(M)]^4 \leqq \left[\left(\frac{1-n}{2}\right)^2 + \left(\frac{3-n}{2}\right)^2 + \cdots + \left(\frac{n-1}{2}\right)^2\right][\gamma_{11}^2 + \gamma_{22}^2 + \cdots + \gamma_{nn}^2].$$

An elementary computation shows that

$$\left(\frac{1-n}{2}\right)^2 + \left(\frac{3-n}{2}\right)^2 + \cdots + \left(\frac{n-1}{2}\right)^2 = \frac{n^3-n}{12}.$$

The result (1.6) now follows by virtue of the fact that $A^*A - AA^* = U\Gamma U^*$ and hence

$$(1.10)\qquad \gamma_{11}^2 + \gamma_{22}^2 + \cdots + \gamma_{nn}^2 \leqq [\varepsilon(\Gamma)]^2 = [\varepsilon(A^*A - AA^*)]^2.$$

Equality holds in (1.6) if and only if equality holds simultaneously in (1.7), (1.9) and (1.10). Equality in (1.10) requires $\Gamma = T^*T - TT^*$ to be diagonal. For an almost diagonal matrix T (as required by equality in (1.7)), Γ is diagonal if and only if all diagonal elements of T are identical. The diagonal elements of Γ then satisfy

$$(1.11)\qquad |m_{k,k+1}|^2 = \gamma_{11} + \gamma_{22} + \cdots + \gamma_{kk}, \qquad k = 1, 2, \ldots, n-1.$$

Equality in (1.9) finally requires that $\gamma_{kk} = a(1+n-2k)$, $k=1, 2, \ldots, n$, where a is a real constant. In view of (1.11) this is the case if and only if for some real $a \geqq 0$

$$|m_{k,k+1}|^2 = a\,k(n-k), \qquad k = 1, 2, \ldots, n-1,$$

i.e., if T has the form indicated in the theorem.

2. Iterated matrices

2.1. *Introduction.* It is a simple consequence of the Schur triangular form that for a *normal* matrix A

$$(2.1)\qquad \sigma(A^r) = \lambda_A^r.$$

Bounds for norms of powers of certain non-normal matrices (arising in finite difference schemes for solving hyperbolic and parabolic differential equations) have been given by LAX and RICHTMYER [*15*] and KATO [*13*]. For arbitrary matrices the problem has been treated by GAUTSCHI [*7, 8*] and OSTROWSKI [*23*]. Denoting by ϱ the norm defined at the beginning of § 1.1, OSTROWSKI shows that for every $\varepsilon > 0$

$$\varrho(A^r) \leqq C(\varepsilon)\,(\lambda_A + \varepsilon)^r, \qquad r = 1, 2, \ldots.$$

Here the function $C(\varepsilon)$ is determined as follows: If J is a Jordan canonical form of $2\varepsilon^{-1}A$, and if $2\varepsilon^{-1}A = S J S^{-1}$, then $C(\varepsilon) = \varrho(S)\,\varrho(S^{-1})$, the ϱ-condition number of S [*2*]. GAUTSCHI's bound similarly requires some knowledge of the Jordan canonical form.

2.2. *A new bound.* In contrast to the above, Theorem 2 below gives an estimate for $\sigma(A^r)$ which depends only on λ_A and $\Delta_\sigma(A)$, and reduces to (2.1) for A normal.

Theorem 2. *If $m = \Delta_\sigma(A)$, and if $\lambda_A > 0$, then*

$$\sigma(A^r) \leqq \lambda_A^r + \binom{r}{1}\lambda_A^{r-1} m + \cdots + \binom{r}{n-1}\lambda_A^{r-n+1} m^{n-1}; \tag{2.2}$$

if $\lambda_A = 0$, then

$$\begin{cases} \sigma(A^r) \leqq m^r, & r = 0, 1, \ldots, n-1, \\ \sigma(A^r) = 0, & r \geqq n. \end{cases} \tag{2.3}$$

Remark. The expression on the right of (2.2) can be estimated further in various ways; for r large it is asymptotically equal to

$$\frac{\lambda_A^{r-n+1}(m\,r)^{n-1}}{(n-1)!}.$$

Proof. Again let $A = U T U^*$, where $T = D + M$ is a Schur triangular form. We then have

$$A^r = U(D + M)^r U^*.$$

Since D and M do not, in general, commute, $(D+M)^r$ cannot be expanded according to the binomial theorem. However, since D is diagonal, it is still true that, if we expand without commuting, any term which contains more than $n-1$ M's is zero. There are exactly $\binom{r}{q}$ terms involving q M's and $r-q$ D's for $q = 0, 1, 2, \ldots, n-1$. Taking the σ-norm and using $\sigma(D) = \lambda_A$, (2.2) and (2.3) follow with m replaced by $\sigma(M)$. Since the result is true for every M, and since the expressions on the right of (2.2) and (2.3) depend continuously on m, the conclusion of the theorem follows immediately.

3. Approximate solutions of linear systems

3.1. *Introduction.* Let A be a non-singular matrix, let b be a given vector, and let $\tilde{x}$ be an alleged solution of $Ax = b$. If we define the residual of $\tilde{x}$ by $r = A\tilde{x} - b$, and if φ is a vector norm, the error $\tilde{x} - A^{-1}b = A^{-1}r$ of $\tilde{x}$ can be estimated as follows:

$$\varphi(\tilde{x} - A^{-1}b) = \varphi(A^{-1}r) \leqq \nu(A^{-1})\,\varphi(r).$$

Here ν denotes a matrix norm compatible with φ. Similarly, if $\tilde{X}$ is an alleged inverse of A, and if ν is any matrix norm, we can calculate a bound for $\nu(\tilde{X}-A^{-1})$ in terms of the norm of the residual matrix $R = A\tilde{X} - I$ as follows:

$$\nu(\tilde{X} - A^{-1}) = \nu(A^{-1}R) \leqq \nu(A^{-1})\,\nu(R).$$

Thus, for both problems we require a bound for $\nu(A^{-1})$. Such a bound is, in principle, easily constructed if we assume that A is similar to a diagonal matrix D:

$$A = S D S^{-1}.$$

Assuming that the norm ν is axis-oriented, we readily find from $A^{-1} = SD^{-1}S^{-1}$ that

(3.1) $$\nu(A^{-1}) = c_\nu(S)\,\lambda_{A^{-1}},$$

where

(3.2) $$c_\nu(S) = \nu(S)\,\nu(S^{-1})$$

denotes the ν-condition number of S introduced by BAUER [2]. If A is normal, then S may be taken unitary, and the spectral condition number of S is 1. Thus

$$\sigma(A^{-1}) \leqq \lambda_{A^{-1}},$$

and in view of (1.1),

(3.3) $$\sigma(A^{-1}) = \lambda_{A^{-1}}$$

for normal matrices A. For non-normal matrices, the bound (3.1), if at all applicable, requires the complete diagonalization of the matrix A, a rather high price for the desired result.

3.2. *A new bound for the norm of the inverse.* For a real variable $x \geqq 0$, let the function f be defined by

(3.4) $$f(x) = x + x^2 + \cdots + x^n.$$

We note that both f and $x^{-1}f$ are monotonically increasing for $x > 0$, and that

(3.5) $$\lim_{x \to 0+} x^{-1} f(x) = 1.$$

We can now state the following Theorem:

Theorem 3. *If A is non-singular and non-normal, and if $x = \lambda_{A^{-1}} \Delta_\sigma(A)$, then*

(3.6) $$\sigma(A^{-1}) \leqq \frac{f(x)}{x}\,\lambda_{A^{-1}}.$$

Remark. For $\lambda_{A^{-1}}$ fixed, as $\Delta_\sigma(A) \to 0$ the bound (3.6) approaches (3.3), in view of (3.5). Bounds for $\nu(A^{-1})$ where ν is arbitrary can be found from (1.2).

Proof. Let $D+M$ be a Schur triangular form of A. Then, for some unitary matrix U,

$$\begin{aligned} A^{-1} &= U(D+M)^{-1}U^* \\ &= U(I + D^{-1}M)^{-1}D^{-1}U^*. \end{aligned}$$

In view of (1.3),

$$(I + D^{-1}M)^{-1} = I - D^{-1}M + \cdots + (-1)^{n-1}(D^{-1}M)^{n-1}.$$

The result (3.6) now follows by taking norms in view of $\sigma(D^{-1}) = \lambda_{A^{-1}}$ and the fact that σ is unitarily invariant.

4. Spectral variation and eigenvalue variation

4.1. *Classical results.* Let the matrix $A=(a_{ij})$ have eigenvalues λ_i, and let $B=(b_{ij})$ have eigenvalues μ_i $(i=1, 2, \ldots, n)$. The quantity

$$s = s_A(B) = \max_{1\leq i\leq n}\{\min_{1\leq j\leq n}|\mu_i - \lambda_j|\} \tag{4.1}$$

is called the *spectral variation* of B with respect to A. No one-to-one correspondence between the eigenvalues of A and those of B is implied. It can be seen from simple examples that $s_A(B) \neq s_B(A)$ in general.

We shall also consider the quantity

$$v = v(A, B) = \min_{\pi}\{\max_{1\leq i\leq n}|\lambda_i - \mu_{\pi(i)}|\} \tag{4.2}$$

called the *eigenvalue variation* of A and B. Here the minimum is taken with respect to all permutations π of the set $(1, 2, \ldots, n)$. Clearly, $v(A, B) = v(B, A)$, and

$$s_A(B) \leq v(A, B) \tag{4.3}$$

for all matrices A and B.

The best available bounds for s and v, due to OSTROWSKI [21] (see also [23], p. 192), are as follows. If $M = \max_{1\leq i,j\leq n}(|a_{ij}|, |b_{ij}|)$, and if the norm α is defined as in section 1.1, then

$$s_A(B) \leq (n+2)\, M\left[\frac{\alpha(A-B)}{M}\right]^{1/n} \tag{4.4}$$

and

$$v(A, B) \leq 2n(n+2)\, M\left[\frac{\alpha(A-B)}{M}\right]^{1/n}. \tag{4.5}$$

It is easily seen by considering an example due to G. E. FORSYTHE (see [27], p. 405) that the exponent $1/n$ in these bounds cannot be improved in general. In special cases, however, improvements are possible. If A is similar to a diagonal matrix D,

$$A = S D S^{-1},$$

and if ν is any axis-oriented lub norm, then BAUER and FIKE [4] showed that

$$s_A(B) \leq c_\nu(S)\, \nu(A - B). \tag{4.6}$$

In particular, if A is normal, we find for any norm ν majorizing the spectral norm

$$s_A(B) \leq \nu(A - B). \tag{4.7}$$

For $\nu = \varepsilon$ and both A and B normal it follows from a result of HOFFMANN and WIELANDT [11] that (4.7) is even valid for the eigenvalue variation:

$$v(A, B) \leq \varepsilon(A - B). \tag{4.8}$$

For A and B real symmetric or hermitian this result has been used frequently [1].

4.2. *A new result on* $s_A(B)$. For arbitrary (diagonalizable) A the result (4.6) suffers from the fact that no sufficiently explicit estimate for the condition number $c_\nu(S)$ is given. The following theorem contains (in view of Theorem 1) an *explicit* estimate as general as (4.4) (and thus more general than (4.6)), which reduces to (4.7) for A normal.

It is convenient to define, for a real variable $y \geqq 0$, the function $g = g(y)$ as the (unique) non-negative solution of the equation

$$g + g^2 + \cdots + g^n = y.$$

The function g is the inverse of the function f defined in section 3.2. For later use we note the relations

$$\lim_{y \to 0+} y^{-1} g(y) = 1, \tag{4.9}$$

$$n^{-1} y \leqq g(y) \leqq y, \qquad 0 \leqq y \leqq n, \tag{4.10}$$

$$g(n) = 1, \tag{4.11}$$

$$(n^{-1} y)^{1/n} \leqq g(y) \leqq y^{1/n}, \qquad y \geqq n, \tag{4.12}$$

$$\lim_{y \to \infty} y^{-1/n} g(y) = 1. \tag{4.13}$$

Theorem 4. *Let A be a non-normal matrix, and let $B - A \neq 0$. If ν is any norm majorizing the spectral norm, and if*

$$y = \frac{\Delta_\nu(A)}{\nu(B-A)},$$

then

$$s_A(B) \leqq \frac{y}{g(y)} \nu(B - A). \tag{4.14}$$

Remarks. The relations (4.10), (4.11), (4.12) may serve to render the bound (4.14) more explicit. Relation (4.9) shows that for $\Delta_\nu(A) \to 0$, and $\nu(B-A)$ bounded away from zero the estimate (4.14) approaches (4.7). (However, we cannot obtain (4.8) in this manner.) Relation (4.13) shows that for a fixed non-normal A and for $B \to A$ the bound (4.14) is of the same order as (4.4); the numerical coefficient may be larger or smaller than the coefficient in (4.4), depending on the departure from normality of A.

Proof of Theorem 4. Let $A = U T U^*$, where $T = D + M$ is a Schur triangular form of A. Setting $E = B - A$,

$$U^* B U = B_1, \qquad U^* E U = F$$

we have

$$B_1 = D + M + F.$$

Let μ be an eigenvalue of B (or B_1) which is not an eigenvalue of A. The matrix $D + M - \mu I$ is then non-singular, and we have

$$\begin{aligned} 0 = \det(B_1 - \mu I) &= \det(D + M - \mu I + F) \\ &= \det(D + M - \mu I) \det[I + (D + M - \mu I)^{-1} F]. \end{aligned}$$

Since $\det(D + M - \mu I) \neq 0$, it follows that -1 is an eigenvalue of the matrix $(D + M - \mu I)^{-1} F$. By the fundamental inequality (1.1) it follows that

$$\sigma\big((D + M - \mu I)^{-1} F\big) \geqq 1$$

or, by (d), since $\sigma(F) = \sigma(E)$,

$$\sigma\big((D + M - \mu I)^{-1}\big) \geqq \frac{1}{\sigma(E)}. \tag{4.15}$$

We shall now estimate $\sigma\left((D+M-\mu I)^{-1}\right)$ from above. Since $D-\mu I$ is non-singular,

$$(4.16)\qquad (D+M-\mu I)^{-1}=[I+(D-\mu I)^{-1}M]^{-1}(D-\mu I)^{-1}.$$

Since $(D-\mu I)^{-1}$ is diagonal, the matrix $(D-\mu I)^{-1}M$ shares with M the property that the elements on and below the main diagonal are zero, and hence that

$$[(D-\mu I)^{-1}M]^r=0 \quad \text{for} \quad r\geqq n.$$

Hence

$$(4.17)\qquad \begin{aligned}[I+(D-\mu I)^{-1}M]^{-1}=I-(D-\mu I)^{-1}M+\cdots\\ +(-1)^{n-1}[(D-\mu I)^{-1}M]^{n-1}.\end{aligned}$$

We set for brevity

$$\sigma\left((D-\mu I)^{-1}\right)=p, \qquad \sigma(M)=m\,(\neq 0), \qquad \sigma(E)=e$$

and have from (4.16) and (4.17) that

$$\sigma\left((D+M-\mu I)^{-1}\right)\leqq p+p^2m+\cdots+p^n m^{n-1}=m^{-1}f(m\,p)$$

where f is defined by (3.4).

Combining the last result with (4.15) we find $f(mp)\geqq me^{-1}$ and thus, by the definition of g, $mp\geqq g(me^{-1})$ or

$$(4.18)\qquad \frac{1}{p}\leqq\frac{m}{g(m\,e^{-1})}\,.$$

Since σ is an axis-oriented lub norm,

$$p=\sigma\left((D-\mu I)^{-1}\right)=\max_{1\leqq j\leqq n}|\lambda_j-\mu|^{-1},$$

and

$$\frac{1}{p}=\min_{1\leqq j\leqq n}|\lambda_j-\mu|\,.$$

Thus, from (4.18),

$$(4.19)\qquad \min_{1\leqq j\leqq n}|\lambda_j-\mu|\leqq\frac{m}{g(m\,e^{-1})}\,.$$

This relation has been proved for an arbitrary eigenvalue $\mu=\mu_i$ of B which is not also an eigenvalue of A. It also holds trivially for μ_i's which *are* eigenvalues of A. Thus we get

$$\begin{aligned}s_A(B)&=\max_{1\leqq i\leqq n}\left\{\min_{1\leqq j\leqq n}|\lambda_j-\mu_i|\right\}\\ &\leqq\frac{m}{g(m\,e^{-1})}=\frac{y_0}{g(y_0)}\,\sigma(B-A), \qquad y_0=\frac{\sigma(M)}{\sigma(B-A)}\,.\end{aligned}$$

This is true for any choice of the Schur triangular form $D+M$. Since the function g is continuous, the statement of Theorem 4 follows for the special norm $\nu=\sigma$.

Let now ν be a norm majorizing σ. Since the function g is non-negative and monotonically increasing, we also have

$$(4.20)\qquad s_A(B)\leqq\frac{\sigma(M)}{g\left(\dfrac{\sigma(M)}{\nu(B-A)}\right)}=\nu(B-A)\,\frac{y_1}{g(y_1)}, \qquad y_1=\frac{\sigma(M)}{\nu(B-A)}\,.$$

Let $0 < y_1 < y_2$, and define $x_i = g(y_i)$, $i = 1, 2$. From the monotonicity of $x^{-1} f(x)$ it follows that

$$\frac{y_1}{g(y_1)} = \frac{f(x_1)}{x_1} < \frac{f(x_2)}{x_2} = \frac{y_2}{g(y_2)}.$$

Thus also the function $y[g(y)]^{-1}$ is monotonically increasing, and we get from (4.20), replacing $\sigma(M)$ by $\nu(M)$,

$$s_A(B) \leqq \frac{y_2}{g(y_2)} \nu(B - A), \qquad y_2 = \frac{\nu(M)}{\nu(B - A)}.$$

The complete statement of Theorem 4 now follows as above by the continuity of g.

4.3. *Numerical example.* We illustrate the numerical performance of the several bounds discussed above by estimating the spectral variation of the two matrices

$$A_1 = \begin{pmatrix} 1 & 10^{-4} \\ 0 & 1 \end{pmatrix}, \qquad A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 - 10^{-4} \end{pmatrix}$$

under the condition that

$$\varepsilon(B_i - A_i) \leqq 10^{-2}, \qquad i = 1, 2. \tag{4.21}$$

We set

$$s_i = \sup s_{A_i}(B_i), \qquad i = 1, 2,$$

where the supremum is taken with respect to all matrices B_i satisfying (4.21). The examples

$$B_1 - A_1 = \begin{pmatrix} 2^{-\frac{1}{2}} 10^{-2} & 0 \\ 0 & -2^{-\frac{1}{2}} 10^{-2} \end{pmatrix}$$

$$B_2 - A_2 = \begin{pmatrix} 0 & 0 \\ -10^{-2} & 0 \end{pmatrix}$$

show that

$$s_1 \geqq .00707, \qquad s_2 \geqq .10000.$$

Using $M = 1.02$, $p_{\alpha\varepsilon} = 2$, the bound (4.4) yields $s_i \leqq .57132$, $i = 1, 2$ in both cases.

The bound (4.6) is not applicable to A_1, since this matrix has a non-linear elementary divisor. The relation $A_2 = S D S^{-1}$ is satisfied for

$$S = \begin{pmatrix} \xi & \eta \\ 0 & -10^{-4} \eta \end{pmatrix},$$

where ξ and η are arbitrary non-zero constants. Choosing ξ and η such that the condition $c_\varepsilon(S)$ is mimimized, (4.6) yields

$$s_2 \leqq 20.00001.$$

Turning to (4.14), we find $\Delta_\varepsilon(A_1) = 10^{-4}$, $y = 10^{-2}$, $g(y) = .009902$, and hence

$$s_1 \leqq .010099.$$

For the diagonalizable matrix A_2, $\Delta_\varepsilon(A_2) = 1$, $y = 10^2$, $g(y) = 9.5125$, and hence

$$s_2 \leqq .105125.$$

It is seen that, in the example discussed above, the bounds given by (4.14) compare favorably with the other bounds. Incidentally, the example also shows that the fact that a matrix has a non-linear elementary divisor does not in itself mean that it will have a large spectral variation.

4.4. *Related results on $v(A, B)$.* For given matrices A, B satisfying the hypotheses of Theorem 4, let δ denote the quantity on the right of (4.14). The statement of the theorem then may be interpreted geometrically by saying that the spectrum of B is contained in the union L_δ of the discs

$$D_i = \{\lambda \mid |\lambda - \lambda_i| \leqq \delta\}, \qquad i = 1, 2, \ldots, n.$$

Since $\delta \to 0$ monotonically as $B \to A$, we may conclude by a well-known continuity argument (see e.g. [*21*]) that each component of L_δ contains as many eigenvalues of B as of A. From this fact we can obtain, again using a well-known argument (see especially the translator's note in [*22*]), the following result:

Theorem 5. *Let A, B, and ν satisfy the hypotheses of Theorem 4, and let y be defined as before. Then*

$$v(A, B) \leqq (2n - 1) \frac{y}{g(y)} \nu(B - A). \tag{4.22}$$

It should be noted that this result does not imply the Hoffmann-Wielandt formula (4.8).

4.5. *A result on approximate eigenvalues.* The following question is frequently considered in matrix computation: Let φ be a vector norm, let x be a vector such that $\varphi(x) = 1$, and let λ be a number such that $\varphi(Ax - \lambda x)$ is small. How close is λ to an eigenvalue of A? (This question was the starting point of WIELANDT's study [*28*] of inclusion domains.) In their paper [*3*], HOUSEHOLDER and BAUER prove a result which as an important special case contains the following: If A is similar to a diagonal matrix D, $A = SDS^{-1}$, and if ν_φ denotes the matrix norm induced by φ, then

$$|\lambda_i - \lambda| \leqq c_{\nu_\varphi}(S)\, \varphi(Ax - \lambda x) \tag{4.23}$$

for at least one eigenvalue λ_i. Here c_{ν_φ} denotes the ν_φ-condition number as defined in section 3.1. Again the above result suffers from the disadvantages that it holds only for diagonalizable matrices, and that no explicit estimate for the condition number is given.

In this direction MORRISON [*18*], by using a method not unlike that used in the proof of Theorem 4, obtained the following result.

Theorem 6 (D. D. MORRISON). *Let A be a non-normal matrix, let φ be a vector norm, and let ν be the matrix norm induced by φ. If x is a vector and λ is a complex number such that $\varphi(x) = 1$, $\varphi(Ax - \lambda x) \neq 0$, and if*

$$y = \frac{\Delta_\nu(A)}{\varphi(Ax - \lambda x)},$$

then there exists an eigenvalue λ_i of A such that

$$|\lambda_i - \lambda| \leqq \frac{y}{g(y)} \varphi(Ax - \lambda x). \tag{4.24}$$

5. Field of values

5.1. *Classical results.* The field of values $F(A)$ of a matrix A was defined in the introduction. A classical result due to TOEPLITZ [*26*] states that for A normal, $F(A)=H(A)$, where $H(A)$ denotes the convex hull of the eigenvalues of A. It is known that the field of values of a non-normal matrix is still convex (HAUSDORFF [*10*]), although it may extend beyond $H(A)$ ([*9*], [*26*]). Crude bounds for the field of values of non-normal matrices have been given by FARNELL [*5*] and PARKER [*24*]. These bounds do not reduce to the convex hull of the eigenvalues if the matrix approaches a normal matrix.

The precise equation of the boundary of the field of values of a non-normal matrix has been given by MURNAGHAN [*19*] and, in more explicit form, by KIPPENHAHN [*14*]. KIPPENHAHN [*14*] also gives bounds for diameter and area of the field of values.

5.2. *The distance of the boundary of the field of values from* $H(A)$. In this section we shall prove two results showing that for certain norms there is a simple connection between the maximum distance of the boundary of $F(A)$ from $H(A)$ and the departure from normality.

Theorem 7. *If ξ is a point of the field of values of a matrix A, then there exists a point η in the convex hull of the eigenvalues of A such that*

$$|\xi-\eta| \leqq \sqrt{\frac{1-n^{-1}}{2}}\,\Delta_\varepsilon(A). \tag{5.1}$$

The constant $[(1-n^{-1})/2]^{\frac{1}{2}}$ cannot be replaced by any smaller constant.*

Proof. Let $A=UTU^*$, where $T=D+M$ is a Schur triangular form of A. Let $\xi=x^*Ax$, where $x^*x=1$. Setting $U^*x=y$, we have $y^*y=1$ and

$$\xi=\eta+y^*My, \tag{5.2}$$

when $\eta=y^*Dy$ is a point in $H(A)$. Hence for this number η,

$$|\xi-\eta|=|y^*My|.$$

Setting $M=(m_{ij})$, $y=(y_i)$, we find by the Cauchy inequality

$$\begin{aligned}|y^*My|^2 &= \Big|\sum_{i<j} m_{ij}\bar{y}_i y_j\Big|^2 \\ &\leqq [\Delta_\varepsilon(A)]^2 \sum_{i<j} |y_i y_j|^2.\end{aligned}$$

Using Cauchy again, we find for the last factor the estimate

$$\begin{aligned}\sum_{i<j}|y_i|^2|y_j|^2 &= \frac{1}{2}\left\{\Big(\sum_{i=1}^n |y_i|^2\Big)^2-\sum_{i=1}^n |y_i|^4\right\} \\ &\leqq \frac{1}{2}\left\{\Big(\sum_{i=1}^n |y_i|^2\Big)^2-\frac{1}{n}\Big(\sum_{i=1}^n |y_i|^2\Big)^2\right\} \\ &= \frac{1}{2}\left(1-\frac{1}{n}\right).\end{aligned}$$

* The bound (5.1) has been found independently by P. J. EBERLEIN (oral communication).

We thus find

$$|y^* M y| \leqq \sqrt{\frac{1-n^{-1}}{2}}\, \Delta_\varepsilon(A),$$

proving (5.1).

In order to show that the equality is attained, consider any matrix of the form $A=\lambda I+M$, where λ is a scalar and $m_{ij}=c$, $i<j$, where $c>0$. Here $\Delta_\varepsilon(A)=\varepsilon(M)=[\frac{1}{2}n(n-1)]^{\frac{1}{2}}c$. The convex hull $H(A)$ reduces to the point λ. Choosing $y_i=n^{-\frac{1}{2}}$ $(i=1,\ldots,n)$ we find that the point y^*Ay of $F(A)$ has the distance

$$y^* M y=\sum_{i<j}\frac{c}{n}=\frac{c}{2}(n-1)=\sqrt{\frac{1-n^{-1}}{2}}\,\Delta_\varepsilon(A),$$

from $H(A)$, as desired.

A similar result also holds for the norm α.

Theorem 8. *With the notation of Theorem 7 the following inequality also holds:*

$$(5.3)\qquad |\xi-\eta|\leqq\tfrac{1}{2}\Delta_\alpha(A).$$

The constant $\frac{1}{2}$ cannot be replaced by any smaller constant.

Proof. We again represent ξ in the form (5.2), but now estimate y^*My as follows:

$$\begin{aligned}|y^* M y| &= \Big|\sum_{i<j} m_{ij}\bar{y}_i y_j\Big|\\ &\leqq \alpha(M)\cdot\max_{i<j}|\bar{y}_i y_j|.\end{aligned}$$

Since $\sum_{i=1}^{n}|y_i|^2=1$, $\max_{i<j}|\bar{y}_i y_j|$ becomes largest, if the vector y has exactly two non-zero components. By the inequality between the arithmetic and the geometric mean, the value of the maximum then is $\frac{1}{2}$. Thus we get

$$(5.4)\qquad |y^* M y|\leqq\tfrac{1}{2}\alpha(M).$$

The inequality (5.3) now follows in view of the fact that (5.4) is true for all M that can appear in a Schur triangular form at A.

In order to show that equality is attained in (5.3) for non-zero $\Delta_\alpha(A)$, let A be any matrix of the special form $A=\lambda I+M$, where λ is a scalar, $m_{12}=m>0$, and all other $m_{ij}=0$. As above, $H(A)$ reduces to the point λ. The unit vector y defined by $y_1=y_2=2^{-\frac{1}{2}}$, $y_i=0$, $i=3,\ldots,n$ yields a point y^*Ay of $F(A)$ for which the distance from $H(A)$ is

$$(5.5)\qquad y^* M y=\tfrac{1}{2}m=\tfrac{1}{2}\alpha(M).$$

Let now $M'=(m'_{ij})$ the off-diagonal part of any other Schur triangular form of A. In view of

$$\alpha(M')\geqq\varepsilon(M')=\varepsilon(M)=m,$$

we have $\Delta_\alpha(A)=m$, and (5.5) can be replaced by

$$y^* M y=\tfrac{1}{2}\Delta_\alpha(A),$$

showing that equality can hold in (5.3).

5.3. *An improved bound for the field of values.* An elementary computation (see also [*14*], [*19*]) shows that for a 2×2 matrix with Schur triangular form

$$T=\begin{pmatrix}\lambda_1 & m_{12}\\ 0 & \lambda_2\end{pmatrix}$$

the field of values is given by the interior and boundary of the ellipse with foci λ_1 and λ_2 and minor semiaxis $\frac{1}{2}|m_{12}|$. Only two points of the field of values (namely, the end points of the minor axis) attain the maximum distance $\frac{1}{2}|m_{12}|$ from $H(A)$ given by the Theorems 7 and 8. All other points have a smaller distance.

The above fact is generalized to matrices of arbitrary order by Theorem 9. Here we denote, for complex λ and μ and real $c\geqq 0$, by $E(\lambda,\mu;c)$ the compact point set in the complex plane bounded by the ellipse with foci λ,μ and minor semiaxis c.

Theorem 9. *Let λ_i, $i=1,2,\ldots,n$ be n complex numbers, and let $a\geqq 0$. Then the field of values of any matrix A with eigenvalues λ_i and $\Delta_\alpha(A)=a$ is contained in the convex hull H_a of the $n(n-1)/2$ sets*

$$E(\lambda_i,\lambda_j,\tfrac{1}{2}a),\qquad 1\leqq i<j\leqq n.$$

Proof. For a matrix A satisfying the hypothesis, let ξ, y_ι, D, and $M=(m_{ij})$ be defined as in the proof of Theorem 7. We set

$$m=\alpha(M)=\sum_{i<j}|m_{ij}|.$$

If $m=0$, the theorem reduces to the theorem by Toeplitz mentioned in § 5.1. We may therefore suppose $m>0$. Setting $\vartheta_{ij}=m_{ij}/|m_{ij}|$ for $m_{ij}\neq 0$, $\vartheta_{ij}=0$ for $m_{ij}=0$, we obtain the representation

$$\begin{aligned}\xi&=\sum_{k=1}^{n}\lambda_k|y_k|^2+\sum_{i<j}m_{ij}\bar{y}_i y_j\\&=\sum_{i<j}|m_{ij}|\left\{\vartheta_{ij}\bar{y}_i y_j+\frac{1}{m}\sum_{k=1}^{n}\lambda_k|y_k|^2\right\}\\&=\frac{1}{m}\sum_{i<j}|m_{ij}|\left\{m\,\vartheta_{ij}\bar{y}_i y_j+\sum_{k=1}^{n}\lambda_k|y_k|^2\right\},\end{aligned}$$

showing that ξ belongs to the convex hull of the points

$$\eta_{ij}=m\,\vartheta_{ij}\bar{y}_i y_j+\sum_{k=1}^{n}\lambda_k|y_k|^2,\qquad 1\leqq i<j\leqq n.$$

We shall show next that the points η_{ij} belong to H_m, the convex hull of the sets

$$E(\lambda_k,\lambda_l;\tfrac{1}{2}m),\qquad 1\leqq k<l=n.$$

This is obvious if $|y_i|^2+|y_j|^2=0$. If $|y_i|^2+|y_j|^2=1$, then $y_k=0$, $k\neq i,j$, and η_{ij} belongs to $E(\lambda_i,\lambda_j;\frac{1}{2}m)$ and thus to H_m. If $0<|y_i|^2+|y_j|^2<1$, we define the variables

$$z_k=\frac{y_k}{\sqrt{|y_i|^2+|y_j|^2}},\qquad k=i,j,$$

$$z_k=\frac{y_k}{\sqrt{1-|y_i|^2-|y_j|^2}},\qquad k\neq i,j$$

and can write

$$\eta_{ij} = \left(|y_i|^2 + |y_j|^2\right)\left(\lambda_i |z_i|^2 + \lambda_j |z_j|^2 + \vartheta_{ij} m \bar{z}_i z_j\right) + \left(1 - |y_i|^2 - |y_j|^2\right) \sum_{\substack{k=1 \\ k \neq i,j}}^{n} \lambda_k |z_k|^2 .$$

The last relation shows that η_{ij} belongs to the convex hull of the set $E(\lambda_i, \lambda_j; \frac{1}{2}m)$ and the points λ_k $(k \neq i, j)$, which all belong to H_m. Thus again η_{ij} belongs to H_m. It follows that also ξ belongs to H_m.

The above is true for any matrix M that can appear in a Schur triangular form of A. Thus ξ is in the intersection of all possible sets H_m. Since the sets H_m are closed, and since they depend continuously on m, it follows that ξ belongs to $H_{\inf m}$, that is, to H_a.

A consideration similar to that used in proving that equality can hold in (5.3) shows that for every point ξ contained in a set $E(\lambda_k, \lambda_l; \frac{1}{2}m)$, a matrix A with eigendiagonal D and $\Delta_\alpha(A) = m$ can be constructed such that $\xi \in F(A)$. (It suffices to make $m_{kl} = m$, $m_{ij} = 0$ for $|i-k| + |j-l| > 0$ in the off-diagonal part of the Schur triangular form.) However, an example suggested to the author by W. G. Strang* shows that the union of the fields of values of all matrices A with given eigenvalues λ_i $(i = 1, 2, \ldots, n)$ and given $\Delta_\alpha(A) = m$ does not always fill out the set H_m.

References

[1] Bargmann, V., D. Montgomery and J. von Neumann: Solution of linear systems of high order. Report prepared under contract NORD 9596 with the Bureau of Ordnance, Navy Department, Institute for Advanced Study, 1946.

[2] Bauer, F. L.: On the definition of condition numbers and their relation to closed methods for solving linear systems. Proceedings of the International Conference on Information Processing, Unesco, Paris 1960, pp. 109–110.

[3] —, and A. S. Householder: Moments and characteristic roots. Numerische Math. **2**, 42–53 (1960).

[4] —, and C. T. Fike: Norms and exclusion theorems. Numerische Math. **2**, 137–141 (1960).

[5] Farnell, A. B.: Limits for the field of values of a matrix. Amer. Math. Monthly **52**, 488–493 (1945).

[6] Forsythe, G. E.: Singularity and near singularity in numerical analysis. Amer. Math. Monthly **65**, 229–240 (1958).

[7] Gautschi, We.: The asymptotic behavior of powers of matrices. Duke Math. J. **20**, 127–140 (1953).

[8] — The asymptotic behavior of powers of matrices. II. Duke Math. J. **20**, 375–379 (1953).

[9] Givens, W.: Fields of values of a matrix. Proc. Amer. Math. Soc. **3**, 206–209 (1952).

[10] Hausdorff, E.: Der Wertevorrat einer Bilinearform. Math. Z. **3**, 314–316 (1919).

[11] Hoffmann, A. J., and H. W. Wielandt: The variation of the spectrum of a normal matrix. Duke Math. J. **20**, 37–39 (1953).

[12] Householder, A. S.: The approximate solution of matrix problems. J. Assoc. Comput. Mach. **5**, 204–243 (1958).

[13] Kato, T.: Estimation of iterated matrices, with application to the von Neumann condition. Numerische Math. **2**, 22–29 (1960).

[14] Kippenhahn, R.: Über den Wertevorrat einer Matrix. Math. Nachrichten **6**, 193–228 (1951).

* Oral communication.

[15] LAX, P. D., and R. D. RICHTMYER: Survey of the stability of linear finite difference equations. Comm. Pure Appl. Math. **9**, 267—293 (1956).
[16] MCRAE, V. V.: On the unitary similarity of matrices. Ph. D. Dissertation, Catholic University of America, 1955.
[17] MIRSKY, L.: An introduction to linear algebra. Oxford: Oxford University Press 1955.
[18] MORRISON, D. D.: Errors in the solution of eigenvalue problems by finite difference methods. Ph. D. Dissertation, University of California, Los Angeles, 1961.
[19] MURNAGHAN, F. D.: On the field of values of a square matrix. Proc. Nat. Acad. Sci. U.S.A. **18**, 246—248 (1932).
[20] OSTROWSKI, A.: Über Normen von Matrizen. Math. Z. **63**, 2—18 (1955).
[21] — Über die Stetigkeit von charakteristischen Wurzeln in Abhängigkeit von den Matrizenelementen. Jahresber. Dtsch. Math. Verein. **60**, Abt. 1, 40—42 (1957).
[22] — On the continuity of characteristic roots in their dependence on the matrix elements (translation of [21]). Translated from the German and annotated by G. E. BUMP. Technical Report No. 2, Project 225(37), Applied Mathematics and Statistics Laboratories, Stanford University, 1959.
[23] — Solution of equations and systems of equations. New York and London: Academic Press 1960.
[24] PARKER, W. V.: Characteristic roots and fields of values of a matrix. Bull. Amer. Math. Soc. **57**, 103—108 (1951).
[25] STONE, B. J.: Best possible ratios of certain matrix norms. Numer. Math., to appear.
[26] TOEPLITZ, O.: Das algebraische Analogon zu einem Satz von FEJÉR. Math. Z. **2**, 187—197 (1918).
[27] WHITE, P. A.: The computation of eigenvalues and eigenvectors of a matrix. J. Soc. Indust. Appl. Math. **6**, 393—437 (1958).
[28] WIELANDT, H.: Inclusion theorems for eigenvalues. Nat. Bur. Standards Appl. Math. Ser. **29**, 75—78 (1951).

Department of Mathematics
University of California at Los Angeles
Los Angeles 24, California

(Received March 9, 1961)