Off-Diagonal Elements of Normal Matrices*

Robert Grone**

(February 4, 1977)

Let \( A \) be an \( n \times n \) complex matrix, and let \( W(A) \) denote the numerical range of \( A \). In this paper, results of Parker and Mirsky are shown to be a consequence of the following fact. If \( A \) is normal, then the maximum value of \( |(Ax, y)| \) as \( x \) and \( y \) run over all orthonormal pairs in \( C^n \) coincides with the radius of the smallest closed disk in \( C \) which contains \( W(A) \).

Key words: Constrained extrema of quadratic forms; numerical range.

1. Introduction

For an \( n \times n \) complex matrix, Parker [1] has proven that the maximum value, for fixed \( A \), of

\[
|\langle Ax, y \rangle| \tag{1}
\]

subject to the constraint \( ||x||=||y||=1 \) is the maximum singular value, or Hilbert norm, \( s_1(A) \) of \( A \). In fact, it is clear that the maximum value in (1) occurs exactly when \( x \) is a unit eigenvector in \( C^n \) corresponding to the maximum eigenvalue of \( A^*A \), and \( y=Ax/||Ax|| \).

If in (1) we require that \( x=y \), then the maximum value of (1) is the numerical radius of \( A \), \( r(A) \). A moment's reflection yields that

\[
r(A)=\max_{U\text{ unitary}} |(U^*AU)_{1,1}|. \tag{2}
\]

Relations between \( s_1(A) \) and \( r(A) \) have been studied. See for example [2,3].

Parker in [1] also investigated the maximum value of (1) when \( A \) is Hermitian and \( x,y \) range over all orthonormal pairs. In this case, the maximum value of (1) is \( 1/2(\lambda_1(A)-\lambda_n(A)) \), where \( \lambda_1(A) \geq \ldots \geq \lambda_n(A) \) are the eigenvalues of \( A \). Mirsky later obtained [4] that if \( A \) is normal with eigenvalues \( \lambda_1, \ldots, \lambda_n \), then

\[
\max_{x, y \text{ orthonormal}} |\langle Ax, y \rangle| \leq \max_{i,j} |\lambda_i-\lambda_j| \cdot \frac{1}{\sqrt{3}}. \tag{3}
\]

We remark that for any \( A \), the left-hand side of (3) is identical with

\[
\max_{U \text{ unitary}} |(U^*AU)_{1,2}|. \tag{4}
\]

In this paper we completely characterize this maximum value for \( A \), and obtain the results of Parker and Mirsky as corollaries.

2. Notation

(i) \( A \) will denote an \( n \times n \) complex matrix with eigenvalues \( \lambda_1, \ldots, \lambda_n \).

(ii) \( W(A)=\{ \langle Ax, x \rangle | ||x||=1 \} \) is the numerical range of \( A \). It is well known that if \( A \) is normal, then \( W(A)=H(\lambda_1, \ldots, \lambda_n) \), the convex hull of \( \lambda_1, \ldots, \lambda_n \).

*An invited paper. The work of this author was supported in part by the Air Force Office of Scientific Research.

**Present address: Department of Mathematics, University of California, Santa Barbara, Calif. 93106.

1 Figures in brackets indicate the literature references at the end of this paper.
(iii) \( S(A) \) will denote \( \{(Ax,y)|x,y \text{ orthonormal}\} \). A theorem of Marcus and Robin [5] states that \( S(A) \) is a closed disk centered at 0 in the complex plane. We will use the notation \( \rho(A) \) for the radius of \( S(A) \).

The following lemma is an immediate consequence of definitions (ii) and (iii).

**Lemma:** Let \( \alpha, \beta \in \mathbb{C} \). Then

(i) \( S(\alpha A + \beta I_n) = |\alpha| S(A) \)

(ii) \( W(\alpha A + \beta I_n) = \alpha W(A) + \beta \).

We are now in a position to state the following:

**Theorem:** If \( A \) is normal, then \( \rho(A) \) is the radius of the smallest closed disc containing \( W(A) \).

**Proof:** By the lemma, there is no loss of generality in assuming that the smallest closed disk containing \( W(A) \) is the unit circle. Since \( A \) is normal this implies that the Hilbert norm of \( A \) is 1, and by Cauchy-Schwarz we have

\[
\max_{x,y \text{ orthonormal}} |(Ax,y)| \leq \max_{||x||=1} ||Ax|| = 1.
\]

Hence \( \rho(A) \leq 1 \), and to complete the proof we need only exhibit an orthonormal pair \( x,y \) for which \( |(Ax,y)| = 1 \). Since \( W(A) \) is closed and convex, and the unit disk is the (unique) smallest disk containing \( W(A) \), we must have at least one of the following holding:

(i) There exist extreme points \( \lambda_1, \lambda_2 \) of \( W(A) \) such that \( \lambda_1 \lambda_2 \) is a diameter of the unit disk.

(ii) There exist extreme points \( \lambda_1, \lambda_2, \lambda_3 \) of \( W(A) \) such that the triangle \( \triangle(\lambda_1, \lambda_2, \lambda_3) \) is an acute triangle inscribed in the unit circle.

Since \( A \) is normal, these extreme points of \( W(A) \) are eigenvalues of \( A \). If (i) holds, then we have eigenvalues \( \lambda_1, -\lambda_1 \) of \( A \) with \( |\lambda_1| = 1 \). Since \( A \) is normal there exists an orthonormal pair of eigenvectors \( u, v \) corresponding to \( \lambda_1, -\lambda_1 \) respectively. The pair

\[
x = \frac{u+v}{\sqrt{2}}, \quad y = \frac{u-v}{\sqrt{2}}
\]

is orthonormal, and we calculate that

\[
|(Ax, y)| = \frac{1}{2} |\lambda_1 - (-\lambda_1)| = 1,
\]

and the proof is finished in this case.

If case (ii) holds, let \( u, v, w \) be a set of orthonormal eigenvectors corresponding to \( \lambda_1, \lambda_2, \lambda_3 \) respectively. Also, let \( a, b, c \) be the unique positive numbers which satisfy

\[
a + b + c = 1
\]

\[
a \lambda_1 + b \lambda_2 + c \lambda_3 = 0.
\]

Then, let

\[
x = \sqrt{a} u + \sqrt{b} v + \sqrt{c} w.
\]

\[
y = \sqrt{a} \lambda_1 u + \sqrt{b} \lambda_2 v + \sqrt{c} \lambda_3 w.
\]

Again, trivial calculations yield that \( x, y \) is an orthonormal pair, and that \( (Ax, y) = 1 \).

**Corollary 1:** If \( W(A) \) is a line segment with endpoints \( \lambda_1, \lambda_n \), then \( \rho(A) = \frac{1}{2} |\lambda_1 - \lambda_n| \).
**Proof:** If $W(A)$ is a line segment, then $A$ is essentially Hermitian and we may apply the theorem.

**Corollary 2:** Let $A$ be a normal matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then

$$\frac{\sqrt{3}}{2} \rho(A) \leq \max_{i,j} \frac{|\lambda_i - \lambda_j|}{2} \leq \rho(A).$$

**Proof:** This follows from the theorem and the geometric fact that if a non-obtuse triangle with longest side $s$ is inscribed in a circle of radius 1, then $\sqrt{3} \leq s \leq 2$.

In light of the theorem it is of interest to be able to compute the radius of the smallest circle containing a triangle $\Delta$ with sides of length $a$, $b$, and $c$. For non-obtuse triangles we may use the following proposition. The proof, which is omitted, uses the formula for $\cos(\alpha + \beta)$ and the law of cosines.

**Proposition 1:** Let the triangle $\Delta$ have sides of length $a$, $b$, $c$. Then the radius of the circle circumscribed about $\Delta$ is

$$abc \overline{E_2(a^2, b^2, c^2)} - (a^4 + b^4 + c^4),$$

where $E_2$ denotes the second elementary symmetric function.

The theorem may be restated as follows: "For $A$ normal,

$$\rho(A) = \min_{\alpha} ||A - \alpha I_n|| = \min_{\alpha} \max_{||x|| = 1} ||Ax - \alpha x||."

The question of whether this result could hold in general is put in a somewhat different perspective by the following proposition.

**Proposition 2:** For any $A$,

$$\rho(A) = \max_{||x|| = 1} \min_{\alpha} ||Ax - \alpha x||.$$

**Proof:**

$$\rho(A) = \max_{x, y \text{ o.n.}} |(Ax, y)|$$

$$= \max_{||x|| = 1} \max_{||y|| = 1} |(Ax, y)|$$

$$= \max_{||x|| = 1} ||Ax - (Ax, x)x||$$

$$= \max_{||x|| = 1} \min_{\alpha} ||Ax - \alpha x||.$$

Certainly we have

$$\max_{||x|| = 1} \min_{\alpha} ||Ax - \alpha x|| \leq \min_{\alpha} \max_{||x|| = 1} ||Ax - \alpha x||,$$

and the question of whether equality holds remains open.
Finally it should be noted that the theorem is false in general, even for convexoid $A$ ($A$ is convexoid if $W(A) = \mathcal{H}(\lambda_1, \ldots, \lambda_n)$), as the following matrix shows. Let

\[
A = \begin{bmatrix}
1+i & 0 & 0 & 0 & 0 \\
0 & 1-i & 0 & 0 & 0 \\
0 & 0 & -1+i & 0 & 0 \\
0 & 0 & 0 & -1-i & 0 \\
0 & 0 & 0 & 0 & 2 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

Then the radius of the smallest circle containing $W(A)$ is $\sqrt{2}$, while $\rho(A) = ||A|| = 2$.

3. References


(Paper 81B 1 & 2-460)