Normality and the Numerical Range

Charles R. Johnson*
Institute for Fluid Dynamics and Applied Mathematics
University of Maryland
College Park, Maryland 20742
and
Applied Mathematics Division
National Bureau of Standards
Washington, D. C. 20234

Submitted by Hans Schneider

ABSTRACT

It is well known that if A is an n by n normal matrix, then the numerical range of A is the convex hull of its spectrum. The converse is valid for $n \leqslant 4$ but not for larger n. In this spirit a characterization of normal matrices is given only in terms of the numerical range. Also, a characterization is given of matrices for which the numerical range coincides with the convex hull of the spectrum. A key observation is that the eigenvectors corresponding to any eigenvalue occurring on the boundary of the numerical range must be orthogonal to eigenvectors corresponding to all other eigenvalues.

We shall assume throughout that $A \in M_n(C)$, the n by n complex matrices. We denote by $\sigma(A)$ the spectrum, or set of eigenvalues, of A, and by

$$W(A) \equiv \{x^*Ax : x^*x = 1, x \in C^n\}$$

the numerical range of A. It is well known that $W(A)$ is a closed, bounded convex subset of the complex plane, and we shall denote its boundary by $\partial W(A)$. For any subset S of the complex plane, we let Co (S) be the closed convex hull of S.

*A portion of the preparation of this manuscript was carried out while the author was a Visiting Staff Member with the T-7 Group at the Los Alamos Scientific Laboratories, Los Alamos, New Mexico, under ERDA Contract W-7405-ENG-36.

It is well known that if A is normal, then $W(A)=\mathrm{Co}\ (\sigma(A))$. It is noted in [5] that the converse is valid for $n \leqslant 4$. The value 4 is critical, since for $n>4$, counterexamples to the converse may be obtained via the direct sum of B_1 and B_2, where $B_1 \in M_3(C)$ is normal with three noncollinear eigenvalues and $B_2 \in M_k(C)$, $k \geqslant 2$, is not normal and satisfies $W(B_2) \subseteq W(B_1)$. Then $W(B_1 \oplus B_2) = W(B_1) = \mathrm{Co}\ (\sigma(B_1 \oplus B_2))$, but $B_1 \oplus B_2$ is not normal. It follows from the observations of this note that, except for unitary equivalence, all counterexamples are of this type. We also give a precise characterization of normal matrices in terms only of the numerical range. The primary observation, contained in Theorem 1, is that eigenvectors belonging to any eigenvalue lying on $\partial W(A)$ are orthogonal to eigenvectors of all other eigenvalues.

We assume the reader is familiar with most elementary facts concerning the numerical range (e.g. see [2, 4, 6]) and we shall make frequent use of them. In addition, the following sequence of more special lemmas will be of assistance.

LEMMA 1 [1,3]. *If $A \in M_2(C)$, then $W(A)$ is a (possibly degenerate) ellipse whose foci are the eigenvalues of A. Furthermore, $W(A)$ is a line segment if and only if A is normal.*

LEMMA 2. *If $A \in M_2(C)$ satisfies*

$$\partial W(A) \cap \sigma(A) \neq \phi,$$

then A is normal.

Proof. Since $W(A)$ is an ellipse at least one of whose foci occurs on the boundary [by virtue of the assumption $\partial W(A) \cap \sigma(A) \neq \phi$], it follows that $W(A)$ is a line segment (or, perhaps, a point) and that A is normal. ∎

LEMMA 3. *If $A=\begin{pmatrix} \alpha & x \\ y & \beta \end{pmatrix}$ is normal and $\alpha \in \sigma(A)$, then $x=y=0$.*

Proof. Since $\alpha \in \sigma(A)$ and $\mathrm{Tr}\ (A)=\alpha+\beta$, it follows that $\sigma(A)=\{\alpha,\beta\}$. Therefore, $\det A=\alpha\beta$, so that $xy=0$, i.e., $x=0$ or $y=0$, and A is triangular. Then the normality of A and a standard computation show that A is diagonal and $x=y=0$. ∎

LEMMA 4 [3]. *For any $A \in M_n(C)$, $\sigma(A) \subseteq W(A)$. Furthermore, if A_0 is a principal submatrix of A, then $W(A_0) \subseteq W(A)$.*

We may now give our first main result which concerns eigenvalues lying on $\partial W(A)$.

THEOREM 1. *Suppose* $\alpha \in \partial W(A) \cap \sigma(A)$ *and* $Ax = \alpha x$, $x^*x = 1$. *We then have*: (i) *if* α *has algebraic multiplicity* m *in* $\sigma(A)$, *then the dimension of the eigenspace for* α *is* m; (ii) *for any* $\alpha \neq \lambda \in \sigma(A)$, $Ay = \lambda y$, $y^*y = 1$, *it follows that* $x^*y = 0$; *and* (iii) A *is unitarily equivalent to* $\alpha I \oplus B$ *where* $\alpha \notin \sigma(B)$.

Proof. By a unitary equivalence, we may assume without loss of generality that

$$x = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad A = \left[\begin{array}{c|c} \alpha & a_1 \\ \hline a_2 & A_1 \end{array}\right].$$

From $Ax = \alpha x$, it follows that $a_2 = 0$. Suppose now that

$$A_2 = \begin{bmatrix} \alpha & z \\ 0 & \beta \end{bmatrix}$$

is a 2 by 2 principal submatrix of A determined by the indices 1 and i, $2 \leqslant i \leqslant n$. Clearly $\alpha \in \sigma(A_2)$, and, by Lemma 4, $\alpha \in \partial W(A_2)$ also. From Lemmas 2 and 3 it then follows that $z = 0$. We thus conclude that $a_1 = 0$ and

$$A = \left[\begin{array}{c|c} \alpha & 0 \\ \hline 0 & A_1 \end{array}\right].$$

To demonstrate (i) and (iii), suppose that now the multiplicity of α is $m \geqslant 2$. Then $\alpha \in \sigma(A_1)$, and, because of Lemma 4, $\alpha \in \partial W(A_1)$. Arguing on A_1 as we argued on A before, it follows that A is unitarily equivalent to

$$\left[\begin{array}{cc|c} \alpha & 0 & \multirow{2}{*}{0} \\ 0 & \alpha & \\ \hline \multicolumn{2}{c|}{0} & A_3 \end{array}\right].$$

Repeating this argument as often as necessary, we conclude that A is unitarily equivalent to

$$\left[\begin{array}{c|c} \alpha I_m & 0 \\ \hline 0 & A_{m+1} \end{array}\right],$$

from which it follows that the eigenspace for α has dimension m.

To demonstrate (ii) suppose $\lambda \neq \alpha$ is any other element of $\sigma(A)$ with associated eigenvector y. Clearly $\lambda \in \sigma(A_1)$, and, since $\lambda \neq \alpha$, the first component of y is 0. Thus $x^* y = 0$ (x and y are orthogonal). ■

Conclusion (iii) of the above theorem may also be deduced from Theorem 1 of [5], and a similar fact is employed in [7].

Definition. For $A \in M_n(C)$, we know that $\sigma(A) \subseteq W(A)$. We shall say that A satisfies the *boundary property* if $\sigma(A) \subseteq \partial W(A)$ also.

Corollary 1. *If A satisfies the boundary property, then A is normal.*

Proof. By Theorem 1, A has a complete set of orthogonal eigenvectors and is, therefore, normal. ■

We may now characterize normal matrices in terms of the numerical range.

Theorem 2. *A is normal if and only if A is unitarily equivalent to $A_1 \oplus \cdots \oplus A_k$, where A_i satisfies the boundary property, $i = 1, \ldots, k$.*

Proof. If A is normal, then A is unitarily equivalent to a diagonal matrix and therefore satisfies the stated condition. On the other hand, since normality is preserved by unitary equivalence and direct summation, the converse follows from Corollary 1. ■

We may next characterize elements of $M_n(C)$ for which $W(A)$ and Co $(\sigma(A))$ coincide. It will be seen that the conditions of Theorems 2 and 3 coincide for $n \leqslant 4$ only.

Theorem 3. *We have $W(A) = \text{Co}\,(\sigma(A))$ if and only if A is normal or A is unitarily equivalent to*

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

where A_1 is normal and $W(A_2)\subseteq W(A_1)$.

Proof. If A is unitarily equivalent to $A_1\oplus A_2$, then

$$W(A)=\mathrm{Co}\,(W(A_1)\cup W(A_2))=\mathrm{Co}\,(W(A_1))=\mathrm{Co}\,(W(A_1)\cup \sigma(A_2))$$
$$=\mathrm{Co}\,(\sigma(A_1)\cup\sigma(A_2))=\mathrm{Co}\,(\sigma(A)).$$

On the other hand, if $W(A)=\mathrm{Co}\,(\sigma(A))$, let $\{\lambda_1,\dots,\lambda_k\}$ be that subset of $\sigma(A)$ which determines $\partial W(A)$. If A is not normal, we must have by virtue of Theorem 1 that $3\leqslant k<n-1$. If we let $A_1=\mathrm{diag}\,\{\lambda_1,\dots,\lambda_k\}$, it also follows from Theorem 1 that A is unitarily equivalent to $A_1\oplus A_2$. Since $W(A_1)=\mathrm{Co}\,(\sigma(A_1))=W(A)=\mathrm{Co}\,(W(A_1)\cup W(A_2))$, it follows that $W(A_2)\subseteq W(A_1)$. Since A_1 is diagonal, it is normal, and the proof is complete. ■

It is an immediate corollary of Theorem 3 that

COROLLARY 2. *If $A\in M_n(C)$, $n\leqslant 4$, then A is normal if and only if $W(A)=\mathrm{Co}\,(\sigma(A))$.*

We also note that two other well known facts follow from the observations of this note. If $W(A)$ is a subset of the real line, then A is Hermitian (from Theorem 1). If $A\in M_n\,(C)$ and $\partial W(A)$ is not smooth at as many as $n-1$ points, then $W(A)$ is polygonal and A is normal.

We close with some observations regarding the vectors which give rise to elements of $W(A)$ lying on line segments joining elements of $\sigma(A)$ whose eigenvectors are orthogonal. For example, the numerical ranges of Hermitian matrices or, more generally, linear portions of $\partial W(A)$ which join vertices of $\partial W(A)$ are such line segments. We suppose

$$Ax=\alpha x,\qquad x^*x=1,$$
$$Ay=\beta y,\qquad y^*y=1,$$

and

$$x^*y=0.$$

We then know that the line segment

$$\theta\alpha+(1-\theta)\beta,\qquad 0\leqslant\theta\leqslant 1,$$

is contained in $W(A)$. Suppose that γ lies on this line segment. How may we give a solution z to

$$z^*Az=\gamma, \qquad z^*z=1?$$

If $\gamma=\theta_0\alpha+(1-\theta_0)\beta$, then

$$z=\theta_0^{1/2}x+(1-\theta_0)^{1/2}y$$

satisfies $z^*z=1$ and $z^*Az=\theta_0 x^*Ax+(1-\theta_0)y^*Ay+\theta_0^{1/2}(1-\theta_0)^{1/2}[x^*Ay+y^*Ax]=\theta_0\alpha+(1-\theta_0)\beta+\theta_0^{1/2}(1-\theta_0)^{1/2}[\beta x^*y+\alpha y^*x]=\gamma+\theta_0^{1/2}(1-\theta_0)^{1/2}[0]=\gamma$. Thus z is such a solution.

The general problem of solving for z, $z^*z=1$, in

$$z^*Az=\theta\alpha+(1-\theta)\beta, \qquad 0\leqslant\theta\leqslant 1,$$

where we only assume $\alpha,\beta\in W(A)$, is as far as the author knows unsolved and deserves further study.

REFERENCES

1 W. F. Donoghue, Jr., On the numerical range of a bounded operator, *Mich. Math. J.* 4 (1962), 261–263.

2 R. Kippenhahn, Über den Wertevorrat einer Matrix, *Math. Nachr.* **6** (1951), 193–228.

3 C. R. Johnson, Computation of the field of values of a 2×2 matrix, *J. Res. NBS* **78B** (1974), 105–108.

4 C. R. Johnson, Numerical determination of the field of values of a general complex matrix, in preparation.

5 B. Moyls and M. Marcus, Field convexity of a square matrix, *Proc. AMS* **6** (1955), 981–983.

6 F. Murnaghan, On the field of values of a square matrix, *Proc. NAS* (1932), 246–248.

7 G. Zwas and M. Goldberg, Inclusion relations between certain sets of matrices, Linear Multilinear Algebra, to be published.

Received 11 August 1975; revised 7 October 1975