PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 50, July 1975

# POWERS OF MATRICES WITH POSITIVE DEFINITE REAL PART[1]

CHARLES R. JOHNSON

ABSTRACT. For $n$ by $n$ complex matrices $A$ the following two facts are proven by elementary techniques: 1. If $A^m$ is never normal, $m \in I^+$, then the equation $xA^m x^* = 0$ has a solution $0 \neq x \in C^n$, $m \in I^+$; 2. If $H(A) = (A + A^*)/2$ is positive definite, then $H(A^m)$ is positive definite for all $m \in I^+$ if and only if $A$ is Hermitian.

Let $M_n(C)$ be the class of $n \times n$ complex matrices and denote by $\Sigma$ the class of Hermitian positive definite elements of $M_n(C)$. Define $H(A) = (A + A^*)/2$ the Hermitian part of $A$, so that $A$ is Hermitian if and only if $H(A) = A$. Let $\Pi = \{A \in M_n(C): H(A) \in \Sigma\}$.

The field of values of an element $A \in M_n(C)$ will be denoted by

$$F(A) = \{xAx^*: xx^* = 1,\ x \in C^n\}. \tag{1}$$

It is well known that

(2) $F(A)$ is convex and compact;

(3) the spectrum $\sigma(A) \subseteq F(A)$; and

(4) $F(UAU^*) = F(A)$ if $U$ is unitary.

When $A$ is normal $F(A)$ coincides with the convex hull of $\sigma(A)$. If $A_0$ is a principal submatrix of $A$, then

$$F(A_0) \subseteq F(A). \tag{5}$$

If $R$ denotes the right complex half-plane, $\mathrm{Re}(z) > 0$, then

(6) $A \in \Pi$ is equivalent to $F(A) \subset R$.

Since $F(A)$ is convex, the condition

Received by the editors March 12, 1973.

*AMS (MOS) subject classifications* (1970). Primary 15–00, 15A24, 15A63.

*Key words and phrases.* Hermitian matrix, positive definite, normal matrix, field of values.

[1]A portion of this work was developed in the author's doctoral thesis under the advisement of Professor Olga Taussky Todd at the California Institute of Technology.

$$0 \notin F(A) \tag{7}$$

is equivalent to the existence of a $\theta \in [0, 2\pi)$ such that

$$e^{i\theta} A \in \Pi \tag{8}$$

as well as to the nonexistence of a nonzero $x \in C^n$ such that

$$xAx^* = 0. \tag{9}$$

We shall prove the following two results concerning integral powers of matrices.

**Theorem 1.** *If no integral power of $A \in M_n(C)$ is normal, then the equation $xA^m x^* = 0$ has a solution $0 \neq x \in C^n$, for some positive integer $m$.*

**Theorem 2.** *Suppose $A \in \Pi$. Then $A^m \in \Pi$ for all $m \in I^+$ if and only if $A \in \Sigma$.*

Since $\Sigma$ is clearly closed under the raising of its elements to powers, it will suffice to establish the fact that *if $A$ is not Hermitian, then there is an $m$ such that $A^m \notin \Pi$*. Since $A \in \Pi$, the condition "$A \in \Sigma$" may also be stated: "$A$ is Hermitian".

The proof of Theorem 2 is our primary goal, but, as a step in the proof, we shall establish Theorem 1, a fact of independent interest. Theorem 1 equivalently gives a sufficient condition which implies that $0 \in F(A^m)$ for some $m$. The requirement that $m$ be positive in Theorems 1 and 2 is of no consequence. An easy computation shows that both conditions $A \in \Pi$ and $0 \in F(A)$ are inherited under inversion. The solvability of $x^* A^m x = 0$ is also considered in [2].

We shall require the aid of a few lemmas. By a straightforward calculation based upon [1] we obtain

**Lemma 1.** *If $A = \left[\begin{smallmatrix} a & c \\ 0 & b \end{smallmatrix}\right]$, $c \neq 0$, then $F(A)$ is* (i) *an ellipse with foci at $a$ and $b$ and eccentricity $e(A) = d/(1 + d^2)^{1/2}$ where $d = |(a - b)/c|$ if $a \neq b$, and* (ii) *a circle of radius $|c/2|$ about $a$ if $a = b$.*

**Lemma 2.** *If*

$$A = \begin{bmatrix} a & 0 & \cdots & 0 & c \\ * & 0 & & & * \\ & \ddots & & & \vdots \\ & & \ddots & & \vdots \\ 0 & & & * & * \\ & & & & b \end{bmatrix} \in M_k(C)$$

*where $a, b, c \neq 0$ and $b/a$ is not a nonreal root of unity, then there is a nonzero $x \in C^k$ and an $m \in I^+$ such that $xA^m x^* = 0$.*

**Proof.** Since the condition (9) on $A$ is invariant under nonzero scalar multiplication, we assume without loss of generality that $a = 1$ and $b$ is not a nonreal root of unity. An elementary induction on $m$ then shows that

$$A^m = \begin{bmatrix} 1 & 0 \cdots 0 & c\sum_{j=0}^{m-1} b^j \\ & * \; \ddots \; 0 & * \\ & \ddots & \vdots \\ & 0 \quad * & * \\ & & b^m \end{bmatrix}.$$

We shall show that under our assumptions $0 \in F(A_0^m)$ for some $m \in I^+$ where

$$A_0^m = \begin{bmatrix} 1 & c\sum_{j=0}^{m} b^j \\ 0 & b^m \end{bmatrix}$$

is a $2 \times 2$ principal submatrix of $A^m$. By (5) it then follows that $0 \in F(A^m)$ which is equivalent by (7) and (9) to what we wish to show.

If $b = -1$, then $0 \in F(A_0)$ because of (2) and (3), and we would be finished. If $b = 1$, then

$$A_0^m = \begin{bmatrix} 1 & cm \\ 0 & 1 \end{bmatrix},$$

and by Lemma 1, $F(A_0^m)$ is a circle about 1 of radius $|cm/2|$. Since $c \neq 0$, $|cm/2| \to \infty$ as $m \to \infty$, so that for some $m$ the circle $F(A_0^m)$ would be large enough to capture the origin. In this event we would again be done. Thus we may simply assume that $b \neq 0$ is not a root of unity, and we then have

$$A_0^m = \begin{bmatrix} 1 & c\left(\dfrac{b^m - 1}{b - 1}\right) \\ 0 & b^m \end{bmatrix}.$$

Then by Lemma 1, $F(A_0^m)$ is an ellipse with foci at 1 and $b^m$ and eccentricity

$$e(A_0^m) = \left|\frac{b-1}{c}\right| \left(1 + \left|\frac{b-1}{c}\right|^2\right)^{-1/2} = e(A_0).$$

Since $0 \in F(A_0^m)$ if and only if $0 \in F(A_0^{-m})$, we may assume without loss of generality that $|b| \geq 1$. We thus distinguish two cases: $|b| > 1$ and $|b| = 1$.

*Case* I. Assume $|b| > 1$. Then $b^m$ grows arbitrarily far from 1 while $0 < e(A_0^m) < 1$ is a fixed positive constant. Thus the ellipse $F(A_0^m)$ grows arbitrarily large. Since its eccentricity is fixed and one of its foci is fixed at 1, $F(A_0^m)$ must include any fixed point (in particular 0) for some $m$. Thus the lemma is verified in this case.

*Case* II. Assume $|b| = 1$, $b = e^{i\theta}$. Since $b$ is not a root of 1, $\theta$ is not a rational multiple of $\pi$. As is well known [3] it then follows that there is a sequence of positive integers $m_1, m_2, \cdots, m_k \cdots$ such that $\lim_{k\to\infty} e^{im_k\theta} = -1$. Since $e(A_0^m) < 1$ is constant, it then follows from (2) that $0 \in F(A_0^{m_k})$ for all $k$ greater than some $K$. Thus the verification of the lemma is complete.

We also make the following easily verified observations:

**Lemma 3.** *If*

$$A = \begin{bmatrix} a & 0 \cdots 0 & c \\ & * & 0 \\ & \ddots\ 0 & \vdots \\ 0 & * & 0 \\ & & b \end{bmatrix}$$

*and $a/b$ is a root of unity not equal to* 1, *then some integral power of $A$ is diagonal.*

**Lemma 4.** *If*

$$T = \begin{bmatrix} A_1 & & * \\ & A_2 & \\ & & \ddots \\ 0 & & A_k \end{bmatrix} \in M_n(C)$$

*is block triangular, $A_1, \cdots, A_k$ square, then*

$$T^m = \begin{bmatrix} A_1^m & & * \\ & A_2^m & \\ & & \ddots \\ 0 & & A_k^m \end{bmatrix}.$$

We now turn to the proofs of the two theorems mentioned.

**Proof of Theorem 1.** In view of (3) it suffices to assume $0 \notin \sigma(A)$. We choose a unitary $U$ to reduce $A$ to upper triangular form:

$$UAU^* = T = \begin{bmatrix} t_{11} & & t_{ij} \\ & \ddots & \\ 0 & & t_{nn} \end{bmatrix}.$$

We then have $A^m = U^*T^mU$; $0 \in F(A^m)$ if and only if $0 \in F(T^m)$; and $A^m$ is normal if and only if $T^m$ is diagonal.

We will show that either $0 \in F(T^m)$ for some $m \in I^+$ or $T^m$ is diagonal for some $m \in I^+$, which proves Theorem 1. Suppose there are $q_0 \leq (n^2 - n)/2$ nonzero entries in $T$ above the diagonal. If $q_0 = 0$, we are done. If $q_0 > 0$, we may choose a principal submatrix $T_0$ of $T$ determined by the set of consecutive indices $j, j+1, \cdots, k$, $k \geq 1 + j$ of the form

$$T_0 = \begin{bmatrix} t_{jj} & 0 \cdots 0 & c \\ & \ddots \quad 0 & 0 \\ & & \vdots \\ 0 & & 0 \\ & & t_{kk} \end{bmatrix}.$$

Now there are two possibilities. If $t_{kk}/t_{jj}$ is not a nonreal root of unity, then by Lemma 2, $0 \in F(T_0^{m_1})$ for some $m_1 \in I^+$. In this event it then follows from Lemma 4 and (5) that $0 \in F(T^{m_1})$, and we would be finished. If, alternatively, $t_{kk}/t_{jj}$ is a root of unity not equal to 1, then $T_0^{m_1}$ is diagonal for some $m_1 \in I^+$ by Lemma 3. Let $q_1$ now count the nonzero entries above the diagonal of $T^{m_1}$. Because of Lemma 4, $q_1 \leq (n(n-1) - (k-j)(k-j+1))/2$. If $q_1 = 0$, we are finished. If not, we next argue on $T^{m_1}$ as we just did on $T$. Pick a principal submatrix $T_{0,m_1}$ of $T^{m_1}$ of the same form as $T_0$. Then determine $m_2 \in I^+$ such that either $0 \in F(T^{m_1 m_2})$ or $T_{0,m_1}^{m_2}$ is diagonal. Continue this process to produce a sequence $m_1, m_2, \cdots, m_t \in I^+$. Because of Lemma 4 this process must terminate $(t \leq (n^2 - n)/2)$ with either $0 \in F(T^m)$, $m = m_1 \cdots m_t$, or $q_t = 0$ (which means $T^m$ is diagonal). In either event $m = m_1 \cdots m_t$ is the needed integer and the proof is complete.

**Proof of Theorem 2.** We suppose $A$ is not Hermitian, and we show there is an $m \in I^+$ such that $A^m \notin \Pi$, which establishes Theorem 2. If $\sigma(A)$ is not real and positive, it follows from De Moivre's theorem and (3) and (6)

that $A^m \notin \Pi$ for some $m$. Thus we assume $\sigma(A)$ is real and positive, and since our problem is invariant under unitary equivalence, we assume without loss of generality that $A$ is in upper triangular form. Since $A$ is not Hermitian (i.e. not diagonal in this case), this means $A$ has a principal submatrix

$$A_0 = \begin{bmatrix} a & 0 \cdots 0 & c \\ * & 0 & 0 \\ & \ddots & \vdots \\ 0 & * & 0 \\ & & b \end{bmatrix}$$

with $c \neq 0$ and $a$ and $b$ positive. It follows inductively that

$$A_0^m = \begin{bmatrix} a^m & 0 \cdots 0 & c\sum_{j=0}^{m-1} a^j b^{m-j-1} \\ * & & 0 \\ & \ddots & \vdots \\ \bigcirc & * & \\ & & b^m \end{bmatrix}.$$

By Lemma 4, $A_0^m$ is a principal submatrix of $A^m$, and since $c\sum_{j=0}^{m-1} a^j b^{m-j-1}$ is never 0, we have that $A^m$ is never diagonal and thus never normal. From Theorem 1 it then follows that for some $m$, $0 \in F(A^m)$. In view of (7) and (8) this yields that $A^m \notin \Pi$ as was to be shown.

**Example.** That the qualification on the existence of solutions to $xA^m x^* = 0$ given in Theorem 1 is in general needed is seen in the following example. Let

$$A = \begin{bmatrix} 1 & \epsilon \\ 0 & z \end{bmatrix}, \qquad z = e^{2\pi i/3}.$$

Then

$$A^{3k} = I, \qquad A^{3k+1} = A, \qquad A^{3k+2} = \begin{bmatrix} 1 & \epsilon(1+z) \\ 0 & \bar{z} \end{bmatrix}.$$

Thus if $\epsilon$ is sufficiently small, $0 \notin F(A^m)$ for all $m \in I^+$. Since $A^{3k} = I$, the condition of Theorem 1 is not satisfied.

**Remark 1.** It is well known that the containment $F(AB) \subseteq F(A)F(B)$ is not in general valid. Not even the containment $F(A^m) \subseteq F(A)^m$ is in general valid. In fact, according to Theorem 1 the latter containment is very generally invalid in case $0 \notin F(A)$. In this event $0 \notin F(A)^m$ for all $m \in I^+$ while, with the exception of only very special cases, $0 \in F(A^m)$ for some $m \in I^+$.

**Remark 2.** The requirement "for all $m \in I^+$" is necessary in Theorem 2. For any $m \in I^+$, a nonhermitian matrix $A$ may be constructed so that $A^i \in \Pi$ for all $i = 1, \cdots, m$.

**Author's Note (added in proof).** Since the announcement of Theorem 2 in the author's thesis, operator theoretic proofs of the analogous fact for linear operators on Hilbert space have been given by C. R. DePrima and B. K. Richard, and E. Shiu. These proofs are of a decidedly different character and make use of, for example: T. Kato, *Some mapping theorems for the numerical range*, Proc. Japan Acad. 41 (1965), 652–655 (MR 36 # 5743); C. Foiaş and B. Sz.-Nagy, *Harmonic analysis of operators on Hilbert space*, North-Holland, Amsterdam, 1970, pp. 176–181 (MR 43 # 947); and T. Kato, *Perturbation theory for linear operators*, Die Grundlehren der math. Wissenschaften, Bd. 132, Springer-Verlag, New York, 1966, pp. 281–286 (MR 34 # 3324). The author wishes to thank Mr. Shiu for several interesting correspondences regarding results related to Theorem 2.

## REFERENCES

1. W. F. Donoghue, *On the numerical range of a bounded operator*, Michigan Math. J. 4 (1957), 261–263. MR 20 # 2622.

2. C. R. Johnson and M. Newman, *Triangles generated by powers of triplets on the unit circle*, J. Res. Nat. Bur. Standards Sect. B 77B (1973), 137–142.

3. H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann. 77 (1916), 313–352.

APPLIED MATHEMATICS DIVISION, NATIONAL BUREAU OF STANDARDS, WASHINGTON, D. C. 20234

*Current address*: Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland 20742