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, \( m \in \mathbb{I}^+ \), then the equation \( xA^m x^* = 0 \) has a solution \( 0 \neq x \in \mathbb{C}^n, m \in \mathbb{I}^+ \); 2. If \( H(A) = (A + A^*)/2 \) is positive definite, then \( H(A^m) \) is positive definite for all \( m \in \mathbb{I}^+ \) if and only if \( A \) is Hermitian.

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

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

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

It is well known that

1. \( F(A) \) is convex and compact;
2. the spectrum \( \sigma(A) \subseteq F(A) \); and
3. \( F(UAU^*) = F(A) \) 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

\[
F(A_0) \subseteq F(A).
\]

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

4. \( A \in \Pi \) is equivalent to \( F(A) \subseteq R \).

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

\[
\text{Received by the editors March 12, 1973.}
\]


\textit{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.

Copyright © 1975, American Mathematical Society
is equivalent to the existence of a $\theta \in [0, 2\pi)$ such that
\begin{equation}
ei^\theta A \in \Pi
\end{equation}
as well as to the nonexistence of a nonzero $x \in \mathbb{C}^n$ such that
\begin{equation}
x Ax^* = 0.
\end{equation}

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

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

**Theorem 2.** Suppose $A \in \Pi$. Then $A^m \in \Pi$ for all $m \in \mathbb{N}^+$ 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 $A = [\begin{smallmatrix} a & c \\ 0 & b \end{smallmatrix}]$, $c \neq 0$, 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
\begin{equation}
A = \begin{bmatrix}
a & 0 & \cdots & 0 & c \\
* & 0 & & * \\
& \ddots & \ddots & \ddots & \ddots \\
0 & * & \cdots & 0 & b
\end{bmatrix} \in M_k(\mathbb{C})
\end{equation}
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 $x A^m x^* = 0$.

**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

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

We shall show that under our assumptions $0 \in F(A_0^m)$ for some $m \in i^+$ where $A_0^m$ is a $2 \times 2$ principal submatrix of $A^m$. By (5) it then follows that $0 \in F(A_0^m)$ which is equivalent by (7) and (9) to what we wish to show.

If $b = -1$, then $0 \in F(A_0^{-1})$ 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

$$A_0^m = \begin{bmatrix}
1 & \frac{b^m - 1}{b - 1} \\
0 & b^m
\end{bmatrix}.$$ 

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

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

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 $m_1, m_2, \ldots, m_k \ldots$ 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

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

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

Lemma 4. If

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

is block triangular, $A_1, \ldots, A_k$ square, then

$$T^m = \begin{bmatrix} A_1^m & \ast \\ \ast & A_2^m \\ \ddots & \ddots \\ \ast & \ast \\ \ast & \ast \\ 0 & \cdots \\ \end{bmatrix}.$$
We now turn to the proofs of the two theorems mentioned.

**Proof of Theorem 1.** In view of (3) it suffices to assume $0 \not\in \sigma(A)$. We choose a unitary $U$ to reduce $A$ to upper triangular form:

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

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 $q_0 < (n^2 - n)/2$ 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 $j, j + 1, \ldots, k, k \geq 1 + j$ of the form

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

Now there are two possibilities. If $t_{kk}/t_{jj}$ is not a nonreal root of unity, then by Lemma 2, $0 \in F(T_{01}^1)$ for some $m_1 \in I^+$. In this event it then follows from Lemma 4 and (5) that $0 \in F(T^m)$, and we would be finished. If, alternatively, $t_{kk}/t_{jj}$ is a root of unity not equal to 1, then $T_{02}^1$ is diagonal for some $m_2 \in I^+$ by Lemma 3. Let $q_1$ now count the nonzero entries above the diagonal of $T_{02}^1$. 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 $T_{02}^1$ as we just did on $T$. Pick a principal submatrix $T_{0,m_1}^1$ of $T^m$ of the same form as $T_0$. Then determine $m_2 \in I^+$ such that either $0 \in F(T_{0,m_1}^m)$ or $T_{0,m_2}^m$ is diagonal. Continue this process to produce a sequence $m_1, m_2, \ldots, m_t \in I^+$. 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 $A^m \not\in \Pi$, 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

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

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

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


REFERENCES


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