Can. J. Math., Vol. XXX, No. 2, 1978, pp. 419–430

NORMALITY AND THE HIGHER NUMERICAL RANGE

MARVIN MARCUS, BENJAMIN N. MOYLS AND IVAN FILIPPENKO

1. Introduction. Let $M_n(\mathbf{C})$ be the vector space of all n-square complex matrices. Denote by $(\cdot\,,\cdot)$ the standard inner product in the space $\mathbf{C}^n$ of complex n-tuples. For a matrix $A \in M_n(\mathbf{C})$ and an n-tuple $c = (c_1, \ldots, c_n) \in \mathbf{C}^n$, define the *c-numerical range of A* to be the set

$$(1) \qquad W_c(A) = \left\{ \sum_{k=1}^{n} c_k(Ax_k, x_k) | \{x_1, \ldots, x_n\} \text{ is an orthonormal basis of } \mathbf{C}^n \right\}$$

in the complex plane. Denote the eigenvalues of A by $\lambda_1, \ldots, \lambda_n$, and define the *c-eigenpolygon of A* to be the convex hull

$$(2) \qquad P_c(A) = \mathscr{H}\left(\left\{ \sum_{k=1}^{n} c_k \lambda_{\sigma(k)} | \sigma \in S_n \right\}\right),$$

where S_n is the symmetric group of degree n. The matrix A is said to be *c-convex* if $W_c(A) = P_c(A)$.

If

$$m \in \{1, \ldots, n\} \quad \text{and} \quad c = (\overbrace{1, \ldots, 1}^{m}, \overbrace{0, \ldots, 0}^{n-m}),$$

then $W_c(A)$ and $P_c(A)$ are called the *m-th numerical range of A* and the *m-th eigenpolygon of A* respectively, and are denoted by $W_m(A)$ and $P_m(A)$. Thus

$$(3) \qquad W_m(A) = \left\{ \sum_{k=1}^{m} (Ax_k, x_k) | x_1, \ldots, x_m \text{ are } m \text{ orthonormal vectors in } \mathbf{C}^n \right\};$$

evidently $W_1(A)$ is the classical *numerical range*

$$W(A) = \{(Ax, x) | x \in \mathbf{C}^n, ||x|| = 1\}.$$

Designating by $Q_{m,n}$ the set of all strictly increasing sequences of m integers chosen from $\{1, \ldots, n\}$, we have

$$(4) \qquad P_m(A) = \mathscr{H}\left(\left\{ \sum_{k=1}^{m} \lambda_{\omega(k)} | \omega \in Q_{m,n} \right\}\right).$$

It was shown by C. A. Berger [**2**, § 167] that the sets $W_m(A)$ are convex. Since $\sum_{k=1}^{m} \lambda_{\omega(k)} \in W_m(A)$ for all $\omega \in Q_{m,n}$ [**1**], it follows that

$$(5) \qquad W_m(A) \supset P_m(A).$$

Received February 21, 1977 and in revised form April 2, 1977.

The matrix A is said to be *m-convex* if $W_m(A) = P_m(A)$ (in case $m = 1$, A is simply said to be *convex*).

It is known that if $A \in M_n(\mathbf{C})$ is normal, then A is m-convex for $1 \leqq m \leqq n$ [**1**]. In the present paper, we obtain this result as a corollary of a theorem concerning the c-convexity of a matrix. Our main purpose is to discuss the question of a converse: does m-convexity for certain values of m imply normality? Initial results in this direction were previously obtained by two of the authors [**6**], who proved that convexity guarantees normality when $n \leqq 4$ but not when $n \geqq 5$.

2. Statement of results.

THEOREM 1. *Let $A \in M_n(\mathbf{C})$ be a normal matrix, and let $c = (c_1, \ldots, c_n) \in \mathbf{C}^n$. Then*

$$W_c(A) \subset P_c(A).$$

Moreover, if $c = (c_1, \ldots, c_n) \in \mathbf{R}^n$, then

$$W_c(A) = P_c(A)$$

(i.e., *A is c-convex*).

An immediate corollary of this theorem is that if $A \in M_n(\mathbf{C})$ is normal, then A is m-convex for $1 \leqq m \leqq n$.

The following useful result contains the key idea in the proof of Theorem 3.

THEOREM 2. *Let $A \in M_n(\mathbf{C})$, and for any $\theta \in [0, 2\pi)$ set $A_\theta = e^{i\theta}A$. Let $m \in \{1, \ldots, n\}$. Then A is m-convex if and only if*

$$(6) \qquad \sum_{k=1}^{m} \lambda_k\left(\frac{A_\theta + A_\theta^*}{2}\right) = \sum_{k=1}^{m} r_k(A_\theta)$$

for all $\theta \in [0, 2\pi)$, where

$$\lambda_1\left(\frac{A_\theta + A_\theta^*}{2}\right) \geqq \ldots \geqq \lambda_n\left(\frac{A_\theta + A_\theta^*}{2}\right)$$

are the eigenvalues of the hermitian matrix $(A_\theta + A_\theta^*)/2$ and

$$r_1(A_\theta) \geqq \ldots \geqq r_n(A_\theta)$$

are the real parts of the eigenvalues of A_θ.

The principal result of this paper is the

THEOREM 3. *Let $A \in M_n(\mathbf{C})$. Then A is normal if and only if A is m-convex for $1 \leqq m \leqq [n/2]$, where $[\quad]$ designates the greatest integer function.*

We conclude with a class of examples showing that Theorem 3 is, in general, the best possible.

THEOREM 4. *Let m be a fixed positive integer. For a given complex number ϵ, let A be the $(2m+3)$-square complex matrix*

$$A = \operatorname{diag}\,(e^{k\omega i}: k = 0, \ldots, 2m) \dotplus \begin{bmatrix} 0 & \epsilon \\ 0 & 0 \end{bmatrix},$$

where $\omega = 2\pi/(2m+1)$. Then

(i) *A is m-convex if and only if $|\epsilon| \leqq 2 \cos (m\pi/(2m+1))$;*

(ii) *if A is m-convex, then A is j-convex for $j = 1, \ldots, m$;*

(iii) *A is $(m+1)$-convex if and only if A is normal* (i.e., $\epsilon = 0$).

Thus for appropriate $\epsilon \neq 0$, the $(2m+3)$-square matrix A is j-convex for $1 \leqq j \leqq m = [(2m+3)/2] - 1$ without being normal.

The methods employed in the proof of Theorem 4 illustrate the power of Theorem 2 as a computable criterion.

3. Preliminaries. This section contains information which will be used in the proofs in Section 4.

Recall that a matrix $S \in M_n(\mathbf{C})$ is *doubly stochastic* if S is a nonnegative matrix (i.e., $S_{ij} \geqq 0, i, j, = 1, \ldots, n$) all of whose row and column sums are 1. Recall also that a matrix $S \in M_n(\mathbf{C})$ is *orthostochastic* if there exists a unitary matrix $U \in M_n(\mathbf{C})$ such that $S_{ij} = |U_{ij}|^2$, $i, j = 1, \ldots, n$. Although it is clear that every orthostochastic matrix is doubly stochastic, the converse is false [**4**, II.1.4.4].

Of central importance is

BIRKHOFF'S THEOREM [4, II.1.7]. *The set Ω_n of all n-square doubly stochastic matrices is a convex polyhedron in $M_n(\mathbf{R})$ whose vertices are the n-square permutation matrices.*

A characterization is available of main diagonals of normal matrices with prescribed eigenvalues:

LEMMA 1 [**4**, II.4.1.3]. *Let $A \in M_n(\mathbf{C})$ be a normal matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$, and set $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbf{C}^n$. Let*

$$E_1 = \{\mu = ((Ax_1, x_1), \ldots, (Ax_n, x_n)) \in \mathbf{C}^n | x_1, \ldots, x_n \text{ o.n.}\}$$

and

$$E_2 = \{\mu = S\lambda \in \mathbf{C}^n | S \in M_n(\mathbf{C}) \text{ orthostochastic}\}.$$

Then $E_1 = E_2$.

Here and in what follows, "o.n." abbreviates the word "orthonormal".

A considerably more difficult result, due primarily to A. Horn [**3**], provides a characterization of main diagonals of hermitian matrices with prescribed eigenvalues (see also M. Marcus, B. N. Moyls, and R. Westwick [**5**]):

LEMMA 2. *Let $C \in M_n(\mathbf{C})$ be a hermitian matrix with eigenvalues $c_1, \ldots, c_n$, and set $c = (c_1, \ldots, c_n) \in \mathbf{R}^n$. Let*

$$E_1 = \{\mu = ((Cx_1, x_1), \ldots, (Cx_n, x_n)) \in \mathbf{R}^n | x_1, \ldots, x_n \text{ o.n.}\}$$

https://doi.org/10.4153/CJM-1978-036-6 Published online by Cambridge University Press

and

$$E_2 = \{\mu = Sc \in \mathbf{R}^n | S \in \Omega_n\}.$$

Then $E_1 = E_2$.

We will have occasion to use the well-known Elliptical Range Theorem [7]. This states that *if*

$$A = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$$

is a 2-square upper triangular complex matrix, then the numerical range $W(A)$ *is the region bounded by an ellipse with foci at* a *and* b, *minor axis of length* $|c|$, *and major axis of length* $\sqrt{|a - b|^2 + |c|^2}$.

Finally, we remark that if $A \in M_n(\mathbf{C})$ and $m \in \{1, \ldots, n\}$, then

$$W_{n-m}(A) = \operatorname{tr}(A) - W_m(A)$$

and

$$P_{n-m}(A) = \operatorname{tr}(A) - P_m(A),$$

so that A is $(n - m)$-convex if and only if A is m-convex.

4. Proofs.

Proof of Theorem 1. Denote the eigenvalues of A by $\lambda_1, \ldots, \lambda_n$ and set

$$\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbf{C}^n.$$

Let $\{x_1, \ldots, x_n\}$ be any o.n. basis of $\mathbf{C}^n$, and set

$$\mu = ((Ax_1, x_1), \ldots, (Ax_n, x_n)) \in \mathbf{C}^n.$$

By Lemma 1, there exists an n-square doubly stochastic matrix S such that $\mu = S\lambda$. By Birkhoff's Theorem, S is a convex combination of the n-square permutation matrices; say

$$S = \sum_{\sigma \in S_n} \alpha_\sigma P_\sigma$$

where $\alpha_\sigma \geqq 0$ for all $\sigma \in S_n$, $\sum_{\sigma \in S_n} \alpha_\sigma = 1$, and $P_\sigma = [\delta_{i\sigma(j)}]$, $\sigma \in S_n$. Then letting $\bar{c} = (\bar{c}_1, \ldots, \bar{c}_n)$, we have

$$\begin{aligned}
\sum_{k=1}^{n} c_k (Ax_k, x_k) &= (\mu, \bar{c}) \\
&= (S\lambda, \bar{c}) \\
&= \left(\sum_{\sigma \in S_n} \alpha_\sigma P_\sigma \lambda, \bar{c} \right) \\
&= \sum_{\sigma \in S_n} \alpha_\sigma (P_\sigma \lambda, \bar{c}) \\
&= \sum_{\sigma \in S_n} \alpha_\sigma \left(\sum_{k=1}^{n} \lambda_{\sigma^{-1}(k)} c_k \right) \\
&= \sum_{\sigma \in S_n} \alpha_{\sigma^{-1}} \left(\sum_{k=1}^{n} c_k \lambda_{\sigma(k)} \right) \quad \in P_c(A).
\end{aligned}$$

https://doi.org/10.4153/CJM-1978-036-6 Published online by Cambridge University Press

We conclude that $W_c(A) \subset P_c(A)$.

Now assume that $c = (c_1, \ldots, c_n) \in \mathbf{R}^n$. Since $A \in M_n(\mathbf{C})$ is a normal matrix, there exists an o.n. basis $\{u_1, \ldots, u_n\}$ of $\mathbf{C}^n$ such that

$$Au_k = \lambda_k u_k, \quad k = 1, \ldots, n.$$

Let $C \in M_n(\mathbf{C})$ be a hermitian matrix with eigenvalues $c_1, \ldots, c_n$; there exists an o.n. basis $\{y_1, \ldots, y_n\}$ of $\mathbf{C}^n$ such that

$$Cy_k = c_k y_k, \quad k = 1, \ldots, n.$$

Denote by $U_n(\mathbf{C})$ the group of n-square unitary matrices. We compute that

$$\begin{aligned}
W_c(A) &= \left\{\sum_{k=1}^{n} c_k(Ax_k, x_k) | x_1, \ldots, x_n \text{ o.n.}\right\} \\
&= \left\{\sum_{k=1}^{n} c_k(AUy_k, Uy_k) | U \in U_n(\mathbf{C})\right\} \\
&= \left\{\sum_{k=1}^{n} (AUCy_k, Uy_k) | U \in U_n(\mathbf{C})\right\} \\
&= \left\{\sum_{k=1}^{n} (U^*AUCy_k, y_k) | U \in U_n(\mathbf{C})\right\} \\
&= \{\text{tr }(U^*AUC) | U \in U_n(\mathbf{C})\} \\
&= \{\text{tr }(UCU^*A) | U \in U_n(\mathbf{C})\} \\
&= \left\{\sum_{k=1}^{n} (UCU^*Au_k, u_k) | U \in U_n(\mathbf{C})\right\} \\
&= \left\{\sum_{k=1}^{n} \lambda_k(UCU^*u_k, u_k) | U \in U_n(\mathbf{C})\right\} \\
&= \left\{\sum_{k=1}^{n} \lambda_k(Cx_k, x_k) | x_1, \ldots, x_n \text{ o.n.}\right\} \\
&= \{(\lambda, Sc) | S \in \Omega_n\} \quad \text{(by Lemma 2)} \\
&= \left\{\left(\lambda, \sum_{\sigma \in S_n} \alpha_\sigma P_\sigma c\right) | \alpha_\sigma \geqq 0 \text{ for all } \sigma \in S_n, \sum_{\sigma \in S_n} \alpha_\sigma = 1\right\} \\
&\qquad \text{(by Birkhoff's Theorem)} \\
&= \mathscr{H}\left(\left\{\sum_{k=1}^{n} c_k \lambda_{\sigma(k)} | \sigma \in S_n\right\}\right) \\
&= P_c(A).
\end{aligned}$$

This completes the proof.

Proof of Theorem 2. We begin by making some general observations. If $\theta \in [0, 2\pi)$ and $z \in W_m(A_\theta)$, say $z = \sum_{k=1}^{m} (A_\theta x_k, x_k)$ where $x_1, \ldots, x_m$ are o.n. vectors in $\mathbf{C}^n$, then

$$\operatorname{Re} z = \operatorname{Re} \sum_{k=1}^{m} (A_\theta x_k, x_k)$$

$$(7) \qquad = \sum_{k=1}^{m} \left(\frac{A_\theta + A_\theta^*}{2} x_k, x_k\right)$$

$$(8) \qquad \in W_m\left(\frac{A_\theta + A_\theta^*}{2}\right).$$

Since $(A_\theta + A_\theta^*)/2$ is a hermitian matrix,

$$(9) \qquad W_m\left(\frac{A_\theta + A_\theta^*}{2}\right) = P_m\left(\frac{A_\theta + A_\theta^*}{2}\right)$$

is a closed real interval with right endpoint $\sum_{k=1}^{m} \lambda_k((A_\theta + A_\theta^*)/2)$. We conclude from (8) and (9) that for all $z \in W_m(A_\theta)$,

$$(10) \qquad \operatorname{Re} z \leqq \sum_{k=1}^{m} \lambda_k\left(\frac{A_\theta + A_\theta^*}{2}\right).$$

In particular, by choosing o.n. vectors $x_1, \ldots, x_m \in \mathbf{C}^n$ such that

$$r_k(A_\theta) = \operatorname{Re} (A_\theta x_k, x_k), \quad k = 1, \ldots, m,$$

we obtain

$$(11) \qquad \sum_{k=1}^{m} r_k(A_\theta) = \operatorname{Re} \sum_{k=1}^{m} (A_\theta x_k, x_k) \leqq \sum_{k=1}^{m} \lambda_k\left(\frac{A_\theta + A_\theta^*}{2}\right).$$

Now assume that A is m-convex, i.e., that $W_m(A) = P_m(A)$. Fix $\theta \in [0, 2\pi)$ and note that

$$\begin{aligned} W_m(A_\theta) &= W_m(e^{i\theta}A) \\ &= e^{i\theta}W_m(A) \\ &= e^{i\theta}P_m(A) \\ &= P_m(e^{i\theta}A) = P_m(A_\theta). \end{aligned}$$

The vertices of the convex polygon $P_m(A_\theta)$ are sums of m eigenvalues of A_θ, and if $z \in P_m(A_\theta)$ then Re z is at most the largest real part of these vertices. Hence $z \in W_m(A_\theta) = P_m(A_\theta)$ implies

$$(12) \qquad \operatorname{Re} z \leqq \sum_{k=1}^{m} r_k(A_\theta).$$

If $x_1, \ldots, x_m$ are any m o.n. vectors in $\mathbf{C}^n$, it follows from (7) and (12) that

$$(13) \qquad \sum_{k=1}^{m} \left(\frac{A_\theta + A_\theta^*}{2} x_k, x_k\right) = \operatorname{Re} \sum_{k=1}^{m} (A_\theta x_k, x_k) \leqq \sum_{k=1}^{m} r_k(A_\theta).$$

In view of (11), the equality (6) is obtained from (13) by choosing an o.n. basis of eigenvectors of the hermitian matrix $(A_\theta + A_\theta^*)/2$.

To prove sufficiency, assume the equality (6) holds for all $\theta \in [0, 2\pi)$. Let l denote a fixed side of the convex polygon $P_m(A)$. It is easy to see that θ may be chosen so that (i) the side $e^{i\theta}l = l_\theta$ of $e^{i\theta}P_m(A) = P_m(A_\theta)$ is oriented vertically in the complex plane, and (ii) $P_m(A_\theta)$ is contained in the closed left half-plane determined by l_θ. Notice that the real part of a point on l_θ is precisely $\sum_{k=1}^m r_k(A_\theta)$. Then for any $z \in W_m(A_\theta)$, we have

$$\begin{aligned} \operatorname{Re} z &\leqq \sum_{k=1}^{m} \lambda_k\left(\frac{A_\theta + A_\theta^*}{2}\right) \quad \text{(by (10))} \\ &= \sum_{k=1}^{m} r_k(A_\theta) \quad \text{(by (6))}. \end{aligned}$$

Thus $W_m(A_\theta) = e^{i\theta}W_m(A)$ is contained in the closed left half-plane determined by $l_\theta = e^{i\theta}l$, so that $W_m(A)$ is contained in the closed left half-plane determined by l. Since l was a fixed but otherwise unspecified side of $P_m(A)$, it follows that $W_m(A)$ is contained in the intersection of the closed left half-planes determined by the sides of $P_m(A)$. Of course, this intersection is simply $P_m(A)$. Thus

$$W_m(A) \subset P_m(A).$$

By (5) we have

$$W_m(A) \supset P_m(A),$$

and the proof is complete.

Proof of Theorem 3. As has already been observed, the necessity of the conditions is an immediate consequence of Theorem 1.

Assume that A is m-convex for $1 \leqq m \leqq [n/2]$. It follows from a remark in Section 1 that A is m-convex for $1 \leqq m \leqq n$. Note that the m-th numerical range, the m-th eigenpolygon, and the status of normality of a complex matrix are invariant under transformation of the matrix by a unitary similarity. We may therefore assume (by the Schur triangularization theorem) that the given matrix A is upper triangular, with eigenvalues $\lambda_1, \ldots, \lambda_n$ arranged down the main diagonal to satisfy $\operatorname{Re} \lambda_i \leqq \operatorname{Re} \lambda_j$ for $1 \leqq i \leqq j \leqq n$. The proof will be completed by showing that A is in fact diagonal.

Suppose A has a nonzero off-diagonal element $\epsilon = A_{ij}$ $(i < j)$. Set

$$B = A - \lambda_j I \quad \in M_n(\mathbf{C})$$

(I is the n-square identity matrix). It is clear that

(i) B has eigenvalues

$$\mu_k = \lambda_k - \lambda_j, \quad k = 1, \ldots, n$$

with

$$\operatorname{Re} \mu_i \leqq 0 \quad \text{and} \quad \mu_j = 0;$$

(ii) B is upper triangular with $B_{ij} = \epsilon \neq 0$; and

(iii) B is m-convex for $1 \leqq m \leqq n$.

https://doi.org/10.4153/CJM-1978-036-6 Published online by Cambridge University Press

Now let m_0 be the number of eigenvalues of B having positive real part: obviously $0 \leqq m_0 \leqq n-2$. Choose $\omega \in Q_{m_0,n}$ so that $\mu_{\omega(1)}, \ldots, \mu_{\omega(m_0)}$ are these eigenvalues (if $m_0 = 0$, then ω is the "empty sequence" which assumes no values). Set

$$x_k = e_{\omega(k)}, \quad k = 1, \ldots, m_0$$

where e_t denotes the t-th standard basis vector in $\mathbf{C}^n$ (1 in position t, 0's elsewhere). We have

$$(14) \quad (Bx_k, x_k) = \mu_{\omega(k)}, \quad k = 1, \ldots, m_0.$$

By the Elliptical Range Theorem,

$$W\left(\begin{bmatrix} \mu_i & \epsilon \\ 0 & \mu_j \end{bmatrix}\right)$$

is the region bounded by an ellipse with foci at μ_i and $\mu_j = 0$ whose minor axis has length $|\epsilon|$. Since $|\epsilon| > 0$, it follows that there exists $z \in W\left(\begin{bmatrix} \mu_i & \epsilon \\ 0 & \mu_j \end{bmatrix}\right)$ for which $\operatorname{Re} z > 0$. Hence there exists a unit vector $x_{m_0+1} \in \mathbf{C}^n$, having nonzero components only in positions i and j, such that

$$(15) \quad \operatorname{Re}\,(Bx_{m_0+1}, x_{m_0+1}) < 0.$$

Observe that since $\omega(k) \neq i, j$ for $k = 1, \ldots, m_0$, the vectors $x_1, \ldots, x_{m_0}, x_{m_0+1}$ in $\mathbf{C}^n$ are o.n.

By virtue of the fact that B has precisely m_0 eigenvalues with positive real part $(\mu_{\omega(1)}, \ldots, \mu_{\omega(m_0)})$ and at least one eigenvalue with zero real part $(\mu_j = 0)$, in the notation of Theorem 2 we compute

$$\begin{aligned}
\sum_{k=1}^{m_0+1} r_k(B) &= \sum_{k=1}^{m_0} \operatorname{Re} \mu_{\omega(k)} \\
&= \sum_{k=1}^{m_0} \operatorname{Re}\,(Bx_k, x_k) \quad (\text{by } (14)) \\
&< \sum_{k=1}^{m_0+1} \operatorname{Re}\,(Bx_k, x_k) \quad (\text{by } (15)) \\
&= \sum_{k=1}^{m_0+1} \left(\frac{B+B^*}{2} x_k, x_k\right) \\
(16) \qquad &\in W_{m_0+1}\left(\frac{B+B^*}{2}\right).
\end{aligned}$$

Since $(B + B^*)/2$ is a hermitian matrix,

$$(17) \quad W_{m_0+1}\left(\frac{B+B^*}{2}\right) = P_{m_0+1}\left(\frac{B+B^*}{2}\right)$$

is a closed real interval with right endpoint $\sum_{k=1}^{m_0+1} \lambda_k((B + B^*)/2)$. Hence from (16) and (17),

$$\sum_{k=1}^{m_0+1} r_k(B) < \sum_{k=1}^{m_0+1} \left(\frac{B+B^*}{2} x_k, x_k\right) \leqq \sum_{k=1}^{m_0+1} \lambda_k\left(\frac{B+B^*}{2}\right).$$

In view of (iii) above, this contradicts Theorem 2. We conclude that the upper triangular matrix A can have no nonzero off-diagonal element ϵ, completing the proof.

Proof of Theorem 4. For a given $\theta \in [0, 2\pi)$ we have

$$A_\theta = e^{i\theta}A = \operatorname{diag}\,(e^{(\theta+k\omega)i}: k = 0, \dots, 2m) \dotplus \begin{bmatrix} 0 & e^{i\theta}\epsilon \\ 0 & 0 \end{bmatrix}$$

and

$$\frac{A_\theta + A_\theta^*}{2} = \operatorname{diag}\,(\cos(\theta + k\omega): k = 0, \dots, 2m) \dotplus \frac{1}{2}\begin{bmatrix} 0 & e^{i\theta}\epsilon \\ e^{-i\theta}\bar{\epsilon} & 0 \end{bmatrix}.$$

Let $\mathscr{L}_\theta$ denote the set of eigenvalues of $(A_\theta + A_\theta^*)/2$, and let $\mathscr{R}_\theta$ denote the set of real parts of the eigenvalues of A_θ. Then

$$\mathscr{L}_\theta = \{\cos(\theta + k\omega) | k = 0, \dots, 2m\} \cup \left\{\frac{|\epsilon|}{2}, -\frac{|\epsilon|}{2}\right\}$$

and

$$\mathscr{R}_\theta = \{\cos(\theta + k\omega) | k = 0, \dots, 2m\} \cup \{0, 0\}.$$

Now assume that A is m-convex. If m is even, set $\theta = 0$, while if m is odd, set $\theta = \pi$. In either situation the regular odd-order convex polygon $P_1(A_\theta)$ has precisely $m + 1$ vertices with positive real part, and these are symmetrically positioned with respect to the real axis. When m is even, a minimal positive real part occurs for $k = m/2$ and has the value

$$\begin{aligned} \cos(\theta + k\omega) &= \cos\left(0 + \frac{m}{2}\frac{2\pi}{2m+1}\right) \\ &= \cos\left(\frac{m\pi}{2m+1}\right). \end{aligned}$$

When m is odd, a minimal positive real part occurs for $k = (m + 1)/2$ and again has the value

$$\begin{aligned} \cos(\theta + k\omega) &= \cos\left(\pi + \frac{m+1}{2}\frac{2\pi}{2m+1}\right) \\ &= \cos\left(\frac{(2m+1)\pi + (m+1)\pi}{2m+1}\right) \\ &= \cos\left(\frac{(3m+2)\pi}{2m+1}\right) \\ &= \cos\left(2\pi - \frac{m\pi}{2m+1}\right) \\ &= \cos\left(-\frac{m\pi}{2m+1}\right) \\ &= \cos\left(\frac{m\pi}{2m+1}\right) \end{aligned}$$

Let $k_0, \ldots, k_{2m}$ be a permutation of the integers $0, \ldots, 2m$ such that

$$\cos(\theta + k_0\omega) \geqq \ldots \geqq \cos(\theta + k_{m-1}\omega) = \cos(\theta + k_m\omega) = \cos\left(\frac{m\pi}{2m+1}\right).$$

If $|\epsilon| > 2\cos(m\pi/(2m+1))$, then the largest sum of m elements of $\mathscr{L}_\theta$ is

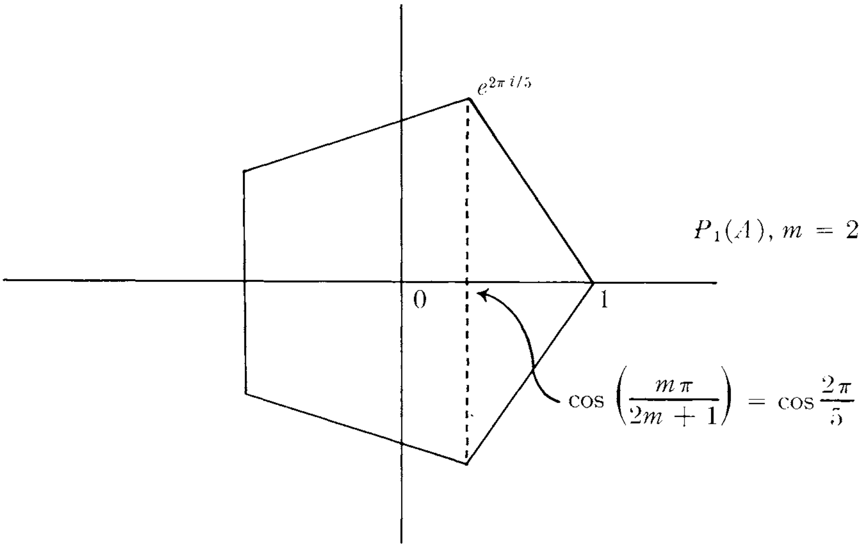
$$\sum_{i=0}^{m-2} \cos(\theta + k_i\omega) + \frac{|\epsilon|}{2},$$

while the largest sum of m elements of $\mathscr{R}_\theta$ is only

$$\sum_{i=0}^{m-2} \cos(\theta + k_i\omega) + \cos\left(\frac{m\pi}{2m+1}\right).$$

In view of Theorem 2, this contradicts our assumption that A is m-convex and establishes the "necessity" portion of (i).

Next, assume that $|\epsilon| \leqq 2\cos(m\pi/(2m+1))$. It is not hard to observe that for any $\theta \in [0, 2\pi)$, $P_1(A_\theta)$ has either m or $m+1$ vertices with real part at least $\cos(m\pi/(2m+1))$. (Rotation through successive angles θ of $P_1(A)$ for $m = 2$ may prove illuminating.)



Upon inspection of the sets $\mathscr{L}_\theta$ and $\mathscr{R}_\theta$ subject to the indicated bound on $|\epsilon|$, we conclude that for any integer $j \in \{1, \ldots, m\}$, the largest sum of j elements of $\mathscr{L}_\theta$ equals the largest sum of j elements of $\mathscr{R}_\theta$. It follows from Theorem 2 that A is j-convex for $j = 1, \ldots, m$. This proves the "sufficiency" portion of (i) and, combined with the "necessity" portion of (i), establishes (ii).

To prove (iii), we first remark trivially that if A is normal, then A is $(m+1)$-convex by Theorem 3. Conversely, assume that A is $(m+1)$-convex. If m is even, set $\theta = \pi$, while if m is odd, set $\theta = 0$. In either situation the regular odd-order convex polygon $P_1(A_\theta)$ has precisely m vertices with positive real part: let $0 \leqq k_1 < \ldots < k_m \leqq 2m$ be the integers k for which $\cos(\theta + k\omega) > 0$. Then the largest sum of $m+1$ elements of $\mathscr{L}_\theta$ is

$$(18) \quad \sum_{i=1}^{m} \cos(\theta + k_i\omega) + \frac{|\epsilon|}{2},$$

while the largest sum of $m+1$ elements of $\mathscr{R}_\theta$ is

$$(19) \quad \sum_{i=1}^{m} \cos(\theta + k_i\omega) \quad (+0).$$

By Theorem 2, the sum (18) must not exceed the sum (19). Hence $\epsilon = 0$ and A is normal.

5. Examples. Our first example indicates that for a normal matrix $A \in M_n(\mathbf{C})$ and a nonreal n-tuple $c = (c_1, \ldots, c_n) \in \mathbf{C}^n$, the proper inclusion

$$W_c(A) \subsetneqq P_c(A)$$

may obtain (see Theorem 1).

I. Let

$$A = \operatorname{diag}(i, 1, 0) \in M_3(\mathbf{C})$$

and

$$c = (1, i, 0) \in \mathbf{C}^3.$$

Then

$$P_c(A) = \mathscr{H}(2i, -1, 1, 0, i),$$

and Lemma 1 may be used to compute that no point on the line segment joining $2i$ and -1 (other than the two endpoints) belongs to $W_c(A)$.

We conclude with two concrete illustrations of the content of Theorem 4.

II. Let $m = 1$, so that $2m+1 = 3$ and $2m+3 = 5$. Then $\omega = 2\pi/3$ and

$$A = \operatorname{diag}(1, e^{2\pi i/3}, e^{4\pi i/3}) \dotplus \begin{bmatrix} 0 & \epsilon \\ 0 & 0 \end{bmatrix}.$$

From Theorem 4 we see that A is convex if and only if $|\epsilon| \leqq 2\cos\pi/3 = 1$, and A is 2-convex if and only if A is normal ($\epsilon = 0$).

III. Let $m = 2$, so that $2m+1 = 5$ and $2m+3 = 7$. Then $\omega = 2\pi/5$ and

$$A = \operatorname{diag}(1, e^{2\pi i/5}, e^{4\pi i/5}, e^{6\pi i/5}, e^{8\pi i/5}) \dotplus \begin{bmatrix} 0 & \epsilon \\ 0 & 0 \end{bmatrix}.$$

From Theorem 4 we see that A is 2-convex if and only if $|\epsilon| \leqq 2 \cos 2\pi/5 \doteq .618$; if A is 2-convex then it is convex; and A is 3-convex if and only if A is normal ($\epsilon = 0$). This particular example was the point of departure for our investigation.

References

1. P. A. Fillmore and J. P. Williams, *Some convexity theorems for matrices*, Glasgow Math. J. *12* (1971), 110–117.
2. P. R. Halmos, *A Hilbert space problem book* (Van Nostrand Co., Inc., 1967).
3. A. Horn, *Doubly stochastic matrices and the diagonal of a rotation matrix*, Amer. J. Math. *76* (1954), 620–630.
4. M. Marcus and H. Minc, *A survey of matrix theory and matrix inequalities* (Prindle, Weber and Schmidt, Inc., Boston, 1964).
5. M. Marcus, B. N. Moyls, and R. Westwick, *Some extreme value results for indefinite hermitian matrices, II*, Illinois J. Math. *2* (1958), 408–414.
6. B. N. Moyls and M. Marcus, *Field convexity of a square matrix*, Proc. Amer. Math. Soc. *6* (1955), 981–983.
7. F. D. Murnaghan, *On the field of values of a square matrix*, Proc. Nat. Acad. Sci. *18* (1932), 246–248.

University of California,
Santa Barbara, California