3×3 Orthostochastic Matrices and the Convexity of Generalized Numerical Ranges

Yik-Hoi Au-Yeung and Yiu-Tung Poon
University of Hong Kong
Hong Kong

Submitted by Olga Taussky Todd

ABSTRACT

Let $\mathcal{U}_3$ be the set of all 3×3 unitary matrices, and let A and B be two 3×3 complex normal matrices. In this note, the authors first give a necessary and sufficient condition for a 3×3 doubly stochastic matrix to be orthostochastic and then use this result to consider the structure of the sets $\mathcal{W}(A)=\{\operatorname{Diag} UAU^*: U\in\mathcal{U}_3\}$ and $W(A,B)=\{\operatorname{Tr} UAU^*B: U\in\mathcal{U}_3\}$, where * denotes the transpose conjugate.

1. INTRODUCTION

Let A and B be two $n\times n$ complex matrices, and let $\mathcal{U}_n$ be the set of all $n\times n$ unitary matrices. Define $\mathcal{W}(A)=\{\operatorname{Diag} UAU^*: U\in\mathcal{U}_n\}$ and $W(A,B)=\{\operatorname{Tr} UAU^*B: U\in\mathcal{U}_n\}$, where * denotes the transpose conjugate. Horn [3] proved that if A is Hermitian, then $\mathcal{W}(A)$ is convex. Au-Yeung and Sing [1] proved that if A is normal, then $\mathcal{W}(A)$ is convex if and only if the eigenvalues of A are collinear. Williams [7] characterized the structure of $\mathcal{W}(A)$ for a 3×3 normal matrix A. Westwick [6] (in an equivalent form) proved that if A is normal and the eigenvalues of A are collinear, then $W(A,B)$ is convex. He also gave an example of two 3×3 normal matrices A and B such that $W(A,B)$ is not convex.

An $n\times n$ doubly stochastic (d.s.) matrix (a_{ij}) is said to be orthostochastic (o.s.) if there exists $(u_{ij})\in\mathcal{U}_n$ such that $a_{ij}=|u_{ij}|^2$. The purpose of this note is (1) to give a necessary and sufficient condition for a 3×3 d.s. matrix to be o.s., (2) to give another characterization of the structure of $\mathcal{W}(A)$ for a normal 3×3 matrix A and (3) to give a necessary and sufficient condition for the convexity of $W(A,B)$ in terms of the eigenvalues of A and B for 3×3 normal matrices A and B.

 0024-3795/79/050069+11$01.75

2. ORTHOSTOCHASTIC MATRICES AND THE CONVEXITY OF GENERALIZED NUMERICAL RANGES

We first give a necessary and sufficient condition for a d.s. matrix to be o.s.

THEOREM 1. *Let (a_{ij}) be a 3×3 real matrix such that $\Sigma_{j=1}^{3}a_{ij}=1$ $(i=1,2,3)$ and $\Sigma_{i=1}^{3}a_{ij}=1$ $(j=1,2,3)$. Then*

(1) *if (a_{ij}) is o.s., then for any $j\neq j'$ and for any l*

$$\sqrt{a_{lj}a_{lj'}} \leqslant \sum_{\substack{i=1\\ i\neq l}}^{3} \sqrt{a_{ij}a_{ij'}} \; ; \tag{*}$$

(2) *conversely, if there exist $j\neq j'$ such that $a_{ij}\geqslant 0$, $a_{ij'}\geqslant 0$ $(i=1,2,3)$ and for any l, the inequality (*) holds, then (a_{ij}) is o.s.*

Proof. Suppose (a_{ij}) is o.s.; then there exist real numbers θ_{ij} $(i,j=1,2,3)$ such that $(\sqrt{a_{ij}}\, e^{\sqrt{-1}\,\theta_{ij}})$ is unitary. Hence for any $j\neq j'$

$$\sum_{i=1}^{3} \sqrt{a_{ij}a_{ij'}}\; e^{\sqrt{-1}\,(\theta_{ij}-\theta_{ij'})} = 0,$$

and consequently the inequality (*) follows.

Conversely, suppose there exist $j\neq j'$ such that the inequality (*) holds for any l. For definiteness, we assume $j=1$ and $j'=2$. Then the nonnegative numbers $\sqrt{a_{11}a_{12}}$, $\sqrt{a_{21}a_{22}}$, $\sqrt{a_{31}a_{32}}$ form the lengths of the three sides of a triangle. Hence there exist real numbers θ and ψ such that

$$\sqrt{a_{11}a_{12}} + \sqrt{a_{21}a_{22}}\; e^{\sqrt{-1}\,\theta} + \sqrt{a_{31}a_{32}}\; e^{\sqrt{-1}\,\psi} = 0.$$

Let $u_{i1}=\sqrt{a_{i1}}$ $(i=1,2,3)$ and $u_{12}=\sqrt{a_{12}}$, $u_{22}=\sqrt{a_{22}}\, e^{\sqrt{-1}\,\theta}$, $u_{32}=\sqrt{a_{32}}\, e^{\sqrt{-1}\,\psi}$, and (u_{13},u_{23},u_{33}) be any unit vector orthogonal to (u_{11}, u_{21},u_{31}) and (u_{12},u_{22},u_{32}). Then (u_{ij}) is unitary and $a_{ij}=|u_{ij}|^2$. ■

In the following we shall use A and B to denote two complex normal matrices with eigenvalues $\lambda_1,\lambda_2,\lambda_3$ and μ_1,μ_2,μ_3 respectively. It follows from the definitions that $\mathcal{W}(A)=\{(\lambda_1,\lambda_2,\lambda_3)(a_{ij}):(a_{ij})$ is a 3×3 o.s. matrix$\}$ and $W(A,B)=\{(\lambda_1,\lambda_2,\lambda_3)(a_{ij})(\mu_1,\mu_2,\mu_3)^T:(a_{ij})$ is a 3×3 o.s. matrix$\}$, where T denotes the transpose. From Theorem 1, we have

COROLLARY 1. *$(\gamma_1,\gamma_2,\gamma_3)\in \mathcal{W}(A)$ ($\gamma\in W(A,B)$ respectively) if and only if $(\gamma_1,\gamma_2,\gamma_3)=(\lambda_1,\lambda_2,\lambda_3)(a_{ij})$ ($\gamma=(\lambda_1,\lambda_2,\lambda_3)(a_{ij})(\mu_1,\mu_2,\mu_3)^T$ respectively), where (a_{ij}) is a d.s. matrix satisfying (*) for some $j\neq j'$ and for any l.*

Obviously, if $(\gamma_1,\gamma_2,\gamma_3)\in\mathcal{W}(A)$, then each γ_i $(i=1,2,3)$ is a convex combination of λ_1, λ_2 and λ_3 and $\gamma_1+\gamma_2+\gamma_3=\lambda_1+\lambda_2+\lambda_3$. The following theorem gives a characterization of $\mathcal{W}(A)$.

THEOREM 2. *Suppose λ_1, λ_2 and λ_3 are not collinear and $\gamma_1=\alpha_1\lambda_1+\alpha_2\lambda_2+\alpha_3\lambda_3$ ($\alpha_i\geqslant 0$, $\alpha_1+\alpha_2+\alpha_3=1$). Then $(\gamma_1,\gamma_2,\gamma_3)\in\mathcal{W}(A)$, where $\gamma_2=x\lambda_1+y\lambda_2+z\lambda_3$, $x,y,z\geqslant 0$, $x+y+z=1$ and $\gamma_3=\operatorname{Tr}A-(\gamma_1+\gamma_2)$, if and only if*

(i) $x\leqslant\alpha_2+\alpha_3$ and
(ii) $(\sqrt{\alpha_1\alpha_2 x}-\sqrt{\alpha_3\alpha_0})^2\leqslant(\alpha_2+\alpha_3)^2 y\leqslant(\sqrt{\alpha_1\alpha_2 x}+\sqrt{\alpha_3\alpha_0})^2$, where $\alpha_0=\alpha_2+\alpha_3-x$.

Proof. We first observe that

$$(\gamma_1,\gamma_2,\gamma_3)=(\lambda_1,\lambda_2,\lambda_3)\begin{bmatrix}\alpha_1 & x & 1-\alpha_1-x\\ \alpha_2 & y & 1-\alpha_2-y\\ \alpha_3 & z & 1-\alpha_3-z\end{bmatrix}$$

Now if $(\gamma_1,\gamma_2,\gamma_3)\in\mathcal{W}(A)$, then there exists an o.s. matrix (a_{ij}) such that

$$(\lambda_1,\lambda_2,\lambda_3)\begin{bmatrix}\alpha_1 & x & 1-\alpha_1-x\\ \alpha_2 & y & 1-\alpha_2-y\\ \alpha_3 & z & 1-\alpha_3-z\end{bmatrix}=(\lambda_1,\lambda_2,\lambda_3)(a_{ij}).$$

Since λ_1, λ_2 and λ_3 are not collinear, by comparing the coefficients we see that

$$\begin{bmatrix}\alpha_1 & x & 1-\alpha_1-x\\ \alpha_2 & y & 1-\alpha_2-y\\ \alpha_3 & z & 1-\alpha_3-z\end{bmatrix}=(a_{ij}).$$

Consequently, by Theorem 1, $(\gamma_1,\gamma_2,\gamma_3)\in\mathcal{W}(A)$ if and only if all the following three inequalities hold:

(1) $\sqrt{\alpha_1 x}\leqslant\sqrt{\alpha_2 y}+\sqrt{\alpha_3 z}$,
(2) $\sqrt{\alpha_2 y}\leqslant\sqrt{\alpha_1 x}+\sqrt{\alpha_3 z}$,
(3) $\sqrt{\alpha_3 z}\leqslant\sqrt{\alpha_1 x}+\sqrt{\alpha_2 y}$.

If $\alpha_2+\alpha_3=0$, from (1), $x=0$ and y can take any value between 0 and 1. So we may assume $\alpha_2+\alpha_3>0$ and notice that (1), (2) and (3) together are equivalent to

$$
\begin{aligned}
& (\sqrt{\alpha_1 x} - \sqrt{\alpha_3 z})^2 \leqslant \alpha_2 y \leqslant (\sqrt{\alpha_1 x} + \sqrt{\alpha_3 z})^2 \\
\Leftrightarrow\quad & -2\sqrt{\alpha_1\alpha_3 xz} \leqslant \alpha_2 y - \alpha_1 x - \alpha_3 z \leqslant 2\sqrt{\alpha_1\alpha_3 xz} \\
\Leftrightarrow\quad & [\alpha_2 y - \alpha_1 x - \alpha_3(1-x-y)]^2 \leqslant 4\alpha_1\alpha_3 x(1-x-y) \\
\Leftrightarrow\quad & (\alpha_2+\alpha_3)^2 y^2 - 2[\alpha_1\alpha_2 x + \alpha_3(\alpha_2+\alpha_3-x)]\, y + [\alpha_1 x - \alpha_3(1-x)]^2 \leqslant 0 \\
& \qquad\qquad\qquad\qquad\qquad\qquad (\because\quad \alpha_1+\alpha_2+\alpha_3=1) \\
\Leftrightarrow\quad & [(\alpha_2+\alpha_3)^2 y]^2 - 2[\alpha_1\alpha_2 x + \alpha_3(\alpha_2+\alpha_3-x)][(\alpha_2+\alpha_3)^2 y] \\
& \quad + [\alpha_1\alpha_2 x - \alpha_3(1-x) + \alpha_1\alpha_3]^2 \leqslant 0 \qquad (\because\quad \alpha_1+\alpha_2+\alpha_3=1).
\end{aligned}
$$

Putting $t=(\alpha_2+\alpha_3)^2 y$, then the above inequality holds for nonnegative real numbers t if and only if

$$\alpha_3\alpha_0 \geqslant 0 \quad\text{and}\quad \left(\sqrt{\alpha_1\alpha_2 x} - \sqrt{\alpha_3\alpha_0}\right)^2 \leqslant t \leqslant \left(\sqrt{\alpha_1\alpha_2 x} + \sqrt{\alpha_3\alpha_0}\right)^2,$$

which in turn are equivalent to (i) and (ii), since if $\alpha_3=0$, then

$$t = \alpha_2^2 y = \alpha_1\alpha_2 x \quad\Rightarrow\quad x = \alpha_2(x+y) \leqslant \alpha_2. \qquad \blacksquare$$

The following theorem shows that the matrix

$$C_0 = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

plays an important role in the consideration of 3×3 o.s. matrices.

THEOREM 3.[1] *A convex combination of a* 3×3 *o.s. matrix* (a_{ij}) *and* C_0 *is an o.s. matrix. Furthermore, the matrix* C_0 *is the unique o.s. matrix with this property.*

[1]Theorem 3 and Corollary 2 were also obtained by M. Goldberg and E. Straus (private communication). The authors are thankful to Straus for giving the second statement of Theorem 3 with a proof which is different from the one given here.

Proof. Let $0 \leqslant \alpha \leqslant 1$. Then for any l

$$\left[\sum_{\substack{i=1\\ i\neq l}}^{3} \sqrt{\left(\alpha a_{i1}+\frac{1-\alpha}{3}\right)\left(\alpha a_{i2}+\frac{1-\alpha}{3}\right)}\right]^2$$

$$=\sum_{\substack{i=1\\ i\neq l}}^{3}\left(\alpha a_{i1}+\frac{1-\alpha}{3}\right)\left(\alpha a_{i2}+\frac{1-\alpha}{3}\right)$$

$$+2\sqrt{\left(\alpha a_{i1}+\frac{1-\alpha}{3}\right)\left(\alpha a_{i2}+\frac{1-\alpha}{3}\right)\left(\alpha a_{i'1}+\frac{1-\alpha}{3}\right)\left(\alpha a_{i'2}+\frac{1-\alpha}{3}\right)}$$

$$(1\leqslant i<i'\leqslant 3,\quad i,i'\neq l)$$

$$\geqslant \sum_{\substack{i=1\\ i\neq l}}^{3}\left[\alpha^2 a_{i1}a_{i2}+\frac{\alpha(1-\alpha)}{3}(a_{i1}+a_{i2})+\left(\frac{1-\alpha}{3}\right)^2\right]+2\alpha^2\sqrt{a_{i1}a_{i2}a_{i'1}a_{i'2}}$$

$$(1\leqslant i<i'\leqslant 3,\quad i,i'\neq l)$$

$$=\alpha^2\left[\sum_{\substack{i=1\\ i\neq l}}^{3}\sqrt{a_{i1}a_{i2}}\right]^2+\frac{\alpha(1-\alpha)}{3}\sum_{\substack{i=1\\ i\neq l}}^{3}(a_{i1}+a_{i2})+2\left(\frac{1-\alpha}{3}\right)^2$$

$$\geqslant \alpha^2 a_{l1}a_{l2}+\frac{\alpha(1-\alpha)}{3}\sum_{\substack{i=1\\ i\neq l}}^{3}(a_{i1}+a_{i2})+\left(\frac{1-\alpha}{3}\right)^2 \qquad [\text{by } (*)]$$

$$=\alpha^2 a_{l1}a_{l2}+\frac{\alpha(1-\alpha)}{3}[2-(a_{l1}+a_{l2})]+\left(\frac{1-\alpha}{3}\right)^2$$

$$\geqslant \alpha^2 a_{l1}a_{l2}+\frac{\alpha(1-\alpha)}{3}(a_{l1}+a_{l2})+\left(\frac{1-\alpha}{3}\right)^2 \qquad (a_{l1}+a_{l2}\leqslant 1)$$

$$=\left[\sqrt{\left(\alpha a_{l1}+\frac{1-\alpha}{3}\right)\left(\alpha a_{l2}+\frac{1-\alpha}{3}\right)}\right]^2.$$

Hence, by Theorem 1, $\alpha(a_{ij})+(1-\alpha)C_0$ is o.s. for any $0\leqslant\alpha\leqslant 1$.

For uniqueness, call any o.s. matrix with such property a center. Let $C=(c_{ij})$ be a center and I the 3×3 identity matrix. Then, for any $0\leqslant\alpha\leqslant 1$, the matrix $(1-\alpha)I+\alpha C$ is o.s., and by Theorem 1 we have

$$\sqrt{\alpha c_{12}[1+\alpha(c_{11}-1)]}\leqslant\sqrt{\alpha c_{21}[1+\alpha(c_{22}-1)]}+\sqrt{\alpha^2 c_{31}c_{32}}$$

and

$$\sqrt{\alpha c_{21}[1+\alpha(c_{22}-1)]}\leqslant\sqrt{\alpha c_{12}[1+\alpha(c_{11}-1)]}+\sqrt{\alpha^2 c_{31}c_{32}}\,.$$

Hence

$$\left(\sqrt{c_{21}[1+\alpha(c_{22}-1)]}-\sqrt{c_{12}[1+\alpha(c_{11}-1)]}\right)^2\leqslant\alpha c_{31}c_{32}$$

for any $0<\alpha\leqslant 1$. This implies $c_{21}=c_{12}$. It is obvious that if C is a center, then for any permutation matrices P_1 and P_2, P_1CP_2 is also a center. Therefore, by the above argument we have $c_{ij}=\frac{1}{3}$ for $i,j=1,2,3$. ■

COROLLARY 2. *For any* $u\in\mathcal{W}(A)$ ($x\in W(A,B)$ *respectively*) *and any* $0\leqslant\alpha\leqslant 1$, $\alpha(\gamma,\gamma,\gamma)+(1-\alpha)u\in\mathcal{W}(A)$, *where* $\gamma=\frac{1}{3}(\lambda_1+\lambda_2+\lambda_3)$ $((\alpha/3)(\lambda_1+\lambda_2+\lambda_3)(\mu_1+\mu_2+\mu_3)+(1-\alpha)x\in W(A,B)$ *respectively*).

Let M_+ (M_-) denote the set of all 3×3 even (odd) permutation matrices. Define $\mathcal{V}_+=\{(\lambda_1,\lambda_2,\lambda_3)P:P\in M_+\}$, $\mathcal{V}_-=\{(\lambda_1,\lambda_2,\lambda_3)P:P\in M_-\}$, $V_+=\{(\lambda_1,\lambda_2,\lambda_3)P(\mu_1,\mu_2,\mu_3)^T:P\in M_+\}$, $V_-=\{(\lambda_1,\lambda_2,\lambda_3)P(\mu_1,\mu_2,\mu_3)^T:P\in M_-\}$. A permutation matrix is o.s. For a convex combination of two permutation matrices, we have the following theorems.

THEOREM 4. *For any* $P_1\in M_+$, $P_2\in M_-$ *and any* $0\leqslant\alpha\leqslant 1$, $\alpha P_1+(1-\alpha)P_2$ *is o.s.*

Proof. Without loss of generality, we may assume P_1 to be the identity matrix (otherwise we consider PP_1 and PP_2, where P is a permutation matrix). Then P_2 is obtained from P_1 by transposing two rows of P_1. For definiteness we assume

$$P_2=\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}.$$

Then obviously,

$$\alpha\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}+(1-\alpha)\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}=\begin{bmatrix}1&0&0\\0&\alpha&1-\alpha\\0&1-\alpha&\alpha\end{bmatrix}$$

is o.s. ∎

COROLLARY 3. *For any* $u\in\mathcal{V}_+$ $(x\in V_+$ *respectively*$)$, $v\in\mathcal{V}_-$ $(y\in V_-)$ *and any* $0\leqslant\alpha\leqslant 1$, *we have* $\alpha u+(1-\alpha)v\in\mathcal{W}(A)$ $(\alpha x+(1-\alpha)y\in W(A,B))$.

THEOREM 5. *For any distinct* P_1 *and* P_2 *in* M_+ *(or in* M_-*) and any* $0<\alpha<1$, $\alpha P_1+(1-\alpha)P_2$ *is not an o.s. matrix.*

Proof. Without loss of generality, we may assume P_1 to be the identity matrix. For definiteness, we assume

$$P_2=\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}.$$

Then for any $0<\alpha<1$,

$$\alpha\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}+(1-\alpha)\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}=\begin{bmatrix}\alpha&1-\alpha&0\\0&\alpha&1-\alpha\\1-\alpha&0&\alpha\end{bmatrix},$$

which, by Theorem 1, is obviously not o.s. ∎

Lerer [4] gave an example of a unitary matrix U such that $\mathcal{W}(U)$ is not convex. But by applying Theorem 5 and comparing coefficients, we have the following result.

COROLLARY 4. *If* $\lambda_1,\lambda_2,\lambda_3$ *are not collinear, then for any distinct* $u,v\in\mathcal{V}_+$ *(or* $\mathcal{V}_-$*) and any* $0<\alpha<1$, $\alpha u+(1-\alpha)v\notin\mathcal{W}(A)$.

For any two distinct complex numbers x and y, we shall denote by $L(x,y)$ the line passing x and y.

COROLLARY 5. *If* x,y *are two distinct points in* V_+ $(V_-$ *respectively*$)$ *such that all the points in* V_- $(V_+$ *respectively*$)$ *lie on one side (the open half plane) of* $L(x,y)$, *then* $\alpha x+(1-\alpha)y\notin W(A,B)$ *for any* $0<\alpha<1$.

Proof. Suppose there exist $x, y \in V_+$ (or V_-) and $0<\alpha<1$ such that $\alpha x+(1-\alpha)y \in W(A,B)$. Then there exists an o.s. matrix (a_{ij}) such that

$$\alpha x + (1-\alpha)y = (\lambda_1,\lambda_2,\lambda_3)(a_{ij})(\mu_1,\mu_2,\mu_3)^T.$$

By Birkhoff's theorem (for example, see [5]), (a_{ij}) is a convex combination of permutation matrices. Since all the points in V_- lie on one side of $L(x,y)$, and since the triangles $\mathcal{C}(V_+)$ and $\mathcal{C}(V_-)$, [where $\mathcal{C}(X)$ is the convex hull of X] have the same center $c_0=\frac{1}{3}(\lambda_1+\lambda_2+\lambda_3)(\mu_1+\mu_2+\mu_3)$, the third point z in V_+ (or V_- respectively) also lies on the same open half plane with the points in V_- (or V_+). Consequently, we have $(a_{ij})=\alpha P_1+(1-\alpha)P_2$, where P_1 and P_2 are in M_+ (or M_-), contradicting Theorem 5. ■

THEOREM 6. *$W(A,B)$ is not convex if and only if there exist distinct x and y in V_+ (or in V_-) such that all points in V_- (or V_+ respectively) lie on one side (the open half plane) of $L(x,y)$.*

Proof. For any two distinct complex numbers x and y, we denote by $S(x,y)$ the line segment joining x and y. It is known [2] that $\mathcal{C}(W(A,B))=\mathcal{C}(V_+\cup V_-)$. By Corollary 3, we see that if $x\in V_+$ and $y\in V_-$, then $S(x,y)\subset W(A,B)$, and by Corollary 2, if $x\in W(A,B)$, then $S(x,c_0)\subset W(A,B)$, where

$$c_0 = \frac{1}{3}(\lambda_1+\lambda_2+\lambda_3)(\mu_1+\mu_2+\mu_3) = \frac{1}{6}\sum_{x\in V_+\cup V_-} x = \frac{1}{3}\sum_{x\in V_+} x = \frac{1}{3}\sum_{x\in V_-} x.$$

Therefore, if $W(A,B)$ is not convex, then there exist distinct x and y in V_+ (or in V_-) and $0<\alpha<1$ such that $\alpha x+(1-\alpha)y\notin W(A,B)$. The third point z in V_+ (in V_- respectively) cannot lie on $L(x,y)$; otherwise, $c_0\in L(x,y)$ and consequently $S(x,y)\subset W(A,B)$. Now all points in V_- (in V_+ respectively) must lie on the same side with z (equivalently with c_0) with respect to $L(x,y)$, since if there exists x_0 in V_- (in V_+ respectively) such that x_0 lies on $L(x,y)$ or on the other side of $L(x,y)$, then $S(c_0,w)\subset W(A,B)$ for all $w\in S(x_0,x)\cup S(x_0,y)$ and consequently $S(x,y)\subset W(A,B)$.

The other part of the theorem is a consequence of Corollary 5. So the proof of the theorem is completed. ■

3. EXAMPLES

In the following figures, we use ○ to denote points in V_+ and × to denote points in V_-.

EXAMPLE 1 (see Fig. 1).

$$A = \begin{bmatrix} 0 & & \\ & 1 & \\ & & e^{\sqrt{-1}\,\pi/3} \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & & \\ & 1 & \\ & & e^{\sqrt{-1}\,2\pi/3} \end{bmatrix},$$

$$V_+ = \left\{ e^{\sqrt{-1}\,\pi/3}, e^{\sqrt{-1}\,2\pi/3}, 0 \right\},$$

$$V_- = \left\{ 1, -1, \sqrt{-3}\, \right\}.$$

$W(A,B)$ is convex.

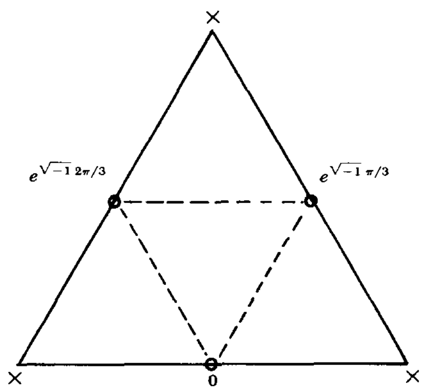

FIG. 1.

EXAMPLE 2 (see Fig. 2).

$$A = \begin{bmatrix} 0 & & \\ & 1 & \\ & & \tfrac{1}{2}e^{\sqrt{-1}\,\pi/12} \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & & \\ & 1 & \\ & & e^{\sqrt{-1}\,\pi/3} \end{bmatrix},$$

$$V_+ = \left\{ \tfrac{1}{2}e^{\sqrt{-1}\,\pi/12}, e^{\sqrt{-1}\,\pi/3}, 1 + \tfrac{1}{2}e^{\sqrt{-1}\,5\pi/12} \right\},$$

$$V_- = \left\{ 1, \tfrac{1}{2}e^{\sqrt{-1}\,5\pi/12}, \tfrac{1}{2}e^{\sqrt{-1}\,\pi/12} + e^{\sqrt{-1}\,\pi/3} \right\}.$$

$W(A,B)$ is convex.

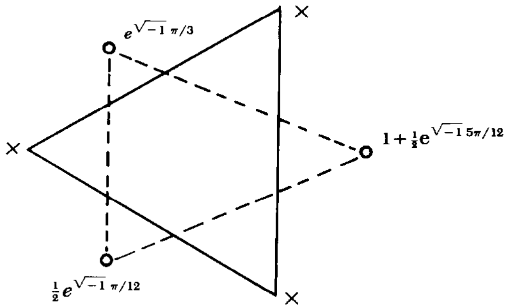

FIG. 2.

EXAMPLE 3 (see Fig. 3).

$$A = \begin{bmatrix} 0 & & \\ & 1 & \\ & & \alpha \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & & \\ & 1 & \\ & & \bar{\alpha} \end{bmatrix}, \qquad \alpha \text{ is not real,}$$

$$V_+ = \{\alpha, \bar{\alpha}, 1+\alpha\bar{\alpha}\},$$

$$V_- = \{1, \alpha\bar{\alpha}, \alpha+\bar{\alpha}\}.$$

$W(A,B)$ is not convex.

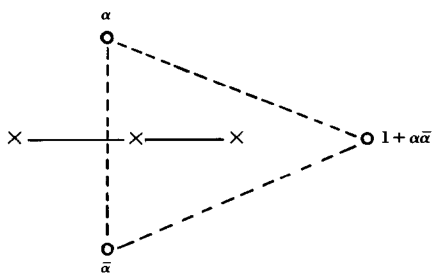

FIG. 3.

EXAMPLE 4 (see Fig. 4).

$$A = \begin{bmatrix} 0 & & \\ & 1 & \\ & & \sqrt{-1} \end{bmatrix} \quad B = \begin{bmatrix} 0 & & \\ & 1 & \\ & & \sqrt{-1} \end{bmatrix},$$

$$V_+ = \{\sqrt{-1}\,, \sqrt{-1}\,, 0\},$$

$$V_- = \{1, -1, 2\sqrt{-1}\,\}.$$

$W(A,B)$ is not convex. (Westwick [6] has considered this example.)

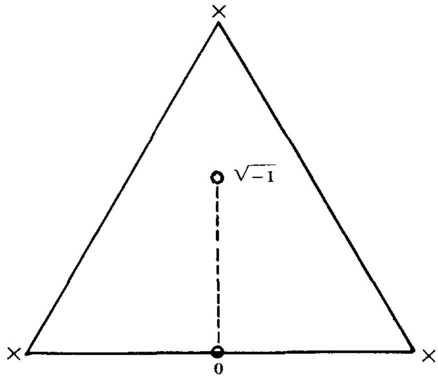

FIG. 4.

REFERENCES

1 Y. H. Au-Yeung and F. Y. Sing, A remark on the generalized numerical range of a normal matrix, *Glasgow Math. J.* 18:179–180 (1977).

2 P. A. Fillmore and J. P. Williams, Some convexity theorems for matrices, *Glasgow Math. J.* 12:110–116 (1971).

3 A. Horn, Doubly stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math.* 76:620–630 (1954).

4 L. E. Lerer, On the diagonal elements of normal matrices (in Russian), *Mat. Issled.* 2:156–163 (1967).

5 L. Mirsky, Proofs of two theorems on doubly-stochastic matrices, *Proc. Amer. Math. Soc.* 9:371–374 (1958).

6 R. Westwick, A theorem on numerical range, *Linear and Multilinear Algebra* 2:311–315 (1975).

7 J. P. Williams, On compressions of matrices, *J. London Math. Soc.* 3:526–532 (1971).

Received 4 November 1978; revised 18 December 1978