A REMARK ON A CONJECTURE OF MARCUS ON THE GENERALIZED NUMERICAL RANGE

by YIK-HOI AU-YEUNG and KAM-CHUEN NG

(Received 10 April, 1982)

1. Introduction. Let A be an $n\times n$ complex matrix and $c=(c_1,\ldots,c_n)\in\mathbb{C}^n$. Define the c-numerical range of A to be the set $W_c(A)=\left\{\sum_{j=1}^{n} c_jx_jAx_j^*:\{x_1,\ldots,x_n\}\text{ is an orthonormal set in }\mathbb{C}^n\right\}$, where * denotes the conjugate transpose. Westwick [**8**] proved that if $c_1,\ldots,c_n$ are collinear, then $W_c(A)$ is convex. (Poon [**6**] gave another proof.) But in general for $n\geqq 3$, $W_c(A)$ may fail to be convex even for normal A (for example, see Marcus [**4**] or Lemma 3 in this note) though it is star-shaped (Tsing [**7**]). In the following, we shall assume that A is normal. Let $\mathcal{W}(A)=\{\operatorname{diag} UAU^*:U\text{ is unitary}\}$. Horn [**3**] proved that if the eigenvalues of A are collinear, then $\mathcal{W}(A)$ is convex. Au-Yeung and Sing [**2**] showed that the converse is also true. Marcus [**4**] further conjectured (and proved for $n=3$) that if $W_c(A)$ is convex for all $c\in\mathbb{C}^n$, then the eigenvalues of A are collinear. Let $\lambda=(\lambda_1,\ldots,\lambda_n)\in\mathbb{C}^n$. We denote by $\bar\lambda$ the vector $(\bar\lambda_1,\ldots,\bar\lambda_n)$ and by $[\lambda]$ the diagonal matrix with $\lambda_1,\ldots,\lambda_n$ lying on its diagonal. Since, for any unitary matrix U, $W_c(A)=W_c(UAU^*)$, the Marcus conjecture reduces to: if $W_c([\lambda])$ is convex for all $c\in\mathbb{C}^n$, then $\lambda_1,\ldots,\lambda_n$ are collinear. For the case $n=3$, Au-Yeung and Poon [**1**] gave a complete characterization on the convexity of the set $W_c([\lambda])$ in terms of the relative position of the points $\sum_{j=1}^{3} c_j\lambda_{\sigma(j)}$, where $\sigma\in S_3$, the permutation group of order 3. As an example they showed that if λ_1, λ_2, λ_3 are not collinear, then $W_{\bar\lambda}([\lambda])$ is not convex (Lemma 3 in this note gives another proof). We shall show that for the case $n=4$, $W_{\bar\lambda}([\lambda])$ is not convex if λ_1, λ_2, λ_3, λ_4 are not collinear. Thus for $n=3$, 4 the Marcus conjecture is answered and improved.

2. The convexity of c-numerical range. A real nonnegative $n\times n$ matrix $S=(s_{ij})$ is said to be *doubly-stochastic* if every row and column sum of S is 1 and orthostochastic (o.s.) if there exists a unitary matrix $U=(u_{ij})$ such that $s_{ij}=|u_{ij}|^2$, $i,j=1,\ldots,n$.

The following two lemmas are obvious.

LEMMA 1. *$W_c([\lambda])=\{\lambda Sc^T:S$ is o.s.$\}$, where T denotes the transpose.*

LEMMA 2. *Let $\mu=\alpha\lambda+\beta(1,1,\ldots,1)$, where $\alpha,\beta\in\mathbb{C}$ and $\alpha\neq 0$. Then, for any $c\in\mathbb{C}^n$, $W_c([\lambda])$ is convex if and only if $W_c([\mu])$ is convex.*

We shall denote $W_{\bar\lambda}([\lambda])$ by $W(\lambda)$. Let $\mathcal{P}_n$ be the set of all $n\times n$ permutation matrices and $\mathcal{P}(\lambda)=\{\lambda P\lambda^*:P\in\mathcal{P}_n\}$. Then, by Lemma 1 and Birkhoff's Theorem (for example, see Mirsky [**5**]), we have $\mathcal{P}(\lambda)\subset W(\lambda)\subset$ convex hull of $\mathcal{P}(\lambda)$. Let $m(\lambda)=\lambda I\lambda^*$, where I is the identity matrix. Then $m(\lambda)$ is positive if $\lambda\neq 0$ and is of largest magnitude

https://doi.org/10.1017/S0017089500005280 Published online by Cambridge University Press

among points in $\mathcal{P}(\lambda)$. It is obvious that both $\mathcal{P}(\lambda)$ and $W(\lambda)$ are symmetric about the real axis. Suppose $\lambda_1, \ldots, \lambda_n$ are not collinear. Consider the lines joining a point in $\mathcal{P}(\lambda)\setminus\{m(\lambda)\}$ and $m(\lambda)$; we see that there are two such lines L (in fact one is the reflection of the other about the real axis) such that all points in $\mathcal{P}(\lambda)$ lie on or on one side of L. For $n=4$, we shall show that if $m \in \mathcal{P}(\lambda)\setminus\{m(\lambda)\}$ is a point on L which is nearest to $m(\lambda)$, then there are points in the open line segment $(m, m(\lambda))$ which are not in $W(\lambda)$. Before we prove our result, we need the following lemma.

LEMMA 3. *If* $\lambda = (1, \alpha, 0, 0, \ldots, 0) \in \mathbb{C}^n$, *where* α *is not real, then* $W(\lambda)$ *is not convex for* $n \geqq 3$.

Proof. Obviously, we have $m(\lambda) = 1 + \alpha\bar{\alpha}$ and the non-real points in $\mathcal{P}(\lambda)$ are α and $\bar{\alpha}$. Let $V(\xi) = \{P \in \mathcal{P}_n : \lambda P \lambda^* = \xi\}$, where $\xi \in \mathbb{C}$. If $(p_{ij}) \in V(\alpha)$, then $p_{11} = p_{12} = 0$ and $p_{21} = 1$ and if $(p_{ij}) \in V(1+\alpha\bar{\alpha})$, then $p_{11} = p_{22} = 1$. Suppose that $0 < t < 1$ and $z = (1-t)\alpha + t(1+\alpha\bar{\alpha}) \in W(\lambda)$. Then, by Lemma 1, there exists an o.s. matrix S such that $z = \lambda S \lambda^*$. By Birkhoff's theorem, S is a convex combination of permutation matrices. Since z is on the open line segment $(\alpha, 1+\alpha\bar{\alpha})$ and all other points in $\mathcal{P}(\lambda)$ lie on one side of the line joining α and $1+\alpha\bar{\alpha}$, we have

$$S = (1-t)\Big(\sum_{p \in V(\alpha)} t_P P\Big) + t\Big(\sum_{P \in V(1+\alpha\bar{\alpha})} t_P P\Big),$$

where $\sum_{P \in V(\alpha)} t_P = \sum_{P \in V(1+\alpha\bar{\alpha})} t_P = 1$ and $t_P \geqq 0$ for all $P \in V(\alpha) \cup V(1+\alpha\bar{\alpha})$.

Considering the first two columns of S, we see that S cannot be o.s. Hence $W(\lambda)$ is not convex.

THEOREM 4. *Let* $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4) \in \mathbb{C}^4$. *If* $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ *are not collinear, then* $W(\lambda)$ *is not convex.*

Proof. There are at most 24 points in $\mathcal{P}(\lambda)$ and since there are 10 symmetric permutation matrices in $\mathcal{P}_4$, there are at most 7 points in $\mathcal{P}(\lambda)$ lying on the upper (lower) half-plane. By Lemma 2 and Lemma 3, we may assume $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are distinct. Then I is the only permutation matrix corresponding to $m(\lambda)$. Using Lemma 2 and the fact that $W(\lambda) = W(\bar{\lambda})$, we may assume $\lambda = (r_1, r_2, r_3 e^{i\theta}, -r_4 e^{i\theta})$ or $\lambda = (r_1, -r_2, r_3 e^{i\theta}, -r_4 e^{i\theta})$, where $r_1 \geqq 0$, $r_2 \geqq 0$, $r_3 > 0$, $r_4 > 0$ and $0 < \theta \leqq \pi/2$ according as the convex hull of $\{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ has 3 or 4 vertices.

Case 1. $\lambda = (r_1, r_2, r_3 e^{i\theta}, -r_4 e^{i\theta})$

In this case, for definiteness, we assume $r_1 > r_2$. Then the 7 possible points in $\mathcal{P}(\lambda)$ and on the upper half-plane are:

$$m_1 = \lambda P_1 \lambda^* = r_3^2 + r_1 r_2 - (r_1 r_4 + r_2 r_4)\cos\theta + i(r_1 r_4 - r_2 r_4)\sin\theta,$$
$$m_2 = \lambda P_2 \lambda^* = r_2^2 - r_3 r_4 + (r_1 r_3 - r_1 r_4)\cos\theta + i(r_1 r_3 + r_1 r_4)\sin\theta,$$
$$m_3 = \lambda P_3 \lambda^* = r_4^2 + r_1 r_2 + (r_1 r_3 + r_2 r_3)\cos\theta + i(r_1 r_3 - r_2 r_3)\sin\theta,$$
$$m_4 = \lambda P_4 \lambda^* = r_1^2 - r_3 r_4 + (r_2 r_3 - r_2 r_4)\cos\theta + i(r_2 r_3 + r_2 r_4)\sin\theta,$$

https://doi.org/10.1017/S0017089500005280 Published online by Cambridge University Press

$$m_5 = \lambda P_5 \lambda^* = (r_1 + r_2)(r_3 - r_4)\cos\theta + i(r_1 - r_2)(r_3 + r_4)\sin\theta,$$
$$m_6 = \lambda P_6 \lambda^* = r_1 r_2 - r_3 r_4 + (r_1 r_3 - r_2 r_4)\cos\theta + i(r_1 r_3 + r_2 r_4)\sin\theta,$$
$$m_7 = \lambda P_7 \lambda^* = r_1 r_2 - r_3 r_4 + (r_2 r_3 - r_1 r_4)\cos\theta + i(r_2 r_3 + r_1 r_4)\sin\theta.$$

The permutations corresponding to the permutation matrices $P_1, P_2, \dots, P_7$ are (142), (143), (123), (243), (1423), (1243), (1432) respectively. Let $m \in \{m_k : k = 1, \dots, 7\}$ be the point on L which is nearest to $m(\lambda)$. Suppose that $0 < t \leqq \frac{1}{2}$ and $z = (1-t)m(\lambda) + tm \in W(\lambda)$; then using similar argument as in the proof of Lemma 3, there exists an o.s. matrix S such that $z = \lambda S \lambda^*$ and

$$S = t_1 P_1 + t_2 P_2 + \dots + t_7 P_7 + t_8 I$$
$$= \begin{bmatrix} t_4 + t_8 & t_3 + t_6 & 0 & t_1 + t_2 + t_5 + t_7 \\ t_1 + t_7 & t_2 + t_8 & t_3 + t_5 & t_4 + t_6 \\ t_2 + t_3 + t_5 + t_6 & t_4 + t_7 & t_1 + t_8 & 0 \\ 0 & t_1 + t_5 & t_2 + t_4 + t_6 + t_7 & t_3 + t_8 \end{bmatrix}$$

where $\frac{1}{2} \leqq 1 - t \leqq t_8 < 1$, $\sum_{j=1}^{8} t_j = 1$ and $t_j \geqq 0$ $(j = 1, \dots, 8)$. Since $t_8 \geqq \frac{1}{2} \geqq t_7$, from the first and the third columns of S, we have $t_2 = t_3 = t_5 = t_6 = 0$ and then from the first and the second columns of S, we have $t_1 = t_7 = 0$. Now columns 3 and 4 give $t_4 = 0$. This is a contradiction.

Case 2. $\lambda = (r_1, -r_2, r_3 e^{i\theta}, -r_4 e^{i\theta})$

In this case we cannot use the method as in case 1. We first have to eliminate two points in $\mathscr{P}(\lambda)$ and on the upper (closed) half-plane that cannot lie on L. Replace r_2 by $-r_2$ in m_k $(k = 1, \dots, 7)$ in case 1 and still denote them by m_k. Since we are considering points in the upper half-plane, we take $\bar{m}_4$ instead of m_4. If $r_1 r_3 - r_2 r_4 < 0$, we take $\bar{m}_6$ and if $r_1 r_4 - r_2 r_3 < 0$, we take $\bar{m}_7$. By comparing the slopes of the lines joining $m(\lambda)$ and m_5, $m(\lambda)$ and m_l (or $\bar{m}_l$) $(l = 6, 7)$, and by direct calculation, we see that m_6, $\bar{m}_6$, m_7 and $\bar{m}_7$ cannot lie on L. So the possible points in $\mathscr{P}(\lambda)$ and on the upper half-plane that lie on L are m_1, m_2, m_3, $\bar{m}_4$ and m_5. Let $m \in \{m_1, m_2, m_3, \bar{m}_4, m_5\}$ be the point on L which is nearest to $m(\lambda)$. Suppose $0 < t < 1$ and $z = (1-t)m(\lambda) + tm \in W(\lambda)$. Then, as in case 1, there is an o.s. matrix S such that $z = \lambda S \lambda^*$ and

$$S = t_1 P_1 + t_2 P_2 + t_3 P_3 + t_4 P_4^{\mathrm{T}} + t_5 P_5 + t_6 I$$
$$= \begin{bmatrix} t_4 + t_6 & t_3 & 0 & t_1 + t_2 + t_5 \\ t_1 & t_2 + t_6 & t_3 + t_4 + t_5 & 0 \\ t_2 + t_3 + t_5 & 0 & t_1 + t_6 & t_4 \\ 0 & t_1 + t_4 + t_5 & t_2 & t_3 + t_6 \end{bmatrix},$$

where $0 < 1 - t \leqq t_6 < 1$, $\sum_{j=1}^{6} t_j = 1$ and $t_j \geqq 0$ $(j = 1, \dots, 6)$. From the first and the third columns of S, we have

$$(t_2 + t_3 + t_5)(t_1 + t_6) = t_1(t_3 + t_4 + t_5). \tag{1}$$

https://doi.org/10.1017/S0017089500005280 Published online by Cambridge University Press

Since $t_6 > 0$, we have $t_4 \geqq t_2$. From the second and the fourth columns of S, we have

$$(t_1 + t_4 + t_5)(t_3 + t_6) = t_3(t_1 + t_2 + t_5).$$

Since $t_6 > 0$, we have $t_2 \geqq t_4$. Hence $t_2 = t_4$. From (1) $t_2 = t_3 = t_4 = t_5 = 0$. Now column 1 and 2 give $t_1 = 0$. This is a contradiction.

Hence in both cases $W(\lambda)$ is not convex.

Added in Proof. Very recently Au-Yeung and Tsing by using less constructive method have proved that for arbitrary n if the coordinates of λ are not collinear, then there is a very small portion of the line L which is not in $W(\lambda)$ and consequently have proved Theorem 4 for general n. This will be published in due course.

REFERENCES

1. Y. H. Au-Yeung and Y. T. Poon, 3×3 orthostochastic matrices and the convexity of generalized numerical ranges, *Linear Algebra and Appl.* **27** (1979), 69–79.

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

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

4. M. Marcus, Some combinatorial aspects of numerical range, *Ann. New York Acad. Sci.* **319** (1979), 368–376.

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

6. Y. T. Poon, Another proof of a result of Westwick, *Linear and Multilinear Algebra* **9** (1980), 35–37.

7. N. K. Tsing, On the shape of the generalized numerical ranges, *Linear and Multilinear Algebra* **10** (1981), 173–182.

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

Department of Mathematics
University of Hong Kong