AN EXTENSION OF THE HAUSDORFF-TOEPLITZ THEOREM ON THE NUMERICAL RANGE

YIK-HOI AU-YEUNG AND NAM-KIU TSING

ABSTRACT. Let $\mathcal{H}_n$ be the set of all $n \times n$ hermitian matrices and $\mathcal{U}_n$ the set of all $n \times n$ unitary matrices. For any $c = (c_1, \dots, c_n) \in \mathbf{R}^n$ and $A_1, A_2, A_3 \in \mathcal{H}_n$, let $W(A, A, A)$ denote the set

$$\{(\operatorname{tr}[c]UA_1U^*, \operatorname{tr}[c]UA_2U^*, \operatorname{tr}[c]UA_3U^*): U \in \mathcal{U}_n\},$$

where $[c]$ is the diagonal matrix with $c_1, \dots, c_n$ as diagonal entries. In this present note, the authors prove that if $n > 2$, then $W_c(A_1, A_2, A_3)$ is always convex. Equivalent statements of this result, in terms of definiteness and inclusion relations, are also given. These results extend the theorems of Hausdorff-Toeplitz, Finsler and Westwick on numerical ranges, respectively.

1. Introduction. Let $\mathcal{H}_n$ be the set of all $n \times n$ (complex) hermitian matrices and $\mathcal{U}_n$ the set of all $n \times n$ unitary matrices. For any $c = (c_1, \dots, c_n) \in \mathbf{R}^n$ and $A_1, \dots, A_p \in \mathcal{H}_n$, let $W_c(A_1, \dots, A_p) = \{(\operatorname{tr}[c]UA_1U^*, \dots, \operatorname{tr}[c]UA_pU^*): U \in \mathcal{U}_n\}$, where 'tr' means the trace and $[c]$ is the diagonal matrix with $c_1, \dots, c_n$ as diagonal entries. In the 1910's Toeplitz and Hausdorff [12, 10] proved (in an equivalent form) that $W_c(A_1, A_2)$ is convex for $c = (1, 0, \dots, 0)$. In the 1960's Halmos [9] conjectured and Berger [3] proved that if k coordinates of c are 1 and the rest are 0, then $W_c(A_1, A_2)$ is convex. In the 1970's, it was Westwick [13] who proved that $W_c(A_1, A_2)$ is convex for any $c \in \mathbf{R}^n$ (for another proof see Poon [11]). By using a result of Bohnenblust [4] (see Friedland and Loewy [7] for another proof) it is not difficult to see (for example, see [1]) that $W_c(A_1, \dots, A_p)$ is convex for $c = (1, 0, \dots, 0)$, $p = 3$ and $n > 2$ and this result is best possible in the sense that the result fails to hold if $p > 3$ or $n = 2$. As from Toeplitz-Hausdorff to Westwick it is quite natural to ask whether $W_c(A_1, A_2, A_3)$ is convex for $n > 2$ and for any $c \in \mathbf{R}^n$. In this note, we shall give an affirmative answer to this question.

2. Convexity of $W_c(A_1, A_2, A_3)$. For any $b, c \in \mathbf{R}^n$, we write $b \ll c$ if there is a doubly-stochastic matrix D (i.e. D is nonnegative and all row sums and column sums are equal to 1) such that $b = cD$. The proof of the following lemma is similar to that of Theorem 1 in [2], although there we consider only the case $p = 2$.

LEMMA 1. *For any $c = (c_1, \dots, c_n) \in \mathbf{R}^n$ and for any $A_1, \dots, A_p \in \mathcal{H}_n$, the following conditions are equivalent:*

- (i) $W_c(A_1, \dots, A_p)$ is convex.
- (ii) If $(r_1, \dots, r_p) \in \mathbf{R}^p \setminus W_c(A_1, \dots, A_p)$, then there exist $\alpha_1, \dots, \alpha_p \in \mathbf{R}$ such that

$$\operatorname{tr}\left([c]U\left(\sum_{j=1}^p \alpha_j A_j\right)U^*\right) > \sum_{j=1}^p \alpha_j r_j, \quad \text{for all } U \in \mathcal{U}_n.$$

Received by the editors November 23, 1982.

1980 *Mathematics Subject Classification.* Primary 47A12; Secondary 15A60.

Key words and phrases. Numerical range, convexity, doubly-stochastic matrix, convex hull.

©1983 American Mathematical Society
0002-9939/83 \$1.00 + \$.25 per page

(iii) If $(r_1, \dots, r_p) \in \mathbf{R}^p \setminus W_c(A_1, \dots, A_p)$, then there exist $\alpha_1, \dots, \alpha_p \in \mathbf{R}$ such that

$$\sum_{i=1}^n c_i \lambda_{\sigma(i)} > \sum_{j=1}^p \alpha_j r_j,$$

for all $\sigma \in S_n$, where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $\sum_{j=1}^p \alpha_j A_j$ and S_n is the permutation group of order n .

(iv) $W_b(A_1, \dots, A_p) \subset W_c(A_1, \dots, A_p)$ for all $b \ll c$.

An $n \times n$ doubly-stochastic matrix $Q = (q_{ij})$ is called a pinching matrix if for some $1 \leq k < l \leq n$ and $0 \leq \mu \leq 1$,

$$q_{ij} = \begin{cases} \mu, & \text{if } (i, j) = (k, k) \text{ or } (l, l), \\ 1 - \mu, & \text{if } (i, j) = (k, l) \text{ or } (l, k), \\ 1, & \text{if } i = j \neq k, l, \\ 0, & \text{otherwise.} \end{cases}$$

The following lemma is well known (for example, see [8]).

LEMMA 2. For any $b, c \in \mathbf{R}^n$, we have $b \ll c$ if and only if $b = cQ$ where Q is a finite product of pinching matrices.

The idea used in the proof of the following theorem has been used in [2].

THEOREM. If $c \in \mathbf{R}^n$, $A_1, A_2, A_3 \in \mathcal{H}_n$ and $n > 2$, then the following conditions hold:

(i) $W_c(A_1, A_2, A_3)$ is convex.

(ii) If $(r_1, r_2, r_3) \in \mathbf{R}^3 \setminus W_c(A_1, A_2, A_3)$, then there exist $\alpha_1, \alpha_2, \alpha_3 \in \mathbf{R}$ such that

$$\text{tr} \left([c] U \left(\sum_{j=1}^3 \alpha_j A_j \right) U^* \right) > \sum_{j=1}^3 \alpha_j r_j,$$

for all $U \in \mathcal{U}_n$.

(iii) If $(r_1, r_2, r_3) \in \mathbf{R}^3 \setminus W_c(A_1, A_2, A_3)$, then there exist $\alpha_1, \alpha_2, \alpha_3 \in \mathbf{R}$ such that

$$\sum_{i=1}^n c_i \lambda_{\sigma(i)} > \sum_{j=1}^3 \alpha_j r_j,$$

for all $\sigma \in S_n$, where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $\sum_{j=1}^3 \alpha_j A_j$ and S_n is the permutation group of order n .

(iv) $W_b(A_1, A_2, A_3) \subset W_c(A_1, A_2, A_3)$ for all $b \ll c$.

PROOF. By Lemmas 1 and 2, it suffices to prove that

$$W_b(A_1, A_2, A_3) \subset W_c(A_1, A_2, A_3),$$

where

$$\begin{cases} b_k = \mu c_k + (1 - \mu) c_l, \\ b_l = (1 - \mu) c_k + \mu c_l, \\ b_i = c_i \text{ for all } i \neq k, l, \end{cases}$$

for some $1 \leq k < l \leq n$ and $0 \leq \mu \leq 1$. We may also assume, without loss of generality, that $(k, l) = (1, 2)$. Let Λ_n denote the set of all (ordered) orthonormal

bases in $\mathbf{C}^n$. Then $W_c(A_1, A_2, A_3)$ is equal to the set

$$\left\{ \left(\sum_{i=1}^n c_i x_i A_1 x_i^*, \sum_{i=1}^n c_i x_i A_2 x_i^*, \sum_{i=1}^n c_i x_i A_3 x_i^* \right) : (x_1, \dots, x_n) \in \Lambda_n \right\}.$$

Let $(r_1, r_2, r_3) \in W_b(A_1, A_2, A_3)$. Then there is $(e_1, \dots, e_n) \in \Lambda_n$ such that

$$\begin{aligned} r_j &= \sum_{i=1}^n b_i e_i A_j e_i^* \\ &= \frac{1}{2}(b_1 + b_2)(e_1 A_j e_1^* + e_2 A_j e_2^*) + \frac{1}{2}(b_1 - b_2)(e_1 A_j e_1^* - e_2 A_j e_2^*) + \sum_{i=3}^n b_i e_i A_j e_i^* \end{aligned}$$

for $j = 1, 2, 3$. For any $\theta, \phi \in \mathbf{R}$ and $(x_1, \dots, x_n) \in \Lambda_n$, define $y_1 = \cos \theta x_1 + \sin \theta e^{\phi\sqrt{-1}} x_2$, $y_2 = -\sin \theta x_1 + \cos \theta e^{\phi\sqrt{-1}} x_2$ and $y_i = x_i$ for all $i \geq 3$. Then $(y_1, \dots, y_n) \in \Lambda_n$ and for $j = 1, 2, 3$,

$$\begin{aligned} \sum_{i=1}^n c_i y_i A_j y_i^* &= \frac{1}{2}(c_1 + c_2)(x_1 A_j x_1^* + x_2 A_j x_2^*) \\ &\quad + \frac{1}{2}(c_1 - c_2)[p_j \cos 2\theta + \sin 2\theta(q_j \cos \phi + s_j \sin \phi)] \\ &\quad + \sum_{i=3}^n c_i x_i A_j x_i^* \end{aligned}$$

where

$$p_j = x_1 A_j x_1^* - x_2 A_j x_2^*, \quad q_j = 2\operatorname{Re}(x_1 A_j x_2^*), \quad s_j = 2\operatorname{Im}(x_1 A_j x_2^*)$$

are real scalars. As θ and ϕ vary in $\mathbf{R}$, the locus of the point

$$\left(\sum_{i=1}^n c_i y_i A_1 y_i^*, \sum_{i=1}^n c_i y_i A_2 y_i^*, \sum_{i=1}^n c_i y_i A_3 y_i^* \right)$$

in $\mathbf{R}^3$ is an ellipsoid, which we shall denote by $E_{(x_1, \dots, x_n)}$ and is degenerate (hence convex) if $c_1 = c_2$ or

$$\det \begin{pmatrix} p_1 & q_1 & s_1 \\ p_2 & q_2 & s_2 \\ p_3 & q_3 & s_3 \end{pmatrix} = 0.$$

It is clear that $E_{(x_1, \dots, x_n)} \subset W_c(A_1, A_2, A_3)$ for any $(x_1, \dots, x_n) \in \Lambda_n$. Since $c_1 + c_2 = b_1 + b_2$, $|c_1 - c_2| \geq |b_1 - b_2|$ and $c_i = b_i$ for all $i \geq 3$, $(r_1, r_2, r_3) \in \operatorname{conv} E_{(e_1, \dots, e_n)}$ where "conv" denotes the convex hull. If $|c_1 - c_2| = |b_1 - b_2|$ or $E_{(e_1, \dots, e_n)}$ degenerates, then $(r_1, r_2, r_3) \in E_{(e_1, \dots, e_n)} \subset W_c(A_1, A_2, A_3)$. So we may assume that $E_{(e_1, \dots, e_n)}$ does not degenerate and (r_1, r_2, r_3) lies in the interior of the ellipsoid $E_{(e_1, \dots, e_n)}$. Let $w_1 \in \mathbf{C}^n$ be a unit eigenvector of A_1 . Since $n > 2$, we can choose a unit vector $w_2 \in \mathbf{C}^n$ which is orthogonal to both w_1 and $w_1 A_2$. Consider a continuous function $f: [0, 1] \rightarrow \Lambda_n$ such that $f(0) = (e_1, \dots, e_n)$ and $f(1) = (w_1, \dots, w_n)$, where $(w_1, \dots, w_n)$ is any extension of (w_1, w_2) to an orthonormal basis of $\mathbf{C}^n$. Now (r_1, r_2, r_3) lies within the interior of $E_{f(0)}$, and $E_{f(1)}$ degenerates (hence convex) as $w_1 A_j w_2^* = 0$ for $j = 1, 2$. If $(r_1, r_2, r_3) \in E_{f(1)}$, then (r_1, r_2, r_3) is automatically in $W_c(A_1, A_2, A_3)$. If $(r_1, r_2, r_3) \notin E_{f(1)}$, then, by continuity, there exists $\alpha \in (0, 1)$ such that $(r_1, r_2, r_3) \in E_{f(\alpha)}$. Thus $(r_1, r_2, r_3) \in W_c(A_1, A_2, A_3)$. $\square$

By taking $c = (1, 0, \dots, 0)$ in (iii) of the theorem, we get the counterpart of a result of Finsler [6] (for another proof of Finsler's result see Calabi [5]).

COROLLARY. If $A_1, A_2, A_3 \in \mathcal{H}_n$, where $n > 2$, are such that $\sum_{j=1}^3 (xA_jx^*)^2 \neq 0$ for all nonzero $x \in \mathbb{C}^n$, then there exist $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ such that $\sum_{j=1}^3 \alpha_j A_j$ is positive definite.

REMARK. Statements (i) and (iv) of the theorem are generalizations of results of Westwick [13] and Goldberg-Straus [8], respectively, for $n > 2$.

REFERENCES

1. Y. H. Au-Yeung and Y. T. Poon, *A remark on the convexity and positive definiteness concerning hermitian matrices*, Southeast Asian Bull. Math. **3** (1979), 85–92.
2. Y. H. Au-Yeung and N. K. Tsing, *Some theorems on the generalized numerical ranges*, Linear and Multilinear Algebra (to appear).
3. C. A. Berger, *Normal dilations*, Cornell Univ., Doctoral dissertation, 1963.
4. F. Bohnenblust, *Joint positiveness of matrices*, unpublished manuscript.
5. E. Calabi, *Linear systems of real quadratic forms*, Proc. Amer. Math. Soc. **15** (1964), 844–846.
6. P. Finsler, *Über das Vorkommen definiter und semidefiniter Formen in Scharen quadratischer Formen*, Comment. Math. Helv. **9** (1936/37), 188–192.
7. S. Friedland and R. Loewy, *Subspaces of symmetric matrices containing matrices with a multiple first eigenvalue*, Pacific J. Math. **62** (1976), 389–399.
8. M. Goldberg and E. G. Straus, *Elementary inclusion relations for generalized numerical ranges*, Linear Algebra Appl. **18** (1977), 1–24.
9. P. Halmos, *A Hilbert space problem book*, Von Nostrand, New York, 1967.
10. F. Hausdorff, *Der Wertvorrat einer Bilinearform*, Math. Z. **3** (1919), 314–316.
11. Y. T. Poon, *Another proof of a result of Westwick*, Linear and Multilinear Algebra **9** (1980), 35–37.
12. O. Toeplitz, *Das algebraische Analogon zu einem Satze von Fejér*, Math. Z. **2** (1918), 187–197.
13. R. Westwick, *A theorem on numerical ranges*, Linear and Multilinear Algebra **2** (1975), 311–315.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF HONG KONG, HONG KONG