The Convexity of a Generalized Matrix Range

Yiu Tung Poon
Department of Mathematics
Iowa State University
Ames, Iowa 50011

Submitted by F. Uhlig

ABSTRACT

The purpose of this paper is to generalize the Toeplitz-Hausdorff theorem on the convexity of the classical numerical range to the matrix range.

1. INTRODUCTION

Given positive integers m, n, let $M_{m,n}$ denote the $m \times n$ complex matrices. We write M_n for $M_{n,n}$, and identify $M_{1,n}$ with $\mathbb{C}^n$, the complex n-tuples. Given $A = (a_{ij})$ in $M_{m,n}$, the conjugate transpose of A is $A^* = (\bar{a}_{ji})$ in $M_{n,m}$. A matrix $A \in M_n$ is called hermitian if $A = A^*$. Let H_n denote the hermitian matrices in M_n. A matrix $A \in H_n$ is said to be positive semidefinite if all eigenvalues of A are nonnegative. Let H_n^+ denote the positive semidefinite matrices in H_n. Let $(H_m)^p = \{(B_1, \dots, B_p) : B_i \in H_m \text{ for } i = 1, \dots, p\}$. We identify H_1 with the real numbers $\mathbb{R}$. Then $V = (H_m)^p$ ($\cong \mathbb{R}^{pm^2}$) is a topological vector space. A subset $S \subset V$ is said to be convex if for any two points s_1, s_2 in S the line segment joining s_1 and s_2, $\{\alpha s_1 + (1-\alpha)s_2 : 0 \leqslant \alpha \leqslant 1\}$, is contained in S. For $S \subseteq V$, let conv(S) be the smallest convex set containing S. Let S be a closed subset of V. If the boundary of conv(S) is contained in S, then we say that S has a convex boundary.

Let $A \in M_n$. The (classical) numerical range of A is given by

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

Toeplitz [13] showed that $W(A)$ has a convex boundary, and Hausdorff [8] proved that $W(A)$ is convex. If we write $A = A_1 + iA_2$, $A_1, A_2 \in H_n$, the hermitian decomposition of A, then the above result can be restated as

THEOREM 1 (Toeplitz-Hausdorff). *For all $A_1, A_2 \in H_n$, the set*

$$W(A_1, A_2) = \left\{\left(xA_1x^*, xA_2x^*\right) : x \in \mathbf{C}^n,\ xx^* = 1\right\}.$$

is convex.

Given $A_1, A_2, \ldots, A_p \in H_n$, a natural generalization of $W(A_1, A_2)$ is

$$W\left(A_1, \ldots, A_p\right) = \left\{\left(xA_1x^*, \ldots, xA_px^*\right) : x \in \mathbf{C}^n,\ xx^* = 1\right\}.$$

Hausdorff [8] has pointed out that Toeplitz's method [13] can be used to show that $W(A_1, A_2, A_3)$ has a convex boundary. He also remarks that, in general $W(A_1, A_2, A_3)$ is not convex. However, it is shown by Au-Yeung and Poon [3] that if $n \geqslant 3$, then $W(A_1, A_2, A_3)$ is convex for every $A_1, A_2, A_3 \in H_n$. This result is a special case of the following

THEOREM 2 (Au-Yeung and Poon [3]). *If $1 \leqslant r \leqslant n-1$ and $p < (r+1)^2 - \delta_{n, r+1}$, then, for all $A_1, \ldots, A_p \in H_n$, the set*

$$W^r\left(A_1, \ldots, A_p\right) = \left\{\left(\sum_{i=1}^r x_i A_1 x_i^*, \ldots, \sum_{i=1}^r x_i A_p x_i^*\right) : x_i \in \mathbf{C}^n,\ \sum_{i=1}^r x_i x_i^* = 1\right\}$$

is convex. Here, $\delta_{i,j}$ is the Kronecker delta.

REMARK 3. It is easy to see that $W^r(A_1, \ldots, A_p)$ is convex iff for every $y_j \in \mathbf{C}^n$, $j = 1, \ldots, N$, such that $\sum_{j=1}^N y_j y_j^* = 1$, there exist $x_i \in \mathbf{C}^n$, $i = 1, \ldots, r$, such that $\sum_{i=1}^r x_i x_i^* = 1$ and $\sum_{i=1}^r x_i A_k x_i^* = \sum_{j=1}^N y_j A_k y_j^*$ for all $1 \leqslant k \leqslant p$. Theorem 2 is shown [3] to be equivalent to a result of Bohnenblust on joint positiveness of matrices [4]. By the latter result, the bound for p is best possible in the sense that if $p \geqslant (r+1)^2 - \delta_{n, r+1}$, then there exist $A_1, \ldots, A_p$ in H_n such that $W^r(A_1, \ldots, A_p)$ is not convex.

In the next section, we will give a generalization of Theorem 2 for the matrix range. For an explanation of this term, the reader should refer to Remark 19. This has a close connection with completely positive maps between matrix algebras, from which we get our motivation and techniques (see [6, 10]). In Section 3, we will discuss this connection and list some open questions.

2. CONVEXITY IN THE MATRIX RANGE

For each $\underline{B} = (B_1, \dots, B_p) \in (H_m)^p$ and $X \in M_m$, let $X\underline{B}X^* = (XB_1X^*, \dots, XB_pX^*)$. A subset S of $(H_m)^p$ is said to be *matricially convex* if for every $S_1, \dots, S_N$ in S, we have $\sum_{i=1}^N X_i S_i X_i^* \in S$ for every $X_1, \dots, X_N$ in M_m such that $\sum_{i=1}^N X_i X_i^* = I_m$, the $m \times m$ identity matrix. A matricial convex subset is convex. In fact, for subsets of $(H_1)^p$ $(\cong \mathbb{R}^p)$, matricial convexity is the same as the usual convexity. However, for $m > 1$, a convex subset of $(H_m)^p$ need not be matricially convex. For example, for $m > 1$ and $p = 1$ take $S = \{I_{m-1} \oplus 0\}$.

The main result in this paper is the following generalization of Theorem 2.

THEOREM 4. *If* $1 \leqslant r \leqslant mn - 1$ *and* $m^2(p+1) - 1 < (r+1)^2 - \delta_{mn, r+1}$, *then for all* $A_1, \dots, A_p \in H_n$, *the set*

$$W_m^r(A_1, \dots, A_p)$$

$$= \left\{ \left(\sum_{i=1}^r X_i A_1 X_i^*, \dots, \sum_{i=1}^r X_i A_p X_i^* \right) : X_i \in M_{m,n} \text{ and } \sum_{i=1}^r X_i X_i^* = I_m \right\}$$

is matricially convex.

The proof of Theorem 4 is obtained by reducing to the case when $m = 1$ and applying Theorem 2. To simplify notation in subsequent arguments, we need some definitions.

DEFINITION 5. For each $m \geqslant 1$ and $1 \leqslant j, k \leqslant m$ let F_{jk}^m be the matrix in M_m with 1 as the (j, k)th entry and 0 elsewhere. Define for $1 \leqslant j, k \leqslant m$

$$E_{jk}^m = \begin{cases} \left(F_{jk}^m + F_{kj}^m\right) & \text{if } 1 \leqslant j < k \leqslant m, \\ F_{jj}^m & \text{if } 1 \leqslant j = k \leqslant m, \\ \sqrt{-1}\left(F_{jk}^m - F_{kj}^m\right) & \text{if } 1 \leqslant k < j \leqslant m. \end{cases}$$

Then each E_{jk}^m is in H_m, and $\{E_{jk}^m : 1 \leqslant j, k \leqslant m\}$ is a basis of M_m over $\mathbb{C}$.

DEFINITION 6. Suppose $X = (x_{ij}) \in M_{m,n}$. Define $v(X) \in M_{1,mn}$ by

$$v(X) = \frac{1}{\sqrt{n}}(x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{m1}, \dots, x_{mn}).$$

Conversely, if $x = [\underline{x}^1, \dots, \underline{x}^m] \in M_{1,mn}$ with $\underline{x}^i \in M_{1,n}$ for $1 \leqslant i \leqslant m$, define $V(x) \in M_{m,n}$ by

$$V(x) = \sqrt{n}\begin{bmatrix} \underline{x}^1 \\ \vdots \\ \underline{x}^m \end{bmatrix}.$$

We note that for $x \in M_{1,mn}$ and $X \in M_{m,n}$, $v(V(x)) = x$ and $V(v(X)) = X$.

Let $B = (b_{kl}) \in M_n$, $A = (a_{ij}) \in M_m$; then $B \otimes A$ will denote the Kronecker product of B and A.

LEMMA 7. *Suppose $X_i, Y_j \in M_{m,n}$ for $1 \leqslant i \leqslant N_1$, $1 \leqslant j \leqslant N_2$, and $A \in H_n$. Let $x_i = v(X_i)$, $y_j = v(Y_j)$ for $1 \leqslant i \leqslant N_1$, $1 \leqslant j \leqslant N_2$. Then the following two conditions are equivalent*:

$$(1) \qquad \sum_{i=1}^{N_1} X_i A X_i^* = \sum_{j=1}^{N_2} Y_j A Y_j^*,$$

$$(2) \qquad \sum_{i=1}^{N_1} x_i\left(E_{kl}^m \otimes A\right) x_i^* = \sum_{j=1}^{N_2} y_j\left(E_{kl}^m \otimes A\right) y_j \qquad \text{for} \quad 1 \leqslant k, l \leqslant m.$$

Proof. For $1 \leqslant k \leqslant m$, let e_k be the kth unit vector. Then the (k, l)th entry of $\sum_{i=1}^{N_1} X_i A X_i^*$ is given by

$$e_k\left(\sum_{i=1}^{N_1} X_i A X_i^*\right) e_l^* = n \sum_{i=1}^{N_1} x_i\left(F_{kl}^m \otimes A\right) x_i^*.$$

thus we have

$$\sum_{i=1}^{N_1} X_i A X_i^* = \sum_{j=1}^{N_2} Y_j A Y_j^*$$

$$\Leftrightarrow \quad \sum_{i=1}^{N_1} x_i\left(F_{kl}^m \otimes A\right) x_i^* = \sum_{j=1}^{N_2} y_j\left(F_{kl}^m \otimes A\right) y_j^* \quad \text{for } 1 \leqslant k, l \leqslant m$$

$$\Leftrightarrow \quad \sum_{i=1}^{N_1} x_i\left(E_{kl}^m \otimes A\right) x_i^* = \sum_{j=1}^{N_2} y_j\left(E_{kl}^m \otimes A\right) y_j^* \quad \text{for } 1 \leqslant k, l \leqslant m. \quad \blacksquare$$

Proof of Theorem 4. Suppose $1 \leqslant r \leqslant mn - 1$, $m^2(p+1) - 1 < (r+1)^2 - \delta_{nm, r+1}$, and $A_1, \dots, A_p \in H_n$. It suffices to prove that for every $Y_1, \dots, Y_N \in M_{m,n}$ with $\sum_{j=1}^{N} Y_j Y_j^* = I_m$, there exist $X_1, \dots, X_r \in M_{m,n}$ such that

$$\sum_{i=1}^{r} X_i X_i^* = I_m$$

and

$$\sum_{i=1}^{r} X_i A_k X_i^* = \sum_{j=1}^{N} Y_j A_k Y_j^* \qquad \text{for} \quad 1 \leqslant k \leqslant p.$$

Let $y_j = v(Y_j)$ for $1 \leqslant j \leqslant N$. Then we have $\sum_{j=1}^{N} y_j y_j^* = 1$. Consider the $m^2(p+1) - 1$ hermitian matrices

$$E_{jk}^m \otimes I_n, \qquad 1 \leqslant j, k \leqslant m, \quad (j, k) \neq (m, m),$$

and

$$E_{jk}^m \otimes A_l, \qquad 1 \leqslant j, k \leqslant m, \quad 1 \leqslant l \leqslant p.$$

Since $m^2(p+1) - 1 < (r+1)^2 - \delta_{nm, r+1}$, we can apply Theorem 2 to the above $m^2(p+1) - 1$ matrices and get $x_i \in M_{1, mn}$ for $1 \leqslant i \leqslant r$ such that

$$(3) \quad \sum_{i=1}^{r} x_i x_i^* = 1,$$

$$(4) \quad \sum_{i=1}^{r} x_i\left(E_{jk}^m \otimes I_n\right) x_i^* = \sum_{j=1}^{N} y_j\left(E_{jk}^m \otimes I_n\right) y_j^*$$

$$\text{for all} \quad 1 \leqslant j, k \leqslant m, \quad (j, k) \neq (m, m),$$

$$(5) \quad \sum_{i=1}^{r} x_i\left(E_{jk}^m \otimes A_l\right) x_i^* = \sum_{j=1}^{N} y_j\left(E_{jk}^m \otimes A_l\right) y_j^*$$

$$\text{for all} \quad 1 \leqslant j, k \leqslant m, \quad 1 \leqslant l \leqslant p.$$

Since

$$1 = \sum_{i=1}^{r} x_i x_i^* = \sum_{k=1}^{m} \sum_{i=1}^{r} x_i \left(E_{kk}^m \otimes I_n \right) x_i^*,$$

we have

$$\begin{aligned}\sum_{i=1}^{r} x_i \left(E_{mm}^m \otimes I_n \right) x_i^* &= 1 - \sum_{k=1}^{m-1} \sum_{i=1}^{r} x_i \left(E_{kk}^n \otimes I_n \right) x_i^* \\ &= \sum_{j=1}^{N} y_j y_j^* - \sum_{k=1}^{m-1} \sum_{i=1}^{N} y_j \left(E_{kk}^m \otimes I_n \right) y_j^* \\ &= \sum_{j=1}^{N} y_j \left(E_{mm}^m \otimes I_n \right) y_j^*.\end{aligned}$$

Thus, condition (4) also holds for $(j, k) = (m, m)$. Let $X_i = V(x_i)$ for $1 \leqslant i \leqslant r$. By Lemma 7, we have

$$\sum_{i=1}^{r} X_i X_i^* = I_m$$

and

$$\sum_{i=1}^{r} X_i A_l X_i^* = \sum_{j=1}^{N} Y_j A_l Y_j^* \qquad \text{for} \quad 1 \leqslant l \leqslant p. \qquad \blacksquare$$

In [3], Au-Yeung and Poon proved the following result which is closely related to Theorem 2.

THEOREM 8. *Let* $1 \leqslant r \leqslant n-1$ *and* $p < (r+1)^2$. *Then for all* $A_1, \dots, A_p \in H_n$, *the set*

$$\hat{W}^r\left(A_1, \dots, A_p\right) = \left\{ \left(\sum_{i=1}^{r} x_i A_1 x_i^*, \dots, \sum_{i=1}^{r} x_i A_p x_i^* \right) : x_i \in \mathbf{C}^n \right\}$$

is convex.

REMARK 9. The bound $(r+1)^2$ in Theorem 8 is also best possible. (See Remark 3.)

Using arguments similar to the proof of Theorem 4, we have

THEOREM 10. *If* $1 \leqslant r \leqslant nm-1$ *and* $m^2p < (r+1)^2$, *then for every* $A_1, \ldots, A_p \in H_n$, *the set*

$$\hat{W}_m^r(A_1, \ldots, A_p) = \left\{ \left(\sum_{i=1}^r X_i A_1 X_i^*, \ldots, \sum_{i=1}^r X_i A_p X_i^* \right) : X_i \in M_{m,n} \right\}$$

is matricially convex.

REMARK 11. *Let* $A_1, \ldots, A_p \in H_n$. *Since the convexity of* $W^r(A_1, \ldots, A_p)$ *implies the convexity of* $\hat{W}^r(A_1, \ldots, A_p)$, *Theorem 8 follows immediately from Theorem 4 except when* $r = n-1$. *However, for* $m > 1$, *the matricial convexity of* $\hat{W}_m^r(A_1, \ldots, A_p)$ *does not follow from that of* $W_m^r(A_1, \ldots, A_p)$.

Let S_n be the real $n \times n$ matrices. Theorem 4 and 10 also hold (see [3]) for $A_1, \ldots, A_p \in S_n$ [with $x_i \in \mathbb{R}^n$ and $(r+1)^2$ replaced by $r(r+1)/2$]. Let $M_{n,m}(\mathbb{R})$ be the real $n \times m$ matrices. In the following two theorems, we use real matrices for the definition of matricial convexity.

THEOREM 12. *If*

$$1 \leqslant r \leqslant nm-1 \quad \textit{and} \quad \frac{m(m+1)}{2}(p+1) - 1 < \frac{r(r+1)}{2} - \delta_{nm,r+1},$$

then for all $A_1, \ldots, A_p \in S_n$, *the set*

$$\left\{ \left(\sum_{i=1}^r X_i A_1 X_i^t, \ldots, \sum_{i=1}^r X_i A_p X_i^t \right) : X_i \in M_{m,n}(\mathbb{R}),\ \sum_{i=1}^r X_i X_i^t = I_m \right\}$$

is matricially convex.

THEOREM 13. *If* $1 \leqslant r \leqslant nm-1$ *and* $m(m+1)p < r(r+1)$, *then for all* $A_1, \ldots, A_p \in S_n$, *the set*

$$\left\{ \left(\sum_{i=1}^r X_i A_1 X_i^t, \ldots, \sum_{i=1}^r X_i A_p X_i^t \right) : X_i \in M_{n,m}(\mathbb{R}) \right\}$$

is matricially convex.

3. CONNECTIONS WITH COMPLETELY POSITIVE MAPS

Given a complex linear map $\Phi : M_n \to M_m$, we define, for each $N \geqslant 1$, $\Phi_N : M_{nN} \to M_{mN}$ by

$$\Phi_N\big(A_{ij}\big) = \big(\Phi\big(A_{ij}\big)\big),$$

where the matrix A in M_{nN} is partitioned into $n \times n$ blocks A_{ij}, $1 \leqslant i, j \leqslant N$. The map Φ is said to be N-positive if $\Phi_N(H_{nN}^+) \subseteq H_{mN}^+$, and completely positive if Φ is N-positive for every $N \geqslant 1$. Let $\mathrm{CP}(n, m)$ denote the set of all completely positive maps from M_n to M_m. For $m = 1$, every 1-positive map is completely positive. For $m > 1$, there exist maps that are $N - 1$-positive but not N-positive (see Choi [5]), and we have

PROPOSITION 14 (Choi [6]). *Let Φ be a linear map from M_n to M_m. Then Φ is completely positive if and only if there exist $X_1, \dots, X_r \in M_{m,n}$ such that*

$$\Phi(A) = \sum_{i=1}^{r} X_i A X_i^* \qquad \textit{for all} \quad A \in M_n.$$

For each $r \geqslant 1$, let $\mathrm{CP}^r(n, m)$ be the set of all Φ in $\mathrm{CP}(m, n)$ such that there exist $X_1, \dots, X_r \in M_{m,n}$ satisfying $\Phi(A) = \sum_{i=1}^{r} X_i A X_i^*$ for all $A \in M_n$. Given $A_1, \dots, A_p$ in H_n, it is easy to see that $\hat{W}_m^r(A_1, \dots, A_p)$ is matricially convex if and only if for every $\Phi \in \mathrm{CP}(n, m)$ there exists $\Psi \in \mathrm{CP}^r(n, m)$ such that $\Psi(A_i) = \Phi(A_i)$ for all $1 \leqslant i \leqslant p$. Following Remark 9, for fixed n, m, p, we are interested in finding the smallest possible $r = r(n, m, p)$ satisfying

(6) For every subspace $\mathscr{A}$ of H_n with $\dim \mathscr{A} = p$ and $\Phi \in \mathrm{CP}(n, m)$, there exists $\Psi \in \mathrm{CP}^r(n, m)$ such that $\Psi(A) = \Phi(A)$ for all $A \in \mathscr{A}$.

In [9], Narcowich and Ward showed that if $r = \left[m\sqrt{p}\right]$, then (6) is satisfied. Here, $[x]$ denotes the smallest integer less than or equal to x. The bound $\left[m\sqrt{p}\right]$ can also be obtained from Theorem 10. When $I_n \in \mathscr{A}$, Theorem 4 gives a slightly lower bound:

PROPOSITION 15. *Suppose $p < n^2$ and $r = \left[\sqrt{m^2 p - 1}\right]$. Let $\mathscr{A}$ be a p-dimensional subspace of H_n containing I_n. Then for every $\Phi \in \mathrm{CP}(n, m)$ there exists $\Psi \in \mathrm{CP}^r(n, m)$ such that $\Psi(A) = \Phi(A)$ for all $A \in \mathscr{A}$.*

Proof. Let $\{A_1, \dots, A_{p-1}, I_n\}$ be a basis of $\mathscr{A}$. When $p < n^2$ and $r = \left[\sqrt{m^2 p - 1}\right]$, we have $r < nm - 1$ and $m^2(p - 1) < (r + 1)^2$. So we can apply Theorem 4 to $A_1, \dots, A_{p-1}$. ∎

REMARK 16. If $p = n^2$, it follows from a result of Choi [6, Remark 4] that $r(n, m, p) = nm$.

REMARK 17. From Remark 2.2 in [9], we have

$$m\left[\sqrt{p}\right] \leqslant r(n, m, p) \leqslant \left[m\sqrt{p}\right].$$

Thus, when p is a perfect square, $r(n, m, p) = m\sqrt{p}$.

REMARK 18. Except for the above results and some special cases, the best bounds for r in Theorems 4 and 10 remain unknown.

REMARK 19. The notion of completely positive maps on operator space is due to Stinespring [11]. Since then, it has been recognized that completely positive maps are the natural generalization of positive linear functionals. (See Stinespring [11], Størmer [12], and Arveson [1, 2]). Let $\mathscr{H}$ be a (possibly infinite dimensional) Hilbert space, and $A \in \mathscr{B}(\mathscr{H})$ the set of all bounded linear operators on $\mathscr{H}$. For each m, let

$$W_m(A) = \{\Phi(A) : \Phi \text{ is a completely positive map from } \mathscr{B}(\mathscr{H}) \text{ to } M_m, \\ \Phi(I) = I_m\}.$$

The sequence $\{W_m(A) : m = 1, 2, \ldots\}$ is called the matrix range of A. This definition is due to Arveson in [2], where he proves that for irreducible compact A, the matrix range is a completely invariant for unitary equivalence. This is part of the motivation for our study of $W_m^r(A_1, \ldots, A_p)$. A very detailed list of references for completely positive maps and the matrix range can be found in Paulsen [10] and Farenick [7].

REFERENCES

1 W. B. Arveson, Subalgebras of C^*-algebra, *Acta Math.* 123:141–224 (1969).
2 ——, Subalgebras of C^*-algebra II, *Acta Math.* 128:721–308 (1972).
3 Y. H. Au-Yeung and Y. T. Poon, A remark on the convexity and positive definiteness concerning Hermitian matrices, *Southeast Asian Bull. Math.* 3:85–92 (1979).
4 F. Bohnenblust, Joint positiveness of matrices, unpublished manuscript.
5 M. D. Choi, Positive linear maps on C^*-algebras, *Canad. J. Math.* 24:520–529 (1972).

6 Completely positive linear maps on complex matrices, *Linear Algebra Appl.* 10:285–290 (1975).

7 D. R. Farenick, The Matricial Spectrum and Range and C^*-Convex Sets, Ph.D. Thesis, Univ. of Toronto, 1990.

8 F. Hausdorff, Der Wertvorrat einer Bilinearform, *Math* Z. 3:314–316 (1919).

9 F. J. Narcowich and J. D. Ward, A Toeplitz-Hausdorff theorem for matrix ranges, *J. Operator Theory* 6:87–101 (1981).

10 V. I. Paulsen, *Completely Bounded Maps and Dilations*, Pitman Res. Notes Math. Ser. 146, Longman Scientific and Technical, 1986.

11 W. F. Stinespring, Positive functions on C^*-algebras, *Proc. Amer. Math. Soc.* 6:211–216 (1955).

12 E. Størmer, Positive linear maps of operator algebras, *Acta Math.* 110:233–278 (1963).

13 O. Toeplitz, Das algebraische Analogon zu einem Satze von Fejér, *Math. Z.* 2:187–197 (1918).

Received 29 June 1990; final manuscript accepted 12 December 1990