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 $(H_m)^p = \{(B_1, \ldots, B_p) : \exists i \in H_m \text{ for } i = 1, \ldots, p\}$. We identify $H_1$ with the real numbers $\mathbb{R}$. Then $V = (H_m)^p (\equiv \mathbb{R}^{pm^2})$ 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 $s_2$, $[\alpha s_1 + (1 - \alpha)s_2 : 0 \leq \alpha \leq 1]$, is contained in $S$. For $S \subseteq V$, let $\text{conv}(S)$ be the smallest convex set containing $S$. Let $S$ be a closed subset of $V$. If the boundary of $\text{conv}(S)$ 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

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

is convex.

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

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

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 $n \geq 3$, 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 $1 \leq r \leq n - 1$ and $p < (r + 1)^2 - \delta_{n,r+1}$, then, for all $A_1, \ldots, A_p \in H_n$, the set

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

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 $y_j \in \mathbb{C}^n$, $j = 1, \ldots, N$, such that $\sum_{j=1}^N y_j y_j^* = 1$, there exist $x_i \in \mathbb{C}^n$, $i = 1, \ldots, r$, 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 $1 \leq k \leq p$. 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 $p \geq (r + 1)^2 - \delta_{n,r+1}$, 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.
GENERALIZED MATRIX RANGE

(see [6, 10]). In Section 3, we will discuss this connection and list some open questions.

2. CONVEXITY IN THE MATRIX RANGE

For each $B = (B_1, \ldots, B_p) \in (H_m)^p$ and $X \in M_m$, let $XBX^* = (XB_1X^*, \ldots, XB_pX^*)$. A subset $S$ of $(H_m)^p$ is said to be matricially convex if for every $S_1, \ldots, S_N$ in $S$, we have $\Sigma_{i=1}^N X_i S_i X_i^* \in S$ for every $X_1, \ldots, X_N$ in $M_m$ such that $\Sigma_{i=1}^N X_i X_i^* = I_m$, the $m \times m$ identity matrix. A matricial convex subset is convex. In fact, for subsets of $(H_1)^p (\equiv \mathbb{R}^p)$, 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 $1 \leq r \leq mn - 1$ and $m^2(p + 1) - 1 < (r + 1)^2 - \delta_{mn, r+1}$, then for all $A_1, \ldots, A_p \in H_n$, the set

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

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

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 $m \geq 1$ and $1 \leq j, k \leq m$ let $F_{jk}^m$ be the matrix in $M_m$ with 1 as the $(j, k)$th entry and 0 elsewhere. Define for $1 \leq j, k \leq m$

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

Then each $E_{jk}^m$ is in $H_m$, and $\{ E_{jk}^m : 1 \leq j, k \leq m \}$ 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

\[
v(X) = \frac{1}{\sqrt{n}} \left( x_{11}, \ldots, x_{1n}, x_{21}, \ldots, x_{2n}, \ldots, x_{m1}, \ldots, x_{mn} \right).
\]

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

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

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 \( 1 \leq i \leq N_1 \), \( 1 \leq j \leq N_2 \), and \( A \in H_n \). Let \( x_i = v(X_i) \), \( y_j = v(Y_j) \) for \( 1 \leq i \leq N_1 \), \( 1 \leq j \leq N_2 \). Then the following two conditions are equivalent:

1. \[
\sum_{i=1}^{N_1} X_i A X_i^* = \sum_{j=1}^{N_2} Y_j A Y_j^*,
\]

2. \[
\sum_{i=1}^{N_1} x_i (E_{kl} \otimes A) x_i^* = \sum_{j=1}^{N_2} y_j (E_{kl} \otimes A) y_j \quad \text{for} \quad 1 \leq k, l \leq m.
\]

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

\[
e_k \left( \sum_{i=1}^{N_1} X_i A X_i^* \right) e_l^* = n \sum_{i=1}^{N_1} x_i (F_{kl} \otimes A) x_i^*.
\]
thus we have

\[
\sum_{i=1}^{N_1} X_i AX_i^* = \sum_{j=1}^{N_2} Y_j AY_j^* \\
\Leftrightarrow \sum_{i=1}^{N_1} x_i (F_{kl}^m \otimes A) x_i^* = \sum_{j=1}^{N_2} y_j (F_{kl}^m \otimes A) y_j^* \quad \text{for} \quad 1 \leq k, l \leq m
\]

\[
\Leftrightarrow \sum_{i=1}^{N_1} x_i (E_{kl}^m \otimes A) x_i^* = \sum_{j=1}^{N_2} y_j (E_{kl}^m \otimes A) y_j^* \quad \text{for} \quad 1 \leq k, l \leq m. \quad \square
\]

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

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

and

\[
\sum_{i=1}^r X_i A_k X_i^* = \sum_{j=1}^N Y_j A_k Y_j^* \quad \text{for} \quad 1 \leq k \leq p.
\]

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

\[
E_{jk}^m \otimes I_n, \quad 1 \leq j, k \leq m, \quad (j, k) \neq (m, m),
\]

and

\[
E_{jk}^m \otimes A_l, \quad 1 \leq j, k \leq m, \quad 1 \leq l \leq p.
\]

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 \( 1 \leq i \leq r \) such that

\[
\sum_{i=1}^r X_i x_i^* = 1, \tag{3}
\]

\[
\sum_{i=1}^r x_i (E_{jk}^m \otimes I_n) x_i^* = \sum_{j=1}^N y_j (E_{jk}^m \otimes I_n) y_j^* \quad \text{for all} \quad 1 \leq j, k \leq m, \quad (j, k) \neq (m, m), \tag{4}
\]

\[
\sum_{i=1}^r x_i (E_{jk}^m \otimes A_l) x_i^* = \sum_{j=1}^N y_j (E_{jk}^m \otimes A_l) y_j^* \quad \text{for all} \quad 1 \leq j, k \leq m, \quad 1 \leq l \leq p. \tag{5}
\]
Since
\[ 1 = \sum_{i=1}^{r} x_i x_i^* = \sum_{k=1}^{m} \sum_{i=1}^{r} x_i (E_k^{m} \otimes I_n) x_i^* , \]
we have
\[ \sum_{i=1}^{r} x_i (E_{mm}^{m} \otimes I_n) x_i^* = 1 - \sum_{k=1}^{m-1} \sum_{i=1}^{r} x_i (E_k^{m} \otimes I_n) x_i^* \]
\[ = \sum_{j=1}^{N} y_j y_j^* - \sum_{k=1}^{m-1} \sum_{i=1}^{N} y_j (E_k^{m} \otimes I_n) y_j^* \]
\[ = \sum_{j=1}^{N} y_j (E_{mm}^{m} \otimes I_n) y_j^* . \]

Thus, condition (4) also holds for \((j, k) = (m, m)\). Let \(X_i = V(x_i)\) for \(1 \leq i \leq r\). By Lemma 7, we have
\[ \sum_{i=1}^{r} X_i X_i^* = I_m \]
and
\[ \sum_{i=1}^{r} X_i A_i X_i^* = \sum_{j=1}^{N} Y_j A_j Y_j^* \quad \text{for} \quad 1 \leq l \leq p. \quad \square \]

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

**Theorem 8.** Let \(1 \leq r \leq n - 1\) and \(p < (r + 1)^2\). Then for all \(A_1, \ldots, A_p \in H_n\), the set
\[ \hat{W}^r(A_1, \ldots, A_p) = \left\{ \left( \sum_{i=1}^{r} x_i A_i x_i^*, \ldots, \sum_{i=1}^{r} x_i A_p x_i^* \right) : x_i \in \mathbb{C}^n \right\} \]
is convex.

**Remark 9.** The bound \((r + 1)^2\) in Theorem 8 is also best possible. (See Remark 3.)
GENERALIZED MATRIX RANGE

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

**THEOREM 10.** If $1 \leq r \leq nm - 1$ and $m^2 p < (r + 1)^2$, 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 $\hat{W}_m^r(A_1, \ldots, A_p)$, 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

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

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 $1 \leq r \leq nm - 1$ 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 \( N \geq 1 \), \( \Phi_N : M_{nN} \to M_{mN} \) by

\[
\Phi_N(A_{ij}) = \begin{pmatrix} \Phi(A_{ij}) \end{pmatrix},
\]

where the matrix \( A \) in \( M_{nN} \) is partitioned into \( n \times n \) blocks \( A_{ij}, 1 \leq i, j \leq N \). The map \( \Phi \) is said to be \( N \)-positive if \( \Phi_N(H^+_{nN}) \subseteq H^+_{mN} \), and completely positive if \( \Phi \) is \( N \)-positive for every \( N \geq 1 \). Let \( \text{CP}(n, m) \) 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 \( \Phi \) be a linear map from \( M_n \) to \( M_m \). Then \( \Phi \) is completely positive if and only if there exist \( X_1, \ldots, X_r \in M_{m,n} \) such that

\[
\Phi(A) = \sum_{i=1}^{r} X_i AX_i^* \quad \text{for all} \quad A \in M_n.
\]

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

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

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

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

**Proof.** Let \( \{A_1, \ldots, A_{p-1}, I_n\} \) be a basis of \( \mathcal{A} \). When \( p < n^2 \) and \( r = \left\lceil \sqrt{m^2 p - 1} \right\rceil \), we have \( r < nm - 1 \) and \( m^2(p - 1) < (r + 1)^2 \). So we can apply Theorem 4 to \( A_1, \ldots, A_{p-1} \).
**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 
\[ m \left\lfloor \sqrt{p} \right\rfloor \leq r(n, m, p) \leq \left[ m \sqrt{p} \right]. \]
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 \( \mathcal{H} \) be a (possibly infinite dimensional) Hilbert space, and \( A \in \mathcal{B}(\mathcal{H}) \) the set of all bounded linear operators on \( \mathcal{H} \). For each \( m \), let 
\[ W_m(A) = \{ \Phi(A) : \Phi \text{ is a completely positive map from } \mathcal{B}(\mathcal{H}) \text{ to } M_m, \Phi(I) = I_m \}. \]

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 \( W_m(A_1, \ldots, A_p) \). A very detailed list of references for completely positive maps and the matrix range can be found in Paulsen [10] and Farenick [7].

**REFERENCES**

4. F. Bohnenblust, Joint positiveness of matrices, unpublished manuscript.


*Received 29 June 1990; final manuscript accepted 12 December 1990*