

Convexity of the Joint Numerical Range

Chi-Kwong Li ^{*} and Yiu-Tung Poon [†]

October 26, 2004

Dedicated to Professor Yik-Hoi Au-Yeung on the occasion of his retirement.

Abstract

Let $A = (A_1, \dots, A_m)$ be an m -tuple of $n \times n$ Hermitian matrices. For $1 \leq k \leq n$, the k th joint numerical range of A is defined by

$$W_k(A) = \{(\operatorname{tr}(X^* A_1 X), \dots, \operatorname{tr}(X^* A_m X)) : X \in \mathbf{C}^{n \times k}, X^* X = I_k\}.$$

We consider linearly independent families of Hermitian matrices $\{A_1, \dots, A_m\}$ so that $W_k(A)$ is convex. It is shown that m can reach the upper bound $2k(n - k) + 1$. A key idea in our study is relating the convexity of $W_k(A)$ to the problem of constructing rank k orthogonal projections under linear constraints determined by A . The techniques are extended to study the convexity of other generalized numerical ranges and the corresponding matrix construction problems.

1 Introduction

Let $A = (A_1, \dots, A_m)$ be an m -tuple of $n \times n$ Hermitian matrices. For $1 \leq k \leq n$, the k th joint numerical range of A is defined as

$$W_k(A) = \{(\operatorname{tr}(X^* A_1 X), \dots, \operatorname{tr}(X^* A_m X)) : X \in \mathbf{C}^{n \times k}, X^* X = I_k\}.$$

When $k = 1$, it reduces to the usual joint numerical range of A that are useful in the study of various pure and applied subjects (see [2, 3, 4, 5, 7]). In particular, in the study of structured singular values arising in robust stability (see [6, 7, 15]), it is important that $W_1(A)$ is convex. Unfortunately, $W_1(A)$ is not always convex if $m > 3$ (e.g., see [1]). We modify the example in [1] to show that the same conclusion holds for $W_k(A)$ in the following.

Example 1.1 Let $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \oplus 0_{n-2}$, $A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \oplus 0_{n-2}$, $A_3 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \oplus 0_{n-2}$, $A_4 = I_2 \oplus 2I_{k-1} \oplus 0_{n-k-1}$, $A_j = I_n$ for all other $j \leq m$. Then $W_k(A)$ is not convex.

^{*}Department of Mathematics, College of William & Mary, P.O. Box 8795, Williamsburg, Virginia 23187-8795 (ckli@math.wm.edu).

[†]Department of Mathematics, Iowa State University, Ames, Iowa 50011 (poon@math.iastate.edu).

Proof. Suppose $W_k(A)$ is convex. Note that $2k - 1$ is the sum of the k largest eigenvalues of A_4 . If $X \in \mathbf{C}^{n \times k}$ with $X^*X = I_k$ such that $\text{tr}(X^*A_4X) = 2k - 1$, then (see e.g. [10]) the column space of X must be spanned by k eigenvectors of A_4 corresponding to the k largest eigenvalues. Thus, there exists an $k \times k$ unitary matrix V such that XV has columns $\alpha e_1 + \beta e_2, e_3, \dots, e_{k+1}$, where $|\alpha|^2 + |\beta|^2 = 1$. Then one can check that the subset

$$S = \{(a_1, \dots, a_m) \in W_k(A) : a_4 = 2k - 1\} = \{(a, b, c, 2k - 1, k, \dots, k) : a^2 + b^2 + c^2 = 1\}$$

of $W_k(A)$ is not convex, which is a contradiction. \square

Apart from these negative results, one would still hope to get the convexity conclusion if A has some special structure. In this paper, we consider linearly independent families $\{A_1, \dots, A_m\}$ so that $W_k(A)$ is convex. In particular, we show that the maximum value of m is $2k(n - k) + 1$, which is much larger than 3 ensured by the general theorem [2].

A key idea in our study is to view $W_k(A)$ as the image of the set of all rank k projections under the linear map

$$\phi(P) = (\text{tr } A_1P, \dots, \text{tr } A_mP).$$

To make this claim precise, denote by

$$\mathcal{U}(C) = \{U^*CU : U^*U = I_n\}$$

the *unitary similarity orbit* of a given Hermitian matrix C . If

$$J_k = I_k \oplus 0_{n-k},$$

then $\mathcal{U}(J_k)$ is the set of rank k orthogonal projections. Since a matrix P belongs to $\mathcal{U}(J_k)$ if and only if $P = XX^*$ for some $X \in \mathbf{C}^{n \times k}$ satisfying $X^*X = I_k$, we have

$$W_k(A) = \{(\text{tr } A_1P, \dots, \text{tr } A_mP) : P \in \mathcal{U}(J_k)\}.$$

So, $W_k(A)$ is just the image of $\mathcal{U}(J_k)$ under the linear projection onto the linear space spanned by $\{A_1, \dots, A_m\}$. With this new formulation of $W_k(A)$, its convex hull can be written as:

$$\text{conv } W_k(A) = \{(\text{tr } A_1X, \dots, \text{tr } A_mX) : X \in \text{conv } \mathcal{U}(J_k)\}.$$

Consequently, we have the following.

Theorem 1.2 *Let $A = (A_1, \dots, A_m)$ be an m -tuple of $n \times n$ Hermitian matrices. Then $W_k(A)$ is convex if and only if for any $X \in \text{conv } \mathcal{U}(J_k)$ there exists $P \in \mathcal{U}(J_k)$ so that $\text{tr } A_jP = \text{tr } A_jX$ for all $j = 1, \dots, m$.*

This rather simple observation turns out to be very useful in our study. Moreover, if A_1, \dots, A_m are chosen from the standard basis for $n \times n$ Hermitian matrices, then the constraints $\text{tr } A_jP = \text{tr } A_jX$ for all $j = 1, \dots, m$, are just specification of certain entries of P . Thus the problem reduces to construction of a rank k orthogonal projection P with some specified entries.

We shall use $\{E_{11}, E_{12}, \dots, E_{nn}\}$ to denote the standard basis for $\mathbf{C}^{n \times n}$ in our discussion. Moreover, the following observations will be used frequently:

- (1) $W_k(A) = W_k(V^*A_1V, \dots, V^*A_mV)$ for any unitary V .
- (2) $W_k(A)$ is convex if and only if $W_k(I, A_1, \dots, A_m)$ is convex.
- (3) Let $\{B_1, \dots, B_s\}$ be a basis for $\text{span}\{A_1, \dots, A_m\}$. Then $W_k(A)$ is convex if and only if $W_k(B_1, \dots, B_s)$ is convex.
- (4) Suppose $W_k(A)$ is convex. Then $W_k(B_1, \dots, B_s)$ is convex if $B_j \in \text{span}\{I, A_1, \dots, A_m\}$ for $1 \leq j \leq s$.

2 The k th Numerical Range and Orthogonal Projections

The focus of this section is to study linearly independent families $\{A_1, \dots, A_m\}$ for which $W_k(A_1, \dots, A_m)$ is convex. We call such a family a *linearly independent convex family* for the k th numerical range. In particular, we would like to study maximal (in the set inclusion sense) linearly independent convex families.

We begin with the following result on the completion of a certain partial Hermitian matrix to a rank k orthogonal projection.

Theorem 2.1 *Let $1 \leq k < n$ and $X \in \mathbf{C}^{k \times (n-k)}$. There exists a rank k orthogonal projection of the form $\begin{pmatrix} * & X \\ X^* & * \end{pmatrix}$ if and only if $\|X\| \leq 1/2$.*

Proof. Suppose X is an off-diagonal submatrix of a rank k orthogonal projection. Then $\|X\| \leq 1/2$ by the result in [12]. For the converse, let $m = \min\{k, n - k\}$. Suppose $2X$ has singular value decomposition UDV , where U and V are unitary, and

$$D_{rs} = \begin{cases} \sin t_r & \text{if } r = s \leq m, \\ 0 & \text{otherwise,} \end{cases}$$

with $\pi/2 \geq t_1 \geq \dots \geq t_m \geq 0$. Set

$$P = \begin{cases} U \text{diag}(1 + \cos t_1, \dots, 1 + \cos t_k) U^* / 2 & \text{if } k \leq n/2, \\ U(\text{diag}(1 + \cos t_1, \dots, 1 + \cos t_m) \oplus I_{2k-n}) U^* / 2 & \text{otherwise,} \end{cases}$$

and

$$Q = \begin{cases} V^*(\text{diag}(1 - \cos t_1, \dots, 1 - \cos t_k) \oplus 0_{n-2k}) V / 2 & \text{if } k \leq n/2, \\ V^* \text{diag}(1 - \cos t_1, \dots, 1 - \cos t_m) V / 2 & \text{otherwise.} \end{cases}$$

One easily checks that $\begin{pmatrix} P & X \\ X^* & Q \end{pmatrix}$ is a rank k orthogonal projection. \square

Clearly, the set of $k \times (n - k)$ matrices X satisfying $\|X\| \leq 1/2$ is convex. By observations (1) and (2), we have the following result.

Theorem 2.2 *Let $A = (A_1, \dots, A_m)$ be an m -tuple of $n \times n$ Hermitian matrices. Suppose $1 \leq k < n$ and*

$$\mathcal{S} = \left\{ \alpha I + \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix} : \alpha \in \mathbf{R}, X \in \mathbf{C}^{k \times (n-k)} \right\}.$$

*If there exists a unitary U such that $U^*A_jU \in \mathcal{S}$ for all j , then $W_k(A)$ is convex.*

Let $1 \leq k < n$. Since all matrices in $\mathcal{U}(J_k)$ have Frobenius norm equal to \sqrt{k} , the set $\mathcal{U}(J_k)$ is highly non-convex in the sense that no three points in $\mathcal{U}(J_k)$ are collinear [18]. It is somewhat surprising that the projection of $\mathcal{U}(J_k)$ to a subspace of dimension $2k(n-k) + 1$ can be convex. In particular, if $n = 2k$ then $2k(n-k) + 1 = n^2/2 + 1$, which is more than half of the dimension of the space of $n \times n$ Hermitian matrices! In any event, we have the following result showing that $2k(n-k) + 1$ is indeed the upper limit.

Theorem 2.3 *The unitary orbit $\mathcal{U}(J_k)$ is a homogeneous manifold, and the tangent space at J_k equals*

$$\{i(HJ_k - J_kH) : H = H^*\} = \left\{ \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix} : X \in \mathbf{C}^{k \times (n-k)} \right\},$$

which has dimension $2k(n-k)$. Consequently, if $A = (A_1, \dots, A_m)$ is such that

$$\dim(\text{span}\{I, A_1, \dots, A_m\}) > 2k(n-k) + 1,$$

then $W_k(A)$ is not convex.

Proof. The set $\mathcal{U}(J_k)$ is the orbit of J_k under the group action $(U, C) \mapsto U^*CU$ for any unitary U and Hermitian C . Thus $\mathcal{U}(J_k)$ is a homogeneous manifold.

Since every unitary matrix admits a representation of the form e^{iH} for some Hermitian H , every smooth path of unitary matrices $U(t)$, $-1 \leq t \leq 1$, $U(0) = I$, is of the form $U(t) = e^{iH(t)}$ where $H(t)$, $-1 \leq t \leq 1$ is a smooth path of Hermitian matrices such that $H(0) = 0$. Thus we have

$$\left. \frac{d(e^{iH(t)} J_k e^{-iH(t)})}{dt} \right|_{t=0} = i(H'(0)J_k - J_k H'(0))$$

and the tangent space of $\mathcal{U}(J_k)$ at J_k is equal to

$$\{i(HJ_k - J_kH) : H = H^*\} = \left\{ \begin{pmatrix} 0 & X \\ X^* & 0 \end{pmatrix} : X \in \mathbf{C}^{k \times (n-k)} \right\},$$

which has dimension $2k(n-k)$.

Let $A = (A_1, \dots, A_m)$, and let $\mathcal{S} = \text{span}\{I, A_1, \dots, A_m\}$ satisfy $\dim \mathcal{S} > 2k(n-k) + 1$. We prove that $W_k(A)$ is not convex. Suppose $W_k(A)$ is indeed convex. By observation (4), we can find a linearly independent subset $\{I, B_1, \dots, B_s\}$ of $\{I, A_1, \dots, A_m\}$ with $s = 2k(n-k) + 1$ such that $W_k(B_1, \dots, B_s)$ is also convex.

We claim that $W_k(B_1, \dots, B_s)$ has nonempty interior in \mathbf{R}^s . If it is not true, then the convex set $W_k(B_1, \dots, B_s)$ must lie in a certain hyperplane

$$\mathcal{P} = \{v \in \mathbf{R}^s : v^t w = d\},$$

for some unit vector $w = (w_1, \dots, w_s)^t \in \mathbf{R}^s$ and $d \in \mathbf{R}$. But then we have

$$\text{tr} \left(\left(\sum_{j=1}^s w_j B_j \right) X \right) = d$$

for all $X \in \mathcal{U}(J_k)$, i.e., $W_k(\sum_{j=1}^s w_j B_j) = \{d\}$. It follows that (see [10]) $\sum_{j=1}^s w_j B_j = dI/k$, contradicting the fact that $\{I, B_1, \dots, B_s\}$ is linearly independent.

Now $\mathcal{U}(J_k)$ is a homogeneous manifold of (real) dimension $2k(n-k)$ and $\phi(\mathcal{U}(J_k)) = W_k(B_1, \dots, B_s)$ with

$$\phi(P) = (\text{tr } B_1 P, \dots, \text{tr } B_s P).$$

The set $W_k(B_1, \dots, B_s) \subseteq \mathbf{R}^s$ with $s = 2k(n-k) + 1$ cannot have non-empty interior. So, the assumption that $\dim \mathcal{S} > 2k(n-k) + 1$ cannot be true. \square

By the last two theorems, we see that a basis for the tangent space of $\mathcal{U}(J_k)$ at any point together with the identity matrix form a maximal linearly independent convex family for the k th generalized numerical range.

Next we turn to other linearly independent families $\{A_1, \dots, A_m\}$ so that $W_k(A)$ is convex. By the result in [14], we have the following.

Theorem 2.4 *Let $1 \leq k \leq n/2$ and $P \in \mathbf{C}^{k \times k}$ be Hermitian. Then there exists a rank k orthogonal projection of the form $\begin{pmatrix} P & * \\ * & * \end{pmatrix}$ if and only if all eigenvalues of P lie in $[0, 1]$.*

Clearly, the set of matrices $P \in \mathbf{C}^{k \times k}$ with eigenvalues lying in $[0, 1]$ is convex. Using Theorem 1.2, observations (1) and (2), we obtain

Theorem 2.5 *Let $A = (A_1, \dots, A_m)$ be an m -tuple of Hermitian matrices. Suppose*

$$\mathcal{S} = \{P \oplus \alpha I_{n-k} : \alpha \in \mathbf{R}, P \text{ is } k \times k \text{ Hermitian}\}.$$

If there exists a unitary U such that $U^ A_j U \in \mathcal{S}$ for all j , then $W_k(A)$ is convex.*

3 Convex Families for the First Joint Numerical Range

We first identify a maximal linearly independent convex family for the first joint numerical range which is different from those constructed in the previous section.

Theorem 3.1 *Let $A = (A_1, \dots, A_m)$ be an m -tuple of Hermitian matrices. Suppose*

$$\mathcal{S} = \text{span} (\{E_{jj} : 1 \leq j \leq n\} \cup \{E_{j,j+1} + E_{j+1,j} : 1 \leq j \leq n-1\}).$$

If there exists a unitary U such that $U^ A_j U \in \mathcal{S}$ for all j , then $W_1(A)$ is convex.*

Proof. Let $A = (A_1, \dots, A_{2n-1})$ so that $A_j = E_{jj}$ for $1 \leq j \leq n$ and $A_{n+j} = (E_{j,j+1} + E_{j+1,j})/2$ for $1 \leq j \leq n-1$. Then $(x_1, \dots, x_{2n-1}) \in W_1(A)$ if and only if there is a unit vector $v = (v_1, \dots, v_n)^t \in \mathbf{C}^n$ such that

$$\begin{aligned} x_j &= |v_j|^2 && \text{if } 1 \leq j \leq n, \\ x_{n+j} &= v_j \bar{v}_{j+1} + \bar{v}_j v_{j+1} && \text{if } 1 \leq j < n. \end{aligned}$$

These conditions hold if and only if (x_1, \dots, x_{2n-1}) satisfies

$$\sum_{j=1}^n x_j = 1, \quad x_j \geq 0 \text{ for } 1 \leq j \leq n, \quad \text{and}$$

$$|x_{n+j}|^2 \leq x_j x_{j+1} \text{ for } 1 \leq j \leq n-1.$$

Now suppose $x = (x_1, \dots, x_{2n-1}), y = (y_1, \dots, y_{2n-1}) \in W_1(A)$ and $z = (z_1, \dots, z_{2n-1})$ equals $(x+y)/2$. Clearly, we have

$$\sum_{j=1}^n z_j = 1, \quad z_j \geq 0 \text{ for } j = 1, \dots, n.$$

Moreover, for $j = 1, \dots, n-1$,

$$4|z_{n+j}|^2 = |x_{n+j} + y_{n+j}|^2 \leq (\sqrt{x_j x_{j+1}} + \sqrt{y_j y_{j+1}})^2 \leq (x_j + y_j)(x_{j+1} + y_{j+1}).$$

Hence $z \in W_1(A)$. Since $W_1(A)$ is closed, we conclude that it is a convex set. \square

A result of Horn [8] implies that $W_1(E_{11}, E_{22}, \dots, E_{nn})$ is convex. Theorem 3.1 strengthens this statement.

Note that a maximal linearly independent convex family of the first joint numerical range may not have $2(n-1) + 1$ elements as shown in the following example.

Example 3.2 Suppose $(k, n) = (1, 3)$. Then $\{I_3, A_1, A_2, A_3\}$ with

$$A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

is a maximal linearly independent convex family for $W_1(A)$.

Proof. Suppose there exists A_4 such that $\{I, A_1, \dots, A_4\}$ is linearly independent, and $W_1(I_3, A_1, \dots, A_4)$ is convex. One may replace A_4 by a matrix of the form $A_4 - (a_0 I_3 + a_1 A_1 + a_2 A_2 + a_3 A_3)$ so that the leading 2×2 principal submatrix is zero. Since $W_1(I_3, A_1, \dots, A_4)$ is convex if and only if $W_1(A)$ is convex with $A = (A_1, \dots, A_4)$, we can focus on $W_1(A) \subseteq \mathbf{R}^4$. Using the unit vectors $(1, 0, 0)^t$ and $(0, 1, 0)^t$, we see that $(1, 0, 0, 0), (-1, 0, 0, 0) \in W_1(A)$. By convexity, we see that $(0, 0, 0, 0) \in W_1(A)$. One easily checks that a unit vector giving rise to the point $(0, 0, 0, 0) \in W_1(A)$ must be of the form $(0, 0, \mu) \in \mathbf{C}^3$. As a result, the $(3, 3)$ entry of A_4 must be 0. Let U be a unitary matrix of the form $U = \hat{U} \oplus [1]$ so that $\hat{A}_4 = U^* A_4 U = \mu(E_{13} + E_{31})$. Then

$$\text{span}\{A_1, A_2, A_3, \hat{A}_4\} = \text{span}\{U^* A_j U : 1 \leq j \leq 4\}.$$

Thus $W_1(A)$ is convex if and only if $W_1(\hat{A})$ is, where $\hat{A} = (A_1, A_2, A_3, \hat{A}_4)$. Note that

$$W_1(\hat{A}) = \{(|v_1|^2 - |v_2|^2, v_1 \bar{v}_2 + v_2 \bar{v}_1, i(v_2 \bar{v}_1 - v_1 \bar{v}_2), v_1 \bar{v}_3 - v_3 \bar{v}_1) : v_1, v_2, v_3 \in \mathbf{C}, \sum_{j=1}^3 |v_j|^2 = 1\}.$$

Using the unit vectors $(1/2, \pm 1/2, 1/\sqrt{2})^t$, we see that $(0, \pm 1/2, 0, 1/\sqrt{2}) \in W_1(\hat{A})$. However, their mid-point $(0, 0, 0, 1/\sqrt{2}) \notin W_1(\hat{A})$, which is a contradiction. \square

In connection with the above discussion, it would be interesting to solve:

Problem 3.3 Characterize maximal linearly independent convex families for the first (or the k th) joint numerical range.

Problem 3.4 Determine maximal linearly independent convex families with smallest number of elements.

4 A Theorem on Non-convexity

Let $n \geq 3$. Then there exists a 4-tuple A of $n \times n$ Hermitian matrices such that $W_1(A)$ is not convex as mentioned in the introduction. Suppose A_1, A_2, A_3 are $n \times n$ Hermitian matrices such that $\{I, A_1, A_2, A_3\}$ is linearly dependent, say, $\text{span}\{I, A_1, A_2, A_3\} = \text{span}\{I, A_1, A_2\}$. Then for any Hermitian matrix A_4 , $W(A_1, A_2, A_4)$ is convex, and so is $W(I, A_1, A_2, A_3, A_4)$ by observations (2) and (3).

In the following, we show that as long as $\{I, A_1, A_2, A_3\}$ is linearly independent, one can find A_4 so that $W_1(A_1, A_2, A_3, A_4)$ is not convex. We first establish the following auxiliary result, which is of independent interest.

Theorem 4.1 *Let A_1, A_2, A_3 be $n \times n$ Hermitian matrices. If $\{I, A_1, A_2, A_3\}$ is linearly independent, then there exists $X \in \mathbf{C}^{n \times 2}$ with $X^*X = I_2$ such that $\{I_2, X^*A_1X, X^*A_2X, X^*A_3X\}$ is linearly independent.*

Proof. We assume that $n > 2$ to avoid trivial consideration. Since I, A_1, A_2 are linearly independent, the complex matrix $A_1 + iA_2$ cannot be written as $\alpha H + \beta I$ for any $\alpha, \beta \in \mathbf{C}$ and Hermitian matrix H . By [14, Theorem 3.5], there exists $X \in \mathbf{C}^{n \times 2}$ such that $X^*(A_1 + iA_2)X$ is not normal. By [9, 1.3.1], we may assume that $X^*(A_1 + iA_2)X$ has equal diagonal entries $\gamma_1 + i\gamma_2$ with $\gamma_1, \gamma_2 \in \mathbf{R}$. Replace A_1 and A_2 by suitable linear combinations of $A_1 - \gamma_1 I$ and $A_2 - \gamma_2 I$, we may assume that

$$X^*A_1X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad X^*A_2X = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

If $\{I_2, X^*A_1X, X^*A_2X, X^*A_3X\}$ is linearly independent, then we are done. If not, we may replace A_3 by a suitable linear combination of I, A_1, A_2, A_3 so that $X^*A_3X = 0$ and $\{I, A_1, A_2, A_3\}$ is still linearly independent. Since $A_3 \neq 0$, there exists $Y = [X|y] \in \mathbf{C}^{n \times 3}$ such that $Y^*Y = I_3$ and

$$\hat{A}_1 = Y^*A_1Y = \begin{pmatrix} 0 & 1 & a_1 \\ 1 & 0 & a_2 \\ \bar{a}_1 & \bar{a}_2 & a_3 \end{pmatrix}, \quad \hat{A}_2 = Y^*A_2Y = \begin{pmatrix} 0 & i & b_1 \\ -i & 0 & b_2 \\ \bar{b}_1 & \bar{b}_2 & b_3 \end{pmatrix},$$

$$\text{and} \quad \hat{A}_3 = Y^*A_3Y = \begin{pmatrix} 0 & 0 & c_1 \\ 0 & 0 & c_2 \\ \bar{c}_1 & \bar{c}_2 & c_3 \end{pmatrix} \neq 0.$$

We may assume that $c_2 \geq |c_1|$; otherwise, we may replace \hat{A}_3 by $D^*\hat{A}_3D$ with $D = \text{diag}(1, \mu, 1)$ or $D = \begin{pmatrix} 0 & \mu \\ 1 & 0 \end{pmatrix} \oplus [1]$ for some suitable $\mu \in \mathbf{C}$ with $|\mu| = 1$. One can then

replace \hat{A}_1 and \hat{A}_2 by suitable linear combinations of $D^*\hat{A}_1D$ and $D^*\hat{A}_2D$, and assume that \hat{A}_1 and \hat{A}_2 are still of the form in the display.

Now consider $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \end{pmatrix}^t$ with $\theta \in (0, \pi/2)$. Then

$$P^*A_1P = \sin \theta \begin{pmatrix} 0 & \cot \theta + a_1 \\ \cot \theta + \bar{a}_1 & (a_2 + \bar{a}_2) \cos \theta + a_3 \sin \theta \end{pmatrix},$$

$$P^*A_2P = \sin \theta \begin{pmatrix} 0 & i \cot \theta + b_1 \\ -i \cot \theta + \bar{b}_1 & (b_2 + \bar{b}_2) \cos \theta + b_3 \sin \theta \end{pmatrix},$$

$$P^*A_3P = \sin \theta \begin{pmatrix} 0 & c_1 \\ \bar{c}_1 & 2c_2 \cos \theta + c_3 \sin \theta \end{pmatrix}.$$

We claim that there exists $\theta \in (0, \pi/2)$ such that $\{I_2, P^*\hat{A}_1P, P^*\hat{A}_2P, P^*\hat{A}_3P\}$ is linearly independent, and hence the conclusion of the theorem follows. Since all the $(1, 1)$ entries of $P^*\hat{A}_1P, P^*\hat{A}_2P$ and $P^*\hat{A}_3P$ equal 0, I_2 is not a linear combination of these 3 matrices. To establish our claim, we need only show that there exists $\theta \in (0, \pi/2)$ such that $P^*\hat{A}_1P, P^*\hat{A}_2P$ and $P^*\hat{A}_3P$ are linearly independent. To this end, construct the following matrix B , whose rows contain the real and imaginary parts of the entries of $P^*\hat{A}_jP$, etc.:

$$B = \sin \theta \begin{pmatrix} \cot \theta + (a_1 + \bar{a}_1)/2 & i(\bar{a}_1 - a_1)/2 & (a_2 + \bar{a}_2) \cos \theta + a_3 \sin \theta \\ (b_1 + \bar{b}_1)/2 & \cot \theta + i(\bar{b}_1 - b_1)/2 & (b_2 + \bar{b}_2) \cos \theta + b_3 \sin \theta \\ (c_1 + \bar{c}_1)/2 & i(\bar{c}_1 - c_1)/2 & 2c_2 \cos \theta + c_3 \sin \theta \end{pmatrix}.$$

Now,

$$\det(B)/\sin^4 \theta = \det \begin{pmatrix} \cot \theta + (a_1 + \bar{a}_1)/2 & i(\bar{a}_1 - a_1)/2 & (a_2 + \bar{a}_2) \cot \theta + a_3 \\ (b_1 + \bar{b}_1)/2 & \cot \theta + i(\bar{b}_1 - b_1)/2 & (b_2 + \bar{b}_2) \cot \theta + b_3 \\ (c_1 + \bar{c}_1)/2 & i(\bar{c}_1 - c_1)/2 & 2c_2 \cot \theta + c_3 \end{pmatrix}$$

can be viewed as $p(\cot \theta)$ for some real polynomial p of degree 3. If $\det(B) = 0$ for all $\theta \in (0, \pi/2)$, then the coefficient of $\cot^3 \theta$ in $p(\cot \theta)$ is 0, which is just $2c_2$ by expanding $\det(B)/\sin^4 \theta$. Since $c_2 \geq |c_1|$, we see that $c_1 = 0$ as well. Now, consider the coefficient of $\cot^2 \theta$ in $p(\cot \theta)$, which is just c_3 . Again, it has to be 0. It follows that $\hat{A}_3 = 0$, which is a contradiction. \square

Now we are ready to state the non-convexity result of the joint numerical range.

Theorem 4.2 *Suppose A_1, A_2, A_3 are Hermitian matrices such that $\{I, A_1, A_2, A_3\}$ is linearly independent. Then there exists a Hermitian matrix A_4 such that $W_1(A_1, A_2, A_3, A_4)$ is not convex.*

Proof. By the previous theorem, there exists $X \in \mathbf{C}^{n \times 2}$ such that $X^*X = I_2$ and $\{I_2, X^*A_1X, X^*A_2X, X^*A_3X\}$ is linearly independent. Let $A_4 = XX^*$. Then $(a, b, c, 1) \in W_1(A_1, A_2, A_3, A_4)$ if and only if $(a, b, c) \in W_1(X^*A_1X, X^*A_2X, X^*A_3X)$, which is not convex [1]. The result follows. \square

We remark that Theorem 4.1 and its proof can be easily modified to deal with infinite dimensional self-adjoint operators A_1, A_2, A_3 . In general, it is interesting to solve the following problem.

Problem 4.3 *If a linearly independent family $\{A_1, \dots, A_m\}$ of (finite or infinite dimensional) self-adjoint operators is given, where $(r-1)^2 < m \leq r^2$, does there exist an X such that $X^*X = I_r$ and $\{X^*A_jX : j = 1, \dots, m\}$ is linearly independent.*

By private communication, Doug Farenick pointed out that this problem can be studied in the context of unital completely positive maps on C^* -algebras.

5 Related results and questions

There are many variations of our problems. We mention a few of them in the following.

Real symmetric matrices

In applications, one often has to consider real symmetric matrices instead of complex Hermitian matrices. One can modify the results and proofs on complex Hermitian matrices and get the following analogs for real symmetric matrices.

Theorem 5.1 *Let $A = (A_1, \dots, A_m)$ be an m -tuple of real symmetric matrices. Suppose $1 \leq k \leq n/2$ and*

$$\mathcal{S} = \left\{ \alpha I + \begin{pmatrix} 0_k & X \\ X^t & 0_{n-k} \end{pmatrix} : \alpha \in \mathbf{R}, X \in \mathbf{R}^{k \times (n-k)} \right\},$$

or

$$\mathcal{S} = \{P \oplus \alpha I_{n-k} : \alpha \in \mathbf{R}, P \text{ is } k \times k \text{ real symmetric}\}.$$

If there is a real orthogonal Q such that $Q^t A_j Q \in \mathcal{S}$, then $W_k(A)$ is convex.

Theorem 5.2 *The orthogonal similarity orbit $\mathcal{O}(J_k)$ is a homogeneous manifold, and the tangent space at J_k equals*

$$\{(KJ_k - J_kK) : K = -K^t\} = \left\{ \begin{pmatrix} 0 & X \\ X^t & 0 \end{pmatrix} : X \in \mathbf{R}^{k \times (n-k)} \right\},$$

which has dimension $k(n-k)$. Consequently, if $A = (A_1, \dots, A_m)$ such that

$$\dim(\text{span}\{I, A_1, \dots, A_m\}) > k(n-k) + 1,$$

then $W_k(A)$ is not convex.

By this theorem and a result of Horn [8], we have the following corollary.

Corollary 5.3 *The elements of $A = (E_{11}, E_{22}, \dots, E_{nn})$ form a maximal linearly independent convex family for $W_1(A)$.*

Rectangular matrices

Suppose $A = (A_1, \dots, A_m)$, where A_1, \dots, A_m are $n \times r$ matrices over $\mathbf{F} = \mathbf{R}$ or \mathbf{C} . To be specific, we assume that $n \leq r$ in our discussion. Otherwise, consider the transposes of A_1, \dots, A_m . For $1 \leq k \leq n$, define

$$V_k(A) = \{(\sum_{j=1}^k x_j^* A_1 y_j, \dots, \sum_{j=1}^k x_j^* A_m y_j) : \{x_1, \dots, x_k\} \text{ is an orthonormal set of } \mathbf{F}^n, \\ \{y_1, \dots, y_k\} \text{ is an orthonormal set of } \mathbf{F}^r\},$$

which is a subset of \mathbf{F}^m . Let \mathcal{V}_k be the collection of $n \times r$ matrices X such that X^*X is a rank k orthogonal projection. It is not hard to see that

$$V_k(A) = \{(\text{tr } X^* A_1, \dots, \text{tr } X^* A_m) : X \in \mathcal{V}_k\}.$$

It was shown in [13, Theorems 14 and 41] that $V_k(A)$ is not convex in general if $m > 2$. Again, if A_1, \dots, A_m belong to a certain subset, then one can get convexity for a much larger m as shown in the following result.

Theorem 5.4 *Let $A = (A_1, \dots, A_m)$ be an m -tuple of $n \times r$ matrices. Suppose*

$$\mathcal{S} = \left\{ \begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix} \in \mathbf{F}^{n \times r} : X \in \mathbf{F}^{p \times q} \right\},$$

where $\min\{p, q\} \leq k \leq n + r - p - q$. If there exist square matrices U and V with $U^*U = I_n$ and $V^*V = I_r$ such that $UA_jV \in \mathcal{S}$ for all j , then $V_k(A)$ is convex.

The proof of this theorem depends on

Theorem 5.5 *Let $(p, q) \leq (n, r)$ in the entrywise sense, and $\min\{p, q\} \leq k \leq n + r - p - q$. Then a $p \times q$ matrix X can be embedded in an $n \times r$ matrix Y so that Y^*Y is a rank k orthogonal projection if and only if $\|X\| \leq 1$. Hence, the collection of all such $X \in \mathbf{F}^{p \times q}$ is a convex set.*

Proof. By a result of Thompson [17] and our assumptions on the positive integers n, r, p, q, k , we see that a $p \times q$ matrix X with singular values $s_1 \geq \dots \geq s_k$ can be embedded in an $n \times r$ matrix with the k largest singular values equal to 1 and the rest equal to 0 if and only if $s_i \leq 1$. The result follows. \square

Let $R_k = E_{11} + E_{22} + \dots + E_{kk} \in \mathbf{F}^{n \times r}$. In the complex case, the tangent space of the manifold \mathcal{V}_k at R_k is

$$\begin{aligned} \mathcal{T}(R_k) &= \{i(R_k H + G R_k) : H = H^*, G = G^*\} \\ &= \left\{ \begin{pmatrix} iX & Y \\ Z & 0_{(n-k) \times (r-k)} \end{pmatrix} : X = X^* \in \mathbf{C}^{k \times k} \right\}, \end{aligned}$$

and has real dimension $k^2 + 2k(n + r - 2k)$. In the real case, the tangent space of the manifold \mathcal{V}_k at R_k is

$$\begin{aligned} \mathcal{T}(R_k) &= \{R_k H + G R_k : H = -H^t, G = -G^t\} \\ &= \left\{ \begin{pmatrix} X & Y \\ Z & 0_{(n-k) \times (r-k)} \end{pmatrix} : X = -X^t \in \mathbf{R}^{k \times k} \right\}, \end{aligned}$$

and has real dimension $k(k - 1)/2 + k(n + r - 2k)$. It would be interesting to see whether one can construct a maximal linearly independent convex family for $V_k(A)$ of this size.

Complex symmetric matrices

Let $A = (A_1, \dots, A_m)$, where A_1, \dots, A_m are $n \times n$ complex symmetric matrices. For $1 \leq k \leq n$, let $J_k = I_k \oplus 0_{n-k}$, and define

$$W_k^t(A) = \{(\operatorname{tr} J_k U^t A_1 U, \dots, \operatorname{tr} J_k U^t A_m U) : U^* U = I\}.$$

Then Theorem 5.1 is valid for the complex case. However, the tangent space of the manifold at J_k is

$$\begin{aligned} \mathcal{U}^t(J_k) &= \{i(J_k H + H^t J_k) : H = H^*\} \\ &= \left\{ i \begin{pmatrix} X & Y \\ Y^t & 0_{n-k} \end{pmatrix} : X = X^t \in \mathbf{R}^{k \times k}, Y \in \mathbf{C}^{k \times (n-k)} \right\}, \end{aligned}$$

and has real dimension $k(k+1)/2 + 2k(n-k)$. It is interesting to know whether one can find a maximal linearly independent convex family for $W_k^t(A)$ of this size.

Note that one may consider $W_k^t(A)$ for an m -tuple of complex matrices $A = (A_1, \dots, A_m)$. However, it is easy to check that $W_k^t(A)$ is the same as $W_k^t(\hat{A}_1, \dots, \hat{A}_m)$, where \hat{A}_j is the symmetric part of A_j for all j . For skew-symmetric matrices, one needs a different treatment as shown in the next subsection.

Skew-symmetric matrices

Suppose A_1, \dots, A_m are skew-symmetric matrices over $\mathbf{F} = \mathbf{R}$ or \mathbf{C} . For $1 \leq k \leq n/2$, let $T_k = \begin{pmatrix} 0_k & I_k \\ -I_k & 0_k \end{pmatrix} \oplus 0_{n-2k}$, and define

$$W_k^t(A) = \{(\operatorname{tr} T_k U^t A_1 U, \dots, \operatorname{tr} T_k U^t A_m U) : U^* U = I\}.$$

We have the following convexity theorem.

Theorem 5.6 *Let $A = (A_1, \dots, A_m)$ be an m -tuple of $n \times n$ skew-symmetric matrices. Suppose*

$$\begin{aligned} \mathcal{S} &= \{E_{j,k+j} : j = 1, \dots, k\}, \quad \text{or} \\ \mathcal{S} &= \left\{ \begin{pmatrix} 0_p & 0 & X \\ 0 & 0 & 0 \\ -X^t & 0 & 0_q \end{pmatrix} : X \in \mathbf{F}^{p \times q}, \|X\| \leq 1 \right\} \end{aligned}$$

for some positive integers p and q such that $n - (p+q) \geq k \geq \min\{p, q\}$. If there exists a square matrix U with $U^* U = I_n$ such that $U^t A_j U \in \mathcal{S}$ for all j , then $W_k^t(A)$ is convex.

The proof of this theorem depends on following two theorems. First of all, we have the following (see [16] and its references).

Theorem 5.7 *Let $(d_1, \dots, d_k) \in \mathbf{F}^k$. There exists U with $U^* U = I_n$ such that the $(1, k+1), (2, k+2), \dots, (k, 2k)$ entries of $U^t T_k U$ equal d_1, \dots, d_k if and only if $|d_j| \leq 1$ for all j . Hence, the collection of such (d_1, \dots, d_k) form a convex set.*

Note that the singular values of the matrix $\begin{pmatrix} 0_k & X \\ -X^t & 0_{n-k} \end{pmatrix}$ are just two copies of those of X . Applying the result on rectangular matrices to the $k \times (n - k)$ right top corner of a skew-symmetric matrix, we have the following result.

Theorem 5.8 *Let p and q be positive integers with $n - (p + q) \geq k \geq \min\{p, q\}$. Then a $p \times q$ matrix X can be embedded in the right top corner of an $n \times n$ skew-symmetric matrix with the $2k$ largest singular values equal to 1 and the rest equal to 0 if $\|X\| \leq 1$. Hence, the collection of all such X is a convex set.*

Acknowledgement

Research of the first author was partially supported by an NSF grant. This paper was done while he was visiting the University of Toronto in the academic year 1998-1999 supported by a Faculty Research grant of the College of William and Mary. He would like to express his thanks to Professor M.D. Choi for making the visit possible, and the staff of the University of Toronto for their warm hospitality.

Note added in proof

Professor M.D. Choi pointed out that the answer of Problem 4.3 is negative. For instance, the 10×10 Hermitian matrices $A_j = E_{1j} + E_{j1}$, $j = 2, \dots, 10$, are linearly independent, but it is impossible to have 10×3 matrix V with $V^*V = I_3$ so that $\{V^*A_jV : j = 2, \dots, 10\}$ is linearly independent. A modified problem would be: Given linearly independent Hermitian operators A_1, \dots, A_m , find the smallest positive integer r so that there exists V with $V^*V = I_r$ for which the set $\{V^*A_jV : j = 1, \dots, m\}$ is still linearly independent.

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, An extension of the Hausdorff-Toeplitz theorem on the numerical range, *Proc. Amer. Math. Soc.* 89 (1983), 215–218.
- [3] P. Binding, D. Farenick and C.K. Li, A dilation and norm in several variable operator theory, *Canad. J. Math.* 47 (1995), 449-461.
- [4] P. Binding and C.K. Li, Joint ranges of Hermitian matrices and simultaneous diagonalization, *Linear Algebra Appl.* 151 (1991), 157-168.
- [5] M. Cho and M. Takaguchi, Some classes of commuting m -tuples of operators, *Studia Math.* 80 (1984), 245-259.
- [6] J.C. Doyle, Analysis of feedback systems with structured uncertainties, *Proc. IEEE-D* 129, 6 (1982), 242-250.

- [7] M. Fan and A. Tits, m -form numerical range and the computation of the structured singular values, *IEEE Trans. Automat. Control* AC-33 (1988), 284-289.
- [8] A. Horn, Doubly Stochastic matrices and the diagonal of a rotation matrix, *Amer. J. Math.* 76 (1954), 620-630.
- [9] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*, Cambridge University Press, 1991.
- [10] C.K. Li, The C -convex matrices, *Linear and Multilinear Algebra* 21 (1987), 303-312.
- [11] C.K. Li, Matrices with some extremal properties, *Linear Algebra Appl.* 101 (1988), 255-267.
- [12] C.K. Li and R. Mathias, Inequalities on the singular values of an off-diagonal block of a Hermitian matrix, to appear in *J. of Inequalities and Applications*. Preprint available at <http://www.math.wm.edu/~ckli/pub.html>.
- [13] C.K. Li and T.Y. Tam, Numerical ranges arising from simple Lie algebras – convexity and inclusion relation. Preprint available at <http://www.auburn.edu/~tamtiny/pub.html>.
- [14] C.K. Li and N.K. Tsing, On the k^{th} matrix numerical range, *Linear and Multilinear Algebra* 28 (1991), 229-239.
- [15] Y.T. Poon, On the convex hull of the multiform numerical range, *Linear and Multilinear Algebra* 37 (1994), 221-224.
- [16] T.Y. Tam, Partial superdiagonal elements and singular values of a complex skew-symmetric matrix, *SIAM J. Matrix Analysis Appl.* 19 (1998), 737-754.
- [17] R.C. Thompson, Principal submatrices IX: Interlacing inequalities for singular values of submatrices, *Linear Algebra Appl.* 5 (1972), 1-12.
- [18] R.C. Thompson, Research problem: The matrix numerical range, *Linear and Multilinear Algebra* 21 (1987), 321-323.