Canad. J. Math. Vol. **52** (1), 2000 pp. 141–171

Numerical Ranges Arising from Simple Lie Algebras

Dedicated to Professor Y. H. Au-Yeung on the occasion of his retirement

Chi-Kwong Li and Tin-Yau Tam

Abstract. A unified formulation is given to various generalizations of the classical numerical range including the c-numerical range, congruence numerical range, q-numerical range and von Neumann range. Attention is given to those cases having connections with classical simple real Lie algebras. Convexity and inclusion relation involving those generalized numerical ranges are investigated. The underlying geometry is emphasized.

1 Introduction

The (classical) numerical range of $A \in \mathbb{C}^{n\times n}$ is defined by

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

This concept and its many generalizations have been studied heavily in the last few decades because of their connections and applications to many pure and applied areas (see *e.g.* [10], [11], [14]). One of the interesting results, perhaps the most fascinating, about the classical numerical range is the celebrated Toeplitz-Hausdorff theorem [38], [12] asserting that the numerical range is always a convex subset of $\mathbb{C}$. In fact, the convexity has often been a concern when different generalizations are considered. For example, given $C \in \mathbb{C}^{n\times n}$ with $C = C^*$, Au-Yeung and Tsing [3] considered the (joint) C-numerical range of several Hermitian matrices $A_1, \ldots, A_p \in \mathbb{C}^{n\times n}$ defined by

$$(1) \qquad W_C(A_1, \ldots, A_p) = \{(\operatorname{tr} CU^*A_1U, \ldots, \operatorname{tr} CU^*A_pU) : U \in U(n)\},$$

where $U(n)$ is the unitary group, and studied the convexity and several other related problems involving $W_C(A_1, \ldots, A_p)$. The C-numerical range embraces various generalizations of the classical numerical range including the joint numerical range $W(A_1, \ldots, A_p)$ considered by Brickman [5], the k-numerical range considered by Halmos and Berger [11], [4], and the c-numerical range considered by Westwick and Poon [41], [24]. (More results on the C-numerical range will be given in the next few sections.) Actually, Au-Yeung and Tsing [3] also studied the C-numerical range of $A_1, \ldots, A_p$, for real symmetric or real quaternion Hermitian matrices $C, A_1, \ldots, A_p$. In these cases, the set $U(n)$ in (1) is replaced by the set of $n \times n$ matrices X over the real field $\mathbb{R}$ or the skew-field of real quaternions $\mathbb{H}$ satisfying $X^*X = I_n$.

Received by the editors March 12, 1998; revised October 14, 1999.
Research of the first author was partially supported by an NSF grant.
AMS subject classification: 15A60, 17B20.
Keywords: numerical range, convexity, inclusion relation.

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

Inspired by the study of Au-Yeung and Tsing, we consider the C-numerical range in the following setting. (In most cases, we will not use new notation for the different kinds of C-numerical range in the following discussion, but will make the definition clear in each case in the context). Let $\mathbf{V}$ be a matrix space (or any finite dimensional linear space) equipped with a real inner product (X, Y) which is invariant under a compact group G of operators acting on $\mathbf{V}$, *i.e.*, $(gX, gY) = (X, Y)$ for all $g \in G$ and $X, Y \in \mathbf{V}$. For a given $C \in \mathbf{V}$, define the (joint) C-numerical range of $A_1, \ldots, A_p \in \mathbf{V}$ by

$$(2) \qquad W_C(A_1, \ldots, A_p) = \{\big((A_1, Z), \ldots, (A_p, Z)\big) : Z \in G(C)\},$$

where

$$G(C) = \{g(C) : g \in G\}$$

is the orbit of C under G. Evidently, one can regard $W_C(A_1, \ldots, A_p)$ as the image of the orbit $G(C)$ under the linear map $Z \mapsto \big((A_1, Z), \ldots, (A_p, Z)\big)$. Since (X, Y) is G-invariant, one easily verifies that

$$W_C(A_1, \ldots, A_p) = \{\big((X_1, C), \ldots, (X_p, C)\big) : (X_1, \ldots, X_p) \in G(A_1, \ldots, A_p)\},$$

where $G(A_1, \ldots, A_p) = \{\big(g(A_1), \ldots, g(A_p)\big) : g \in G\}$ is the joint orbit of $A_1, \ldots, A_p$ under the group G. Thus, $W_C(A_1, \ldots, A_p)$ can also be viewed as the image of a linear map on the joint orbit $G(A_1, \ldots, A_p)$. Furthermore, $W_C(A_1, \ldots, A_p)$ covers many other types of generalized numerical ranges in the literature. We describe a few of them in the following.

Thompson [37] introduced the C-congruence numerical range of a complex $n \times n$ matrix A: $W_C^T(A) = \{\operatorname{tr} CU^TAU : U \in U(n)\}$, where C is a given $n \times n$ complex symmetric matrix. He proved that $W_C^T(A)$ is a circular disk centered at the origin when $n > 1$ and is a circle when $n = 1$. Then the complex skew symmetric case was studied in [26]. It is convex except for $n = 2$ in which case the range is a circle (may be a point). Then Tam and Tsing [34] conjectured and Choi *et al.* [6] proved that $W_C^T(A)$ is convex whenever $n > 2$ for general complex matrices A and C (the case $n = 1$ is trivial). Clearly, $W_C^T(A)$ can be viewed as $W_{C^*}(A, iA)$ in (2) if we let $G(X) = \{U^TXU : U \in U(n)\}$ and $(X, Y) = \operatorname{Re} \operatorname{tr}(XY^*)$ on $\mathbb{C}^{n \times n}$.

Next, let $G(X) = \{UXV : U, V \in U(n)\}$ and $(X, Y) = \operatorname{Re} \operatorname{tr}(XY^*)$ on $\mathbb{C}^{n \times n}$. This setting covers two other generalizations of the classical numerical range. First, for any $n \times n$ complex matrices C and A, $W_{C^*}(A, iA)$ reduces to the set $\{\operatorname{tr} CUAV : U, V \in U(n)\}$ considered by von Neumann [22]. The von Neumann range is a circular disk centered at the origin when $n > 1$ and hence convex; and it is a circle when $n = 1$.

The q-numerical range of an $n \times n$ complex matrix A, $q \in \mathbb{C}$ satisfying $|q| \le 1$, is the set $W(q : A) = \{y^*Ax : x, y \in \mathbb{C}^n, x^*x = y^*y = 1, y^*x = q\}$. Evidently, $W(1 : A) = W(A)$. Tsing [39] proved that $W(q : A)$ is convex. See [19] for a shorter proof, and [20] for further results and references. One can obtain $W(q : A)$ by fixing the third and the fourth coordinates of the set $W_C(A, -iA, I, -iI)$, *i.e.*, $\operatorname{Re} y^*x = \operatorname{Re} q$ and $\operatorname{Im} y^*x = \operatorname{Im} q$.

Our definition of $W_C(A_1, \ldots, A_p)$ also covers the notion of numerical range in the context of compact connected Lie groups studied in [31] recently (see the next section for the definition and the convexity result). In this paper, we consider the study of $W_C(A_1, \ldots, A_p)$ in connection to classical simple real Lie algebras. The convexity of $W_C(A_1, \ldots, A_p)$ is our main concern.

Following Au-Yeung and Tsing [3] (see also [25], [31]), we relate the convexity problem to inclusion relations for $W_C(A_1,\dots,A_p)$ (see Section 3). The underlying geometry of the orbit $G(C)$ will be emphasized. Some Lie theory background will be given in Section 2. Connection between the convexity and inclusion relation together with some technical lemmas are given in Section 3. In Sections 4–11, we consider $W_C(A_1,\dots,A_p)$ arising from real classical simple Lie algebras. Some concluding remarks are given in Section 12.

2 The Formulations in Lie Setting

Let G be a semisimple compact connected Lie group, let $\mathfrak{g}$ be its Lie algebra with the Killing form $B(\cdot,\cdot)$. For a given $C \in \mathfrak{g}$, we define the C-numerical range of $A_1,\dots,A_p \in \mathfrak{g}$ by

$$W_C(A_1,\dots,A_p) = \{\big(B(A_1,Z),\dots,B(A_p,Z)\big) : Z \in O(C)\},$$

where $O(C) = \{\mathrm{Ad}(g)C : g \in G\}$ is the orbit of C in $\mathfrak{g}$ under the adjoint action of G. Since the Killing form is negative definite, one sees that up to a suitable scalar multiplication the C-numerical range associated with a compact connected Lie group G defined above can be viewed as a special case of the C-numerical range defined in (2). The Lie group numerical range was studied in [31] and the following result was proved.

Theorem 2.1 *The Lie group numerical range $W_C(A_1,A_2)$ is convex.*

Indeed Theorem 2.1 is true for general compact connected Lie groups. It is because for every compact connected Lie group G, G is the commuting product G_sZ_0 and $\mathfrak{g} = \mathfrak{g}_s + \mathfrak{z}$ where G_s is the analytic subgroup of G with semisimple [13, p. 132] Lie algebra $\mathfrak{g}_s = [\mathfrak{g},\mathfrak{g}]$ and Z_0 is the identity component of the center Z of G, whose Lie algebra is $\mathfrak{z}$. Now $\mathrm{Ad}(Z)$ is trivial and $\mathrm{Ad}(G)$ acts trivially on $\mathfrak{z}$. So for any $X = X_s + Y$ where $X_s \in \mathfrak{g}_s$, $Y \in \mathfrak{z}$, $O_G(X) = O_{G_s}(X_s) + Y$ where $O_G(\cdot)$ denotes the orbit under the adjoint action of G.

We remark that Theorem 2.1 is very useful in handling the numerical ranges associated with the realifications of classical (exceptional as well) complex simple Lie algebras discussed in the next few sections. Here is another result that will be used in our later study.

Proposition 2.2 *Let G_1 and G_2 be connected Lie groups such that $\psi\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is an isomorphism.*

1. *If $C \in \mathfrak{g}_1$, then $\psi\big(O_1(C)\big) = O_2\big(\psi(C)\big)$, where $O_i(\cdot)$ denotes the adjoint orbit corresponding to G_i, $i = 1, 2$.*
2. *If $C, A_1,\dots,A_p \in \mathfrak{g}_1$, then $W^1_C(A_1,\dots,A_p) = W^2_{\psi(C)}\big(\psi(A_1),\dots,\psi(A_p)\big)$, where W^i denotes the numerical range corresponding to G_i, $i = 1, 2$.*

Proof (1) Suppose G_1 is simply connected. Then there exists a homomorphism $\varphi\colon G_1 \to G_2$ onto G_2 such that $d\varphi_e = \psi$ [40, pp. 100–101]. Since $d\varphi_e \cdot \mathrm{Ad}(g) = \mathrm{Ad}\big(\varphi(g)\big) \cdot d\varphi_e$ for any $g \in G_1$ [13, p. 110, p. 127], $\psi\big(O_1(C)\big) = O_2\big(\psi(C)\big)$.

If G_1 is not simply connected, let G_1' be a simply connected Lie group with the same Lie algebra $\mathfrak{g}_1$. Then we have $O_1(C) = O_1'(C)$. In other words, the orbit is invariant under different choices of Lie groups with the same Lie algebra and we have the desired result.

(2) Notice that $\operatorname{ad}\big(\psi(C)\big) = \psi \operatorname{ad} C\psi^{-1}$ for any $C \in \mathfrak{g}_1$. Thus for any $X, Y \in \mathfrak{g}_1$, $B_1(X,Y) = B_2\big(\psi(X), \psi(Y)\big)$ and

$$\begin{aligned} W_C^1(A_1,\dots,A_p) &= \{\big(B(A_1,Z),\dots,B(A_p,Z)\big) : Z \in O_1(C)\} \\ &= \Big\{\Big(B\big(\psi(A_1),\psi(Z)\big),\dots,B\big(\psi(A_p),\psi(Z)\big)\Big) : \psi(Z) \in \psi\big(O_1(C)\big)\Big\} \\ &= \Big\{\Big(B\big(\psi(A_1),\psi(Z)\big),\dots,B\big(\psi(A_p),\psi(Z)\big)\Big) : \psi(Z) \in \Big(O_2\big(\psi(C)\big)\Big)\Big\} \\ &= W^2_{\psi(C)}\big(\psi(A_1),\dots,\psi(A_p)\big). \end{aligned}$$ ■

While the Lie group numerical range embraces many types of generalized numerical ranges, and has nice convexity property (see [31]), it is not adequate to cover all kinds of generalized numerical ranges mentioned in the introduction. For instance, it does not cover the C-numerical range on real symmetric matrices $A_1,\dots,A_p$ considered by Au-Yeung and Tsing [2]. To correct this, we need to consider numerical ranges arising from real semi-simple Lie algebras.

Let G be an analytic group associated with the real semisimple Lie algebra $\mathfrak{g}$. Let $K \subset G$ (it is unique once we fix G [13, p. 112]) be the analytic group of $\mathfrak{k}$, and let $\mathfrak{g} = \mathfrak{k}+\mathfrak{p}$ be a given Cartan decomposition of $\mathfrak{g}$, here $\mathfrak{p}$ is the orthogonal complement of $\mathfrak{k}$ in $\mathfrak{g}$ with respect to the Killing form $B(\cdot,\cdot)$. For $A_1,\dots,A_p, C \in \mathfrak{p}$, the C-numerical range of $(A_1,\dots,A_p)$ is defined [31] as the following set in $\mathbb{R}^p$:

$$W_C(A_1,\dots,A_p) = \{\big(B(A_1,Z),\dots,B(A_p,Z)\big) : Z \in O(C)\},$$

where $O(C) = \{\operatorname{Ad}(k)C : k \in K\}$ is the orbit of C in $\mathfrak{p}$ under the adjoint action of K. In the following, we show that once we identify the Lie algebra $\mathfrak{g}$, the C-numerical range is independent of the choice of analytic group associated with it.

Proposition 2.3 *Let $C \in \mathfrak{p}$. The orbit $O(C)$ is independent of the choice of the analytic group G and so is the C-numerical range.*

Proof Let G' be a simply connected Lie group whose Lie algebra is also $\mathfrak{g}$. Consider the trivial isomorphism $\operatorname{id}: \mathfrak{g} \to \mathfrak{g}$. Then there is a unique analytic homomorphism $\pi: G' \to G$ [40, p. 101] such that $d\pi_e = \operatorname{id}$. Let K' (K) be the analytic subgroup of G' (G) with Lie algebra $\mathfrak{k}$. The group K is generated by the elements $\exp(Z)$, $Z \in \mathfrak{k}$. Likewise, the group $\pi(K')$ is generated by $\pi(\exp Z) = \exp d\pi_e(Z) = \exp(Z)$, $Z \in \mathfrak{k}$. It follows that $K = \pi(K')$. Now using $\operatorname{Ad}_G\big(\pi(k)\big) \cdot d\pi_e = d\pi_e \cdot \operatorname{Ad}_{G'}(k)$, $k \in K'$, we have $O_K(C) = O_{K'}(C)$, $C \in \mathfrak{p}$. ■

By Proposition 2.3, we can choose any analytic group of $\mathfrak{g}$ when we consider the corresponding numerical range associated with a given Cartan decomposition. Next, we show that there is a nice relation between the generalized numerical ranges arising from two isomorphic semisimple real Lie algebras, and hence one can transfer convexity (or non-convexity) results between them. Let $\mathfrak{g}_1 = \mathfrak{k}_1 + \mathfrak{p}_1$ and $\mathfrak{g}_2 = \mathfrak{k}_2 + \mathfrak{p}_2$ be Cartan decompositions of two isomorphic semisimple real Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$. Let $\phi: \mathfrak{g}_1 \to \mathfrak{g}_2$ be an isomorphism. Thus $\mathfrak{g}_2 = \phi(\mathfrak{k}_1) + \phi(\mathfrak{p}_1)$ is also a Cartan decomposition of $\mathfrak{g}_2$. There exists [13, p. 183] $\sigma \in \operatorname{Int}(\mathfrak{g}_2)$ satisfying $\sigma\big(\phi(\mathfrak{k}_1)\big) = \mathfrak{k}_2$ and $\sigma\big(\phi(\mathfrak{p}_1)\big) = \mathfrak{p}_2$.

Proposition 2.4 *With the above notations, let $\varphi = \sigma \cdot \phi$.*

1. *For any $C \in \mathfrak{p}_1$, $\varphi\big(O_{K_1}(C)\big) = O_{K_2}\big(\varphi(C)\big)$ where K_i is the analytic subgroup of G_i for $\mathfrak{k}_i$, $i = 1, 2$.*
2. *$W^1_C(A_1, \dots, A_p) = W^2_{\varphi(C)}\big(\varphi(A_1), \dots, \varphi(A_p)\big)$ where W^i denotes the numerical range corresponding to the given Cartan decomposition, $i = 1, 2$.*

Proof Let G_1 (we assume that G is simply connected because of Proposition 2.3) and G_2 be analytic groups of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ respectively. There is an analytic homomorphism $\pi\colon G_1 \to G_2$ onto G_2 such that $d\pi_e = \varphi$. Since $d\pi_e \cdot \operatorname{Ad}(k) = \operatorname{Ad}\big(\pi(k)\big) \cdot d\pi_2$, we have $\varphi\big(O_{K_1}(C)\big) = O_{\pi(K_1)}\big(\varphi(C)\big)$. Since [13, p. 110] $\pi(e^{k_1}) = e^{d\pi_e k_1} = e^{\varphi(k_1)}$ where $k_1 \in K_1$, $\mathfrak{k}_2$ has $\pi(K_1) \subset G_2$ as an analytic subgroup which is K_2 [13, p. 112]. So $\varphi\big(O_{K_1}(C)\big) = O_{K_2}\big(\varphi(C)\big)$. The rest follows from a similar argument as in the proof of Proposition 2.2. ■

Thus we will fix a Cartan decomposition of $\mathfrak{g}$ when we study $W_C(A_1, \dots, A_p)$.

The classical real simple Lie algebras are isomorphic to one of the real forms $\mathfrak{h} \subset \mathfrak{g}$ and $\mathfrak{g}^{\mathbb{R}}$ (the realification of $\mathfrak{g}$) in [23, p. 233]. We will use the special isomorphisms between the classical real Lie algebras of different series [13, pp. 519–520], [23, p. 235].

Since the Cartan decomposition for a compact real form $\mathfrak{h}$ is trivial, *i.e.*, $\mathfrak{k} = \mathfrak{h}$ and $\mathfrak{p} = 0$, the corresponding numerical range is trivial, *i.e.*, $\{0\}$. For any classical complex simple Lie algebra $\mathfrak{g}$, if $\mathfrak{h}$ is a compact real form of $\mathfrak{g}$, then $\mathfrak{g}^{\mathbb{R}} = \mathfrak{h} + i\mathfrak{h}$ is a Cartan decomposition. The corresponding numerical range is always convex by Theorem 2.1.

The Killing forms of the classical complex simple Lie algebras are well known [13, pp. 186–190] and that of $\mathfrak{g}^{\mathbb{R}}$ is given by $B_{\mathfrak{g}^{\mathbb{R}}}(X, Y) = 2 \operatorname{Re} B_{\mathfrak{g}}(X, Y)$ for all $X, Y \in \mathfrak{g}$, and for the other real forms $\mathfrak{h}$, $B_{\mathfrak{h}}(X, Y) = B_{\mathfrak{g}}(X, Y)$ for all $X, Y \in \mathfrak{h}$ [13, p. 180].

As mentioned in Section 1, we will consider the convexity problem of $W_C(A_1, \dots, A_p)$ associated with noncompact classical simple Lie algebras.

3 Convexity and Inclusion Relation

Using the idea in [24] and [3] (see also [31]), we can prove the following result relating the convexity and inclusion relations for the generalized numerical ranges corresponding to a group G defined in (2).

Proposition 3.1 *The C-numerical range $W_C(A_1, \dots, A_p)$ defined in (2) is convex if and only if $W_D(A_1, \dots, A_p) \subset W_C(A_1, \dots, A_p)$ for all $D \in \operatorname{conv} G(C)$.*

Proof By the discussion after the definition of $W_C(A)$, where $A = (A_1, \dots, A_p)$, we see that $W_C(A)$ is the image of $G(C)$ under the linear map $\phi\colon \mathbf{V} \to \mathbb{R}^p$ defined by $\phi(Z) = \big((A_1, Z), \dots, (A_p, Z)\big)$. Thus, we have $\phi\big(G(C)\big) \subset \operatorname{conv}\Big(\phi\big(G(C)\big)\Big) = \phi\Big(\operatorname{conv}\big(G(C)\big)\Big)$. Consequently, $\phi\big(G(C)\big)$ is convex if and only if $\phi\Big(\operatorname{conv}\big(G(C)\big)\Big) \subset \phi\big(G(C)\big)$, *i.e.*, $W_D(A) = \phi\big(G(D)\big) \subseteq \phi\big(G(C)\big) = W_C(A)$ for any $D \in \operatorname{conv} G(C)$. ■

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

For $W_C(A_1,\dots,A_p)$ associated with a real semisimple Lie algebra $\mathfrak{g}$ with the maximal abelian subalgebra $\mathfrak{a}$, we can further the result. It is known that $O(C)\cap\mathfrak{a}_+\neq\phi$ where $\mathfrak{a}_+$ is a (closed) fundamental Weyl chamber of the maximal abelian subalgebra $\mathfrak{a}$ in $\mathfrak{p}$. So we can assume that C and one of A_i's are in $\mathfrak{a}_+$ since the Killing form is G-invariant.

The famous Kostant's convexity theorem [16] asserts that the orthogonal projection of the orbit $O(C)$ onto $\mathfrak{a}$ is the convex hull of the orbit of $C'\in O(C)\cap\mathfrak{a}$ under the action of the Weyl group W of the pair $(\mathfrak{g},\mathfrak{a})$. The orthogonal projection $\pi\colon\mathfrak{p}\to\mathfrak{a}$ can be thought as $(\pi_1,\dots,\pi_m)$ (m is the dimension of $\mathfrak{a}$) where π's are the components of π. Now $W_C(A_1,\dots,A_p)$ can be viewed as the collections of p-tuples of functional values of p arbitrary real linear functionals of $\mathfrak{p}$ (represented by $A_1,\dots,A_p$) acting on the orbit $O(C)$. Using the Kostant's convexity theorem and Proposition 3.1 we can deduce the following corollary (also see [33]).

Corollary 3.2 ([**31**]) *Let $X_1,\dots,X_p$ be elements in $\mathfrak{p}$ and let $Y\in\mathfrak{a}_+$. Then $W_Y(X_1,\dots,X_p)$ is convex if and only if $W_Z(X_1,\dots,X_p)\subset W_Y(X_1,\dots,X_p)$ whenever $Z\in\operatorname{conv}W(Y)$ and $Z\in\mathfrak{a}_+$.*

Corollary 3.2 is very useful for establishing convexity or nonconvexity of numerical range via inclusion relation. We will demonstrate this repeatedly in the forthcoming sections.

Next, we consider some more concepts and lemmas that are useful in studying the inclusion relations $W_D(A_1,\dots,A_p)\subset W_C(A_1,\dots,A_p)$ for $D\in\operatorname{conv}W(C)$. As we will see in later sections, the lemmas help us to reduce the proofs of the inclusion relations to low dimensions, *e.g.*, $n=2$ or 3.

Let $x,y\in\mathbb{R}^n$. We say that x is *weakly majorized* by y, denoted by $x\prec_w y$ if the sum of the k largest entries of x is not larger than that of y for $y=1,\dots,n$. If in addition that the sum of the entries of x is the same as that of y, we say that x is *majorized* by y, denoted by $x\prec y$. The relation $Z\in\operatorname{conv}W(Y)$ is related to either $\prec$ [for $\mathfrak{sl}_n(\mathbb{F})$] or $\prec_w$ (for others classical simple Lie algebras, except the cases $\mathfrak{so}_{n,n}$ and $\mathfrak{so}(2n)$ which are more difficult to deal with. In the latter cases, we need the Thompson's partial ordering $x\ll y$ requiring that x lying in the convex hull of the set $\{Py : P$ is a diagonal special orthogonal matrix$\}$, see [35] and [27] for details). A pinching matrix P is an $n\times n$ matrix such that for some $1\le i<j\le n$,

$$P[i,j\mid i,j]=\begin{pmatrix}\alpha & 1-\alpha\\ 1-\alpha & \alpha\end{pmatrix},$$

where $0\le\alpha\le 1$, and the complementary submatrix $P(i,j\mid i,j)=I_{n-2}$.

Lemma 3.3 ([**7**]) *Let $x,y\in\mathbb{R}^n$. Then $y\prec_w x$ if and only if $y\le P_1\cdots P_k x$ for some pinching matrices $P_1,\dots,P_k$. Hence, if $x,y\in\mathbb{R}^n_+$, then $y\prec_w x$ if and only if $y=\Gamma P_1\cdots P_k x$ for some pinching matrices $P_1,\dots,P_k$ and $\Gamma=\operatorname{diag}(\gamma_1,\dots,\gamma_n)$ with $0\le\gamma_i\le 1$, $i=1,\dots,n$.*

The following lemma is related to Question 2 of [30].

Lemma 3.4 *Suppose $b\ll c$ be such that $b_1\ge\cdots\ge b_{n-1}\ge|b_n|$ and $c_1\ge\cdots\ge c_{n-1}\ge|c_n|$, where $n\ge 4$. Then there exists a sequence of vectors $b=v_{n-2}\ll v_{n-3}\ll\cdots\ll v_1\ll v_0=c$ in $\mathbb{R}^n$ so that for $i=1,\dots,n-3$,*

1. *v_i and v_{i+1} differ in at most 2 entries, and*
2. *one can remove $n-3$ common entries from both v_i and v_{i+1} to obtain $\tilde v_i, \tilde v_{i+1} \in \mathbf{R}^3$ so that $\tilde v_{i+1} \ll \tilde v_i$.*

Proof One may assume that $c_1 \geq \cdots \geq c_n \geq 0$. Otherwise, apply the arguments to the vectors $(c_1, \ldots, c_{n-1}, -c_n)$ and $(b_1, \ldots, b_{n-1}, -b_n)$, and change the signs of the entries with the smallest magnitude in v_i's in the final step.

Our assertion follows from a careful study of the proof of Lemma 6 in [35]. Using the proof of Thompson, one can construct a sequence of vectors so that $v_0 = c$, and for $i > 1$,

(a) v_i is generated from v_{i-1} with by changing at most 2 entries such that condition 1 holds, and
(b) v_i has $b_1, \ldots, b_i$ as entries.

For our purpose, we can stop after getting v_{n-3}, and set $v_{n-2} = b$. We need to prove that the vectors also satisfy condition 2. To this end, let us take a close look at the construction from v_0 to v_1 using the idea in Lemma 6 of [35]. In Thompson's proof, one has to change c_i and c_{i+1} to b_1 and t for a suitable construction of t, where i is the smallest integer satisfying $c_i \geq b_i \geq c_{i+1}$. To prove condition 2, we consider 2 cases. If $i = 1$, then we keep the entries c_1, c_2, c_3 in v_1, and keep the entries b_1, t, c_3 in v_2 so that $(b_1, t, c_3) \ll (c_1, c_2, c_3)$ by the construction. If $i > 1$, we keep the entries c_1, c_i, c_{i+1} of v_1 and c_1, b_1, t of v_2 so that $(c_1, b_1, t) \ll (c_1, c_i, c_{i+1})$ by the construction.

To prove condition 2 holds for $i = 1$, we can focus on the $n-1$ entries v_1 excluding b_1, and the entries $b_2, \ldots, b_n$, and proceed to construct v_2. Inductively, we get the desired conclusion. ■

The following geometrical result is clear (see *e.g.* [25], [31]).

Lemma 3.5 *Let A be an $m \times n$ real matrix and let k be the rank of A. Let S^{n-1} be the unit sphere in $\mathbb{R}^n$.*

1. *If $k < n$, then $A(S^{n-1})$ is a $(k-1)$-ellipsoid with the interior.*
2. *If $k = n$ $(\leq m)$, then $A(S^{n-1})$ is an $(n-1)$-ellipsoid.*

4 The $\mathfrak{sl}_n(\mathbb{F})$ Case

The Cartan decomposition of $\mathfrak{sl}_n(\mathbb{F})$ is $\mathfrak{sl}_n(\mathbb{F}) = \mathfrak{k} + \mathfrak{p}$ where $\mathfrak{p}$ is the space of traceless (trace zero) real symmetric, Hermitian and quaternion Hermitian matrices, where $\mathbb{F} = \mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ respectively. The group K is $\mathrm{SU}_n(\mathbb{F})$. Let $C \in \mathfrak{p}$. The C-numerical range of $A_1, \ldots, A_p \in \mathfrak{p}$, associated with $\mathfrak{sl}_n(\mathbb{F})$ (after a translation and disregarding the constant $4n$ when $\mathbb{F} = \mathbb{C}$; $2n$ when $\mathbb{F} = \mathbb{R}$ or $\mathbb{H}$) is

$$W_C^{\mathbb{F}}(A_1, \ldots, A_p) = \{(\operatorname{tr} CU^*A_1U, \ldots, \operatorname{tr} CU^*A_pU) : U \in \mathrm{SU}_n(\mathbb{F})\},$$

where $C, A_1, \ldots, A_p$ are real symmetric, Hermitian, and quaternion Hermitian matrices when $\mathbb{F} = \mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$ respectively. This is the c-numerical range of $(A_1, \ldots, A_p)$ when $C = \operatorname{diag}(c_1, \ldots, c_n)$ and c's are real. It is a well-studied object and we summarize the result in the following (see [3], [2], [8], [25], [41] for details).

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

Theorem 4.1 *Let $\mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{H}$. Suppose C, A_1, A_2, A_3 are $n \times n$ matrices over $\mathbb{F}$ such that $C = C^*$, $A_i = A_i^*$, $i = 1, 2, 3$.*

1. *Unless $\mathbb{F} = \mathbb{R}$ and $n = 2$, $W_C^{\mathbb{F}}(A_1, A_2)$ is convex. When $n = 2$, $W_C^{\mathbb{R}}(A_1, \dots, A_p)$ is an ellipse satisfying* $\operatorname{conv} W_C^{\mathbb{R}}(A_1, A_2) = W_C^{\mathbb{C}}(A_1, A_2)$.
2. *If $n > 2$ and $\mathbb{F} \neq \mathbb{R}$, then $W_C^{\mathbb{F}}(A_1, A_2, A_3)$ is convex. When $n = 2$, $W_C^{\mathbb{C}}(A_1, A_2, A_3)$ is an ellipsoid in $\mathbb{R}^3$.*

The above results are best possible in the sense that $W_C^{\mathbb{F}}(A_1, \dots, A_p)$ fails to be convex if

(i) $p > 3$ or $(n, p) = (2, 3)$ when $\mathbb{F} = \mathbb{C}$ or $\mathbb{H}$ [1], [25]; or
(ii) $p > 2$ or $(n, p) = (2, 2)$ when $\mathbb{F} = \mathbb{R}$. One may see [25] for a unified treatment of the above three numerical ranges and related results.

Often times $\mathfrak{sl}_n(\mathbb{H})$ is identified with $\mathfrak{su}^*(2n)$ via the standard isomorphism $\mathbb{H}^n \to \mathbb{C}^{2n}$ [15, pp. 26–27]. There $K = \mathrm{Sp}(n)$ and

$$\mathfrak{p} = \left\{ \begin{pmatrix} X & Y \\ -\overline{Y} & \overline{X} \end{pmatrix} : X^* = X, \operatorname{tr} X = 0, Y^T = -Y \right\}.$$

Then the C-numerical range of $A_1, \dots, A_p \in \mathfrak{p}$ will be written in the form:

$$W_C(A_1, \dots, A_p) = \{(\operatorname{tr} CW^*A_1W, \dots, \operatorname{tr} CW^*A_pW) : W \in \mathrm{Sp}(n)\}.$$

5 The $\mathfrak{su}_{p,q}$ Case

It is known that

$$K = \mathrm{SU}(p, q) = \left\{ \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} : U \in U(p), V \in U(q), \det U \det V = 1 \right\},$$

$$\mathfrak{p} = \left\{ \begin{pmatrix} 0 & Y \\ Y^* & 0 \end{pmatrix} : Y \in \mathbb{C}^{p \times q} \right\}, \quad \mathfrak{a} = \bigoplus_{1 \le j \le p} \mathbb{R}(E_{j,p+j} + E_{p+j,j}).$$

The range associated with $\mathfrak{su}_{p,q}$ (after disregarding a suitable constant) is

$$W_C(A_1, \dots, A_m) = \{(\operatorname{Re} \operatorname{tr} C^*UA_1V, \dots, \operatorname{Re} \operatorname{tr} C^*UA_mV) : U \in U(p), V \in U(q)\},$$

where $C, A_1, \dots, A_m$ are given $p \times q$ complex matrices and is symmetric about the origin.

Proposition 5.1 *Let C, A_1, A_2, A_3 be $p \times q$ complex matrices and suppose* $\min\{p, q\} \ge 2$. *Then $W_B(A_1, A_2, A_3) \subset W_C(A_1, A_2, A_3)$ if $b \prec c$ where b and c denote the vectors of singular values of B and C respectively.*

Proof We may assume that $p \le q$. It is sufficient to consider the case [2] that $(b_1, b_2) \prec (c_1, c_2)$ and $b_i = c_i$, $i = 3, \dots, p$. In order to avoid trivial case, we assume $|c_1 - c_2| >$

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

$|b_1 - b_2|$. Let $(r_1, r_2, r_3) = (\mathrm{Re}\sum_{i=1}^p b_i y_i^* A_1 x_i, \mathrm{Re}\sum_{i=1}^p b_i y_i^* A_2 x_i, \mathrm{Re}\sum_{i=1}^p b_i y_i^* A_3 x_i) \in W_B(A_1, A_2, A_3)$. For any $\theta, \phi \in [0, 2\pi]$, define

$$u_1 = e^{-i\phi}\cos\theta x_1 + e^{i\phi}\sin\theta x_2, \quad v_1 = e^{-i\phi}\cos\theta y_1 + e^{i\phi}\sin\theta y_2,$$
$$u_2 = -e^{-i\phi}\sin\theta x_1 + e^{i\phi}\cos\theta x_2, \quad v_2 = -e^{-i\phi}\sin\theta y_1 + e^{i\phi}\cos\theta y_2,$$

and $u_i = x_i$, $i = 3, \ldots, q$ and $v_i = y_i$, $i = 3, \ldots, p$. Since $c_1 + c_2 = b_1 + b_2$,

$$\mathrm{Re}\sum_{i=1}^p b_i v_i^* A_j u_i = \frac{1}{2}(b_1 - b_2)[p_j\cos 2\theta + \sin 2\theta(q_j\cos 2\phi + s_j\sin 2\phi)]$$
$$+ \frac{1}{2}(c_1 + c_2)\,\mathrm{Re}(y_1^* A_j x_1 + y_2^* A_j x_2) + \mathrm{Re}\sum_{i=3}^p c_i y_i^* A_j x_i,$$

where for $i = 1, 2, 3$,

$$p_j = \mathrm{Re}(y_1^* A_j x_1 - y_2^* A_j x_2), \quad q_j = \mathrm{Re}(y_2^* A_j x_1 + y_1^* A_j x_2), \quad s_j = \mathrm{Im}(y_2^* A_j x_1 - y_1^* A_j x_2).$$

As θ and ϕ vary from 0 to 2π, we have an ellipsoid $E_{x,y,b}$ centered at 0. Now $(r_1, r_2, r_3) \in E_{x,y,b} \subset \mathrm{conv}\, E_{x,y,c}$ since $c_1 + c_2 = b_1 + b_2$ and $|c_1 - c_2| > |b_1 - b_2|$. If $E_{x,y,c}$ is degenerated, we have $(r_1, r_2, r_3) \in E_{x,y,c} \subset W_C(A_1, A_2, A_3)$. So we assume that it is not degenerated. For any 2×2 complex matrix A, there exist $U, V \in U(2)$ such that $UAV = \mathrm{diag}(is_1, is_2)$ where s_1 and s_2 are singular values of A. This implies that we can find orthonormal y_1', y_2' in the span of y_1 and y_2 and orthonormal x_1', x_2' in the span of x_1 and x_2 such that the corresponding $p_1' = q_1' = s_1' = 0$. Set $x_i' = x_i$, $i = 3, \ldots, q$, $y_i' = y_i$, $i = 3, \ldots, p$. In other words, the ellipsoid $E_{x',y',c}$ is degenerated.

Now, consider a continuous map $t \mapsto \big(x(t), y(t)\big)$ with $t \in [0, 1]$, where $x(t) = \big(x_1(t), x_2(t)\big)$ (resp., $y(t) = \big(y_1(t), y_2(t)\big)$) is an orthonormal pair of vectors in the span of $\{x_1, x_2\}$ (resp. $\{y_1, y_2\}$), so that $x(0) = (x_1, x_2)$, $x(1) = (x_1', x_2')$, $y(0) = (y_1, y_2)$ and $y(1) = (y_1', y_2')$. Then $E_{x(t),y(t),c}$ will change continuously from $E_{x,y,c}$ to $E_{x',y',c}$. Thus, (r_1, r_2, r_3) will be included in one of the $E_{x(t),y(t),c}$. ■

We remark that the continuity argument in the above proof has been used in [2] and [25], and will be used repeatedly in the next few sections.

Proposition 5.2 *Let $C, A_1, \ldots, A_m$ be $p \times q$ complex matrices where* $\min\{p, q\} = 1$. *Let* $r = \max\{p, q\}$ *and let* $k = \mathrm{rank}\, A$ *where A is the $m \times 2r$ real matrix*

$$A = \begin{pmatrix} \mathrm{Re}\, a_{11} & -\mathrm{Im}\, a_{11} & \cdots & \mathrm{Re}\, a_{1r} & -\mathrm{Im}\, a_{1r} \\ \mathrm{Re}\, a_{21} & -\mathrm{Im}\, a_{21} & \cdots & \mathrm{Re}\, a_{2r} & -\mathrm{Im}\, a_{2r} \\ & \cdots & \cdots & \cdots & \cdots \\ \mathrm{Re}\, a_{m1} & -\mathrm{Im}\, a_{m1} & \cdots & \mathrm{Re}\, a_{mr} & -\mathrm{Im}\, a_{mr} \end{pmatrix}$$

and

$$A_j = \begin{cases} (a_{j1} \cdots a_{jq}) & \text{if } p = 1 \\ (a_{j1} \cdots a_{jp})^T & \text{if } q = 1, \end{cases} \qquad j = 1, \ldots, m.$$

The numerical range $W_C(A_1, \ldots, A_m)$ is

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

1. *an* $(k-1)$*-ellipsoid with the interior embedding in* $\mathbb{R}^m$ *when* $k < 2r$ *and hence convex;*
2. *an* $(2r-1)$*-ellipsoid embedding in* $\mathbb{R}^m$ *when* $k = 2r$.

Proof Assume $p = 1$ for definiteness. We may further assume that $C = (c\,0\cdots 0)$ where $c \geq 0$. Let $A_j = (a_{j1}\cdots a_{jq})$. Then

$$W_C(A_1,\ldots,A_m) = \{\big(\operatorname{Re} c(a_{11}\cdots a_{1q})u,\ldots,\operatorname{Re} c(a_{m1}\cdots a_{mq})u\big) : u \in \mathbb{C}^q, u^*u = 1\},$$

which is the image of the sphere cS^{2q-1} under the map A. By Lemma 3.5, we are done. ■

Corollary 5.3 *Let* C, A_1, A_2, A_3 *be* $1 \times q$ *complex matrices, where* $q = 2, 3, 4$. *Then the numerical range*

$$\begin{aligned} &W_C(A_1,A_2,A_3)\\ &\quad = \{(\operatorname{Re}\operatorname{tr} C^*UA_1V, \operatorname{Re}\operatorname{tr} C^*UA_2V, \operatorname{Re}\operatorname{tr} C^*UA_3V) : U \in U(1), V \in U(q)\}\end{aligned}$$

is an ellipsoid with interior in $\mathbb{R}^3$.

Proof By Proposition 5.2 and the fact that $k \leq m = 3 < 4 \leq 2r$. ■

Theorem 5.4 *Let* C, A_1, A_2, A_3, A_4 *be* $p \times q$ *complex matrices such that* $p \neq q$. *Then* $W_C(A_1, A_2, A_3)$ *is convex. Moreover,* $W_C(A_1, A_2, A_3, A_4)$ *is not convex in general.*

Proof By Proposition 5.2, it suffices to consider the case $\min\{p, q\} \geq 2$. Assume $p < q$ for definiteness. By Proposition 5.1, it remains to show that $W_B(A_1, A_2, A_3) \subset W_C(A_1, A_2, A_3)$ if $0 \leq b_1 < c_1$ and $b_i = c_i$, $i = 2, \ldots, p$. Let

$$(r_1, r_2, r_3) = \Big(\operatorname{Re}\sum_{i=1}^{p} b_i y_i^* A_1 x_i, \operatorname{Re}\sum_{i=1}^{p} b_i y_i^* A_2 x_i, \operatorname{Re}\sum_{i=1}^{p} b_i y_i^* A_3 x_i\Big) \in W_B(A_1, A_2, A_3).$$

For $U \in U(2)$, set $(u_1\ u_q) = U(x_1\ x_q)$ and $u_i = x_i$, $i = 2, \ldots, q-1$, *i.e.*, u_1 and u_q are orthonormal pair from the span of x_1 and x_q. Now $\operatorname{Re}\sum_{i=1}^p b_i y_i^* A_j u_i = \operatorname{Re} b_1 y_1^* A_j u_1 + \operatorname{Re}\sum_{i=2}^p c_i y_i^* A_j x_i$. Then the locus of the above point in $\mathbb{R}^3$ is an ellipsoid E_b with the interior when U varies over $U(2)$ by Corollary 5.3. Clearly $(r_1, r_2, r_3) \in E_b \subset E_c \subset W_C(A_1, A_2, A_3)$ since $b_1 < c_1$.

Assume $p < q$, $B = [\hat{B} \mid 0]$ where $\hat{B} = I_{p-2} \oplus 3I_2$ and $C = [\hat{C} \mid 0]$ where $\hat{C} = I_{p-2} \oplus \operatorname{diag}(4, 2)$. Let $A_i = [\hat{A}_i \mid 0]$ for $i = 1, 2, 3, 4$, such that

$$\hat{A}_1 = I_p, \quad \hat{A}_2 = I_{p-2} \oplus \operatorname{diag}(1, -1), \quad \hat{A}_3 = I_{p-2} \oplus \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{A}_4 = I_{p-2} \oplus \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}.$$

Then $(p+4, p-2, p-2, p-2) \in W_B(A_1, A_2, A_3, A_4) \setminus W_C(A_1, A_2, A_3, A_4)$ because of the following reason. If $\operatorname{tr} B^*U^*A_1V = \operatorname{tr} C^*U^*A_1V = p+4$, then by extremal properties the first $(p-2)$ columns of U (resp. V) must be the left (resp. right) singular vectors of A_1 corresponding to the singular values 1, and the $p-1$ and p-th columns of U (respectively, of V) must be the singular vectors of A_1 corresponding to the singular values 3. Thus U is of the form $U_1 \oplus U_2 \in U(p)$, where $U_2 \in U(2)$, and V is of the form $U_1 \oplus U_2 \oplus V_3 \in U(q)$. However $(\operatorname{Re}\operatorname{tr} C^*U^*A_2V, \operatorname{Re}\operatorname{tr} C^*U^*A_3V, \operatorname{Re}\operatorname{tr} C^*U^*A_4V)$ cannot be $(p-2, p-2, p-2)$. Thus the inclusion relation fails, and hence $W_C(A_1, \ldots, A_4)$ is not convex. ■

Theorem 5.5 *Let C, A_1, A_2, A_3 be $n \times n$ complex matrices. Then $W_C(A_1, A_2)$ is convex if $n > 1$. It is an ellipse if $n = 1$. Moreover, $W_C(A_1, A_2, A_3)$ is not convex in general.*

Proof Suppose $n > 1$. Then $W_C(A_1, A_2)$ is equal to the set

$$\Big\{\Big(\mathrm{Re}\sum_{i=1}^{n} c_i y_i^* A_1 x_i, \mathrm{Re}\sum_{i=1}^{n} c_i y_i^* A_2 x_i\Big) : (x_1 \cdots x_n), (y_1 \cdots y_n) \in U(p)\Big\}.$$

By Corollary 3.2 and Lemma 3.3, it suffices to prove $W_B(A_1, A_2) \subset W_C(A_1, A_2)$ when

Case 1 $0 \le b_1 < c_1$, and $b_i = c_i$, $i = 1, \ldots, n$.

Let $(r_1, r_2) = (\mathrm{Re}\sum_{i=1}^n b_i y_i^* A_1 x_i, \mathrm{Re}\sum_{i=1}^n b_i y_i^* A_2 x_i) \in W_B(A_1, A_2)$. For any $\theta \in [0, 2\pi]$, we consider $x_1' = e^{i\theta} x_1$ and $x_i' = x_i$, $i = 1, \ldots, n$. Then for $j = 1, 2$, we have

$$\mathrm{Re}\sum_{i=1}^{n} b_i y_i^* A_j x_i' = b_1(\cos\theta\,\mathrm{Re}\, y_1^* A_j x_1 - \sin\theta\,\mathrm{Im}\, y_1^* A_j x_1) + \mathrm{Re}\sum_{i=2}^{n} b_i y_i^* A_j x_i.$$

As θ varies in $[0, 2\pi]$, the locus of the point $(\mathrm{Re}\sum_{i=1}^n b_i y_i^* A_1 x_i', \mathrm{Re}\sum_{i=1}^n b_i y_i^* A_2 x_i')$ traces out an ellipse $E_{X,b}$, where X denotes the unitary matrix $(x_1 \cdots x_n)$. Similarly we have $E_{X,c}$ and obviously $E_{X,b} \subset \mathrm{conv}\, E_{X,c}$ $(0 \le b_1 < c_1)$. If $E_{X,c}$ is degenerated, then $(r_1, r_2) \in \mathrm{conv}\, E_{X,c} = E_{X,c}$. So we assume that $E_{X,c}$ is not degenerated. Let $u_1 \in \mathbb{C}^n$ be a unit vector such that $y_1^* A_1 u_1 = 0$. Extend u_1 to an orthonormal basis $\{u_1, \ldots, u_n\}$ of $\mathbb{C}^n$. Evidently $E_{U,c}$ is a line segment or a point. Let H_U and H_X be the skew Hermitian matrices such that $\exp(H_U) = U$ and $\exp(H_X) = X$ respectively. Now consider the curve $f\colon [0, 1] \to U(n)$ defined by $f(t) = \exp\big(tH_U + (1 - t)H_X\big)$. So $E_{X,c} = E_{f(1),c}$ and $E_{U,c} = E_{f(0),c}$. Now $(r_1, r_2) \in E_{X,b} \subset \mathrm{conv}\, E_{X,c}$. By continuity, there is $0 \le t < 1$ such that $(r_1, r_2) \in E_{f(t),c} \subset W_C(A_1, A_2)$.

Case 2 $b \prec c$. It follows from Proposition 5.1 by setting $A_3 = 0$.

When $n = 1$, the image of the unit sphere in $\mathbb{R}^2$ (the unit circle) is clearly an ellipse. This is just a special case of the second part of Proposition 5.2. However, $W_C(A_1, A_2, A_3)$ is an ellipsoid in $\mathbb{R}^3$ by Proposition 5.2 and hence not convex in general.

Let $B = I_{n-1} \oplus (1/3)$, $C = I_{n-1} \oplus (1/2)$, $A_1 = I_{n-1} \oplus (0)$, $A_2 = I_{n-1} \oplus (i)$, $A_3 = I_n$. Then we claim that $W_B(A_1, A_2, A_3)$ is not a subset of $W_C(A_1, A_2, A_3)$ and hence $W_C(A_1, A_2, A_3)$ is not convex. Now $(n - 1, n - 1, n - 1 + 1/3) = (\mathrm{Re}\,\mathrm{tr}\, BA_1, \mathrm{Re}\,\mathrm{tr}\, BA_2, \mathrm{Re}\,\mathrm{tr}\, BA_3) \in W_B(A_1, A_2, A_3)$ and we claim that this point does not belong to $W_C(A_1, A_2, A_3)$. Suppose $(n-1, n-1, x) \in W_C(A_1, A_2, A_3)$. Then $\mathrm{Re}\,\mathrm{tr}\, CU^* A_1 V = n-1$ for some unitary U, V, and the sum of the first $n - 1$ diagonal entries of $U^* A_1 V$ is $n - 1$, which is the sum of the $n - 1$ singular values of the matrix $U^* A_1 V$. It follows from Corollary 3.2 in [17] that $U^* A_1 V = A_1$. Thus the first $n - 1$ columns of U are identical to those of V and the last columns of U and V are scalar multiple to each other, *i.e.*, $u_n = e^{i\theta} v_n$. Now $\mathrm{Re}\,\mathrm{tr}\, CU^* A_2 V = n - 1$. So $e^{i\theta} = \pm 1$. Hence $\mathrm{Re}\,\mathrm{tr}\, CU^* A_3 V$ cannot be $n - 1 + 1/3$. Thus, the inclusion relation fails though $s(B) \prec_w s(C)$, and so $W_C(A_1, A_2, A_3)$ is not convex. ■

Corollary 5.6 *The set $\{\mathrm{tr}\, CUAV : U, V \in U(n)\}$ is a circular disk centered at the origin when $n > 1$ and is a circle when $n = 1$.*

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

Now we consider $C = \operatorname{diag}(1, 0, \dots, 0)$. Convexity will then be established for the corresponding numerical range.

Theorem 5.7 *Let* $C = \operatorname{diag}(1, 0, \dots, 0)$ *and let* $A_1, A_2, A_3 \in \mathbb{C}^{n\times n}$*, where* $n \geq 2$*. Then the numerical range* $W_C(A_1, A_2, A_3)$ *is convex.*

Proof By Corollary 3.2, Lemma 3.3 and Proposition 5.1, it is sufficient to show that $W_B(A_1, A_2, A_3) \subset W_C(A_1, A_2, A_3)$ when $B = \operatorname{diag}(\beta, 0, \dots, 0)$, $0 \leq \beta \leq 1$. Let $r = (r_1, r_2, r_3) \in W_B(A_1, A_2, A_3)$, *i.e.*, $r_j = \beta y^* A_j x$ where $x, y \in \mathbb{C}^n$ are unit vectors. Consider $r' = (r_1', r_2', r_3')$ where $r_j' = \beta y^* A_j u$, $j = 1, 2, 3$. As u runs over the unit sphere of $\mathbb{C}^n$, the locus of r' is then $E_\beta = W_{B'}(A_1', A_2', A_3')$ where $A_j' = y^* A_j \in \mathbb{C}^{1\times n}$ and $B' = (\beta\, 0 \cdots 0) \in \mathbb{C}^{1\times n}$. Hence by Proposition 5.2 ($m = 3$, $r = n$, $k < 4 \leq 2r$), E_β is an ellipsoid with interior centered at the origin and clearly $r \in E_\beta \subset E_1 \subset W_C(A_1, A_2, A_3)$. ∎

Corollary 5.8 *Let* $A_1, A_2 \in \mathbb{C}^{n\times n}$*, and let* $q \in \mathbb{C}$ *satisfy* $|q| \leq 1$*. Then*

$$\{(\operatorname{Re} y^* A_1 x, \operatorname{Re} y^* A_2 x, \operatorname{Re} y^* x) : x, y \in \mathbb{C}^n, x^* x = y^* y = 1\}$$

and

$$\{y^* A_1 x : x, y \in \mathbb{C}^n, \operatorname{Re} y^* x = q\} = \bigcup \{W(q' : A_1) : q' \in \mathbb{C}, \operatorname{Re} q' = q\}.$$

are convex.

6 The $\mathfrak{so}_n(\mathbb{C})$ Case

The range of $A_1, \dots, A_p \in \mathfrak{so}_n$, after disregarding a suitable constant is

$$W_C(A_1, \dots, A_p) = \{(\operatorname{tr} CO^T A_1 O, \dots, \operatorname{tr} CO^T A_p O) : O \in \mathrm{SO}(n)\},$$

which is symmetric about the origin when n is odd but it is not true for the even case.

Theorem 6.1 **([31])** *Let* C, A_1, A_2 *be* $n \times n$ *real skew symmetric matrices. Then the numerical range* $W_C(A_1, A_2) = \{(\operatorname{tr} CO^T A_1 O, \operatorname{tr} CO^T A_2 O) : O \in \mathrm{SO}(n)\}$ *is convex.*

The following result settles Question 1 in [30].

Theorem 6.2 *Let* C, A_1, A_2, A_3, A_4 *be* $n \times n$ *real skew symmetric matrices.*

1. *If* $n \geq 5$*, then* $W_C(A_1, A_2, A_3)$ *is always convex in* $\mathbb{R}^3$*. Moreover,* $W_C(A_1, A_2, A_3, A_4)$ *is not convex in general.*
2. *If* $n = 4$*,* C, A_1, A_2, A_3 *are* 4×4 *real skew symmetric matrices, then* $W_C(A_1, A_2, A_3)$ *is generally not convex.*
3. *If* $n = 3$*, then* $W_C(A_1, A_2, A_3)$ *is an ellipsoid (perhaps degenerated) in* $\mathbb{R}^3$*.*
4. *If* $n = 2$*, then* $W_C(A_1, A_2, A_3)$ *is a point in* $\mathbb{R}^3$*.*

Proof (1) Due to [30], it is sufficent to consider the even case $2n \times 2n$. Given a $2n \times 2n$ real skew-symmetric matrix X with singular values $s_1 = s_1 \geq s_2 = s_2 \geq \cdots \geq s_n = s_n$, let $s(X) = (s_1, \ldots, s_n)$.

Suppose $n = 3$. It is known [13, p. 521] $\mathfrak{su}_4 \cong \mathfrak{so}(6)$. By Proposition 2.2 and Theorem 4.1, $W_C(A_1, A_2, A_3)$ is convex when $n = 3$, and equivalently, $W_B(A_1, A_2, A_3) \subset W_C(A_1, A_2, A_3)$ whenever $b \ll c$ where $b = s(B)$ and $c = s(C)$. Now, suppose $n \geq 4$. We can assume that B and C are in canonical form and $s(B) \ll s(C)$. By Lemma 3.4, we can construct $b = b_{n-1} \ll \cdots \ll b_1 = c$ satisfying the two conditions of the lemma. Let B_j be the real skew symmetric matrices corresponding to b_j, $j = 1, \ldots, n-1$ ($B_1 = C$ and $B_{n-1} = B$) in cannocial form and the 2×2 blocks can be permuted as we please. If $(x_1, x_2, x_3) \in W_{B_j}(A_1, A_2, A_3)$, $j = 2, \ldots, k$, then there exists $O \in \mathrm{SO}(2n)$ such that $x_i = \operatorname{tr} B_j O^T A_i O$, $i = 1, 2, 3$. Let $B_j = P \oplus R$ and $B_{j-1} = Q \oplus R$ where P and Q are 6×6 such that $s(P) \ll s(Q)$. Now let D_i be the leading 6×6 submatrix of $O^T A_i O$, $i = 1, 2, 3$. Thus we can find a $2n \times 2n$ real orthogonal matrix of the form $U = O(U_1 \oplus I_{n-6})$ so that $(x_1, x_2, x_3) = (\operatorname{tr} B_j U^T A_1 U, \operatorname{tr} B_j U^T A_2 U, \operatorname{tr} B_j U^T A_3 U) \in W_{B_{j-1}}(A_1, A_2, A_3)$. So we have the inclusions $W_B(A_1, A_2, A_3) \subset \cdots \subset W_C(A_1, A_2, A_3)$ and hence the convexity.

The result for the odd case is best possible in the sense that if $p > 3$, there are $(2n+1) \times (2n+1)$ ($n \geq 2$) real skew symmetric matrices $C, A_1, \ldots, A_p$ such that $W_C(A_1, \ldots, A_p)$ is not convex [31]. The example in [31] also works for even case.

(2) Notice that [13, p. 240] $\mathfrak{su}_2 \oplus \mathfrak{su}_2 \cong \mathfrak{so}(4)$. This yields that $W_C(A_1, A_2, A_3)$ is generally not convex when $C, A_1, A_2, A_3 \in \mathfrak{so}(4)$. The result follows from Proposition 2.2 and an example in [1] or Theorem 4.1.

(3) The isomorphism $\mathfrak{so}(3) \cong \mathfrak{su}(2)$ explains the common ellipsoid phenomenon for the numerical ranges associated with $\mathfrak{sl}_2(\mathbb{C})^{\mathbb{R}}$ and $\mathfrak{so}_3(\mathbb{C})^{\mathbb{R}}$ when $p = 3$ (see Theorem 4.1).

(4) It is trivial. ∎

Remark 6.3 If $\mathrm{SO}(k)$ is replaced by $O(k)$, denote the corresponding set by $\tilde{W}_C(A_1, \ldots, A_p)$. When $k = 2n+1$, $\tilde{W}_C(A_1, \ldots, A_p) = W_C(A_1, \ldots, A_p)$. However, if $k = 2n$, then

$$\tilde{W}_C(A_1, \ldots, A_p) = W_C(A_1, \ldots, A_p) \cup W_{C'}(A_1, \ldots, A_p),$$

where $C' = DCD$ and $D = \operatorname{diag}(1, \ldots, 1, -1)$. If C is singular, *i.e.*, the rank of C is less than or equal to $2(n-1)$, then $\tilde{W}_C(A_1, \ldots, A_p) = W_C(A_1, \ldots, A_p)$. When $p = 2$, and suppose C is nonsingular, then $\tilde{W}_C(A_1, A_2)$ is the union of two convex sets [31] and is not convex in general. We have the following example: Let

$$X = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad C = A_1 = X \oplus \cdots \oplus X, \quad A_2 = X \oplus \cdots \oplus X \oplus (-X).$$

Then $(-2n, -2n-4)$ and $(-2n+4, -2n) \in \tilde{W}_C(A_1, A_2)$ and the midpoint of the two points is $(-2n+2, -2n+2)$. If it were in $\tilde{W}_C(A_1, A_2)$, then there would exist $U \in O(2n)$ such that $\operatorname{tr} A_1 U^T C U = \operatorname{tr} A_2 U^T C U = -2n+2$. Let $B = U^T C U$. So $\sum_{i=1}^{n} b_{2i-1,2i} = \sum_{i=1}^{n-1} b_{2i-1,2i} - b_{2n-1,2n} = n-1$. Thus $n-1 = \sum_{i=1}^{n-1} b_{2i-1,2i}$ and $b_{2n-1,2n} = 0$. However, we have $n-1 = |\sum_{i=1}^{n-1} b_{2i-1,2i}| \leq \sum_{i=1}^{n-1} |b_{2i-1,2i}| = \sum_{i=1}^{n-1} |b_{2i-1,2i}| - |b_{2n-1,2n}| \leq n-2$ according to a result in [27]. It is a contradiction.

7 The $\mathfrak{sp}_{2n}(\mathbb{C})$ Case

The Cartan decomposition is $\mathfrak{sp}_{2n}(\mathbb{C})^{\mathbb{R}} = \mathfrak{sp}(n) + i\mathfrak{sp}(n)$ where K is the symplectic group

$$\mathrm{Sp}(n) = \begin{pmatrix} U & -\overline{V} \\ V & \overline{U} \end{pmatrix} \in U(2n).$$

The C-numerical range of $A_1, \dots, A_p \in \mathfrak{p}$ will then take the form (after disregarding the constant $-2(n+1)$):

$$\{(\operatorname{tr} CW^*A_1W, \dots, \operatorname{tr} CW^*A_pW) : W \in \mathrm{Sp}(n)\},$$

where $C, A_1, \dots, A_p \in \mathfrak{sp}(n)$. Suppose

$$W = \begin{pmatrix} U & -\overline{V} \\ V & \overline{U} \end{pmatrix} \in \mathrm{Sp}(n), \quad A_j = \begin{pmatrix} A_{j1} & -\overline{A}_{j2} \\ A_{j2} & \overline{A}_{j1} \end{pmatrix} \in \mathfrak{sp}(n), \quad j = 1, \dots, p,$$
$$C = \begin{pmatrix} C_1 & -\overline{C}_2 \\ C_2 & \overline{C}_1 \end{pmatrix},$$

then

$$\begin{aligned} \operatorname{tr} CW^*A_jW = {} & 2 \operatorname{Re} \operatorname{tr} C_1[U^*A_{j1}U - U^*\overline{A}_{j2}V + V^*A_{j2}U + V^*\overline{A}_{j1}V] \\ & - 2 \operatorname{Re} \operatorname{tr} \overline{C}_2[-V^T\overline{A}_{j1}U + V^T\overline{A}_{j2}V + U^TA_{j2}U + U^T\overline{A}_{j1}V]. \end{aligned}$$

If $C \in \mathfrak{sp}(n)$, then there exists $U \in \mathrm{Sp}(n)$ such that $U^*AU = i \operatorname{diag}(c_1, \dots, c_n, -c_1, \dots, -c_n)$, where $c_i \geq 0$, $i = 1, \dots, n$. Denote by c the vector $(c_1, \dots, c_n)$. Hence the j-th component of the numerical range is of the form $2 \operatorname{Re}[\operatorname{tr} CU^*A_{j1}U + \operatorname{tr} CV^*\overline{A}_{j1}V] - 4 \operatorname{Im} \operatorname{tr} CU^*\overline{A}_{j2}V$ (since $A_{j2}^T = A_{j2}$) where $C = i \operatorname{diag}(c_1, \dots, c_n)$, *i.e.*, $-2 \operatorname{Im} \sum_{i=1}^n c_i(u_i^*A_{j1}u_i + v_i^*\overline{A}_{j1}v_i) - 4 \operatorname{Re} \sum_{i=1}^n c_iu_i^*\overline{A}_{j2}v_i$. The numerical range is also symmetric about the origin. Since $\mathrm{Sp}(n)$ is compact connected, by Theorem 2.1 we have

Theorem 7.1 (**[31]**) *Let $C, A_1, A_2 \in \mathfrak{sp}(n)$. Then $W_C(A_1, A_2)$ is convex.*

Proposition 7.2 *Let $C, A_1, A_2, A_3, A_4 \in \mathfrak{sp}(2)$. Then $W_C(A_1, A_2, A_3)$ is convex. Moreover, $W_C(A_1, A_2, A_3, A_4)$ is not convex in general.*

Proof Since $\mathfrak{sp}(2) \cong \mathfrak{so}(5)$, the result follows from Proposition 2.4 and Theorem 6.2 (1). ∎

Proposition 7.3 *Let $C, A_1, A_2, A_3 \in \mathfrak{sp}(n)$ where $n \geq 2$. If $b \prec c$, then $W_B(A_1, A_2, A_3) \subset W_C(A_1, A_2, A_3)$.*

Proof Assume that B and C are in diagonal form, *i.e.*, $C = i\operatorname{diag}(c_1, \dots, c_n, -c_1, \dots, -c_n)$, $c_i \ge 0$. It is sufficient to handle the case that $(b_1, b_2) \prec (c_1, c_2)$ and $c_i = b_i$, $i = 3, \dots, n$. For $j = 1, 2, 3$, let A_j be of the form

$$\begin{pmatrix} A_{j1} & -\overline{A}_{j2} \\ A_{j2} & \overline{A}_{j1} \end{pmatrix} \in \mathfrak{sp}(n).$$

So the elements of $W_C(A_1, A_2, A_3)$ are of the form (x_1, x_2, x_3) where

$$x_j = -2\operatorname{Im}\sum_{i=1}^n c_i(u_i^* A_{j1} u_i + v_i^* \overline{A}_{j1} v_i) - 4\operatorname{Re}\sum_{i=1}^n c_i u_i^* \overline{A}_{j2} v_i,$$

and u's are the columns of U and v's are the columns of V in

$$W = \begin{pmatrix} U & -\overline{V} \\ V & \overline{U} \end{pmatrix} \in \mathrm{Sp}(n).$$

For any $\theta, \phi \in [0, 2\pi]$, define

$$u_1' = e^{-i\phi}\cos\theta u_1 + e^{i\phi}\sin\theta u_2, \quad v_1' = e^{-i\phi}\cos\theta v_1 + e^{i\phi}\sin\theta v_2,$$
$$u_2' = -e^{-i\phi}\sin\theta u_1 + e^{i\phi}\cos\theta u_2, \quad v_2' = -e^{-i\phi}\sin\theta v_1 + e^{i\phi}\cos\theta v_2,$$

and $u_i' = u_i$, $i = 3, \dots, n$ and $v_i' = v_i$, $i = 3, \dots, n$. Since $b_1 + b_2 = c_1 + c_2$, for $j = 1, 2, 3$, we have

$$\begin{aligned} y_j &= -2\operatorname{Im}\sum_{i=1}^n c_i(u_i'^* A_{j1} u_i' + v_i'^* \overline{A}_{j1} v_i') - 4\operatorname{Re}\sum_{i=1}^n c_i u_i'^* \overline{A}_{j2} v_i' \\ &= (c_1 + c_2)[-\operatorname{Im}(u_1^* A_{j1} u_1 + u_2^* A_{j1} u_2 + v_1^* \overline{A}_{j1} v_1 + v_2^* \overline{A}_{j1} v_2) - 2\operatorname{Re}(u_1^* \overline{A}_{j2} v_1 + u_2^* \overline{A}_{j2} v_2)] \\ &\quad + (c_1 - c_2)[p_j \cos 2\theta + \sin 2\theta(q_j \cos 2\phi + s_j \sin 2\phi)] \\ &\quad - 2\operatorname{Im}\sum_{i=3}^n c_i(u_i^* A_{j1} u_i + v_i^* \overline{A}_{j1} v_i) - 4\operatorname{Re}\sum_{i=3}^n c_i u_i^* \overline{A}_{j2} v_i \end{aligned}$$

where

$$p_j = -\operatorname{Im}(u_1^* A_{j1} u_1 - u_2^* A_{j1} u_2 + v_1^* \overline{A}_{j1} v_1 - v_2^* \overline{A}_{j1} v_2) - 2\operatorname{Re}(u_1^* \overline{A}_{j2} v_1 - u_2^* \overline{A}_{j2} v_2),$$
$$q_j = -2\operatorname{Im} u_1^* A_{j1} u_2 - 2\operatorname{Im} v_1^* \overline{A}_{j1} v_2 - 2\operatorname{Re}(u_1^* \overline{A}_{j2} v_2 + u_2^* \overline{A}_{j2} v_1),$$
$$s_j = -2\operatorname{Re} u_1^* A_{j1} u_2 - 2\operatorname{Re} v_1^* \overline{A}_{j1} v_2 - 2\operatorname{Im}(u_2^* \overline{A}_{j2} v_1 - v_1^* \overline{A}_{j2} u_2),$$

$j = 1, 2, 3$. The locus of (y_1, y_2, y_3) is an ellipsoid $E_{c,W}$ when ϕ and θ vary on $[0, 2\pi]$. So for any $x \in W_B(A_1, A_2, A_3)$, there is $W \in \mathrm{Sp}(n)$ and $x \in E_{b,W} \subset \operatorname{conv} E_{c,W}$ since $|b_1 - b_2| \le |c_1 - c_2|$ and $b_1 + b_2 = c_1 + c_2$. We notice that the matrix $R(\theta, \phi) \oplus I_{n-2} \oplus \overline{R}(\theta, \phi) \oplus I_{n-2}$ is an element of $\mathrm{Sp}(n)$ for any θ and ϕ where

$$R(\theta, \phi) = \begin{pmatrix} e^{-i\phi}\cos\theta & e^{i\phi}\sin\theta \\ -e^{-i\phi}\sin\theta & e^{i\phi}\cos\theta \end{pmatrix}.$$

In particular, $R(\theta, \phi) \oplus \overline{R}(\theta, \phi) \in \mathrm{Sp}(2)$. By Proposition 7.2, $\operatorname{conv} E_{c,W} \subset W_C(A_1, A_2, A_3)$ so that $x \in W_C(A_1, A_2, A_3)$. This completes the proof. ■

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

Theorem 7.4 *Let $C, A_1, A_2, A_3, A_4 \in \mathfrak{sp}(n)$. Then $W_C(A_1, A_2, A_3)$ is convex if $n > 1$, and is an ellipsoid (perhaps degenerated) centered at the origin if $n = 1$. In general, $W_C(A_1, A_2, A_3, A_4)$ is not convex.*

Proof First we establish the simplest cases. When $n = 1$, $\mathrm{Sp}(1) = \mathrm{SU}(2)$ and hence by Theorem 4.1, $W_C(A_1, A_2, A_3)$ is an ellipsoid (perhaps degenerate) centered at the origin.

It is sufficient to show that $W_B(A_1, A_2, A_3) \subset W_C(A_1, A_2, A_3)$ when $0 \le b_1 < c_1$ and $c_i = b_i$, $i = 1, \ldots, n$. The elements of $W_B(A_1, A_2, A_3)$ are of the form (x_1, x_2, x_3) where

$$\begin{aligned} x_j &= -2\,\mathrm{Im} \sum_{i=1}^n b_i(u_i^* A_{j1} u_i + v_i^* \overline{A}_{j1} v_i) - 4\,\mathrm{Re} \sum_{i=1}^n b_i u_i^* \overline{A}_{j2} v_i \\ &= -2\,\mathrm{Im}\, b_1(u_1^* A_{j1} u_1 + v_1^* \overline{A}_{j1} v_1) - 4\,\mathrm{Re}\, b_1 u_1^* \overline{A}_{j2} v_1 \\ &\quad - 2\,\mathrm{Im} \sum_{i=2}^n c_i(u_i^* A_{j1} u_i + v_i^* \overline{A}_{j1} v_i) - 4\,\mathrm{Re} \sum_{i=2}^n c_i u_i^* \overline{A}_{j2} v_i, \end{aligned}$$

$j = 1, 2, 3$. Let $(u_1' v_1') = U(u_1 v_1)$ where $U \in \mathrm{Sp}(1)$ and set $u_i' = u_i$, $v_i' = v_i$, $i = 2, \ldots, n$. Similar to the previous treatment, we have an ellipsoid E_{u,v,b_1} as U runs over $\mathrm{Sp}(1)$, by using $n = 1$ case. So we deduce that a point $x \in W_B(A_1, A_2, A_3)$ is contained in $E_{b_1,u,v} \subset \mathrm{conv}\, E_{c_1,u,v}$ since $b_1 < c_1$. Thus $x \in \mathrm{conv}\, E_{c_1,u,v} \subset W_C(A_1, A_2, A_3)$ by Proposition 7.2.

Now we construct nonconvex examples for the more general case. Let

$$\begin{gathered} B = I_{n-2} \oplus 3I_2 \oplus (-I_{n-2}) \oplus (-3I_2), \\ C = I_{n-2} \oplus \mathrm{diag}(4, 2) \oplus (-I_{n-2}) \oplus \mathrm{diag}(-4, -2), \\ A_1 = I_n \oplus (-I_n), \quad A_2 = I_{n-2} \oplus \mathrm{diag}(1, -1) \oplus (-I_{n-2}) \oplus \mathrm{diag}(-1, 1), \\ A_3 = I_{n-2} \oplus \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \oplus (-I_{n-2}) \oplus \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \\ A_4 = I_{n-2} \oplus \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \oplus (-I_{n-2}) \oplus \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}. \end{gathered}$$

We are going to show that

$$\big(2(n-2) + 12, 2(n-2), 2(n-2), 2(n-2)\big) \in W_B(A_1, A_2, A_3, A_4) \setminus W_C(A_1, A_2, A_3, A_4).$$

Consider a set which is larger than $W_C(A_1, A_2, A_3, A_4)$:

$$\begin{aligned} &W_C'(A_1, A_2, A_3, A_4) \\ &\qquad = \{(\mathrm{tr}\, CU^* A_1 U, \mathrm{tr}\, CU^* A_2 U, \mathrm{tr}\, CU^* A_3 U, \mathrm{tr}\, CU^* A_3 U) : U \in U(2n)\}. \end{aligned}$$

Indeed the set is the C-numerical range of (A_1, A_2, A_3, A_4) associated with $\mathfrak{gl}(2n, \mathbb{C})$. Applying the reasoning in the first example of the proof of Theorem 5.4, then $\big(2(n-2) + 12, 2(n-2), 2(n-2), 2(n-2)\big) \notin W_C'(A_1, A_2, A_3, A_4)$. ■

8 The $\mathfrak{sp}_{2n}(\mathbb{R})$ Case

It is known that

$$K = \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix} : A^TA + B^TB = I, A^TB = B^TA, A, B \in \mathbb{R}^{n\times n} \right\},$$

$$\mathfrak{p} = \left\{ \begin{pmatrix} X & Y \\ Y & -X \end{pmatrix} : Y^T = Y, X^T = X, X, Y \in \mathbb{R}^{n\times n} \right\}, \quad \mathfrak{a} = \bigoplus_{1\le j\le n} \mathbb{R}(E_{jj} - E_{n+j,n+j}).$$

Notice that

$$\begin{pmatrix} A & B \\ -B & A \end{pmatrix} \in K$$

if and only if $A + iB \in U(n)$. Hence we identify K with $U(n)$. Similarly we identify $\mathfrak{k}$ with $\mathfrak{u}(n)$. Now we identify $\mathfrak{p}$ with the space of $n \times n$ complex symmetric matrices via the map

$$\begin{pmatrix} X & Y \\ Y & -X \end{pmatrix} \mapsto X + iY, \quad X, Y \in \mathbb{R}^{n\times n}, X^T = X, Y^T = Y.$$

Hence $\mathfrak{a}$ is identified with the space of real diagonal matrices. So the corresponding C-numerical range, after disregarding the constant $2(n + 1)$, takes the form

$$W_C(A_1, \ldots, A_p) = \{(\operatorname{Re}\operatorname{tr} CU^TA_1U, \ldots, \operatorname{Re}\operatorname{tr} CU^TA_pU) : U \in U(n)\}.$$

Clearly the numerical range is symmetric about the origin. We can assume that $C = \operatorname{diag}(c_1, \ldots, c_n)$ where c's are the singular values of C. When $p = 1$, the set $W_C(A)$ is a closed interval [32]. We have the following convexity result when $p = 2$.

Theorem 8.1 *Let C, A_1, A_2, A_3 be $n \times n$ complex symmetric matrices. Then $W_C(A_1, A_2)$ is convex if $n > 1$. It is an ellipse (perhaps degenerated) in $\mathbb{R}^2$ if $n = 1$. Moreover, $W_C(A_1, A_2, A_3)$ is not convex in general.*

Proof The second assertion is trivial since the numerical range is just the image of the unit circle under a linear map from $\mathbb{R}^2$ to $\mathbb{R}^2$. Let $n > 1$. We need to consider the following two cases.

Case 1 $0 \le b_1 < c_1$, and $b_i = c_i$, $i = 2, \ldots, n$.

Let $(r_1, r_2) = (\operatorname{Re}\sum_{i=1}^n b_i x_i^T A_1 x_i, \operatorname{Re}\sum_{i=1}^n b_i x_i^T A_2 x_i) \in W_B(A_1, A_2)$. For any $\theta \in [0, 2\pi]$ we consider $x_1' = e^{i\theta}x_1$ and $x_i' = x_i$, $i = 2, \ldots, n$. Then for $j = 1, 2$, we have

$$\operatorname{Re}\sum_{i=1}^n b_i x_i'^T A_j x_i' = b_1(\cos 2\theta \operatorname{Re} x_1^T A_j x_1 - \sin 2\theta \operatorname{Im} x_1^T A_j x_1) + \operatorname{Re}\sum_{i=2}^n b_i x_i^T A_j x_i.$$

As θ varies on $[0, 2\pi]$, the locus of the point $(\operatorname{Re}\sum_{i=1}^p b_i x_i'^T A_1 x_i', \operatorname{Re}\sum_{i=1}^p b_i x_i^T A_2 x_i')$ traces out an ellipse $E_{X,b}$, where X denotes the unitary matrix $(x_1 \cdots x_n)$. Similarly we have $E_{X,c}$ and obviously $E_{X,b} \subset \operatorname{conv} E_{X,c}$. If $E_{X,c}$ is degenerated, then $(r_1, r_2) \in \operatorname{conv} E_{X,c} = E_{X,c}$. So

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

we assume that $E_{X,c}$ is not degenerated. Let $u_1 \in \mathbb{C}^n$ be unit vector such that $u_1^T A_1 u_1 = 0$ (see Lemma 3 of Thompson [36]). Extend u_1 to an orthonormal basis $\{u_1, \ldots, u_n\}$. Hence the ellipse $E_{U,c}$ is degenerated. Using the continuity argument, we are done.

Case 2 Suppose $(b_1, b_2) \prec (c_1, c_2)$ and $b_i = c_i$, $i = 3, \ldots, n$. Let

$$(r_1, r_2) = \Bigl(\mathrm{Re} \sum_{i=1}^{p} b_i x_i^T A_1 x_i, \mathrm{Re} \sum_{i=1}^{p} b_i x_i^T A_2 x_i\Bigr) \in W_B(A_1, A_2).$$

For any $\theta \in [0, 2\pi]$, define $y_1 = \cos\theta x_1 + \sin\theta x_2$, $y_2 = -\sin\theta x_1 + \cos\theta x_2$, and $y_i = x_i$, $i = 3, \ldots, n$. Then

$$\begin{aligned} \mathrm{Re} \sum_{i=1}^{n} c_i y_i^T A_j y_i &= \frac{1}{2}(c_1 + c_2)\,\mathrm{Re}(x_1^T A_j x_1 + x_2^T A_j x_2) \\ &\qquad + \frac{1}{2}(c_1 - c_2)(p_j \cos 2\theta + q_j \sin 2\theta) + \mathrm{Re} \sum_{i=3}^{n} c_i x_i^T A_j x_i, \end{aligned}$$

where $p_j = \mathrm{Re}(x_1^T A_j x_1 - x_2^T A_j x_2)$ and $q_j = \mathrm{Re}(x_2^T A_j x_1 - x_1^T A_j x_2)$. As θ varies from 0 to 2π, we get an ellipse E_c. Now $(r_1, r_2) \in E_b \subset \mathrm{conv}\, E_c$. The ellipse E_c can also be viewed as the image of a loop in $\mathrm{SU}(2)$ under the above continuous function, namely, the set of rotation matrices. By the simple connectedness of $\mathrm{SU}(2)$, $\mathrm{conv}\, E_c \subset W_C(A_1, A_2)$. Hence $(r_1, r_2) \in W_C(A_1, A_2)$.

The example in the proof of Theorem 5.4 works for this case and the computation is similar. ■

Corollary 8.2 ([37]) *Let C and A be $n \times n$ complex matrices such that $C = C^T$. Then the congruence numerical range $W_C(A) = \{\mathrm{tr}\, CU^T AU : U \in U(n)\}$ is a circular disk if $n > 1$.*

9 The $\mathfrak{sp}_{p,q}$ Case

We may assume that $p \leq q$. It is known that

$$\mathfrak{p} = \left\{ \begin{pmatrix} 0 & X_{12} & 0 & X_{14} \\ X_{12}^* & 0 & X_{14}^T & 0 \\ 0 & \overline{X}_{14} & 0 & -\overline{X}_{12} \\ X_{14}^* & 0 & -X_{12}^T & 0 \end{pmatrix} \right\},$$

$$\mathfrak{a} = \bigoplus_{1 \leq j \leq p} \mathbb{R}(E_{j,p+j} + E_{p+j,j} - E_{p+q+j,2p+q+j} - E_{2p+q+j,p+q+j}),$$

$$K = \left\{ \begin{pmatrix} U_1 & 0 & -\overline{V}_1 & 0 \\ 0 & U_2 & 0 & -\overline{V}_2 \\ V_1 & 0 & \overline{U}_1 & 0 \\ 0 & V_2 & 0 & \overline{U}_2 \end{pmatrix} : \begin{pmatrix} U_1 & -\overline{V}_1 \\ V_1 & \overline{U}_1 \end{pmatrix} \in \mathrm{Sp}(p), \begin{pmatrix} U_2 & -\overline{V}_2 \\ V_2 & \overline{U}_2 \end{pmatrix} \in \mathrm{Sp}(q) \right\}.$$

Given $C \in \mathfrak{sp}_{p,q}$, there exists $W \in \mathrm{Sp}(p,q)$ such that W^*CW is of the form:

$$\begin{pmatrix} 0 & C_1 \\ C_1^T & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & -C_1 \\ -C_1^T & 0 \end{pmatrix},$$

where $C_1 = c_1E_{11} \oplus \cdots \oplus c_pE_{pp}$ with $c_i \geq 0$ for all $i = 1, \ldots, p$. Now the (12)-block of an element of $O(A_j)$ ($A_j \in \mathfrak{p}$) has the form of the (12)-block of the matrix

$$Q = \begin{pmatrix} U_1 & 0 & -\overline{V}_1 & 0 \\ 0 & U_2 & 0 & -\overline{V}_2 \\ V_1 & 0 & \overline{U}_1 & 0 \\ 0 & V_2 & 0 & \overline{U}_2 \end{pmatrix}^* \begin{pmatrix} 0 & A_{12}^j & 0 & A_{14}^j \\ A_{12}^{j\,*} & 0 & A_{14}^{j\,T} & 0 \\ 0 & \overline{A}_{14}^j & 0 & -\overline{A}_{12}^j \\ A_{14}^{j\,*} & 0 & -A_{12}^{j\,T} & 0 \end{pmatrix} \begin{pmatrix} U_1 & 0 & -\overline{V}_1 & 0 \\ 0 & U_2 & 0 & -\overline{V}_2 \\ V_1 & 0 & \overline{U}_1 & 0 \\ 0 & V_2 & 0 & \overline{U}_2 \end{pmatrix},$$

namely, $Q_{12} = U_1^*A_{12}^jU_2 + U_1^*A_{14}^jV_2 + V_1^*\overline{A}_{14}^jU_2 - V_1^*\overline{A}_{12}^jV_2$. Hence the j-th component of the numerical range is $\operatorname{Re}\operatorname{tr} C^TQ_{12} + \operatorname{Re}\operatorname{tr} CQ_{12}^* + \operatorname{Re}\operatorname{tr} CQ_{12}^T + \operatorname{Re}\operatorname{tr} C^T\overline{Q}_{12} = 4\operatorname{Re}\operatorname{tr} C^TQ_{12}$, where $C = c_1E_{11} \oplus \cdots \oplus c_pE_{pp}$. In other words, the j-th component is of the form

$$4\operatorname{Re}\sum_{i=1}^{p} c_i[u_{1i}^*A_{12}^ju_{2i} + u_{1i}^*A_{14}^jv_{2i} + v_{1i}^*\overline{A}_{14}^ju_{2i} - v_{1i}^*\overline{A}_{12}^jv_{2i}],$$

where $U_1 = (u_{11} \cdots u_{1p})$, $V_1 = (v_{11} \cdots v_{1p})$, $U_2 = (u_{21} \cdots u_{2q})$, $V_2 = (v_{21} \cdots v_{2q})$ form an element of K. The numerical range is also symmetric about the origin. By Remark 11.1, we have

Proposition 9.1 *Let $C, A_1, A_2, A_3 \in \mathfrak{sp}_{1,1}$. Then $W_C(A_1, A_2, A_3)$ is an ellipsoid with interior centered at the origin in $\mathbb{R}^3$ and hence is convex.*

Proposition 9.2 *Let $C, A_1, A_2, A_3 \in \mathfrak{sp}_{p,q}$. If $\min\{p,q\} > 1$ and $b \prec c$ $W_B(A_1, A_2, A_3) \subset W_C(A_1, A_2, A_3)$.*

Proof It is sufficient to consider the case $(b_1, b_2) \prec (c_1, c_2)$, $b_i = c_i$, $i = 3, \ldots, p$. Let $(x_1, x_2, x_3) \in W_B(A_1, A_2, A_3)$, *i.e.*, $x_j = 4\operatorname{Re}\sum_{i=1}^p b_i[u_{1i}^*A_{12}^ju_{2i} + u_{1i}^*A_{14}^jv_{2i} + v_{1i}^*\overline{A}_{14}^ju_{2i} - v_{1i}^*\overline{A}_{12}^jv_{2i}]$, $j = 1, 2, 3$. For any $\theta \in [0, 2\pi]$ and $\phi \in [0, 2\pi]$, $k = 1, 2$, define

$$u_{k1}' = e^{-i\phi}\cos\theta u_{k1} + e^{i\phi}\sin\theta u_{k2}, \quad v_{k1}' = e^{-i\phi}\cos\theta v_{k1} + e^{i\phi}\sin\theta v_{k2},$$
$$u_{k2}' = -e^{-i\phi}\sin\theta u_{k1} + e^{i\phi}\cos\theta u_{k2}, \quad v_{k2}' = -e^{-i\phi}\sin\theta v_{k1} + e^{i\phi}\cos\theta v_{k2}.$$

Since $b_1 + b_2 = c_1 + c_2$, for $j = 1, 2, 3$, we have

$$\begin{aligned} y_j &= 4\operatorname{Re}\sum_{i=1}^{p} b_i[u_{1i}'^*A_{12}^ju_{2i}' + u_{1i}'^*A_{14}^jv_{2i}' + v_{1i}'^*\overline{A}_{14}^ju_{2i}' - v_{1i}'^*\overline{A}_{12}^jv_{2i}'] \\ &= 2(c_1 + c_2)\operatorname{Re}[u_{11}^*A_{12}^ju_{21} + u_{12}^*A_{12}^ju_{22} + u_{11}^*A_{14}^jv_{21} + u_{12}^*A_{14}^jv_{22} \\ &\qquad + v_{11}^*\overline{A}_{14}^ju_{21} + v_{12}^*\overline{A}_{14}^ju_{22} - v_{11}^*\overline{A}_{12}^jv_{21} - v_{12}^*\overline{A}_{12}^jv_{22}] \\ &\quad + 2(b_1 - b_2)[p_j\cos 2\theta + (q_j\cos 2\phi + r_j\sin 2\phi)\sin 2\theta] \\ &\quad + 4\operatorname{Re}\sum_{i=3}^{p} c_i[u_{1i}^*A_{12}^ju_{2i} + u_{1i}^*A_{14}^jv_{2i} + v_{1i}^*\overline{A}_{14}^ju_{2i} - v_{1i}^*\overline{A}_{12}^jv_{2i}], \end{aligned}$$

where

$$\begin{aligned} p_j &= 2\,\mathrm{Re}[u_{11}^* A_{12}^j u_{21} - u_{12}^* A_{12}^j u_{22} + u_{11}^* A_{14}^j v_{21} - u_{12}^* A_{14}^j v_{22} \\ &\qquad + v_{11}^* \overline{A}_{14}^j u_{21} - v_{12}^* \overline{A}_{14}^j u_{22} - v_{11}^* \overline{A}_{12}^j v_{21} + v_{12}^* \overline{A}_{12}^j v_{22}] \\ q_j &= 2\,\mathrm{Re}[u_{11}^* A_{12}^j u_{22} + u_{12}^* A_{12}^j u_{21} + u_{11}^* A_{14}^j v_{22} + u_{12}^* A_{14}^j v_{21} \\ &\qquad + v_{11}^* \overline{A}_{14}^j u_{22} + v_{12}^* \overline{A}_{14}^j u_{21} - v_{11}^* \overline{A}_{12}^j v_{22} - v_{12}^* \overline{A}_{12}^j v_{21}] \\ r_j &= 2\,\mathrm{Im}[-u_{11}^* A_{12}^j u_{22} + u_{12}^* A_{12}^j u_{21} - u_{11}^* A_{14}^j v_{22} + u_{12}^* A_{14}^j v_{21} \\ &\qquad - v_{11}^* \overline{A}_{14}^j u_{22} + v_{12}^* \overline{A}_{14}^j u_{21} + v_{11}^* \overline{A}_{12}^j v_{22} - v_{12}^* \overline{A}_{12}^j v_{21}]. \end{aligned}$$

The map which sends u's and v's to u''s and v''s is in $\gamma\big(\mathrm{Sp}(1)\times\mathrm{Sp}(1)\big) \subset K$ where γ denotes the imbedding from $\mathrm{Sp}(p)\times\mathrm{Sp}(q) \to K$ [13, p. 455]. As θ and ϕ vary on $[0, 2\pi]$, the locus of (y_1, y_2, y_3) is an ellipsoid E_b with interior by Proposition 9.1. Since $|b_1 - b_2| \le |c_1 - c_2|$, we have $x \in E_b \subset E_c \subset W_C(A_1, A_2, A_3)$. By a continuity argument, we are done. ■

Theorem 9.3 *Let $C, A_1, A_2, A_3 \in \mathfrak{sp}_{p,q}$. When $\min\{p, q\} > 1$, $W_C(A_1, A_2)$ is convex. Furthermore, $W_C(A_1, A_2, A_3)$ is not convex in general.*

Proof We may assume that $1 < p \le q$. It suffices to show that $W_B(A_1, A_2) \subset W_C(A_1, A_2)$ when $0 \le b_1 < c_1$, $b_i = c_i$, $i = 2, \ldots, p$. Let $(x_1, x_2) \in W_B(A_1, A_2)$, *i.e.*, for $j = 1, 2$, $x_j = 4\,\mathrm{Re}\sum_{i=1}^p b_i[u_{1i}^* A_{12}^j u_{2i} + u_{1i}^* A_{14}^j v_{2i} + v_{1i}^* \overline{A}_{14}^j u_{2i} - v_{1i}^* \overline{A}_{12}^j v_{2i}]$. For any $\theta \in [0, 2\pi]$, let $u'_{11} = e^{i\theta} u_{11}$ and $v'_{11} = e^{i\theta} v_{11}$, $u'_{1i} = u_{1i}$, $v'_{1i} = v_{1i}$, $i = 2, \ldots, p$; $u'_{2i} = u_{2i}$, $v'_{2i} = v_{2i}$, $i = 1, \ldots, q$. Then for $j = 1, 2$,

$$\begin{aligned} y_j &= 4\,\mathrm{Re}\sum_{i=1}^p b_i[u_{1i}'^* A_{12}^j u'_{2i} + {u'_{1i}}^* A_{14}^j v'_{2i} + {v'_{1i}}^* \overline{A}_{14}^j u'_{2i} - {v'_{1i}}^* \overline{A}_{12}^j v'_{2i}] \\ &= 4b_1[p_j \cos\theta + q_j \sin\theta] \\ &\qquad + 4\,\mathrm{Re}\sum_{i=2}^p c_i[u_{1i}^* A_{12}^j u_{2i} + u_{1i}^* A_{14}^j v_{2i} + v_{1i}^* \overline{A}_{14}^j u_{2i} - v_{1i}^* \overline{A}_{12}^j v_{2i}], \end{aligned}$$

where $p_j = \mathrm{Re}[u_{11}^* A_{12}^j u_{21} + u_{11}^* A_{14}^j v_{21} + v_{11}^* \overline{A}_{14}^j u_{21} - v_{11}^* \overline{A}_{12}^j v_{21}]$, $q_j = -\mathrm{Im}[u_{11}^* A_{12}^j u_{21} + u_{11}^* A_{14}^j v_{21} + v_{11}^* \overline{A}_{14}^j u_{21} - v_{11}^* \overline{A}_{12}^j v_{21}]$. The matrix $\mathrm{diag}(e^{i\theta}, e^{-i\theta})$ belongs to $\mathrm{Sp}(1)$ and thus $\gamma\big(\mathrm{diag}(e^{i\theta}, e^{-i\theta}) \oplus I_{p-2}, I_q\big) \in K$. As θ varies on $[0, 2\pi]$, the locus of (y_1, y_2) is an ellipse E_b. Since $0 \le b_1 < c_1$ and $\mathrm{Sp}(1)$ is simply connected, we have $(x_1, x_2) \in E_b \in \mathrm{conv}\, E_c \in W_C(A_1, A_2)$.

The convexity result is best possible. We will work out the $p = q$ case and the $p \ne q$ case is similar. Let $\hat{B} = I_{n-1} \oplus (1/3)$, $\hat{C} = I_{n-1} \oplus (1/2)$, $\hat{A}_1 = I_{n-1} \oplus (0)$, $\hat{A}_2 = I_{n-1} \oplus (i)$, $\hat{A}_3 = I_n$. Set

$$X = \begin{pmatrix} 0 & \hat{X} \\ \hat{X}^* & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & -\overline{\hat{X}} \\ -\hat{X}^T & 0 \end{pmatrix},$$

where $X = B, C, A_i$, $i = 1, 2, 3$. We claim that $W_B(A_1, A_2, A_3) \not\subset W_C(A_1, A_2, A_3)$ and hence $W_C(A_1, A_2, A_3)$ is not convex. Notice that

$$4(n-1, n-1, n-1+1/3) = (\operatorname{Re}\operatorname{tr} BA_1, \operatorname{Re}\operatorname{tr} BA_2, \operatorname{Re}\operatorname{tr} BA_3) \in W_B(A_1, A_2, A_3)$$

and we are going to show that this point does not belong to the set

$$W'_C(A_1, A_2, A_3) = \{\operatorname{tr} CU^*A_1, U, \operatorname{tr} CU^*A_2U, \operatorname{tr} CU^*A_3U) : U \in U(4p)\}$$

and $W_C(A_1, A_2, A_3) \subset W'_C(A_1, A_2, A_3)$. Suppose $4(n-1, n-1, x) \in W_C(A_1, A_2, A_3)$. Then $\operatorname{Re}\operatorname{tr} CU^*A_1V = n-1$. Then using the reasoning in the second example of the proof of Theorem 5.4, we see that $4(n-1, n-1, n-1+1/3) \notin W'_C(A_1, A_2, A_3)$. Hence inclusion relation fails when $s(B) \prec_w s(C)$. Thus $W'_C(A_1, A_2, A_3)$ is not convex. ■

10 The $\mathfrak{so}^*(2n)$ Case

It is known that

$$K = \left\{ \begin{pmatrix} A & B \\ -B & A \end{pmatrix} : A^TA + B^TB = I, A^TB = B^TA, A, B \in \mathbb{R}^{n\times n} \right\},$$

$$\mathfrak{p} = \left\{ \begin{pmatrix} X & Y \\ Y & -X \end{pmatrix} : X^T = -X, Y^T = -Y, X, Y \in i\mathbb{R}^{n\times n} \right\},$$

$$\mathfrak{a} = i\mathbb{R}\big((E_{12} - E_{21}) - (E_{n+1,n+2} - E_{n+2,n+1})\big) \oplus i\mathbb{R}\big((E_{23} - E_{32}) - (E_{n+2,n+3} - E_{n+3,n+2})\big) \oplus \cdots .$$

Analogously to $\mathfrak{sp}_{2n}(\mathbb{R})$ case, we identify K with the unitary group $U(n)$ and the subspace $\mathfrak{p}$ with the space of complex skew symmetric matrices respectively. Then $\mathfrak{a}$ is identified with $i \oplus_{1\le j\le [n/2]} \mathbb{R}(E_{2j-1,2j} - E_{2j,2j-1})$. Then the group K acts on $\mathfrak{p}$ such that $A \to UAU^T$. So the C-numerical range of the complex skew symmetric matrices $A_1, \dots, A_p \in \mathfrak{p}$ is

$$W_C(A_1, \dots, A_p) = \{(\operatorname{Re}\operatorname{tr} CU^TA_1U, \dots, \operatorname{Re}\operatorname{tr} CU^TA_pU) : U \in U(n)\}.$$

The set is symmetric about the origin.

Since $\mathfrak{su}_{1,3} \cong \mathfrak{so}^*(6)$, by Corollary 5.3, we have the following result and one can give a more geometric proof by identifying $O(C)$ with a 5-sphere.

Theorem 10.1 *Let $C, A_1, \dots, A_p$ be 3×3 complex skew symmetric matrices. When $1 \le p \le 5$, $W_C(A_1, \dots, A_p)$ is an ellipsoid with the interior in $\mathbb{R}^p$ and hence a convex set.*

Corollary 10.2 *Let $n \ge 3$ be an odd integer. Suppose B and C are complex skew symmetric matrices with vectors of singular values (nonincreasing order) b and c, respectively such that $c - b \ge 0$. Then $W_B(A_1, \dots, A_p) \subset W_C(A_1, \dots, A_p)$ if $1 \le p \le 5$.*

Theorem 10.3 *Let C, B, A_1, A_2, A_3 be $n \times n$ complex skew symmetric matrices. Let $n \ge 4$ and b and c be the vectors of singular values of B and C respectively. If $b \prec c$, then $W_B(A_1, A_2, A_3) \subset W_C(A_1, A_2, A_3)$.*

Proof It is sufficient to consider the case $(b_1, b_2) \prec (c_1, c_2)$ and $b_i = c_i$, $i = 3, \ldots, n$. Suppose $x = (x_1, x_2, x_3) \in W_B(A_1, A_2, A_3)$, *i.e.*, there exist $e_1, e_2, \ldots, e_n$ orthonormal vectors in $\mathbb{C}^n$ such that for $i = 1, 2, 3$,

$$x_i = -\operatorname{Re}\Big[(b_1 + b_2)(e_1^T A_i e_2 + e_3^T A_i e_4) - (b_1 - b_2)(e_1^T A_i e_2 - e_3^T A_i e_4) - 2\sum_{j=3}^{[n/2]} b_j e_{2j-1}^T A_i e_{2j}\Big].$$

Let f_1, f_2, f_3 and $f_4 \in \mathbb{C}^n$ be the vectors defined by [30]

$$\begin{aligned}
f_1 &= \cos\phi\cos\theta e_1 & -\sin\phi\cos\theta e_2 & \quad -\cos\phi\sin\theta e_3 & +\sin\phi\sin\theta e_4\\
f_2 &= \sin\phi\cos\theta e_1 & +\cos\phi\cos\theta e_2 & \quad -\sin\phi\sin\theta e_3 & -\cos\phi\sin\theta e_4\\
f_3 &= \cos\phi\sin\theta e_1 & +\sin\phi\sin\theta e_2 & \quad +\cos\phi\cos\theta e_3 & +\sin\phi\cos\theta e_4\\
f_4 &= -\sin\phi\sin\theta e_1 & +\cos\phi\sin\theta e_2 & \quad -\sin\phi\cos\theta e_3 & +\cos\phi\cos\theta e_4.
\end{aligned}$$

The matrix which sends (e_1, e_2, e_3, e_4) to (f_1, f_2, f_3, f_4) is an element of $SO(4)$. So f_1, f_2, f_3, $f_4 \in \mathbb{C}^n$ are orthonormal vectors. Direct computation leads to

$$\operatorname{Re}(f_1^T A_j f_2 + f_3^T A_j f_4) = e_1^T A_j e_2 + e_3^T A_j e_4,$$
$$\operatorname{Re}(f_1^T A_j f_2 - f_3^T A_j f_4) = p_j \cos 2\theta + \sin 2\theta(q_j \sin 2\phi + s_j \cos 2\phi), \quad j = 1, 2, 3,$$

where

$$p_j = \operatorname{Re}(e_1^T A_j e_2 - e_3^T A_j e_4), \quad q_j = \operatorname{Re}(e_1^T A_j e_3 - e_2^T A_j e_4), \quad s_j = \operatorname{Re}(-e_2^T A_j e_3 + e_1^T A_j e_4).$$

Then for $i = 1, 2, 3$, y_i is just the real part of the number

$$(b_1 - b_2)[p_j \cos 2\theta + \sin 2\theta(q_j \sin 2\phi + s_j \cos 2\phi)]$$
$$- (c_1 + c_2)(e_1^T A_i e_2 + e_3^T A_i e_4) + 2\sum_{j=3}^{[n/2]} b_j e_{2j-1}^T A_i e_{2j}.$$

As θ and ϕ vary in $\mathbb{R}$, the locus of the point (y_1, y_2, y_3) in $\mathbb{R}^3$ is an ellipsoid (compare [2]) which will be denoted by $E_{b,E}$. Here $E = (e_1 \cdots e_n) \in U(n)$. Notice that $|c_1 - c_2| \geq |b_1 - b_2|$ and hence $(x_1, x_2, x_3) \in E_{b,E} \subset \operatorname{conv} E_{c,E} \subset W_C(A_1, A_2, A_3)$.

Now, given a 4×4 complex skew symmetric matrix A, there exists $U \in U(4)$ such that

$$U^T A U = \begin{pmatrix} 0 & is_1 \\ -is_1 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & is_2 \\ -is_2 & 0 \end{pmatrix}$$

where s_1, s_1, s_2, s_2 are singular values of A. This implies that we can find orthonormal vectors e_1', e_2', e_3', e_4' in the span of e_1, e_2, e_3, e_4 such that $E_{c,E'}$ is degenerated where $E' = (e_1' e_2' e_3' e_4' e_5 \cdots e_n) \in U(n)$. By a continuity argument, the result follows.

Theorem 10.4 *Let C, A_1, A_2, A_3 be $n \times n$ complex skew symmetric matrices.*

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

1. *Then* $W_C(A_1, A_2) = \{(\mathrm{Re}\,\mathrm{tr}\, CU^T A_1 U, \mathrm{Re}\,\mathrm{tr}\, CU^T A_2 U) : U \in U(n)\}$ *is convex when* $n > 2$. *It is an ellipse (perhaps degenerated) if* $n = 2$.
2. *If* n *is even, then* $W_C(A_1, A_2, A_3)$ *is not convex in general. If* $n \geq 3$ *is odd, then* $W_C(A_1, A_2, A_3)$ *is convex.*

Proof Suppose $n > 2$. (1) We notice that $W_C(A_1, A_2)$ is equal to the set

$$\Big\{-2\Big(\mathrm{Re}\sum_{i=1}^{[n/2]} c_i x_{2i-1}^T A_1 x_{2i}, \mathrm{Re}\sum_{i=1}^{[n/2]} c_i x_{2i-1}^T A_2 x_{2i}\Big) : (x_1 \cdots x_n) \in U(n)\Big\}.$$

By Lemma 3.3, Corollary 3.2 and Theorem 10.3, it is sufficient to consider that case that $0 \leq b_1 < c_1$ and $b_i = c_i$, $i = 2, \ldots, [n/2]$. Suppose $x = (x_1, x_2) \in W_B(A_1, A_2)$, *i.e.*, there exist $e_1, e_2, \ldots, e_n \in \mathbb{C}^n$ such that

$$x_i = -2\,\mathrm{Re}\Big(b_1 e_1^T A_i e_2 + \sum_{j=2}^{[n/2]} b_j e_{2j-1}^T A_i e_{2j}\Big), \quad i = 1, 2.$$

Define $f_1 = e^{i\theta} e_1$ and $f_i = e_i$, $i = 2, \ldots, n$. Then for $i = 1, 2$,

$$\begin{aligned} y_i &= -2\,\mathrm{Re}\Big(b_1 f_1^T A_i f_2 + \sum_{j=2}^{[n/2]} b_j f_{2j-1}^T A_i f_{2j}\Big) \\ &= -2\Big(b_1[\cos\theta\,\mathrm{Re}\, e_1^T A_i e_2 - \sin\theta\,\mathrm{Im}\, e_1^T A_i e_2] + \mathrm{Re}\sum_{j=2}^{[n/2]} b_j e_{2j-1}^T A_i e_{2j}\Big). \end{aligned}$$

The locus of the point (y_1, y_2) traces out an ellipse which is denoted by $E_{e,b}$. Now $(x_1, x_2) \in E_{e,b} \subset \mathrm{conv}\, E_{e,c}$. There are orthonormal vectors u_1, u_2 such that $u_1^T A_1 u_2 = 0$ ([29], $n > 2$). Extend u_1, u_2 to an orthonormal basis of $\mathbb{C}^n$, $\{u_1, \ldots, u_n\}$. The corresponding ellipse is degenerated. By continuity argument, we are done.

Suppose $n = 2$. The orbit $O(C)$ is

$$\Big\{U^T \begin{pmatrix} 0 & -c \\ c & 0 \end{pmatrix} U : U \in U(n)\Big\} = \Big\{e^{i\theta} \begin{pmatrix} 0 & -c \\ c & 0 \end{pmatrix} : \theta \in [0, 2\pi]\Big\},$$

by considering the determinant of $U^T C U$, where $C = \left(\begin{smallmatrix} 0 & -c \\ c & 0 \end{smallmatrix}\right)$. Let

$$A_1 = \begin{pmatrix} 0 & -a_1 \\ a_1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -a_2 \\ a_2 & 0 \end{pmatrix}.$$

Then

$$W_C(A_1, A_2) = \{(\mathrm{Re}\, ca_1 \cos\theta - \mathrm{Im}\, ca_1 \sin\theta, \mathrm{Re}\, ca_2 \cos\theta - \mathrm{Im}\, ca_2 \sin\theta) : \theta \in [0, 2\pi]\}$$

is an ellipse.

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

The following example shows that the first part, when n is even, is best possible. Let $X = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Let $B = X \oplus \cdots \oplus X \oplus X/3$, $C = X \oplus \cdots \oplus X \oplus X/2$, $A_1 = X \oplus \cdots \oplus X \oplus O_2$, $A_2 = X \oplus \cdots \oplus X \oplus iX$, $A_3 = X \oplus \cdots \oplus X \oplus X$, where each matrix is of size $2n \times 2n$. Then we claim that $W_B(A_1, A_2, A_3)$ is not a subset of $W_C(A_1, A_2, A_3)$.

Notice that $\big(-2(n-1), -2(n-1), -2(n-1)-2/3\big) = (\operatorname{Re}\operatorname{tr} BA_1, \operatorname{Re}\operatorname{tr} BA_2, \operatorname{Re}\operatorname{tr} BA_3) \in W_B(A_1, A_2, A_3)$. Now if $\big(-2(n-1), -2(n-1), x\big) \in W_C(A_1, A_2, A_3)$, then $\operatorname{Re}\operatorname{tr} CU^T A_1 U = -2(n-1)$ and by extremal properties, we have $U^T A_1 U = A_1$. So $U = U_1 \oplus U_2$ where U_2 is a 2×2 unitary matrix. Now $\operatorname{Re}\operatorname{tr} CU^T A_2 U = -2(n-1)$ implies that $U_2^T X U_2 = \pm X$. Thus $\operatorname{Re}\operatorname{tr} CU^T A_3 U$ cannot be $-2(n-1) - 2/3$. Hence the inclusion relation fails though $s(B) \prec_w s(C)$. So $W_C(A_1, A_2, A_3)$ is not convex.

(2) Let $n = 2m+1$. Similarly, we show that $W_B(A_1, A_2, A_3) \subset W_C(A_1, A_2, A_3)$ where $b_1 < c_1$ and $b_i = c_i$, $i = 2, \ldots, n$. Suppose $x = (x_1, x_2, x_3) \in W_B(A_1, A_2, A_3)$, *i.e.*, there exist orthonormal vectors $e_1, e_2, \ldots, e_{2m+1} \in \mathbb{C}^{2m+1}$ such that $x_i = -2(b_1 e_1^T A_i e_2 + \sum_{j=2}^m b_j e_{2j-1}^T A_i e_{2j})$, $i = 1, 2, 3$.

The point $\gamma = -2b_1(e_1^T A_1 e_2, e_1^T A_2 e_2, e_1^T A_3 e_2)$ belongs to $W_{B'}(A_1', A_2', A_3')$ which is the ellipsoid with interior and centered at the origin by Theorem 10.1. Here

$$A_i' = (E^T A_i E)[1, 2, 2m+1 \mid 1, 2, 2m+1], \quad i = 1, 2, 3,$$

are 3×3 skew symmetric matrices, and $A[\alpha \mid \beta]$ denotes the submatrix of A lying in the rows and columns indexed by the sequence α and β, respectively, and

$$B' = \begin{pmatrix} 0 & b_1 \\ -b_1 & 0 \end{pmatrix} \oplus 0.$$

The ellipsoid with interior is denoted by C_{e,b_1}. Since the 5-sphere $b_1 S^5$ centered at the origin and with radius b_1 in $\mathbb{R}^6$ is contained in the interior of the larger sphere $c_1 S^5$ with radius c_1 $(0 \le b_1 < c_1)$, $(x_1, x_2, x_3) \in C_{e,b_1} \subset C_{e,c_1} \subset W_C(A_1, A_2, A_3)$. ■

Remark 10.5 The $n = 2$ case follows from the isomorphism $\mathfrak{so}^*(4) \cong \mathfrak{su}(2) \oplus \mathfrak{sl}_2(\mathbb{R})$. The numerical range associated with $\mathfrak{su}(2) \oplus \mathfrak{sl}_2(\mathbb{R})$ is indeed the numerical range associated with $\mathfrak{sl}_2(\mathbb{R})$ since $\mathfrak{su}(2)$ is a compact form. Also $\mathfrak{so}^*(8) \cong \mathfrak{so}_{2,6}$ and see Theorem 11.4.

Corollary 10.6 ([**26**]) *Let C be a complex $n \times n$ skew symmetric matrix and let A be an $n \times n$ complex matrix. Then the congruence numerical range $W_C(A) = \{\operatorname{tr} CU^T AU : U \in U(n)\}$ is a circular disk centered at the origin when $n > 2$ or $n = 1$. When $n = 2$, it is a circle centered at the origin.*

11 The $\mathfrak{so}_{p,q}$ Case

Now

$$K = \mathrm{SO}(p) \times \mathrm{SO}(q), \quad \mathfrak{p} = \left\{ \begin{pmatrix} 0 & Y \\ Y^T & 0 \end{pmatrix} : Y \in \mathbb{R}_{p \times q} \right\}, \quad \mathfrak{a} = \bigoplus_{1 \le j \le p} \mathbb{R}(E_{j,p+j} + E_{p+j,j}).$$

The corresponding C-numerical range of $p \times q$ matrices $A_1, \ldots, A_m$, after disregarding the constant $2(p+q-2)$, is

$$W_C(A_1, \ldots, A_m) = \{(\operatorname{tr} C^T U A_1 V, \ldots, \operatorname{tr} C^T U A_m V) : U \in \mathrm{SO}(p), V \in \mathrm{SO}(q)\},$$

where $C, A_1, \ldots, A_m$ are $p \times q$ real matrices. It is clear that when $p \neq q$, say $p < q$, the special orthogonal groups can be replaced by the orthogonal groups and hence the set is symmetric about the origin. It is also symmetric when $p = q = 2n$.

When $m = 1$, the set $W_C(A)$ is evidently a line segment and is fully known [21] and [28]. Let $m = 2$. When $(p, q) = (1, 1)$, the numerical range is a singleton set. When $(p, q) = (1, 2)$ or $(2, 1)$, the numerical range $W_C(A_1, A_2)$ is then the image of the circle centered at the origin under a linear map from $\mathbb{R}^2$ to $\mathbb{R}^2$, *i.e.*, an ellipse and hence not convex. This is certainly the case since $\mathfrak{so}_{1,2} \cong \mathfrak{sl}_2(\mathbb{R})$.

Remark 11.1 When $p = 1$ and $q \geq 3$, $W_C(A_1, A_2)$ is the image of the unit sphere S^{q-1} in $\mathbb{R}^q$ under a linear map from $\mathbb{R}^q$ to $\mathbb{R}^2$. It is an elliptical disk and hence is convex. We already learned the special cases $q = 3$ and $q = 5$ from the isomorphisms $\mathfrak{so}_{1,3} \cong \mathfrak{sl}_2(\mathbb{C})^{\mathbb{R}}$ and $\mathfrak{so}_{1,5} \cong \mathfrak{sl}_2(\mathbb{H})$. Similarly, if $p = 1$ and $q \geq 4$, $W_C(A_1, A_2, A_3)$ is the image of the unit sphere S^{q-1} in $\mathbb{R}^q$ under a linear map from $\mathbb{R}^q$ to $\mathbb{R}^3$. It is an ellipsoid with interior and hence convex. We then conclude that the numerical range $W_C(A_1, A_2, A_3)$ is an ellipsoid with interior centered at the origin in $\mathbb{R}^3$ for $\mathfrak{sp}_{1,1}$ since $\mathfrak{sp}_{1,1} \cong \mathfrak{so}_{1,4}$.

When $(p, q) = (2, 2)$ we have the following example.

Example 11.2 The numerical range $W_C(A_1, A_2)$ is not convex when

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

Proof Clearly the points $(1, 0)$ and $(-1, 0)$ belong to $W_C(A_1, A_2)$. We want to show that their midpoint is not in $W_C(A_1, A_2)$. Suppose $(0, x) \in W_C(A_1, A_2)$, *i.e.*, there exist $P, Q \in \mathrm{SO}(2)$ such that $PQ = PA_1Q = \left(\begin{smallmatrix} 0 & \alpha \\ \beta & \gamma \end{smallmatrix}\right)$. By Theorem 2 of [35], $\gamma = 0$. Since the matrices have the same determinant, *i.e.*, $\det I_2 = \det PQ = 1$ and they have the same singular values, *i.e.*, $\alpha = -\beta$ and $\beta = \pm 1$, we conclude that $PQ = A_2$ or $-A_2$. Let

$$P = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \quad Q = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}.$$

Direct computation on $PQ = \pm A_2$ leads to $\cos(\theta + \phi) = 0$ and $\sin(\theta + \phi) = \pm 1$. This implies that $PA_2Q = -I_2$ and I_2 respectively. In other words, $x = \pm 1$ and hence $W_C(A_1, A_2)$ does not contain the origin. ∎

Remark 11.3 The orbit of $C = \operatorname{diag}(1, 0)$ is merely a part of the sphere $S^3 \subset \mathbb{R}^4$. The real linear map $C' \mapsto (\operatorname{tr} C'A_1, \operatorname{tr} C'A_2)$ does not send $O(C)$ onto an elliptical disk in $\mathbb{R}^2$.

Indeed, by Proposition 2.4 one can deduce the nonconvexity from the isomorphism $\mathfrak{so}_{2,2} \cong \mathfrak{sl}_2(\mathbb{R}) \oplus \mathfrak{sl}_2(\mathbb{R})$. The numerical range corresponding to $\mathfrak{sl}_2(\mathbb{R}) \oplus \mathfrak{sl}_2(\mathbb{R})$ is the sum (pointwise) of two ellipses, *i.e.*, the locus traced by one of the ellipses when its center

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

is moving on the boundary of the other ellipse (when the directions of the axes of the moving ellipse do not change). The figure is then the region between the outer and inner envelopes. In particular, if the two ellipses are circles, we have an annulus. If the two ellipses are degenerated, *e.g.*, two line segments centered at the origin, the numerical range is then a parallelogram with interior and hence convex.

Proposition 11.4 *Let C, A_1, A_2 be $p \times q$ real matrices. If*

(i) $\min\{p, q\} \geq 2$ *and* $p \neq q$, *or*
(ii) $p = q \geq 3$, *then* $W_B(A_1, A_2) \subset W_C(A_1, A_2)$ *when* $b \prec c$.

Proof For definiteness we assume $p \leq q$. Let $(r_1, r_2) \in W_B(A_1, A_2)$, *i.e.*, there exist $x_1, x_2 \in \mathbb{R}^q$ and $y_1, y_2 \in \mathbb{R}^p$ such that for $j = 1, 2$,

$$\begin{aligned} r_j &= \sum_{i=1}^{p} b_i y_i^T A_j x_i \\ &= \frac{1}{2}(b_1 + b_2)(y_1^T A_j x_1 + y_2^T A_j x_2) + \frac{1}{2}(b_1 - b_2)(y_1^T A_j x_i - y_2^T A_j x_2) + \sum_{i=3}^{p} b_i y_i^T A_j x_i. \end{aligned}$$

Let

$$\begin{aligned} u_1 &= \cos\theta x_1 + \sin\theta x_2, \quad & v_1 &= \cos\theta y_1 + \sin\theta y_2, \\ u_2 &= -\sin\theta x_1 + \cos\theta x_2, \quad & v_2 &= -\sin\theta y_1 + \cos\theta y_2, \end{aligned}$$

and $u_i = x_i$ and $v_i = y_i$, $i = 3, \ldots, n$. Then

$$\begin{aligned} \sum_{i=1}^{p} b_i v_i^T A_j u_i = \frac{1}{2}(b_1 + b_2)(y_1^T A_j x_1 + y_2^T A_j x_2) + \frac{1}{2}(b_1 - b_2)[\cos 2\theta(y_1^T A_j x_1 - y_2^T A_j x_2) \\ + \sin 2\theta(y_2^T A_j x_1 + y_1^T A_j x_2)] + \sum_{i=3}^{p} b_i y_i^T A_j x_i. \end{aligned}$$

Let $E_{b,x,y}$ denotes the ellipse which is the locus of the above expressioin as θ varies on $[0, 2\pi]$.

(i) We consider three cases:

(a) If $q > p \geq 3$, then there is a unit vector x_1' in the null space of A_1, *i.e.*, $A_1 x_1' = 0$. Then choose a unit vector $x_2' \in \mathbb{R}^q$ which is orthogonal to $x_1' \in \mathbb{R}^q$, and choose the orthonormal vectors y_1' and y_2' in $\mathbb{R}^p$ such that they are orthogonal to $A_1 x_2' \in \mathbb{R}^p$.

(b) If $q \geq p + 2$, and $q > p \geq 2$, then take x_1', x_2' in the null space of A_1 and set $y_1' = y_1$, $y_2' = y_2$.

https://doi.org/10.4153/CJM-2000-007-2 Published online by Cambridge University Press

(c) It remain to consider $(p,q) = (2,3)$. Given any $A \in \mathbb{R}_{2\times 3}$, there exist $U \in \mathrm{SO}(2)$ and $V \in \mathrm{SO}(3)$ such that

$$UAV = \begin{pmatrix} 0 & a & 0 \\ b & 0 & 0 \end{pmatrix},$$

where $a \geq b \geq 0$ are the singular values of A. Now choose $W = 1 \oplus R(\theta) \in \mathrm{SO}(3)$ where $R(\theta)$ is a rotation matrix such that

$$UAVW = \begin{pmatrix} 0 & -b & c \\ b & 0 & 0 \end{pmatrix}$$

and $b^2 + c^2 = a^2$. This implies that there exist $x_1', x_2' \in \mathbb{R}^3$ and $y_1', y_2' \in \mathbb{R}^2$ such that $y_1' A_1 x_1' = y_2' A_1 x_2' = 0$ and $y_2' A_1 x_1' = -y_1' A_1 x_2'$.

(ii) We consider two cases:

(a) If $p = q \geq 4$, then obviously we can choose two orthonormal vectors $x_1', x_2' \in \mathbb{R}^p$, and two orthonormal vectors $y_1', y_2' \in \mathbb{R}^p$ such that $y_i'^T A_1 x_j' = 0$, where $i, j = 1, 2$.

(b) Suppose $(p,q) = (3,3)$. Let $A \in \mathbb{R}_{3\times 3}$. There exist $U, V \in \mathrm{SO}(3)$ such that $UAV = \mathrm{diag}(s_2, s_1, \delta s_3)$ where δ is the sign of $\det A$ and $s_1 \geq s_2 \geq s_3 \geq 0$ are the singular values of A. Let $R(\theta)$ be a rotation. Then there exists $\theta \in \mathbb{R}$ such that the (1,1) entry of $R^{-1}(\theta)\,\mathrm{diag}(s_1, \delta s_3) R(\theta)$ is s_2. This implies that there exist $x_1', x_2', y_1', y_2' \in \mathbb{R}^3$ such that $y_1' A_1 x_1' = y_2' A_1 x_2'$ and $y_2' A_1 x_1' = y_1' A_1 x_2' = 0$.

Extend $\{x_1', x_2'\}$ and $\{y_1', y_2'\}$ to orthonormal bases $\{x_1', \dots, x_p'\}$ and $\{y_1', \dots, y_q'\}$ of $\mathbb{R}^p$ and $\mathbb{R}^q$ respectively. So the corresponding $E_{x',y',b}$ is a line segment or a point. By continuity argument, the inclusion relation follows. ■

Theorem 11.5 *Let C, A_1, A_2, A_3 be $p \times q$ real matrices. If $\min\{p,q\} \geq 2$ and $p \neq q$, then $W_C(A_1, A_2)$ is convex. Moreover, $W_C(A_1, A_2, A_3)$ is not convex in general.*

Proof It is sufficient to show that $W_B(A_1, A_2) \subset W_C(A_1, A_2)$ if $0 \leq b_1 < c_1$ in view of (i) of Proposition 11.4. Let $(r_1, r_2) = (\sum_{i=1}^p b_i y_i^T A_1 x_i, \sum_{i=1}^p b_i y_i^T A_1 x_i) \in W_C(A_1, A_2)$. Let $x_1' = \cos\theta x_1 + \sin\theta x_q$ and $x_q' = -\sin\theta x_1 + \cos\theta x_q$, $x_i' = x_i$, $i = 2, \dots, q-1$. Then for $j = 1, 2$,

$$\sum_{i=1}^p b_i y_i^T A_j x_i' = b_1 (y_1^T A_j x_1 \cos\theta + y_1^T A_j x_q \sin\theta) + \sum_{i=2}^p b_i y_i^T A_j x_i.$$

The locus of the point $(\sum_{i=1}^p b_i y_i^T A_1 x_i', \sum_{i=1}^p b_i y_i^T A_2 x_i')$ is an ellipse as θ varies on $[0, 2\pi]$, denoted by $E_{x,y,b}$. We have $E_{x,y,b} \subset \mathrm{conv}\, E_{x,y,c}$ since $0 \leq b_1 < c_1$. Let u_1 be a unit vector in the null space of A_1 and extend it to an orthonormal basis $\{u_1, \dots, u_q\}$ of $\mathbb{R}^q$. Then choose a unit vector $v_1 \in \mathbb{R}^p$ which is perpendicular to $A_1 u_2 \in \mathbb{R}^p$ $(p \geq 2)$ and then extend it to an orthonormal basis $\{v_1, \dots, v_p\}$ of $\mathbb{R}^p$. Then $E_{u,v,c}$ is a line segment or a point. Applying the continuity argument will finish the proof.

The convexity is best possible because of the following example. Assume $p < q$ without loss of generality, $B = [\hat{B} \mid 0]$ where $\hat{B} = I_{p-2} \oplus 3I_2$ and $C = [\hat{C} \mid 0]$ where $\hat{C} = I_{p-2} \oplus \operatorname{diag}(4,2)$. Let $A_i = [\hat{A}_i \mid 0]$ for $i = 1,2,3$, such that

$$\hat{A}_1 = I_p, \quad \hat{A}_2 = I_{p-2} \oplus \operatorname{diag}(1,-1), \quad \hat{A}_3 = I_{p-2} \oplus \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then $(p+4, p-2, p-2) \in W_B(A_1,A_2,A_3) \setminus W_C(A_1,A_2,A_3)$ because of the following reason. If $\operatorname{tr} B^T U^T A_1 V = \operatorname{tr} C^T U^T A_1 V = p+4$, then by the same argument in the proof of Theorem 5.4 U is of the form $U_1 \oplus U_2 \in \mathrm{SO}(p)$, where $U_2 \in \mathrm{SO}(2)$, and V is of the form $U_1 \oplus U_2 \oplus V_3 \in \mathrm{SO}(q)$. Now $VC^TU^T = [D \mid 0]^T$ where

$$D = I_{p-2} \oplus \begin{pmatrix} a & c \\ c & d \end{pmatrix}.$$

If $(\operatorname{Re}\operatorname{tr} C^T U^T A_1 V, \operatorname{Re}\operatorname{tr} C^T U^T A_2 V, \operatorname{Re}\operatorname{tr} C^T U^T A_3 V)$ were $(p+4, p-2, p-2)$, then $a+b = 6$, $a-b = 0$, $c = 0$, implying that $a = b = 3$ and $c = 0$ which is impossible. Thus inclusion does not hold, and $W_C(A_1,A_2,A_3)$ is not convex. ■

Remark 11.6 By Proposition 2.4 the convexity result for $\mathfrak{so}_{2,3}$, $\mathfrak{so}_{2,4}$, and $\mathfrak{so}_{2,6}$ can also be deduced from those of $\mathfrak{sp}_4(\mathbb{R})$, $\mathfrak{su}_{2,2}$, and $\mathfrak{so}^*(8)$ respectively, since $\mathfrak{so}_{2,3} \cong \mathfrak{sp}_4(\mathbb{R})$, $\mathfrak{so}_{2,4} \cong \mathfrak{su}_{2,2}$, and $\mathfrak{so}_{2,6} \cong \mathfrak{so}^*(8)$.

The above technique does not apply for the $n \times n$ case ($n \geq 3$) since the condition $Z \in \operatorname{conv} W(Y)$ is not equivalent to $\prec_w$ nor $\prec$. It is Thompson's partial ordering $\ll$. Nevertheless we have the following result.

Theorem 11.7 *Let C, A_1, A_2, A_3 be $n \times n$ real matrices where $n \geq 3$. Then $W_C(A_1,A_2)$ is convex. Moreover, $W_C(A_1,A_2,A_3)$ is not convex in general if $n \geq 2$.*

Proof The proof is similar to Theorem 6.2. From the isomorphism $\mathfrak{so}_{3,3} \cong \mathfrak{sl}_4(\mathbb{R})$ and Theorem 4.1, $W_C(A_1,A_2) = \{(\operatorname{tr} CUA_1V, \operatorname{tr} CUA_2V) : U, V \in \mathrm{SO}(3)\}$ is convex for any 3×3 real matrices C, A_1, A_2. Then apply the arguments in the proof of Theorem 6.2 to finish the proof.

Let $C = I_{n-2} \oplus \operatorname{diag}(1,0)$, $A_1 = I_{n-2} \oplus O_2$, $A_2 = I_{n-2} \oplus \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $A_3 = I_n$. Then we claim that $W_C(A_1,A_2,A_3)$ is not convex. It is clear that the points $(n-2, n-2, n-2 \pm 1/2)$ are in $W_C(A_1,A_2,A_3)$. We are going to show that the mid-point $(n-2, n-2, n-2)$ is not inside. If $(n-2, n-2, n-2) = (\operatorname{tr} CU^T A_1 V, \operatorname{tr} CU^T A_2 V, \operatorname{tr} CU^T A_3 V) \in W_C(A_1,A_2,A_3)$, then by extremal properties [17], we have $U^T A_1 V = A_1$ and hence $U = W \oplus U_1$ and $V = W \oplus V_1$, where $U_1, V_1 \in \mathrm{SO}(2)$. Then consider $\operatorname{tr} CU^T A_2 V$ and $\operatorname{tr} CU^T A_3 V$. It will then reduce to the computation of Example 11.2. So $W_C(A_1,A_2,A_3)$ is not convex. ■

Remark 11.8 If $\mathrm{SO}(n)$ is replaced by $O(n)$ in the above setting, then we have $\tilde{W}_C(A_1,A_2) = \{(\operatorname{tr} CUA_1V, \operatorname{tr} CUA_2V) : U, V \in O(n)\}$. It is the union of the convex sets $W_C(A_1,A_2)$ and $W_{C'}(A_1,A_2)$ where $C' = DC$ and $D = \operatorname{diag}(1,\ldots,1,-1)$. Clearly $\tilde{W}_C(A_1,A_2) = W_C(A_1,A_2)$ when the rank of C is less than n. However the set $\tilde{W}_C(A_1,A_2)$ is not convex in general and we have the following example. Let $C = A_1 = I_n$, $A_2 = D$. Evidently $(n, n-2)$ and $(n-2, n) \in \tilde{W}_C(A_1, A_2)$. If the midpoint $(n-1, n-1)$ were in $\tilde{W}_C(A_1, A_2)$, then we would have $U, V \in O(n)$ such that $\operatorname{tr} A_1UCV = \operatorname{tr} A_2UCV = n-1$. Let $d_1, \dots, d_n$ be the diagonal elements of UCV. So $\sum_{i=1}^n d_i = \sum_{i=1}^{n-1} d_i - d_n = n-1$. Hence $d_n = 0$ and $\sum_{i=1}^{n-1} d_i = n-1$. Then $n-1 = |\sum_{i=1}^{n-1} d_i| \le \sum_{i=1}^{n-1} |d_i| = \sum_{i=1}^{n-1} |d_i| - |d_n| \le n-2$, by Thompson's inequalities [35]. It is absurd.

12 Conclusion

We conclude that $\mathfrak{sl}_2(\mathbb{R})$ is the only one giving nonconvex $W_C(A_1, A_2)$ among simple classical real Lie algebras (up to isomorphism). Concerning the convexity of $W_C(A_1, A_2, A_3)$ we make the following table.

$\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C}),\ n \ge 2$	Yes if $n > 2$ (best possible)
$\mathfrak{h} = \mathfrak{sl}_n(\mathbb{R})$	No
$\mathfrak{h} = \mathfrak{sl}_m(\mathbb{H}),\ n = 2m$	Yes if $n > 2$ (best possible)
$\mathfrak{h} = \mathfrak{su}_{p,q}\ (p = 0, 1, \dots, [n/2],\ p + q = n)$	Yes if $p \ne q$ (best possible). No if $p = q$
$\mathfrak{g} = \mathfrak{so}_{2n+1}(\mathbb{C}),\ n \ge 2$	Yes (best possible)
$\mathfrak{h} = \mathfrak{so}_{p,q}\ (p = 0, 1, \dots, n,\ p + q = 2n + 1)$	No
$\mathfrak{g} = \mathfrak{sp}_n(\mathbb{C}),\ n = 2m,\ m \ge 3$	Yes (best possible)
$\mathfrak{h} = \mathfrak{sp}_n(\mathbb{R}),\ n = 2m$	No
$\mathfrak{h} = \mathfrak{sp}_{p,q},\ (p = 0, 1, \dots, [m/2],\ p + q = m)$	No
$\mathfrak{g} = \mathfrak{so}_{2n}(\mathbb{C}),\ n \ge 4$	Yes (best possible)
$\mathfrak{h} = \mathfrak{so}_{p,q},\ (p = 0, 1, \dots, n,\ p + q = 2n)$	No
$\mathfrak{h} = \mathfrak{so}^*(2n)$	No if n is even. Yes if n is odd.

The following is the only case in the above list we have no answer.

Problem For the case $\mathfrak{so}^*(2n)$ with an odd integer n, what is the largest $m \ge 3$ so that $W_C(A_1, \dots, A_m)$ is always convex?

From the proof of Theorem 10.1, we see that $m \le 5$.

Remark 12.1 The exceptional simple Lie algebras are [23]: 3 for $\mathfrak{g}_2$; 4 for $\mathfrak{f}_4$; 6 for $\mathfrak{e}_6$; 5 for $\mathfrak{e}_7$ and 4 for $\mathfrak{e}_8$. The total number of cases is 22. Among them 5 are compact Lie algebras and the corresponding numerical ranges are trivial. For those 5 complex simple Lie algebras of exceptional type when we consider them as real Lie algebras, Theorem 2.1 yields the convexity of $W_C(A_1, A_2)$. Hence 12 cases are left.

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.* Proc. Amer. Math. Soc. **89**(1983), 215–218.
[3] ______, *Some theorems on the numerical range.* Linear and Multilinear Algebra **15**(1984), 215–218.
[4] C. A. Berger, *Normal Dilations.* Ph. D. dissertation, Cornell University, 1963.
[5] L. Brickman, *On the field of values of a matrix.* Proc. Amer. Math. Soc. **12**(1961), 61–66.
[6] M. D. Choi, C. Laurie, H. Radjavi and P. Rosenthal, *On the congruence numerical range and related functions of matrices.* Linear and Multilinear Algebra **22**(1987), 1–5.
[7] K. M. Chong, *An induction theorem for rearrangements.* Canad. J. Math. **28**(1976), 154–160.
[8] C. Davis, *The Toeplitz-Haudorff theorem explained.* Canad. Math. Bull. **14**(1971), 245–246.
[9] M. Goldberg and E. G. Straus, *Elementary inclusion relations for generalized numerical ranges.* Linear Algebra Appl. **18**(1977), 1–24.
[10] K. E. Gustafson and D. K. M. Rao, *Numerical Range: the field of values of linear operators and matrices.* Springer, New York, 1997.
[11] P. Halmos, *A Hilbert Space Problem Book.* Springer-Verlag, New York, 1982.
[12] F. Hausdorff, *Der Wertvorrat einer Bilinearform.* Math Z. **3**(1919), 314–316.
[13] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces.* Academic Press, New York, 1978.
[14] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.
[15] A. W. Knapp, *Representation Theory of Semisimple Groups.* Princeton University Press, New Jersey, 1986.
[16] B. Kostant, *On convexity, the Weyl group and Iwasawa decomposition.* Ann. Sci. Ecole Norm. Sup. (4) **6**(1973), 413–460.
[17] C. K. Li, *Matrices with some extremal properties.* Linear Algebra Appl. **101**(1988), 255–267.
[18] ______, *C-numerical ranges and C-numerical radii.* Linear and Multilinear Algebra **37**(1994), 51–82.
[19] ______, *Some convexity theorems for the generalized numerical ranges.* Linear and Multilinear Algebra **40**(1996), 235–240.
[20] C. K. Li and H. Nakazato, *Some results on the q-numerical range.* Linear and Multilinear Algebra **43**(1998), 385–410.
[21] H. Miranda and R. C. Thompson, *A supplement to the von Neumann trace inequality for singular values.* Linear Algebra Appl. **248**(1994), 61–66.
[22] J. von Neumann, *Some matrix-inequalities and metrization of matrix-space.* Tomsk. Univ. Rev. **1**(1937), 286–300. In: *Collected Works*, Pergamon, New York, 1962, Vol. 4, 205–219.
[23] A. L. Onishchik and E. B. Vinberg, *Lie groups and algebraic groups.* Springer-Verlag, Berlin, 1990.
[24] Y. T. Poon, *Another proof of a result of Westwick.* Linear and Multilinear Algebra **9**(1980), 35–37.
[25] ______, *Generalized numerical ranges, joint positive definiteness and multiple eigenvalues.* Proc. Amer. Math. Soc. **125**(1997), 1625–1634.
[26] T. Y. Tam, *Note on a paper of Thompson: the congruence numerical range.* Linear and Multilinear Algebra **17**(1985), 107–115.
[27] ______, *Kostant's convexity theorem and the compact classical groups.* Linear and Multilinear Algebra **43**(1997), 87–113.
[28] ______, *Miranda and Thompson's trace inequality and a log convexity result.* Linear Algebra Appl. **262**(1997), 307–325.
[29] ______, *Partial superdiagonal elements and singular values of a complex skew symmetric matrix.* SIAM J. Matrix Anal. Appl. **19**(1998), 737–754.
[30] ______, *An extension of a convexity theorem of the generalized numerical range associated with* $SO(2n+1)$. Proc. Amer. Math. Soc. (1) **127**(1999), 35–44.
[31] ______, *Generalized numerical ranges, numerical radii, and Lie groups.* Manuscript, 1996.
[32] ______, *A Lie theoretic approach of Thompson's theorems on singular values-diagonal elements and some related results.* J. London Math. Soc., to appear.
[33] ______, *Group majorization, Eaton triples and numerical range.* Linear and Multilinear Algebra, to appear.
[34] T. Y. Tam and N. K. Tsing, *Research problem: the congruence numerical range.* Linear and Multilinear Algebra **19**(1986), 405.
[35] R. C. Thompson, *Singular values, diagonal elements and convexity.* SIAM J. Appl. Math. **32**(1977), 39–63.
[36] ______, *Singular values and diagonal elements of complex symmetric matrices.* Linear Algebra Appl. **26**(1979), 65–106.
[37] ______, *The congruence numerical range.* Linear and Multilinear Algebra **8**(1980), 197–206.
[38] O. Toeplitz, *Das algebraische Analogon zu einem Satze von Fejér.* Math. Z. **2**(1918), 187–197.
[39] N. K. Tsing, *The constrainted bilinear form and the C-numerical range.* Linear Algebra Appl. **56**(1984), 195–206.

[40] F. Warmer, *Foundation of Differentiable manifolds and Lie Groups.* Scott Foresman and Company, 1971.
[41] R. Westwick, *A theorem on numerical range.* Linear and Multilinear Algebra **2**(1975), 311–315.

Department of Mathematics
College of William and Mary
P.O. Box 8795
Williamburg, VA 23187-8795
USA
email: chkli@math.wm.edu
website: http://www.math.wm.edu/˜ckli

Department of Mathematics
Auburn University
Auburn, AL 36849-5310
USA
email: tamtiny@mail.auburn.edu
website: http://www.auburn.edu/˜tamtiny