Unspecified Journal
Volume 00, Number 0, Pages 000–000
S ????-????(XX)0000-0

A RANGE ASSOCIATED WITH SKEW SYMMETRIC MATRIX

XUHUA LIU, DAWIT G. TADESSE AND TIN-YAU TAM

Abstract. We study the range

$$S(A) := \{x^T Ay : x, y \text{ are orthonormal in } \mathbb{R}^n\},$$

where A is an $n \times n$ complex skew symmetric matrix. It is a compact convex set. Power inequality $s(A^{2k+1}) \le s^{2k+1}(A)$, $k \in \mathbb{N}$, for the radius $s(A) := \max_{\xi\in S(A)} |\xi|$ is proved. When $n = 3, 4, 5, 6$, relations between $S(A)$ and the classical numerical range and the k-numerical range are given. Axiomatic characterization of $S(A)$ is given. Sharp points and extreme points are studied.

1. Introduction

Let $\mathbb{C}_{n\times n}$ be the set of $n \times n$ complex matrices. The classical numerical range of $A \in \mathbb{C}_{n\times n}$ is

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

It is the image of the unit sphere

$$\mathbb{S}^{n-1} = \{x \in \mathbb{C}^n : x^*x = 1\}$$

under the quadratic map $x \mapsto x^*Ax$. Toeplitz-Hausdorff theorem asserts that $W(A)$ is a compact convex set [10]. It can be written as

$$W(A) = \{(U^*AU)_{11} : U \in \mathrm{U}(n)\},$$

where $\mathrm{U}(n)$ denotes the group of $n \times n$ unitary matrices. The Lie algebra $\mathfrak{u}(n)$ of $\mathrm{U}(n)$ is the set of $n \times n$ skew Hermitian matrices and $\mathfrak{gl}_n(\mathbb{C}) = \mathbb{C}_{n\times n}$ is the complexification of $\mathfrak{u}(n)$.

Similar to the classical numerical range, a different range emerges if one replaces the unitary group $\mathrm{U}(n)$ by the special orthogonal group $\mathrm{SO}(n)$. The Lie algebra $\mathfrak{so}(n)$ of $\mathrm{SO}(n)$ is the set of $n \times n$ real skew symmetric matrices, whose complexification is the Lie algebra $\mathfrak{so}_n(\mathbb{C})$ of $n \times n$ complex skew symmetric matrices. The range of $A \in \mathfrak{so}_n(\mathbb{C})$ associated with $\mathrm{SO}(n)$ is defined to be the set

$$S(A) := \{(O^T AO)_{12} : O \in \mathrm{SO}(n)\} \subset \mathbb{C}. \tag{1.1}$$

2000 Mathematics Subject Classification. Primary 15A60
Keywords. numerical range, skew symmetric matrix, convexity, extreme point, sharp point

It is a compact convex set in $\mathbb{C}$ by [20, Corollary 2.4 (b),(c)] (with $C = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \oplus 0_{n-2}$). Also see [15, 21]. When $n = 2$, i.e., $A = \begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix}$, $S(A) = \{\alpha\}$ is a singleton set. When $n \geq 3$,

$$\begin{aligned} S(A) &= \{x_1^T A x_2 : x_1, x_2 \text{ are two columns of some } O \in \mathrm{O}(n)\} \\ &= \{x^T A y : x, y \text{ are orthonormal in } \mathbb{R}^n\}. \end{aligned}$$

We remark that $S(B)$ may not be convex for general $B \in \mathbb{C}_{n\times n}$, even though $S(B)$ is well defined for all $B \in \mathbb{C}_{n\times n}$.

Example 1.1. Let $B = \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}$. By direct computation,

$$S(B) := \{(O^T B O)_{12} : O \in \mathrm{SO}(2)\} = \{-\cos\theta(\sin\theta - i\cos\theta) : \theta \in \mathbb{R}\}.$$

So $S(B)$ contains the points $\pm\dfrac{1}{2} + \dfrac{i}{2}$ but not their midpoint $\dfrac{i}{2}$.

We briefly describe the contents of the paper. In Section 2, some formulation of $S(A)$ is given. Power inequality $s(A^{2k+1}) \leq s^{2k+1}(A)$, $k \in \mathbb{N}$, for the radius $s(A) := \max_{\xi \in S(A)} |\xi|$ is proved. In Section 3 we study the special cases $n = 3, 4, 5, 6$ in which relations between $S(A)$ and the classical numerical range and the k-numerical range are given. Explicit description of the elliptical disk $S(A)$ is provided when $n = 3$. In Section 4 we provide some basic properties of $S(A)$. In Section 5 axiomatic characterization of $S(A)$ is given. In Section 6 characterization of sharp points and extreme points are obtained. In Section 7 we compare the q-numerical range $W_q(A)$ and a real analog $S_q(A)$.

2. Some formulations of $S(A)$ and power inequality for $s(A)$

We first show that $S(A)$ is an elliptical disk centered at the origin when $A \in \mathfrak{so}_3(\mathbb{C})$ in the following remark.

Remark 2.1. Alike Davis' treatment [4] for $W(A)$ when $A \in \mathbb{C}_{2\times 2}$, there is a geometric way to see that $S(A)$ is an elliptical disk. If

$$A = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} \in \mathfrak{so}_3(\mathbb{C}),$$

direct computation yields

$$x^T A y = a(x_1y_2 - x_2y_1) + b(x_1y_3 - y_1x_3) + c(x_2y_3 - x_3y_2) = (c, -b, a) \cdot (x \times y).$$

Since $\{x \times y : x, y \in \mathbb{R}^3 \text{ are orthonormal}\} = \mathbb{S}^2_{\mathbb{R}}$ (the unit sphere in $\mathbb{R}^3$), $S(A)$ is the image under the linear map $z \in \mathbb{S}^2_{\mathbb{R}} \mapsto (c, -b, a) \cdot z \in \mathbb{C}$. Thus $S(A) \subset \mathbb{C}$ is an elliptical disk centered at the origin.

We will see that the orthogonality restriction on $x, y \in \mathbb{R}^n$ in the definition (1.1) of $S(A)$ can be relaxed. Moreover, $S(A) = W(\hat{A})$ where $\hat{A} := \begin{pmatrix} 0 & A \\ -A & 0 \end{pmatrix}$. In order to do so, we recall a real analogy of the classical numerical range $W(B)$ of $B \in \mathbb{C}_{n\times n}$ [2, 16]:

$$V(B) := \{z^T B z : z \in \mathbb{R}^n, z^T z = 1\},$$

for which we may restrict B to be symmetric since $V(B) = V\left(\frac{B+B^T}{2}\right)$. The corresponding radius is

$$v(B) := \max_{\xi \in V(B)} |\xi|.$$

Theorem 2.2. (McIntosh [16, p.476]) If $B \in \mathbb{C}_{n\times n}$ is symmetric and $n \geq 3$, then $V(B) = W(B)$ and hence $w(B) = v(B)$.

Lemma 2.3. Let $A \in \mathfrak{so}_n(\mathbb{C})$ with $n \geq 3$.

(1) Suppose $x_1, x_2 \in \mathbb{R}^n$ and $y_1, y_2 \in \mathbb{R}^n$ are orthonormal pairs, and $\operatorname{span}\{x_1, x_2\} = \operatorname{span}\{y_1, y_2\}$. Then $y_2^T A y_1 = \pm x_2^T A x_1$.

(2) If $\xi \in S(A)$, then $t\xi \in S(A)$ for $-1 \leq t \leq 1$.

Proof. (1) Notice that $[y_1\ y_2] = [x_1\ x_2]O$, where $O \in \mathrm{O}(2)$, that is,

$$[y_1\ y_2] = [x_1\ x_2]\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

or

$$[y_1\ y_2] = [x_1\ x_2]\begin{pmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Then for the first case

$$\begin{aligned} y_2^T A y_1 &= (\sin\theta x_1 + \cos\theta x_2)^T A(\cos\theta x_1 - \sin\theta x_2) \\ &= \cos^2\theta x_2^T A x_1 - \sin^2\theta x_1^T A x_2 \\ &= x_2^T A x_1 \end{aligned}$$

and for the second case $y_2^T A y_1 = -x_2^T A x_1$.

(2) Because of Remark 2.1 we may assume that $n \geq 4$. Suppose that $\xi = x^T A y$ where $x, y \in \mathbb{R}^n$ are orthonormal. Choose a unit vector z that is orthogonal to y and $y_1 := (\operatorname{Re} A)y \in \mathbb{R}^n$ and $y_2 := (\operatorname{Im} A)y \in \mathbb{R}^n$ since $n \geq 4$. Then choose $\mu \in \mathbb{R}$ so that $w := tx + \mu z$ is a unit vector. Hence y and w are orthonormal and $t\xi = tx^T A y = w^T A y \in S(A)$. □

The following theorem shows that the orthogonality restriction on $x, y \in \mathbb{R}^n$ in the definition (1.1) of $S(A)$ can be removed. Moreover $S(A)$ can be realized as the classical numerical range of some matrix with double size. Thus it yields the convexity of $S(A)$, independent of the approach in [20].

Theorem 2.4. Let $A \in \mathfrak{so}_n(\mathbb{C})$ with $n \geq 3$ and $\hat{A} := \begin{pmatrix} 0 & A \\ -A & 0 \end{pmatrix}$. Then

$$\begin{aligned} S(A) &= \{x^T A y : x, y \in \mathbb{R}^n \text{ are unit vectors}\} \\ &= \{x^T A y : x, y \in \mathbb{R}^n, \|x\|_2^2 + \|y\|_2^2 = 2\} \\ &= V(\hat{A}) \\ &= W(\hat{A}). \end{aligned}$$

Thus $S(A)$ is compact and convex.

Proof. The last equality follows from Theorem 2.2. For arbitrary unit vectors $x, y \in \mathbb{R}^n$, write $y = \alpha x + \beta y'$, where $\alpha, \beta \in \mathbb{R}$ with $|\beta| \leq 1$ and $y' \in x^\perp$ is a unit vector. So $x^T A y = \beta x^T A y' \in S(A)$ by Lemma 2.3(2). Notice that

$$(x^T \; y^T)\hat{A} \begin{pmatrix} x \\ y \end{pmatrix} = 2x^T A y.$$

Thus it is easy to see that

$$\begin{aligned} S(A) &= \{x^T A y : x, y \in \mathbb{R}^n \text{ are unit vectors}\} \\ &\subset \{x^T A y : x, y \in \mathbb{R}^n, \|x\|_2^2 + \|y\|_2^2 = 2\} \\ &\subset \{z^T \hat{A} z : z \in \mathbb{R}^{2n}, z^T z = 1\}. \end{aligned}$$

On the other hand, if $z = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^n$ is a unit vector, then by Lemma 2.3(2),

$$z^T \hat{A} z = 2\|x\|_2 \|y\|_2 (\frac{x}{\|x\|_2})^T A (\frac{y}{\|y\|_2}) \in S(A)$$

since $2\|x\|_2\|y\|_2 \leq \|x\|_2^2 + \|y\|_2^2 = 1$. □

Define

$$s(A) := \max_{\xi \in S(A)} |\xi|.$$

The power inequality [7, p.118–119], [17] for the numerical radius asserts that

$$w(A^k) \leq w^k(A) \tag{2.1}$$

for any $A \in \mathbb{C}_{n\times n}$ and $k \in \mathbb{N}$. Since $A \in \mathfrak{so}_n(\mathbb{C})$, the power A^{2k+1} remains skew symmetric but A^{2k} is symmetric.

Theorem 2.5. Let $A \in \mathfrak{so}_n(\mathbb{C})$ with $n \geq 2$.

(1) When $n = 2$, $s(A^{2k+1}) = s^{2k+1}(A)$.
(2) When $n \geq 3$ and $k \in \mathbb{N}$, $s(A^{2k+1}) \leq s^{2k+1}(A)$.

Proof. (1) It is trivial.

(2) By Theorem 2.4 $s(A) = w(\hat{A})$, where $\hat{A} = \begin{pmatrix} 0 & A \\ -A & 0 \end{pmatrix}$. Observe that $\widehat{A^{2k+1}} = (-1)^k \hat{A}^{2k+1}$. So by (2.1)

$$s(A^{2k+1}) = w(\widehat{A^{2k+1}}) = w(\hat{A}^{2k+1}) \le w(\hat{A})^{2k+1} = s(A)^{2k+1}.$$

□

3. Shape of $S(A)$ when $n = 3, 4, 5, 6$

From Theorem 2.4, $S(A) = W(\hat{A})$ and $\hat{A} = \begin{pmatrix} 0 & A \\ -A & 0 \end{pmatrix} \in \mathbb{C}_{2n\times 2n}$ when $A \in \mathfrak{so}_n(\mathbb{C})$ with $n \ge 3$. We want to study $S(A)$ and its relation to the classical numerical range and k-numerical range when the dimension is low. The k-numerical range of $B \in \mathbb{C}_{n\times n}$, $1 \le k \le n$, is

$$W_k(B) := \{x_1^* B x_1 + \cdots + x_k^* B x_k : x_1, \cdots, x_k \in \mathbb{C}^n \text{ are orthonormal}\}.$$

Berger [1] (also see the remark on [7, p.315–316]) proved the convexity of $W_k(B)$ for any $B \in \mathbb{C}_{n\times n}$.

Theorem 3.1. (1) If $A \in \mathfrak{so}_3(\mathbb{C})$, then $S(A) = W(B)$ for some $B \in \mathfrak{sl}_2(\mathbb{C})$ and thus is an elliptical disk (possibly degenerate) centered at the origin. Conversely, if $B \in \mathfrak{sl}_2(\mathbb{C})$, then there is $A \in \mathfrak{so}_3(\mathbb{C})$ such that $S(A) = W(B)$.

(2) If $A \in \mathfrak{so}_4(\mathbb{C})$, then $S(A) = W(B)+W(C)$ for some $B, C \in \mathfrak{sl}_2(\mathbb{C})$ and thus it is the sum of two elliptical disks (possibly degenerate) centered at the origin. Conversely, if $B, C \in \mathfrak{sl}_2(\mathbb{C})$, then there is $A \in \mathfrak{so}_4(\mathbb{C})$ such that $S(A) = W(B) + W(C)$.

(3) If $A \in \mathfrak{so}_5(\mathbb{C})$, then $S(A) = W_2(B)$ for some $B \in \mathfrak{sp}_2(\mathbb{C}) \subset \mathbb{C}_{4\times 4}$, i.e., $B = \begin{pmatrix} B_1 & B_2 \\ B_3 & -B_1^T \end{pmatrix}$, where $B_1, B_2, B_3 \in \mathbb{C}_{2\times 2}$ and B_2, B_3 are symmetric. Conversely, if $B \in \mathfrak{sp}_2(\mathbb{C})$, then there is $A \in \mathfrak{so}_5(\mathbb{C})$ such that $S(A) = W_2(B)$.

(4) If $A \in \mathfrak{so}_6(\mathbb{C})$, then $S(A) = W_2(B)$ for some $B \in \mathfrak{sl}_4(\mathbb{C})$. Conversely, if $B \in \mathfrak{sl}_4(\mathbb{C})$, then there is $A \in \mathfrak{so}_6(\mathbb{C})$ such that $S(A) = W_2(B)$.

Proof. The following lower dimensional Lie algebra isomorphisms

$$\mathfrak{a}_1 \cong \mathfrak{b}_1, \quad \mathfrak{d}_2 \cong \mathfrak{a}_1 \oplus \mathfrak{a}_1, \quad \mathfrak{b}_2 \cong \mathfrak{c}_2, \quad \mathfrak{a}_3 \cong \mathfrak{d}_3$$

are known [8, p.516] and motivate our results.

Let $K := \mathrm{SO}(n)$ (or any connected Lie group K with Lie algebra $\mathfrak{k} = \mathfrak{so}(n)$). Given $A \in \mathfrak{g} := \mathfrak{k} \oplus i\mathfrak{k} = \mathfrak{so}_n(\mathbb{C})$, consider the orbit of A under the adjoint action of K

$$\mathrm{Ad}\, K \cdot A := \{\mathrm{Ad}\,(k)A : k \in K\}.$$

So

$$S(A) = \{\mathrm{tr}\, CY : Y \in \mathrm{Ad}\, K \cdot A\}$$

where $C = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \oplus 0_{n-2}$. The orbit $\mathrm{Ad}\, K \cdot A$ depends on $\mathrm{Ad}\ K$ which is the analytic subgroup of the adjoint group $\mathrm{Int}\, \mathfrak{k} \subset \mathrm{Aut}\, \mathfrak{k}$ corresponding to $\mathrm{ad}\, \mathfrak{k}$ [8, p.126, p.129]. Thus $\mathrm{Ad}\, K \cdot A$ is independent of the choice of K. In particular we can pick the simply connected $\widetilde{\mathrm{SO}}(n)$.

(1) The following is a Lie algebra isomorphism $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_3(\mathbb{C})$ given by

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} \mapsto \begin{pmatrix} 0 & -2ia & i(b+c) \\ 2ia & 0 & c-b \\ -i(b+c) & b-c & 0 \end{pmatrix}$$

and its restriction $\psi : \mathfrak{su}(2) \to \mathfrak{so}(3)$ is a Lie algebra isomorphism. Thus there is a Lie group isomorphism [23, p.101] $\varphi : \widetilde{\mathrm{SU}}(2) \to \widetilde{\mathrm{SO}}(3)$ so that $d\varphi_e = \psi$ which naturally extends to $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_3(\mathbb{C})$ (indeed there is a double covering $\mathrm{SU}(2) \to \mathrm{SO}(3)$). So we have the relation

$$\begin{aligned} & d\varphi_e\{UAU^{-1} : U \in \mathrm{SU}(2)\} \\ = \ & d\varphi_e\{UAU^{-1} : U \in \widetilde{\mathrm{SU}}(2)\} \\ = \ & \{\varphi(U)(d\varphi_e(A))\varphi^{-1}(U) : U \in \widetilde{\mathrm{SU}}(2)\} \\ = \ & \{O(d\varphi_e(A))O^{-1} : O \in \widetilde{\mathrm{SO}}(3)\} \\ = \ & \{O(d\varphi_e(A))O^{-1} : O \in \mathrm{SO}(3)\}. \end{aligned}$$

where the second equality is due to the fact that

$$d\varphi_e(\mathrm{Ad}\,(g)A) = \mathrm{Ad}\,(\varphi(g))d\varphi_e(A), \quad A \in \mathfrak{sl}_2(\mathbb{C}),$$

and that the adjoint action is conjugation for matrix group. The numerical range map $x \mapsto x^*Ax$, $x \in \mathbb{S}^1$, amounts to $U^{-1}AU \mapsto (U^{-1}AU)_{11}$, $U \in \mathrm{SU}(2)$ and thus corresponds to $-\frac{1}{2i}(\varphi^{-1}(U)d\varphi_e(A)\varphi(U))_{12}$, i.e., $(x,y) \mapsto x^T d\varphi_e(A)y$, where $x, y \in \mathbb{R}^3$ are orthonormal. Thus

$$W\begin{pmatrix} a & b \\ c & -a \end{pmatrix} = -\frac{1}{2i} S \begin{pmatrix} 0 & -2ia & i(b+c) \\ 2ia & 0 & c-b \\ -i(b+c) & b-c & 0 \end{pmatrix}.$$

Since ψ is an isomorphism, the converse is true.

(2) The following is a Lie algebra isomorphism $\psi : \mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_4(\mathbb{C})$:

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} \oplus \begin{pmatrix} d & e \\ f & -d \end{pmatrix} \mapsto \begin{pmatrix} 0 & i(a-d) & \frac{1}{2}(c-b-f+e) & \frac{1}{2}i(b+c-e-f) \\ -i(a-d) & 0 & \frac{1}{2}i(b+c+e+f) & \frac{-1}{2}(c-b+f-e) \\ \frac{-1}{2}(c-b-f+e) & \frac{-1}{2}i(b+c+e+f) & 0 & i(a+d) \\ \frac{-1}{2}i(b+c-e-f) & \frac{1}{2}(c-b+f-e) & -i(a+d) & 0 \end{pmatrix}$$

and thus its restriction $\psi : \mathfrak{su}(2) \oplus \mathfrak{su}(2) \to \mathfrak{so}(4)$ is a Lie algebra isomorphism. Thus there is a Lie group isomorphism [23, p.101] $\varphi : \widetilde{\mathrm{SU}}(2) \times \widetilde{\mathrm{SU}}(2) \to \widetilde{\mathrm{SO}}(4)$ so that $d\varphi_e = \psi$

(indeed there is a double covering map $\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$ [19, p.42]). So we have the relation

$$\begin{aligned} & d\varphi_e\{UAU^{-1} : U \in \mathrm{SU}(2) \times \mathrm{SU}(2)\} \\ = \; & d\varphi_e\{UAU^{-1} : U \in \widetilde{\mathrm{SU}}(2) \times \widetilde{\mathrm{SU}}(2)\} \\ = \; & \{\varphi(U)(d\varphi_e(A))\varphi(U)^{-1} : U \in \widetilde{\mathrm{SU}}(2) \times \widetilde{\mathrm{SU}}(2)\} \\ = \; & \{O(d\varphi_e(A))O^{-1} : O \in \widetilde{\mathrm{SO}}(4)\} \\ = \; & \{O(d\varphi_e(A))O^{-1} : O \in \mathrm{SO}(4)\}. \end{aligned}$$

where the second equality is due to the fact that

$$d\varphi_e(\mathrm{Ad}\,(g)A) = \mathrm{Ad}\,(\varphi(g))d\varphi_e(A), \quad A \in \mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C})$$

and that the adjoint action is conjugation for matrix group. The map $UAU^{-1} \mapsto (UAU^{-1})_{11} + (UAU^{-1})_{44}$, $U \in \mathrm{SU}(2) \times \mathrm{SU}(2)$ corresponds to $\frac{1}{i}(\varphi(U)(d\varphi_e(A))\varphi^{-1}(U))_{12}$, i.e., $(x, y) \mapsto x^T d\varphi_e(A)y$ where $x, y \in \mathbb{R}^4$ are orthonormal. So

$$\begin{aligned} & W\left(\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right) + W\left(\begin{smallmatrix} d & e \\ f & -d \end{smallmatrix}\right) \\ = \; & \frac{1}{i} S \begin{pmatrix} 0 & i(a-d) & \frac{1}{2}(c-b-f+e) & \frac{1}{2}i(b+c-e-f) \\ -i(a-d) & 0 & \frac{1}{2}i(b+c+e+f) & \frac{-1}{2}(c-b+f-e) \\ \frac{-1}{2}(c-b-f+e) & \frac{-1}{2}i(b+c+e+f) & 0 & i(a+d) \\ \frac{-1}{2}i(b+c-e-f) & \frac{1}{2}(c-b+f-e) & -i(a+d) & 0 \end{pmatrix}. \end{aligned}$$

Notice that $W\left(\begin{smallmatrix} a & b \\ c & -a \end{smallmatrix}\right) + W\left(\begin{smallmatrix} d & e \\ f & -d \end{smallmatrix}\right)$ is simply the sum of two elliptical disks centered at the origin. Since ψ is an isomorphism, the converse is true.

(4) Following the idea in Knapp [12, p.162], direct computation yields an isomorphism between $\mathfrak{sl}_4(\mathbb{C})$ and $\mathfrak{so}_6(\mathbb{C})$:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & -a_{11} - a_{22} - a_{33} \end{pmatrix} \mapsto$$

$$\begin{pmatrix} 0 & \frac{(a_{23}+a_{14}-a_{41}-a_{32})}{2} & \frac{(a_{24}-a_{13}+a_{31}-a_{42})}{2} & i(a_{11}+a_{22}) & \frac{i(a_{23}-a_{14}-a_{41}+a_{32})}{2} & \frac{i(a_{24}+a_{13}+a_{31}+a_{42})}{2} \\ -\frac{(a_{23}+a_{14}-a_{41}-a_{32})}{2} & 0 & \frac{a_{34}+a_{12}-a_{21}-a_{43}}{2} & \frac{i(a_{23}+a_{14}+a_{41}+a_{32})}{2} & i(a_{11}+a_{33}) & \frac{i(a_{34}-a_{12}-a_{21}+a_{43})}{2} \\ -\frac{(a_{24}-a_{13}+a_{31}-a_{42})}{2} & -\frac{(a_{34}+a_{12}-a_{21}-a_{43})}{2} & 0 & \frac{i(a_{24}-a_{13}-a_{31}+a_{42})}{2} & \frac{i(a_{34}+a_{12}+a_{21}+a_{43})}{2} & -i(a_{22}+a_{33}) \\ -i(a_{11}+a_{22}) & -\frac{i(a_{23}+a_{14}+a_{41}+a_{32})}{2} & -\frac{i(a_{24}-a_{13}-a_{31}+a_{42})}{2} & 0 & \frac{(a_{23}-a_{14}+a_{41}-a_{32})}{2} & \frac{(a_{24}+a_{13}-a_{31}-a_{42})}{2} \\ -\frac{i(a_{23}-a_{14}-a_{41}+a_{32})}{2} & -i(a_{11}+a_{33}) & -\frac{i(a_{34}+a_{12}+a_{21}+a_{43})}{2} & -\frac{(a_{23}-a_{14}+a_{41}-a_{32})}{2} & 0 & \frac{(a_{34}-a_{12}+a_{21}-a_{43})}{2} \\ -\frac{i(a_{24}+a_{13}+a_{31}+a_{42})}{2} & -\frac{i(a_{34}-a_{12}-a_{21}+a_{43})}{2} & i(a_{22}+a_{33}) & -\frac{(a_{24}+a_{13}-a_{31}-a_{42})}{2} & -\frac{(a_{34}-a_{12}+a_{21}-a_{43})}{2} & 0 \end{pmatrix}$$

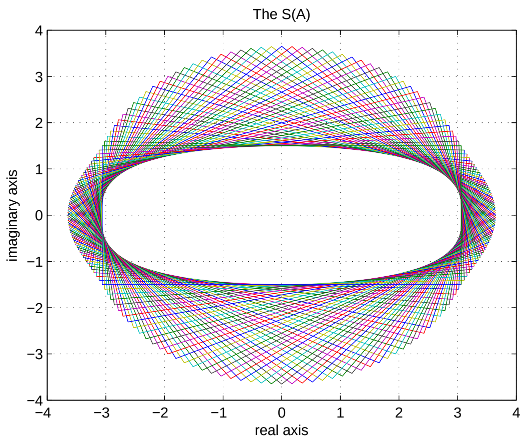
Then apply the same argument in (1) to have the desired result.

(3) When $A \in \mathfrak{sp}_4(\mathbb{C})$, the fifth row and column are zero so that the map in (4) yields an isomorphism of $\mathfrak{sp}_4(\mathbb{C})$ and $\mathfrak{so}_5(\mathbb{C})$. Then apply the argument in (1). $\square$

Example 3.2. Let

$$A = \begin{pmatrix} 0 & -2 & \frac{i}{2} & \frac{1}{2} \\ 2 & 0 & \frac{3}{2} & -\frac{3i}{2} \\ -\frac{i}{2} & -\frac{3}{2} & 0 & 2 \\ -\frac{1}{2} & \frac{3i}{2} & -2 & 0 \end{pmatrix}$$

The following is the plot of $S(A)$:



Indeed by Theorem 3.1(2), it is the sum of the unit disk $W\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ and $W\begin{pmatrix} 2 & 1 \\ 0 & -2 \end{pmatrix}$ which is the elliptical disk with foci ± 2 and the length of minor axis is 1.

The following provides a description of the elliptical disk $S(A)$ in terms of the entries of A.

Proposition 3.3. $S\begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \gamma \\ -\beta & \gamma & 0 \end{pmatrix}$ is an elliptical disk (possibly degenerate) with foci

$$\pm\sqrt{\alpha^2 + \beta^2 + \gamma^2}$$

and minor axis of length

$$\ell := \sqrt{2|\alpha|^2 + 2|\beta|^2 + 2|\gamma|^2 - 2|\alpha^2 + \beta^2 + \gamma^2|}.$$

In particular, it is a line segment (possibly degenerate) if and only if $0, \alpha, \beta, \gamma$ are collinear.

Proof. This follows from the proof of Theorem 3.1(1) and the elliptical range theorem [13] for the classical numerical range which asserts that given $A \in \mathbb{C}_{2\times 2}$, $W(A)$ is an elliptical disk with eigenvalues λ_1, λ_2 as foci and minor axis of length $\sqrt{\mathrm{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}$. The rest is straightforward computation. □

4. Basic properties

The following is a collection of some basic properties of $S(A)$.

Theorem 4.1. Let $A, B \in \mathfrak{so}_n(\mathbb{C})$ and $C \in \mathfrak{so}_m(\mathbb{C})$. Then

(1) $S(A)$ is compact and convex.
(2) When $n \geq 3$, $S(\operatorname{Re} A) = \operatorname{Re} S(A) = [-\sigma_1(\operatorname{Re} A), \sigma_1(\operatorname{Re} A)]$, where $\operatorname{Re} A$ denotes the real part of A and $\operatorname{Re} S(A)$ denotes the real part of the set $S(A)$ and $\sigma_1(\operatorname{Re} A)$ is the largest singular value of $\operatorname{Re} A$. Moreover, $\frac{1}{2}\sigma_1(A) \leq s(A) \leq \sigma_1(A)$.
(3) When $n \geq 3$, $S(A)$ is symmetric about the origin.
(4) $S(\alpha A) = \alpha S(A)$ for all $\alpha \in \mathbb{C}$.
(5) $S(A + B) \subset S(A) + S(B)$.
(6) $S(O^T AO) = S(A)$ for all $O \in \mathrm{SO}(n)$.
(7) $S(A) \subset S(C)$, if A is a principal submatrix of C.
(8) $S(A \oplus C) = \operatorname{conv}\{\pm S(A), \pm S(C)\}$. In particular,
   (a) if $\min\{m, n\} \geq 3$, then $S(A \oplus C) = \operatorname{conv}\{S(A), S(C)\}$.
   (b) if $A = \begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix}$ and $C = \begin{pmatrix} 0 & \beta \\ -\beta & 0 \end{pmatrix}$, then $S(A \oplus C) = \operatorname{conv}\{\pm\alpha, \pm\beta\}$.

Proof. (1) We provide a proof different from [20] and Theorem 2.4. It is based on Theorem 3.1. The cases $n = 3, 4$ are proved in Theorem 3.1. Consider $n > 4$. Let $w_1 = x_1^T A y_1$ and $w_2 = x_2^T A y_2$ be two distinct points in $S(A)$, where x_1, y_1 and x_2, y_2 are orthonormal pairs in $\mathbb{R}^n$. Let $\hat{A} : E \to E$ be the compression of A onto $E := \operatorname{span}\{x_1, x_2, y_1, y_2\}$. Then the matrix $\hat{A}$ is complex skew symmetric and $S(\hat{A})$ contains w_1 and w_2. Since w_1, w_2 are distinct, $2 \leq \dim E \leq 4$.
Case 1: $3 \leq \dim E \leq 4$. By Theorem 3.1, $S(\hat{A})$ is convex and hence contains the line segment $[w_1, w_2]$. So does $S(A)$ since $S(\hat{A}) \subset S(A)$.
Case 2: $\dim E = 2$. Then $\operatorname{span}\{x_1, y_1\} = \operatorname{span}\{x_2, y_2\}$ so that $w_1 = -w_2 \neq 0$ by Lemma 2.3(1). Pick $x_3 \in \mathbb{R}^n$ such that $x_3 \notin \operatorname{span}\{x_1, y_1\}$ since $n > 4$. Apply the previous argument on $E' := \operatorname{span}\{x_1, x_3, y_1\}$ to have the desired result.

(2) Clearly $S(\operatorname{Re} A) = \operatorname{Re} S(A)$. It remains to show that if $D \in \mathfrak{so}(n)$, then

$$\max\{x^T D y : x, y \in \mathbb{R}^n \text{ are orthonormal}\} = \sigma_1(D).$$

This is true because there is $O \in \mathrm{O}(n)$ [9, p.107]

$$O^T D O = \begin{cases} \sigma_1(D) J \oplus \cdots \oplus \sigma_{n/2}(D) J & \text{if } n \text{ is even} \\ \sigma_1(D) J \oplus \cdots \oplus \sigma_{(n-1)/2}(D) J \oplus 0 & \text{if } n \text{ is odd,} \end{cases}$$

where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\sigma(D)$'s are the singular values of D. Clearly $s(A) \leq \sigma_1(A)$ and

$$\sigma_1(A) = \|A\|_2 \leq \|\operatorname{Re} A\|_2 + \|\operatorname{Im} A\|_2 = \sigma_1(\operatorname{Re} A) + \sigma_1(\operatorname{Im} A) = s(\operatorname{Re} A) + s(\operatorname{Im} A) \leq 2s(A),$$

where $\| \cdot \|_2$ denotes the spectral norm of a matrix.

(3), (4), (5), (6) and (7) are straightforward.

(8) Clearly $S(A \oplus C) \supset S(A)$ since $x^T A y = (x^T \ 0^T)(A \oplus C)\begin{pmatrix} y \\ 0 \end{pmatrix}$, and similarly $S(A \oplus C) \supset S(C)$. Thus $S(A \oplus C) \supset \operatorname{conv}\{\pm S(A), \pm S(C)\}$ because $S(A \oplus C)$ is convex and symmetric about the origin.

To establish the converse inclusion, write the orthonormal pair $x, y \in \mathbb{R}^{n+m}$ in partitioned form

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

and then write $y_i = t_i x_i + y_i'$ ($t_i \in \mathbb{R}$) so that $x_i^T y_i' = 0$ for $i = 1, 2$. Then

$$\begin{aligned} x^T(A \oplus C)y &= x_1^T A y_1 + x_2^T C y_2 \\ &= \|x_1\|_2 \|y_1'\|_2 \frac{x_1^T A y_1'}{\|x_1\|_2 \|y_1'\|_2} + \|x_2\|_2 \|y_2'\|_2 \frac{x_2^T C y_2'}{\|x_2\|_2 \|y_2'\|_2}. \end{aligned}$$

Clearly

$$\alpha := \frac{x_1^T A y_1'}{\|x_1\|_2 \|y_1'\|_2} \in S(A), \quad \beta := \frac{x_2^T C y_2'}{\|x_2\|_2 \|y_2'\|_2} \in S(C).$$

By Cauchy-Schwarz's inequality

$$\|x_1\|_2 \|y_1'\|_2 + \|x_2\|_2 \|y_2'\|_2 \le \|x_1\|_2 \|y_1\|_2 + \|x_2\|_2 \|y_2\|_2 \le \|x\|_2 \|y\|_2 = 1.$$

So $x^T(A \oplus C)y \in \operatorname{conv}\{\alpha, \beta, 0\} \subset \operatorname{conv}\{\pm S(A), \pm S(C)\}$. □

5. Axiomatic characterization and continuity of $S(\cdot)$

We now view $S : \mathfrak{so}_n(\mathbb{C}) \to \mathcal{P}(\mathbb{C})$ as a function, where $\mathcal{P}(\mathbb{C})$ denotes the power set of $\mathbb{C}$. We have the following axiomatic characterization of $S(\cdot)$ alike to C.R. Johnson's [11] for $W(\cdot)$. The convexity of $S(A)$ plays a very important role in the proof.

Theorem 5.1. Let $n \ge 3$. Then $S : \mathfrak{so}_n(\mathbb{C}) \to \mathcal{P}(\mathbb{C})$ is the unique function such that

(1) $S(A) \subset \mathbb{C}$ is compact and convex.
(2) $S(\alpha A) = \alpha S(A)$ for all $\alpha \in \mathbb{C}$.
(3) $S(\operatorname{Re} A) = \operatorname{Re} S(A) = [-\sigma_1(\operatorname{Re} A), \sigma_1(\operatorname{Re} A)]$, where $\operatorname{Re} A$ denotes the real part of A and $\operatorname{Re} S(A)$ denotes the real part of the set $S(A)$ and $\sigma_1(\operatorname{Re} A)$ is the largest singular value of $\operatorname{Re} A$.

Proof. Suppose S_1 and S_2 are the two complex set-valued functions on $\mathfrak{so}_n(\mathbb{C})$ satisfying the conditions. Let $A \in \mathfrak{so}_n(\mathbb{C})$. We are going to show that $S_1(A) = S_2(A)$. Suppose on the contrary $S_1(A) \not\subset S_2(A)$, i.e., there is $\beta \in S_1(A) \setminus S_2(A)$. Because of (1), by the separation theorem, there is $\theta \in \mathbb{R}$ such that

$$\operatorname{Re}(e^{i\theta}\beta) \notin \operatorname{Re}(e^{i\theta} S_2(A)) = S_2(\operatorname{Re} e^{i\theta} A) = [-\sigma_1(\operatorname{Re} e^{i\theta} A), \sigma_1(\operatorname{Re} e^{i\theta} A)]$$

where the equalities are from (2) and (3). However

$$\operatorname{Re}(e^{i\theta}\beta) \in S_1(\operatorname{Re} e^{i\theta} A) = [-\sigma_1(\operatorname{Re} e^{i\theta} A), \sigma_1(\operatorname{Re} e^{i\theta} A)],$$

a contradiction. □

Alike to [7, p.117] we now discuss the continuity of S when the range is the set $\mathcal{C}$ of all compact sets of $\mathbb{C}$. We first equip $\mathbb{C}$ with the Hausdorff metric. For each $M \subset \mathbb{C}$ and $\epsilon > 0$, define

$$M + (\epsilon) = \{z + \alpha : z \in M, |\alpha| < \epsilon\}$$

If $M, N \subset \mathbb{C}$ are compact subsets, the Hausdorff metric $d(M, N)$ is the infimum of all positive numbers ϵ such that both $M \subset N + (\epsilon)$ and $N \subset M + (\epsilon)$.

Theorem 5.2. The function $S : \mathfrak{so}_n(\mathbb{C}) \to \mathcal{C}$, where $\mathcal{C}$ is the set of all compact convex sets in $\mathbb{C}$ endowed with the Hausdorff metric, is continuous with respect to the operator norm.

Proof. If $\|A - B\| < \epsilon$, then $|x^T(A - B)y| < \epsilon$ and

$$x^T Ay = x^T By + x^T(A - B)y \in S(B) + (\epsilon).$$

It follows that $S(A) \subset S(B) + (\epsilon)$ and by symmetry $S(B) \subset S(A) + (\epsilon)$. □

6. Sharp points and extreme points of $S(A)$

Let $K \subset \mathbb{C}$ be a compact convex set. A point $\xi \in K$ is called an extreme point if ξ is not in any open line segment that is contained in K. A point $\xi \in K$ is a sharp point if ξ is the intersection point of two distinct supporting lines of K [10, p.50]. Clearly sharp points are extreme points.

Donoghue [5] showed that sharp points of $W(A)$ are eigenvalues of A. Indeed the following is a characterization of the sharp points of $W(A)$.

Theorem 6.1. ([10, p.50-51], [18]) Let $A \in \mathbb{C}_{n\times n}$ and $\xi \in W(A)$. Then ξ is a sharp point if and only if A is unitarily similar to $\xi I \oplus B$ with $\xi \notin W(B)$.

Analogous to Donoghue's result, we are going to study the sharp points of $S(A)$ for $A \in \mathfrak{so}_n(\mathbb{C})$.

Theorem 6.2. Let $A \in \mathfrak{so}_n(\mathbb{C})$. When $n \geq 3$, λ is a sharp point of $S(A)$ if and only if A is orthogonally similar to $\lambda J \oplus \cdots \oplus \lambda J \oplus B$, with $\lambda \notin \operatorname{conv}\{\pm S(B)\}$, where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and B is skew symmetric.

Proof. We may assume that $A \neq 0$, i.e., $S(A)$ is not trivial. Suppose that A is orthogonally similar to $\lambda J \oplus \cdots \oplus \lambda J \oplus B$ with $\lambda \notin \operatorname{conv}\{\pm S(B)\}$. By Theorem 4.1,

$$S(A) = \operatorname{conv}\{\pm\lambda, \pm S(B)\}.$$

Since λ is not contained in the compact convex set $\operatorname{conv}\{\pm S(B)\}$, λ is a sharp point of $S(A)$.

Conversely suppose $\lambda \in S(A)$ is a sharp point and we denote $A = (a_{ij})$. Without loss of generality, we may assume that $\lambda = a_{12}$ by orthogonal similarity. For each $k = 3, \ldots, n$, the 3×3 principal submatrix $A_k := \begin{pmatrix} 0 & \lambda & a_{1k} \\ -\lambda & 0 & a_{2k} \\ -a_{1k} & -a_{2k} & 0 \end{pmatrix}$ of A satisfies $S(A_k) \subset S(A)$ and $S(A_k)$ is an elliptical disk (possibly degenerate). Since $\lambda \neq 0$ is a sharp point of $S(A)$, $S(A_k)$ must be the line segment $[-\lambda, \lambda]$. Hence $a_{1k} = a_{2k} = 0$ by Proposition 3.3. Thus $A = \lambda J \oplus \hat{A}$ with $\hat{A} \in \mathbb{C}_{(n-2)\times(n-2)}$. If $n - 2 \leq 2$ or $\lambda \notin S(\hat{A})$, we are done; otherwise λ must be a sharp point of $S(\hat{A})$, and we can repeat the above argument until we get the desired form. □

Because of Theorem 2.4, we have the following notions.

Definition 6.3. Let $A \in \mathfrak{so}_n(\mathbb{C})$ and $\xi \in S(A)$. Define

$$M_A(\xi) := \{t \begin{pmatrix} x \\ y \end{pmatrix} : t \in \mathbb{R},\ x^T Ay = \xi,\ x, y \in \mathbb{R}^n \text{ are orthonormal}\} \subset \mathbb{R}^{2n}$$

and

$$M'_A(\xi) := \{t \begin{pmatrix} x \\ y \end{pmatrix} : t \in \mathbb{R},\ x^T Ay = \xi,\ x, y \in \mathbb{R}^n \text{ are unit vectors}\} \subset \mathbb{R}^{2n}$$

and

$$\tilde{M}_A(\xi) := \{t \begin{pmatrix} x \\ y \end{pmatrix} : t \in \mathbb{R},\ x^T Ay = \xi,\ x, y \in \mathbb{R}^n, \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|_2^2 = 2\} \subset \mathbb{R}^{2n}$$

Notice that $M_A(\xi)$ is not a subspace in general, but it is homogenous. Moreover

$$M_{OAO^T}(\xi) = OM_A(\xi), \quad O \in \mathrm{O}(n) \tag{6.1}$$

and

$$M_{\alpha A}(\alpha\xi) = M_A(\xi), \quad \alpha \in \mathbb{C}. \tag{6.2}$$

We also remark that $\begin{pmatrix} x \\ y \end{pmatrix} \in M_A(\xi)$ if and only if $\begin{pmatrix} y \\ -x \end{pmatrix} \in M_A(\xi)$. It is the same for $M'_A(\xi)$ and $\tilde{M}_A(\xi)$. We are going to characterize the extreme points of $S(A)$ using $M_A(\xi)$ alike Embry's characterization for the classical numerical range [6]; also see [3].

Lemma 6.4. *Let $A \in \mathfrak{so}(n)$ with $n \geq 3$ and let $\sigma_1(A)$ be the largest singular value of A. If $x^T Ay = \sigma_1(A)$ with $x, y \in \mathbb{S}_{\mathbb{R}}^{n-1}$, then x and y are orthonormal and $x + iy$ is an eigenvector of A corresponding to $i\sigma_1(A)$, i.e., $Ax = -\sigma_1 y$ and $Ay = \sigma_1 x$.*

Proof. If x and y were not orthonormal, then $y = tx + \sqrt{1-t^2}y'$ where $y' \in x^\perp$ is a unit vector and $0 < |t| \leq 1$. Then

$$x^T Ay = x^T A(tx + \sqrt{1-t^2}y') = \sqrt{1-t^2}x^T Ay' < x^T Ay' \leq \sigma_1(A)$$

since $\sigma_1(A) = \max_{x,y\in\mathbb{S}^{n-1}} x^T Ay$. Use the equality case of Cauchy-Schwarz's inequality $\sigma_1(A) = x^T Ay \leq \|x\|_2 \|Ay\|_2 \leq \sigma_1(A)$ to conclude that x and Ay are collinear and thus $Ax = -\sigma_1 y$ and $Ay = \sigma_1 x$. □

Theorem 6.5. Let $A \in \mathfrak{so}_n(\mathbb{C})$ with $n \geq 3$ and $\xi \in S(A)$. Then ξ is an extreme point of $S(A)$ if and only if ξ is a boundary point and $M_A(\xi)$ is a subspace of $\mathbb{R}^{2n}$. In this event $M_A(\xi) = M'_A(\xi) = \tilde{M}_A(\xi)$.

Proof. We may assume that $A \neq 0$. All extreme points of $S(A)$ are boundary points. Suppose that $\xi = \xi_1 + i\xi_2$ ($\xi_1, \xi_2 \in \mathbb{R}$) is an extreme point of $S(A)$. Since $S(A)$ is convex, there is a supporting line L of $S(A)$ at ξ. Let $p := p_1 + ip_2 \in \mathbb{C}$ ($p_1, p_2 \in \mathbb{R}$) be a unit vector perpendicular to L. Project $S(A)$ onto the line $\mathbb{R}p$. Write $A = A_1 + iA_2$, where $A_1, A_2 \in \mathfrak{so}(n)$. Then $\hat{\xi} := p_1\xi_1 + p_2\xi_2 \in \mathbb{R}$ is an end point of the line segment

$$S(\hat{A}) = [-\sigma_1(\hat{A}), \sigma_1(\hat{A})],$$

where

$$\hat{A} := p_1A_1 + p_2A_2 \in \mathfrak{so}(n) \tag{6.3}$$

Clearly we have $M_A(\xi) = M_{\hat{A}}(\hat{\xi})$. For definiteness, we assume $\hat{\xi} = \sigma_1(\hat{A})$. By the spectral theorem for real skew symmetric matrices, there is $O \in \mathrm{O}(n)$ such that

$$O\hat{A}O^T = A' := \begin{cases} \sigma_1(\hat{A})J \oplus \cdots \oplus \sigma_{n/2}(\hat{A})J & \text{if } n \text{ is even,} \\ \sigma_1(\hat{A})J \oplus \cdots \oplus \sigma_{(n-1)/2}(\hat{A})J \oplus 0 & \text{if } n \text{ is odd,} \end{cases}$$

where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and $\sigma(\hat{A})$'s are the singular values of $\hat{A}$. We may assume that $\hat{A} = A'$ since $M_{OAO^T}(\xi) = OM_A(\xi)$. So $x^TAy = \sigma_1(\hat{A})$ implies that $[x\ y] = [e_{2j-1}\ e_{2j}]R(\theta)$, where $R(\theta)$ is some rotation matrix in $\mathrm{SO}(2)$ for some $j = 1, \ldots k$ and k is the multiplicity of $\sigma_1(\hat{A})$. Thus if $\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix} \in M_{\hat{A}}(\hat{\xi})$, i.e., $x^T\hat{A}y = u^T\hat{A}v = \sigma_1(\hat{A})$ with orthonormal pairs $x, y \in \mathbb{R}^n$ and $u, v \in \mathbb{R}^n$, then either

(a) $[u\ v] = [x\ y]R(\theta)$, for some $R(\theta) \in \mathrm{SO}(2)$, or
(b) $k \geq 2$ and $\operatorname{span}\{x, y\} \perp \operatorname{span}\{u, v\}$.

If (a) happens, then for $j = 1, \ldots, k$, $x = e_{2j-1}$ and $y = e_{2j}$ so that $u = x\cos\theta - y\sin\theta$, $v = x\sin\theta + y\cos\theta$. Hence $x + u$ and $y + v$ are orthogonal,

$$\|x + u\|_2^2 = \|y + v\|_2^2 = 2(1 + \cos\theta),$$

and $(x+u)^T\hat{A}(y+v) = 2(1+\cos\theta)x^TAy = 2(1+\cos\theta)\xi$. So $\begin{pmatrix} x+u \\ y+v \end{pmatrix} \in M_{\hat{A}}(\hat{\xi})$ and thus $M_{\hat{A}}(\hat{\xi})$ is a subspace of $\mathbb{R}^{2n}$.

If (b) happens, then $x + u$ and $y + v$ are orthogonal and $\|x + u\|_2^2 = \|y + v\|_2^2 = 2$. Moreover for some $1 \leq j \neq i \leq k$, $x = e_{2j-1}$, $y = e_{2j}$, $u = e_{2i-1}$, $v = e_{2i}$ so that $(x + u)^T\hat{A}(y + v) = 2\xi$. So $\begin{pmatrix} x+u \\ y+v \end{pmatrix} \in M_{\hat{A}}(\hat{\xi})$ and thus $M_{\hat{A}}(\hat{\xi})$ is a subspace of $\mathbb{R}^{2n}$.

Conversely, suppose that $\xi \in S(A)$ is a boundary point and $M_A(\xi)$ is a subspace of $\mathbb{C}^n$. If ξ were not an extreme point, there would exist distinct $\alpha, \beta \in S(A)$ such that

$\xi \in (\alpha, \beta)$. Let $x, y \in \mathbb{R}^n$ and $u, v \in \mathbb{R}^n$ be orthonormal pairs such that $\alpha = x^T Ay$, $\beta = u^T Av$ and let $\hat{A}$ denote the compression of A onto $E := \operatorname{span}\{x, y, u, v\}$ over $\mathbb{C}$. We may assume that $\beta \neq \pm\alpha$. Hence $\dim E = 3$ or 4. Clearly $\xi \in (\alpha, \beta) \subset S(\hat{A}) \subset S(A)$ and ξ is a boundary point of $S(\hat{A})$. We consider the following two cases:

(a) $\dim E = 3$. By Proposition 3.3 and Theorem 4.1(7), $S(\hat{A})$ has to be a line segment (but not a point since $\alpha \neq \beta$) $[-\gamma, \gamma]$ passing through the origin. Moreover $\hat{A}$ is orthogonally similar to $A' = \begin{pmatrix} 0 & \gamma \\ -\gamma & 0 \end{pmatrix} \oplus 0$. Since $M_{O\hat{A}O^T}(\xi) = OM_{\hat{A}}(\xi)$ for any $O \in \mathrm{O}(3)$ we may assume that $\hat{A} = A'$. Since $M_{\alpha\hat{A}}(\alpha\xi) = M_{\hat{A}}(\xi)$ for any $\alpha \in \mathbb{C}$, with appropriate rotation, we may assume that $\gamma > 0$ and $\xi \geq 0$. Notice that

$$e_1^T \hat{A}(\frac{\xi}{\gamma} e_2 \pm \sqrt{1 - (\frac{\xi}{\gamma})^2} e_3) = \xi$$

so that

$$\begin{pmatrix} e_1 \\ \frac{\xi}{\gamma} e_2 \pm \sqrt{1 - (\frac{\xi}{\gamma})^2} e_3 \end{pmatrix} \in M_{\hat{A}}(\xi) \subset M_A(\xi).$$

However their difference $\begin{pmatrix} 0 \\ 2\sqrt{1 - (\frac{\xi}{\gamma})^2} e_3 \end{pmatrix} \notin M_A(\xi)$.

(b) $\dim E = 4$. By Theorem 3.1(2) $S(\hat{A})$ is the sum of an elliptical disk (possibly degenerate) $\mathcal{E} = W\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ and the line segment $W\begin{pmatrix} d & 0 \\ 0 & -d \end{pmatrix}$ $(d \neq 0)$. Moreover $\hat{A}$ is orthogonally similar to

$$A' = \frac{1}{i}\begin{pmatrix} 0 & i(a-d) & \frac{1}{2}(c-b) & \frac{1}{2}i(b+c) \\ -i(a-d) & 0 & \frac{1}{2}i(b+c) & \frac{-1}{2}(c-b) \\ \frac{-1}{2}(c-b) & \frac{-1}{2}i(b+c) & 0 & i(a+d) \\ \frac{-1}{2}i(b+c) & \frac{1}{2}(c-b) & -i(a+d) & 0 \end{pmatrix}.$$

Since $M_{O\hat{A}O^T}(\xi) = OM_{\hat{A}}(\xi)$ for any $O \in \mathrm{O}(4)$ we may assume that $A' = \hat{A}$. With appropriate rotation we may assume that $d > 0$. So $\xi = \xi_1 + \xi_2$, where $\xi_1 \in \partial W\begin{pmatrix} a & b \\ c & -a \end{pmatrix}$ (either the highest or the lowest point of the ellipse $\mathcal{E}$) and ξ_2 is in the open interval $(-d, d)$. Without loss of generality we may assume that $\pm a$ are the highest and lowest points of $\mathcal{E}$ and $\xi_1 = a$. Then

$$\begin{aligned} & (e_1 \sin\theta + e_3 \cos\theta)^T A'(e_2 \sin\theta + e_4 \cos\theta) \\ = \; & a + (\cos^2\theta - \sin^2\theta)d + \frac{b+c}{2}\cos\theta\sin\theta - \frac{b+c}{2}\cos\theta\sin\theta \\ = \; & a + (\cos 2\theta)d. \end{aligned}$$

There are two distinct $\theta_1, \theta_2 \in [0, \pi]$ such that $d \cos 2\theta = \xi$. So

$$z_k := \begin{pmatrix} e_1 \sin\theta_k + e_3 \cos\theta_k \\ e_2 \sin\theta_k + e_4 \cos\theta_k \end{pmatrix} \in M_{\hat{A}}(\xi) \subset M_A(\xi),$$

$k = 1, 2$, but the linear combination

$$z_1 \sin\theta_2 - z_2 \sin\theta_1 = \sin(\theta_2 - \theta_1) \begin{pmatrix} e_3 \\ e_4 \end{pmatrix} \notin M_A(\xi)$$

since $\theta_1 \neq \theta_2$ and $e_3^T \hat{A} e_4 = a + d \neq \xi$.

We already established the equivalence of the two conditions. Suppose $\xi \in S(A)$ is an extreme point. One can see that $M'_A(\xi) = M'_{\hat{A}}(\hat{\xi})$ and $\tilde{M}_A(\xi) = \tilde{M}_{\hat{A}}(\hat{\xi})$, where $\hat{A}$ is given in (6.3) and $\hat{\xi} = \sigma_1(\hat{A})$. If $x, y \in \mathbb{R}^n$ with $\|x\|_2^2 + \|y\|_2^2 = 2$, then by Lemma 2.4

$$\hat{\xi} = x^T \hat{A} y = \|x\|_2 \|y\|_2 (\frac{x}{\|x\|_2})^T \hat{A} (\frac{y}{\|y\|_2}) \leq (\frac{x}{\|x\|_2})^T \hat{A} (\frac{y}{\|y\|_2}) \in S(\hat{A})$$

since $2\|x\|_2\|y\|_2 \leq \|x\|_2^2 + \|y\|_2^2 = 1$. The fact that $\hat{\xi} = \sigma_1(\hat{A})$ forces the inequality being an equality so that $\|x\|_2 = \|y\|_2 = 1$. Write $y = \alpha x + \beta y'$, where $\beta \geq 0$ and y' is a unit vector orthogonal to x with $\alpha^2 + \beta^2 = 1$. Thus $\hat{\xi} = x^T \hat{A} y = \beta x^T \hat{A} y'$ and $x^T \hat{A} y' \in S(\hat{A})$. So $\beta = 1$ and we have $x \perp y$. We just proved $M_A(\xi) = M'_A(\xi) = \tilde{M}_A(\xi)$. □

7. q-NUMERICAL RANGE

We remark that for any $A \in \mathbb{C}_{n\times n}$, $q \in \mathbb{C}$ a fixed number such that $0 \leq |q| \leq 1$, Tsing's result [22] asserts that the q-numerical range of A

$$W_q(A) := \{x^* A y : x^* y = q, \|x\|_2 = \|y\|_2 = 1, x, y \in \mathbb{C}^n\}$$

is convex. It evidently is a generalization of Toeplitz-Hausdorff's theorem. See Li [14] for a simpler proof of Tsing's result. It is known that [22, p.199]

(7.1)
$$W_q(A) = \{qz + r\sqrt{1-|q|^2}\sqrt{\|Ay\|^2 - |z|^2} : y \in \mathbb{C}^n, \|y\| = 1, z = y^* A y \in W(A), r \in \mathbb{C}, |r| \leq 1\}.$$

Let $-1 \leq q \leq 1$ and $A \in \mathfrak{so}_n(\mathbb{C})$. One can define a real analog of q-numerical range

$$S_q(A) := \{x^T A y : x, y \in \mathbb{S}_{\mathbb{R}}^{n-1},\ x^T y = q\}.$$

The following result shows that $S_q(A)$ is convex and that $S(A)$ and $S_q(A)$ has a simpler relation comparing with the complex case (7.1).

Theorem 7.1. Let $-1 \leq q \leq 1$ and let $A \in \mathfrak{so}_n(\mathbb{C})$. When $n \geq 3$, $S_q(A) = \sqrt{1-q^2} S(A)$ and hence is convex and symmetric about the origin.

Proof. Assume that $n \geq 3$. Let $\xi \in S_q(A)$, i.e., $\xi = x^T A y$ with $x^T y = q$, $\|x\|_2 = \|y\|_2 = 1$, $x, y \in \mathbb{R}^n$. Write $y = qx + \sqrt{1-q^2} y'$ where $y' \in x^\perp$ is a unit vector. Since A is skew symmetric, $x^T A x = 0$ and thus

$$x^T A y = \sqrt{1-q^2} x^T A y' \in \sqrt{1-q^2} S(A)$$

so that $S_q(A) \subset \sqrt{1-q^2}S(A)$ and the converse inclusion is true as well. □

References

[1] C.A. Berger, Normal Dilations, Ph.D Dissertation, Cornell University, 1963.

[2] L. Brickman, On the field of values of a matrix, Proc. Amer. Math. Soc. **12** (1961), 61–66.

[3] W.S. Cheung, X. Liu, and T.Y. Tam, Multiplicities, boundary points, and joint numerical ranges, to appear in Operators and Matrices.

[4] C. Davis, The Toeplitz-Hausdorff theorem explained, Canad. Math. Bull. **14** (1971), 245–246.

[5] W.F. Donoghue, Jr., On the numerical range of a bounded operator, Michigan Math. J. **4** (1957), 261–263.

[6] M. R. Embry, The numerical range of an operator, Pacific J. Math. **32** (1970), 647–650.

[7] P.R. Halmos, *A Hilbert Space Problem Book*, Springer-Verlag, New York, 1982.

[8] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York, 1978.

[9] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[10] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[11] C.R. Johnson, Functional characterizations of the field of values and the convex hull of the spectrum, Proc. Amer. Math. Soc. **61** (1976), 201–204.

[12] A.W. Knapp, *Lie Groups Beyond an Introduction*, Birkhäuser, Boston, 1996.

[13] C.K. Li, A simple proof of the elliptical range theorem, Proc. Am. Math. Soc. **124** (1996), 1985–1986.

[14] C.K. Li, Some convexity theorems for the generalized numerical ranges, Linear and Multilinear Algebra **40** (1996), 235–240.

[15] C.K. Li and T.Y. Tam, Numerical ranges arising from simple Lie algebras, J. Canad. Math. Soc., **52** (2000), 141–171.

[16] A. McIntosh, The Toeplitz-Hausdorff theorem and ellipticity conditions, Amer. Math. Monthly **85** (1978), 475–477.

[17] C. Pearcy, An elementary proof of the power inequality for the numerical radius, Michigan Math. J. **13** (1966), 289–291.

[18] J. G. Stampfli, Extreme points of the numerical range of hyponormal operators, Michigan Math. J. **13** (1966), 87–89.

[19] J. Stillwell, *Naive Lie Theory*, Springer, New York, 2008.

[20] T.Y. Tam, Convexity of generalized numerical range associated with a compact Lie group, J. Austral. Math. Soc. **72** (2002), 57–66.

[21] T.Y. Tam, An extension of a convexity theorem of the generalized numerical range associated with SO($2n+1$), Proc. Amer. Math. Soc., **127** (1999), 35–44.

[22] N.K. Tsing, The constrained bilinear form and the C-numerical range, Linear Algebra Appl. **56** (1984), 195–206.

[23] F.W. Warner, *Foundation of Differentiable Manifolds and Lie Groups*, Scott Foresman and Company, 1971.

Department of Mathematics and Statistics, Auburn University, AL 36849–5310, USA

E-mail address: xzl0002@auburn.edu, dgt0001@auburn.edu, tamtiny@auburn.edu