Convexity of Generalized Numerical Ranges Associated with $\mathbf{SU}(n)$ and $\mathbf{SO}(n)$

by

Dawit Gezahegn Tadesse

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Auburn, Alabama
August 9, 2010

Keywords: Convexity, generalized numerical range, orthogonal group, unitary group

Approved by:

Tin-Yau Tam, Chair, Professor of Mathematics
Geraldo S. De Souza, Professor of Mathematics
Huajun Huang, Assistant Professor of Mathematics

Abstract

We give a new proof of a result of Tam on the convexity of the generalized numerical range associated with the classical Lie groups $\mathrm{SO}(n)$. We also provide a connection between the result and the convexity of the classical numerical range.

Acknowledgments

The author would like to thank Dr. Tin-Yau Tam for his excellent teaching and guidance throughout the course of this research, and his parents for their continual support, love, and encouragement.

Table of Contents

Chapter 1

Introduction

A subset $\Omega \subseteq \mathbb{C}$ is said to be convex if $(1-\lambda)x + \lambda y \in \Omega$ whenever $x, y \in \Omega$ and $0 \leq \lambda \leq 1$.

Let H be a Hilbert space over $\mathbb{C}$ and let $B(H)$ denotes the algebra of all bounded linear operators on H. The classical numerical range of $T \in B(H)$ is

$$W(T) := \{(Tx, x) : x \in H, (x, x) = 1\}.$$

Toeplitz-Hausdorff theorem [16, 6] asserts that $W(A)$ is a convex set. See [11] for a simple proof.

The finite dimensional case may be phrased as the following. If $\mathbb{C}_{n\times n}$ denotes the set of $n \times n$ complex matrices, then

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

is a compact convex set, where $A \in \mathbb{C}_{n\times n}$ and $\|x\|^2 = x^*x$. It is the image of the (compact) unit sphere $\mathbb{S}^{n-1} \subseteq \mathbb{C}^n$ under the nonlinear map $x \mapsto x^*Ax$. It is truly remarkable since the unit sphere is very "hallow" but its image under the quadratic map is convex.

When $n = 2$, $W(A)$ is an elliptical disk (possibly degenerate) [8], known as the elliptical range theorem.

Theorem 1.1. [8] Let $A \in \mathbb{C}_{2\times 2}$ with eigenvalues λ_1, λ_2. Then $W(A)$ is an elliptical disk with λ_1, λ_2 as foci and minor axis of length $\sqrt{\operatorname{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}$.

Indeed the convexity of the numerical range of T can be reduced to the 2-dimensional case, i.e., even the most general version of the theorem is equivalent to a statement about 2-dimensional spaces. That is the reason why almost every approach ends up with a quadratic computation that has no merit except correctness.

Let $\xi_1 = x_1^* A x_1$ and $\xi_2 = x_2^* A x_2$ be two points in $W(A)$ where $x_1, x_2 \in H$ are unit vectors. We may assume that x_1, x_2 are linearly independent, otherwise, $\xi_1 = \xi_2$. Consider the compression $\hat{A} : C \to C$ where $C := \operatorname{span}\{x_1, x_2\}$ defined by $\hat{A}x = PAx$, $x \in C$, where $P : H \to C$ is the orthogonal projection onto C. Clearly $\xi_1, \xi_2 \in W(\hat{A})$ and thus the line segment $[\xi_1, \xi_2] \subset W(\hat{A})$ if the theorem holds for the 2-dimensional case. Notice that $W(\hat{A}) \subset W(A)$ since for each unit vector $x \in C$,

$$x^* \hat{A} x = x^* P A x = x^* P A P x = (Px)^* A (Px) = x^* A x$$

as $P^* = P$.

The classical numerical range of $A \in \mathbb{C}_{n\times n}$ may be written as

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

since each unit vector x can be extended to a $U \in \mathrm{U}(n)$ of the form $(x\ U_1)$ where $U_1 \in \mathbb{C}_{n\times(n-1)}$. Here $\mathrm{U}(n)$ 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 of $\mathrm{SO}(n)$ is the set of $n \times n$ real skew symmetric matrices, whose complexification is the algebra of $n \times n$ complex skew symmetric matrices. When $n \geq 3$, the numerical range of a skew symmetric

$A \in \mathbb{C}_{n\times n}$ associated with $\mathrm{SO}(n)$ is defined to be the set

$$S(A) := \{(O^T AO)_{1,2} : O \in \mathrm{SO}(n)\} \subseteq \mathbb{C}$$

and is known to be a compact convex set according to a result of Tam [14] (see [10, 15] for related results). Indeed the result in [14] is more general and is in the context of compact connected Lie group. The method of Tam is the usage of a lemma of Atiyah on symplectic manifold since the adjoint orbit of a Lie algebra element has a natural symplectic structure.

From now on we denote by $\mathfrak{so}_n(\mathbb{C})$ the algebra of $n\times n$ complex skew symmetric matrices.

Theorem 1.2. (Tam [14]) If $A \in \mathfrak{so}_n(\mathbb{C})$, where $n \geq 2$, then $S(A)$ is a compact convex set.

The main goal of this thesis is to provide an elementary proof of Theorem 1.2 and point out some relation between $S(A)$ and $W(A)$ and the k-numerical range when n is small.

When $n = 2$, $S(A)$ is the set $\{\alpha\}$, where $A = \begin{pmatrix} 0 & \alpha \\ -\alpha & 0 \end{pmatrix}$. From now on we assume that $n \geq 3$ to avoid the trivial case. Rewrite

$$S(A) = \{x_1^T Ax_2 : x_1, x_2 \text{ are the two columns of some } O \in \mathrm{SO}(n)\}.$$

In particular,

$$\begin{aligned} S(A) &= \{(O^T AO)_{1,2} : O \in \mathrm{O}(n)\} \\ &= \{x_1^T Ax_2 : x_1, x_2 \text{ are the two columns of some } O \in \mathrm{O}(n)\} \\ &= \{x_1^T Ax_2 : x_1, x_2 \text{ are orthonormal in } \mathbb{R}^n\}. \end{aligned}$$

Clearly $x^T Ay = -y^T Ax$ for all $x, y \in \mathbb{R}^n$, because A is skew symmetric. So $\xi \in S(A)$ if and only if $-\xi \in S(A)$, i.e., $S(A)$ is symmetric about the origin. We remark in general that such symmetry property is not present in $W(A)$, e.g., Theorem 1.1.

Chapter 2

Three simple proofs of the convexity of $W(A)$

We first provide three proofs of the convexity of classical numerical range, namely, Raghavendran's proof [11], Gustafson's proof [3] and the proof of Davis [1].

1. Raghavendran's proof:

Proof. We need to consider only the case where $W(A)$ contains at least two points. Let x_k $(k = 1, 2)$ be any two elements of H with $\|x_k\| = 1$ such that $x_k^* A x_k = w_k$ are two distinct points of $W(A)$. As $x_1 + zx_2 = 0$ for $z \in \mathbb{C}$ will imply that $\bar{z}z = 1$ and then that $w_1 = w_2$, we see that $\|x_1 + zx_2\| \neq 0$ for all $z \in \mathbb{C}$. So the theorem will be proved if we show that, for any given real number t with $0 < t < 1$, there exists at least $z = x + iy \in \mathbb{C}$ (with x, y real) which satisfies the equation

$$(x_1 + zx_2)^* A(x_1 + zx_2) = (tw_1 + (1-t)w_2)\|x_1 + zx_2\|^2.$$

The equation may be rewritten in the form

$$p|z|^2 + qz + r\bar{z} + s = 0, \tag{2.1}$$

where $p = t(w_2 - w_1)$, $s = (1-t)(w_2 - w_1)$ and $q, r \in \mathbb{C}$. Dividing this equation by p, and then separating the real and imaginary parts, we get the two equations

$$x^2 + y^2 + ax + by - \left(\frac{1-t}{t}\right) = 0, \tag{2.2}$$

$$cx + dy = 0, \tag{2.3}$$

where a, b, c, d are some well-defined real numbers such that this pair of equations is equivalent to the single equation (2.1).

Equation (2.2) represents a real circle with a positive radius having the origin in its interior (because the constant term in this equation is negative); and when c, d are not both zero, the straight line represented by the equation (2.3) meets this circle in two real and distinct points. We can, therefore, always find (at least) two distinct complex numbers z_k such that $z = z_k$ satisfy the equation (2.1). This completes the proof. □

We remark that the above proof is actually revealing that the general statement is reduced to the $n = 2$ case.

2. Gustafson's proof:

Proof. Since $W(\mu A + \gamma) = \mu W(A) + \gamma$, for any scalars $\mu, \gamma \in \mathbb{C}$, it suffices to consider the situation

$$(Ax_1, x_1) = 0, \quad (Ax_2, x_2) = 1, \quad \|x_i\| = 1,\ x_i \in H,\ i = 1, 2.$$

Let $x = \alpha x_1 + \beta x_2, \alpha$ and β real, and require

$$\|x\|^2 \equiv \alpha^2 + \beta^2 + 2\alpha\beta \mathrm{Re}\,(x_1, x_2) = 1, \tag{2.4}$$

and desire (for each $0 < \lambda < 1$)

$$(Ax, x) := \beta^2 + \alpha\beta\{(Ax_1, x_2) + (Ax_2, x_1)\} = \lambda. \tag{2.5}$$

Let $B := (Ax_1, x_2) + (Ax_2, x_1)$. If B is real, then the system (2.4) and (2.5) describe an ellipse (intercepts $\pm 1, \pm 1$) and a hyperbola (intercepts $\pm\lambda^{1/2}$) clearly possesses four solutions since $\|\mathrm{Re}\,(x_1, x_2)\| < 1$ by Schwartz's inequality. But B can always be guaranteed real by using an appropriate (scalar multiple of) x_1, i.e., explicitly, use $\tilde{x}_1 = \mu x_1$, where $\mu = a + ib$ satisfies

$$|\mu|^2 \equiv a^2 + b^2 = 1 \tag{2.6}$$

and

$$\operatorname{Im} B(\tilde{x}_1) \equiv a \operatorname{Im} B(x_1) + b \operatorname{Re}\{(Ax_1, x_2) - (Ax_2, x_1)\} = 0, \tag{2.7}$$

a system clearly possessing (two) solutions. □

3. Davis's idea:

Proof. Accordingly, let us assume without loss of generality that $\dim H = 2$. Notice that $x^*Ax = \operatorname{tr}(Axx^*)$ which the key to Davis' proof. Consider the mapping Φ which takes the arbitrary hermitian operator X on H to

$$\Phi(X) = \operatorname{tr}(AX).$$

It is plainly real-linear. Its domain is a real 4-dimensional space, i.e., the space M of matrices

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad (a = \bar{a}, b = \bar{c}, d = \bar{d}.)$$

The range of Φ is a real 2-dimensional space, namely, the complex numbers. The conclusion which is to be proved is that Φ takes the set of one-dimensional orthoprojectors xx^* onto a convex set.

In the matrix representation of M, these orthoprojectors xx^* may be parametrized as follows. It is enough to consider

$$x = \begin{pmatrix} \cos\theta \\ e^{i\delta}\sin\theta \end{pmatrix}, \qquad \theta, \delta \in \mathbb{R}$$

because any other unit vector is a scalar multiple of one of these. These matrices comprise a 2-sphere centered at $\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ and lying in a 3-flat in M (the set of H having trace 1). But the image of a 2-sphere, under a linear map with range in real 2-space, is either an ellipse with interior, or a segment, or a point – in any case it is convex. □

Chapter 3

Basic Lie Theory

A (real) Lie group is a group which is also a finite-dimensional real smooth manifold, and in which the group operations of multiplication and inversion are smooth maps. For example $\mathrm{SU}(n)$, $\mathrm{SO}(n)$, $\mathrm{Sp}(n)$ are compact Lie groups.

A real vector space L with an operation $L \times L \to L$, denoted $(x, y) \mapsto [x, y]$ and called the bracket or commutator of x and y, is called a Lie algebra if

(L1) The bracket operation is bilinear.

(L2) $[x, x] = 0$ for all $x \in L$.

(L3) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in L$.

The condition (L3) is called the Jacobi identity.

There is a Lie algebra $\mathfrak{g}$ associated with G. The following is the description of $\mathfrak{g}$.

1. Let M be a real smooth manifold and denote by $C^\infty(M)$ the ring of all smooth functions $f : M \to \mathbb{R}$. A map $v : C^\infty(M) \to \mathbb{R}$ is called a tangent vector at $p \in M$ if for all $f, g \in C^\infty(M)$, $p \in M$

   (i) $v(f + g) = v(f) + v(g)$ and

   (ii) $v(fg) = v(f)g(p) + f(p)v(g)$.

   Each tangent vector can be thought as a derivative. The set $T_p(M)$ of all tangent vectors at p is a finite dimensional vector space.

2. Denote by $T(M) = \cup_{p \in M} T_p(M)$ the tangent bundle of M. A vector field on any smooth manifold M, $X : M \to T(M)$, is a smooth map such that $X(p) \in T_p(M)$. The

extension of X is the map $\tilde{X} : C^\infty(M) \to C^\infty(M)$ defined by $\tilde{X}(f)(0) = X(p)(f)$. It is known that $\tilde{X} = \tilde{Y}$ if and only if $X = Y$. Now the bracket $[X, Y]$ of two vector fields is defined by

$$[X, Y](p)(f) = X(p)(\tilde{Y}(f)) - Y(p)(\tilde{X}(f))$$

3. The left translations $L_g : G \to G$, $g \in G$ is given by $L_g(h) = gh$, $h \in G$. The left invariant vector fields (vector fields satisfying $dL_g(X) = X \circ L_g$ for every $g \in G$, where dL_g denotes the differential of L_g) on a Lie group is a Lie algebra under the Lie bracket of vector fields, i.e., the bracket of two left invariant vector field is also left invariant.

4. The map $X \to X(1)$ defines a one to one correspondence between the left invariant vector fields and the tangent space $T_e(G)$ at the identity e and therefore makes the tangent space at the identity into a Lie algebra, called the Lie algebra of G, usually denoted by $\mathfrak{g}$.

For example:

1. The special unitary group

$$\mathrm{SU}(n) = \{g \in \mathrm{GL}_n(\mathbb{C}) : g^* g = 1, \det g = 1\}$$

has Lie algebra

$$\mathfrak{su}(n) = \{A \in \mathbb{C}_{n\times n} : A^* = -A,\ \operatorname{tr} A = 0\}$$

which is a real subspace of $\mathbb{C}_{n\times n}$.

2. The complex special linear group

$$\mathrm{SL}_n(\mathbb{C}) = \{g \in \mathrm{GL}_n(\mathbb{C}) : \det g = 1\}$$

has Lie algebra

$$\mathfrak{sl}_n(\mathbb{C}) = \{A \in \mathbb{C}_{n\times n} : \operatorname{tr} A = 0\}$$

.

3. The complex orthogonal group

$$\mathrm{SO}_n(\mathbb{C}) = \{g \in \mathrm{GL}_n(\mathbb{C}) : g^T g = 1, \det g = 1\}$$

has Lie algebra

$$\mathfrak{so}_n(\mathbb{C}) = \{A \in \mathbb{C}_{n\times n} : A^T = -A\}.$$

We now recall some basic notions about adjoint representation $\mathrm{Ad} : G \to \mathrm{Aut}\,\mathfrak{g}$. Let G be a Lie group and let $\mathfrak{g}$ be its Lie algebra (which we identify with $T_e G$, the tangent space to the identity element in G). Define a map

$$\Psi : G \to \mathrm{Aut}\, G$$

by the equation $\Psi(g) = \Psi_g$ for all $g \in G$, where $\mathrm{Aut}\, G$ is the automorphism group of G and the automorphism Ψ_g is defined by

$$\Psi_g(h) = ghg^{-1}$$

for all $h \in G$. It follows that the derivative of Ψ_g at the identity is an automorphism of the Lie algebra $\mathfrak{g}$. We denote this map by $\mathrm{Ad}\, g$:

$$\mathrm{Ad}_g \colon \mathfrak{g} \to \mathfrak{g}.$$

The map

$$\mathrm{Ad} : G \to \mathrm{Aut}\,\mathfrak{g}$$

which sends g to $\mathrm{Ad}\, g$ is called the adjoint representation of G. This is indeed a representation of G since $\mathrm{Aut}\, \mathfrak{g}$ is a Lie subgroup of $\mathrm{GL}\,(\mathfrak{g})$ and the above adjoint map is a Lie group homomorphism. The dimension of the adjoint representation is the same as the dimension of the group G.

One may always pass from a representation of a Lie group G to a representation of its Lie algebra by taking the derivative at the identity. Taking the derivative of the adjoint map

$$\mathrm{Ad}\, : G \to \mathrm{Aut}\, \mathfrak{g}$$

gives the adjoint representation of the Lie algebra $\mathfrak{g}$:

$$\mathrm{ad}\, : \mathfrak{g} \to \mathrm{Aut}\, \mathfrak{g}.$$

The adjoint representation of a Lie algebra is related in a fundamental way to the structure of that algebra. In particular, one can show that

$$\mathrm{ad}\, X(Y) = [X, Y]$$

for all $X, Y \in \mathfrak{g}$.

The adjoint group $\mathrm{Int}\, \mathfrak{g}$ is the analytic group of $\mathrm{ad}\, \mathfrak{g}$, which is contained in $\mathrm{GL}\,(\mathfrak{g})$. It is known that if G is connected, then $\mathrm{Ad}\, G = \mathrm{Int}\, \mathfrak{g}$ [7, p.129].

A Lie algebra homomorphism $\phi : \mathfrak{g} \to \mathfrak{h}$ is a vector space isomorphism that respects bracket: $\varphi[X, Y] = [\varphi X, \varphi Y]$.

Lemma 3.1. The map $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_3(\mathbb{C})$ given by

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

is a Lie algebra isomorphism.

Proof. One can directly verify that it is a Lie algebra isomorphism. Here we establish the result via a double covering $\varphi : \mathrm{SU}(2) \to \mathrm{SO}(3)$.

As we know SU(2) is connected. Thus the image of SU(2) by φ is connected as well, hence contained in the connected component SO(3) of O(3). Let $d\varphi_e : \mathfrak{su}(2) \to \mathfrak{so}(3)$ be the differential of φ at the identity e. Since $\varphi = \mathrm{Ad}$, and $d\varphi_e = \mathrm{ad}$, the kernel of $d\varphi_e$ consists of the $X \in \mathfrak{su}(2)$ satisfying $XY - YX = 0$ for all skew Hermitian $Y \in \mathbb{C}_{2\times 2}$ of zero trace, hence for all skew Hermitian $Y \in \mathbb{C}_{2\times 2}$, hence for all $Y \in \mathbb{C}_{2\times 2}$ (as one sees by writing $Y = iY_1 + Y_2$ with Y_1 and Y_2 skew Hermitian). Thus X is a scalar and therefore equals 0, because $\mathrm{tr}\, X = 0$. Since the kernel of $d\varphi_e$ is $\{0\}$, $d\varphi_e$ is an isomorphism. It follows that φ is a covering. Its kernel consists of all $a \in \mathrm{SU}(2)$ satisfying $aYa^{-1} = Y$ for all skew Hermitian Y, hence are scalar and thus equal to ± 1.

To compute the Lie algebra isomorphism explicitly let

$$X = \begin{pmatrix} ix_3 & -x_1 + ix_2 \\ x_1 + ix_2 & -ix_3 \end{pmatrix}, \; Y = \begin{pmatrix} iy_3 & -y_1 + iy_2 \\ y_1 + iy_2 & -iy_3 \end{pmatrix} \in \mathfrak{su}(2).$$

Then

$$\begin{aligned} & \mathrm{ad}\,(X)Y \\ = \; & XY - YX \\ = \; & \begin{pmatrix} 2i(x_2y_1 - x_1y_2) & -2(x_3y_2 - x_2y_3) + 2i(-x_3y_1 + x_1y_3) \\ 2(x_3y_2 - x_2y_3) + 2i(-x_3y_1 + x_1y_3) & -2i(x_2y_1 - x_1y_2) \end{pmatrix} \\ \leftrightarrow \; & \begin{pmatrix} 2(x_3y_2 - x_2y_3) \\ 2(-x_3y_1 + x_1y_3) \\ 2(x_2y_1 - x_1y_2) \end{pmatrix} = \begin{pmatrix} 0 & 2x_3 & -2x_2 \\ -2x_3 & 0 & 2x_1 \\ 2x_2 & -2x_1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \end{aligned}$$

which gives us the desired isomorphism

$$\begin{pmatrix} ix_3 & -x_1+ix_2 \\ x_1+ix_2 & -ix_3 \end{pmatrix} \mapsto \begin{pmatrix} 0 & 2x_3 & -2x_2 \\ -2x_3 & 0 & 2x_1 \\ 2x_2 & -2x_1 & 0 \end{pmatrix}.$$

Then extend it to $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_3(\mathbb{C})$. □

Lemma 3.2. The map $\mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_4(\mathbb{C})$

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} + \begin{pmatrix} d & e \\ f & -d \end{pmatrix}$$

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

is a Lie algebra isomorphism.

Proof. One can directly verify that it is a Lie algebra isomorphism. However it is good to see how one can get it from the double covering $\varphi : \mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$ [12, p.42]. The differential at the identity $d\varphi_e : \mathfrak{su}_2 \oplus \mathfrak{su}_2 \to \mathfrak{so}(4)$ is the desired Lie algebra isomorphism. Let $\varphi : \mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$ be the map acting on the quaternions

$$\mathbb{H} = \{Q = \begin{pmatrix} \rho_1 + i\rho_2 & -\rho_3 - i\rho_4 \\ \rho_3 - i\rho_4 & \rho_1 - i\rho_2 \end{pmatrix} : \rho_1, \rho_2, \rho_3, \rho_4 \in \mathbb{R}\} \longleftrightarrow \{\overrightarrow{Q} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{pmatrix} : \rho_1, \rho_2, \rho_3, \rho_4 \in \mathbb{R}\}$$

as follows:

$\varphi((U,V))\overrightarrow{Q} = \overrightarrow{UQV^{-1}}$ where $Q \in \mathbb{H}$ and $(U,V) \in \mathrm{SU}(2) \times \mathrm{SU}(2)$. The corresponding Lie map is

$d\varphi_e : \mathfrak{su}_2(\mathbb{C}) \oplus \mathfrak{su}_2(\mathbb{C}) \to \mathfrak{so}(4)$ defined by $d\varphi_e((X,Y))\overrightarrow{Q} = \overrightarrow{-XQ+QY}$.

Using the same idea we used above it can be verified that $d\varphi_e$ is a Lie algebra isomorphism. To compute this Lie algebra isomorphism explicitly. Let

$$X = \begin{pmatrix} ix_3 & -x_1 + ix_2 \\ x_1 + ix_2 & -ix_3 \end{pmatrix}, \quad Y = \begin{pmatrix} iy_3 & -y_1 + iy_2 \\ y_1 + iy_2 & -iy_3 \end{pmatrix} \in \mathfrak{su}(2)$$

and

$$Q = \begin{pmatrix} \rho_1 + i\rho_2 & -\rho_3 - i\rho_4 \\ \rho_3 - i\rho_4 & \rho_1 - i\rho_2 \end{pmatrix} \in \mathbb{H}.$$

Direct computation yields

$$d\varphi_e((X,Y))\overrightarrow{Q} = \overrightarrow{-XQ+QY} = \begin{pmatrix} 0 & x_3 - y_3 & x_1 - y_1 & -x_2 + y_2 \\ -x_3 + y_3 & 0 & -x_2 - y_2 & -x_1 - y_1 \\ -x_1 + y_1 & x_2 + y_2 & 0 & x_3 + y_3 \\ x_2 - y_2 & x_1 + y_1 & -x_3 - y_3 & 0 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{pmatrix}$$

which gives us the Lie algebra isomorphism

$$\begin{pmatrix} ix_3 & -x_1 + ix_2 \\ x_1 + ix_2 & -ix_3 \end{pmatrix} + \begin{pmatrix} iy_3 & -y_1 + iy_2 \\ y_1 + iy_2 & -iy_3 \end{pmatrix}$$

$$\mapsto \begin{pmatrix} 0 & x_3 - y_3 & x_1 - y_1 & -x_2 + y_2 \\ -x_3 + y_3 & 0 & -x_2 - y_2 & -x_1 - y_1 \\ -x_1 + y_1 & x_2 + y_2 & 0 & x_3 + y_3 \\ x_2 - y_2 & x_1 + y_1 & -x_3 - y_3 & 0 \end{pmatrix}$$

Then extend it to $\mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_4(\mathbb{C})$ in order to get the desired isomorphism. $\square$

Lemma 3.3. ([9, p.162]) The map $\mathfrak{sl}_4(\mathbb{C}) \to \mathfrak{so}_6(\mathbb{C})$

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

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

is a Lie algebra isomorphism.

Proof. Let $I_{3,3}$ be the 6-by-6 diagonal matrix given by $I_{3,3} = \operatorname{diag}(1, -1, 1, -1, 1, -1)$ and define $\mathfrak{g} = \{X \in \mathfrak{gl}_6(\mathbb{C}) : X^T I_{3,3} + I_{3,3} X = 0\}$. Let $S = \operatorname{diag}(i, i, i, 1, 1, 1)$. For $X \in \mathfrak{g}$, let $Y = SXS^{-1}$. One easily sees that the map $X \mapsto Y$ is an isomorphism of $\mathfrak{g}$ onto $\mathfrak{so}_6(\mathbb{C})$. Any member of $\mathfrak{sl}_4(\mathbb{C})$ acts on the 6-dimensional complex vector space of alternating tensors of rank 2 by

$$M(e_i \wedge e_j) = Me_i \wedge e_j + e_i \wedge Me_j,$$

where $\{e_i\}_{i=1}^4$ is the standard basis of $\mathbb{C}^4$.

Using the following ordered basis:

$$e_1 \wedge e_2 + e_3 \wedge e_4, e_1 \wedge e_3 - e_2 \wedge e_4, e_1 \wedge e_4 + e_2 \wedge e_3, e_1 \wedge e_2 - e_3 \wedge e_4, e_1 \wedge e_3 + e_2 \wedge e_4, e_1 \wedge e_4 - e_2 \wedge e_3$$

of the exterior space $\wedge^2 \mathbb{C}^4$.

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \in \mathfrak{sl}_4(\mathbb{C})$$

such that $a_{44} = -(a_{11} + a_{22} + a_{33})$. Consider the map $A(e_i \wedge e_j) = Ae_i \wedge e_j + e_i \wedge Ae_j$. We now compute the matrix representation of A with respect the above basis.

$$\begin{aligned}
& A(e_1 \wedge e_2 + e_3 \wedge e_4) \\
= \ & A(e_1 \wedge e_2) + A(e_3 \wedge e_4) \\
= \ & 0 \cdot (e_1 \wedge e_2 + e_3 \wedge e_4) + \frac{1}{2}(a_{32} - a_{14} - a_{23} + a_{41})(e_1 \wedge e_3 - e_2 \wedge e_4) \\
& + \frac{1}{2}(a_{42} + a_{13} - a_{31} - a_{24})(e_1 \wedge e_4 + e_2 \wedge e_3) + (a_{11} + a_{22})(e_1 \wedge e_2 - e_3 \wedge e_4) \\
& + \frac{1}{2}(a_{32} - a_{14} + a_{23} - a_{14})(e_1 \wedge e_3 + e_2 \wedge e_4) + \frac{1}{2}(a_{42} + a_{13} + a_{31} + a_{24})(e_1 \wedge e_4 - e_2 \wedge e_3)
\end{aligned}$$

$$\begin{aligned}
& A(e_1 \wedge e_3 - e_2 \wedge e_4) \\
= \ & A(e_1 \wedge e_3) - A(e_2 \wedge e_4) \\
= \ & \frac{1}{2}(a_{23} + a_{14} - a_{41} - a_{32})(e_1 \wedge e_2 + e_3 \wedge e_4) + 0 \cdot (e_1 \wedge e_3 - e_2 \wedge e_4) \\
& + \frac{1}{2}(a_{43} - a_{12} + a_{21} - a_{34})(e_1 \wedge e_4 + e_2 \wedge e_3) + \frac{1}{2}(a_{23} + a_{14} + a_{41} + a_{32})(e_1 \wedge e_2 - e_3 \wedge e_4) \\
& + (a_{11} + a_{33})(e_1 \wedge e_3 + e_2 \wedge e_4) + \frac{1}{2}(a_{43} - a_{12} - a_{21} + a_{34})(e_1 \wedge e_4 - e_2 \wedge e_3)
\end{aligned}$$

$$
\begin{aligned}
& A(e_1 \wedge e_4 + e_2 \wedge e_3) \\
= \; & A(e_1 \wedge e_4) + A(e_2 \wedge e_3) \\
= \; & \frac{1}{2}(a_{24} - a_{13} + a_{31} - a_{42})(e_1 \wedge e_2 + e_3 \wedge e_4) + \frac{1}{2}(a_{34} + a_{12} - a_{21} - a_{43})e_1 \wedge e_3 - e_2 \wedge e_4) \\
& +0 \cdot (e_1 \wedge e_4 + e_2 \wedge e_3) + \frac{1}{2}(a_{24} - a_{13} - a_{31} + a_{42})(a_{23} + a_{14} + a_{41} + a_{32})(e_1 \wedge e_2 - e_3 \wedge e_4) \\
& +\frac{1}{2}(a_{34} + a_{12} + a_{21} + a_{43})(e_1 \wedge e_3 + e_2 \wedge e_4) - (a_{22} + a_{33})(e_1 \wedge e_4 - e_2 \wedge e_3)
\end{aligned}
$$

$$
\begin{aligned}
& A(e_2 \wedge e_3 - e_3 \wedge e_4) \\
= \; & A(e_1 \wedge e_3) - A(e_2 \wedge e_4) \\
= \; & (a_{11} + a_{22})(e_1 \wedge e_2 + e_3 \wedge e_4) + \frac{1}{2}(a_{32} + a_{14} + a_{23} + a_{41})(e_1 \wedge e_3 - e_2 \wedge e_4) \\
& +\frac{1}{2}(a_{42} - a_{13} - a_{31} + a_{24})(e_1 \wedge e_4 + e_2 \wedge e_3) + 0 \cdot (e_1 \wedge e_2 - e_3 \wedge e_4) \\
& +\frac{1}{2}(a_{32} + a_{14} - a_{23} - a_{14}))(e_1 \wedge e_3 + e_2 \wedge e_4) + \frac{1}{2}(a_{42} - a_{13} + a_{31} - a_{24})(e_1 \wedge e_4 - e_2 \wedge e_3)
\end{aligned}
$$

$$
\begin{aligned}
& A(e_1 \wedge e_3 + e_2 \wedge e_4) \\
= \; & A(e_1 \wedge e_3) + A(e_2 \wedge e_4) \\
= \; & \frac{1}{2}(a_{23} - a_{14} - a_{41} + a_{32})(e_1 \wedge e_2 + e_3 \wedge e_4) + (a_{11} + a_{33})(e_1 \wedge e_3 - e_2 \wedge e_4) \\
& +\frac{1}{2}(a_{43} + a_{12} + a_{21} + a_{34})(e_1 \wedge e_4 + e_2 \wedge e_3) + \frac{1}{2}(a_{23} - a_{14} + a_{41} - a_{32})(e_1 \wedge e_2 - e_3 \wedge e_4) \\
& +0 \cdot (e_1 \wedge e_3 + e_2 \wedge e_4) + \frac{1}{2}(a_{43} + a_{12} + a_{21} - a_{34})(e_1 \wedge e_4 - e_2 \wedge e_3)
\end{aligned}
$$

$$
\begin{aligned}
& A(e_1 \wedge e_4 - e_2 \wedge e_3) \\
= \ & A(e_1 \wedge e_4) - A(e_2 \wedge e_3) \\
= \ & \frac{1}{2}(a_{24} + a_{13} + a_{31} + a_{42})(e_1 \wedge e_2 + e_3 \wedge e_4) + \frac{1}{2}(a_{34} - a_{12} - a_{21} + a_{43})(e_1 \wedge e_3 - e_2 \wedge e_4) \\
& -(a_{22} + a_{33})(e_1 \wedge e_4 + e_2 \wedge e_3) + \frac{1}{2}(a_{24} + a_{13} - a_{31} - a_{42})(e_1 \wedge e_2 - e_3 \wedge e_4) \\
& +\frac{1}{2}(a_{34} - a_{12} + a_{21} - a_{43})(e_1 \wedge e_3 + e_2 \wedge e_4) + 0 \cdot (e_1 \wedge e_4 - e_2 \wedge e_3)
\end{aligned}
$$

So we get the following 6×6 matrix

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

in $\mathfrak{g}$ which gives us the desired Lie algebra isomorphism after conjugation $A \mapsto SAS^{-1}$ where $S = \text{diag}\,(i, i, i, 1, 1, 1)$. □

Theorem 3.4. ([17, p.101])

1. Let G and H be Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ and with G simply connected. Let $\psi : \mathfrak{g} \to \mathfrak{h}$ be a homomorphism. Then there is a unique homomorphism $\varphi : G \to H$ such that $d\varphi_e = \psi$.

2. For each Lie algebra $\mathfrak{g}$, there is a simply connected Lie group G with Lie algebra $\mathfrak{g}$.

Chapter 4

The connection between $W(A)$ and $S(A)$ for small n

We want to study $S(A)$ and its relation to the classical numerical range when the dimension is small, i.e., $2 \leq n \leq 6$. The k-numerical range of $B \in \mathbb{C}_{n\times n}$, $1 \leq k \leq n$, is

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

Halmos [4] asked if $W_k(B)$ is convex and Berger [5, p.110-111] provided an affirmative answer for any $B \in \mathbb{C}_{n\times n}$.

Theorem 4.1. (Berger) The k-numerical range of $B \in \mathbb{C}_{n\times n}$ is convex.

Lemma 4.2. Let Ω_1 and Ω_2 be convex subsets of $\mathbb{C}$. The sum of of Ω_1 and Ω_2, denoted by

$$\Omega_1 + \Omega_2 := \{a + b : a \in \Omega_1, b \in \Omega_2\},$$

is a convex subset of $\mathbb{C}$.

Proof. Let $a_1 + b_1, a_2 + b_2 \in \Omega_1 + \Omega_2$. Then $(1-\lambda)(a_1+b_1) + \lambda(a_2+b_2) = ((1-\lambda)a_1 + \lambda a_2) + ((1-\lambda)b_1 + \lambda b_2) \in \Omega_1 + \Omega_2$
□

Lemma 4.3. Let Ω be a convex subset of $\mathbb{C}$, and α be a scalar (either real number or complex number), then $\alpha\Omega := \{\alpha a : a \in \Omega\}$ is a convex subset of $\mathbb{C}$.

Proof. Let $a_1, a_2 \in \Omega$. Then $\alpha a_1, \alpha a_2 \in \alpha\Omega$, and $(1-\lambda)(\alpha a_1) + \lambda(\alpha a_2) = \alpha((1-\lambda)a_1 + \lambda a_2)) \in \alpha\Omega$. □

Theorem 4.4. 1. If $A \in \mathfrak{so}_3(\mathbb{C})$, then $S(A)$ is equal to $W(B)$ for some $B \in \mathfrak{sl}_2(\mathbb{C})$ and thus is an elliptical disk (possibly degenerate) centered at the origin.

2. If $A \in \mathfrak{so}_4(\mathbb{C})$, then $S(A)$ is the sum of $W(B)$ and $W(C)$, where $B, C \in \mathfrak{sl}_2(\mathbb{C})$. Thus $S(A)$ is the sum of two elliptical disks (possibly degenerate) centered at the origin.

3. If $A \in \mathfrak{so}_5(\mathbb{C})$, then $S(A) = iW_2(B)$ 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.

4. If $A \in \mathfrak{so}_6(\mathbb{C})$, then $S(A) = iW_2(B)$ for some $B \in \mathfrak{sl}_4(\mathbb{C})$.

Proof. Let K be a connected Lie group 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 $\mathrm{SO}(n)$

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

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 Ad K which is the analytic subgroup of the adjoint group Int $\mathfrak{k} \subset$ Aut $\mathfrak{k}$ corresponding to ad $\mathfrak{k}$ [7, 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) By Lemma 3.1 the following is a Lie algebra isomorphism $\mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_3(\mathbb{C})$

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

and thus its restriction $\psi : \mathfrak{su}(2) \to \mathfrak{so}(3)$ is a Lie algebra isomorphism. By Theorem 3.4 there is a Lie group isomorphism $\varphi : \widetilde{\mathrm{SU}}(2) \to \mathrm{SO}(3)$ 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

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

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 quadratic 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

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

(2) By Lemma 3.2 the following is a Lie algebra isomorphism $\psi : \mathfrak{sl}_2(\mathbb{C}) \oplus \mathfrak{sl}_2(\mathbb{C}) \to \mathfrak{so}_4(\mathbb{C})$:

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

and thus its restriction $\psi : \mathfrak{su}(2) \oplus \mathfrak{su}(2) \to \mathfrak{so}(4)$ is a Lie algebra isomorphism. By Theorem 3.4 there is a Lie group isomorphism $\varphi : \widetilde{\mathrm{SU}}(2) \times \widetilde{\mathrm{SU}}(2) \to \mathrm{SO}(4)$ so that $d\varphi_e = \psi$ (indeed there is a double covering map $\mathrm{SU}(2) \times \mathrm{SU}(2) \to \mathrm{SO}(4)$ [12, p.42]). So we have the relation

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

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 $U^{-1}AU \mapsto (U^{-1}AU)_{11} + (U^{-1}AU)_{44}$, $U \in \mathrm{SU}(2) \times \mathrm{SU}(2)$ corresponds to $i(\varphi^{-1}(U)(d\varphi_e(A))\varphi(U))_{12}$, i.e., $(x, y) \mapsto$

$x^T d\varphi_e(A)y$ where $x, y \in \mathbb{R}^4$ are orthonormal. So

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

Notice that $W\begin{pmatrix} a & b \\ c & -a \end{pmatrix} + W\begin{pmatrix} d & e \\ f & -d \end{pmatrix}$ is simply the sum of two elliptical disks centered at the origin. By Lemma 4.2 it is convex.

(4) By Lemma 3.3

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

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

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). ∎

Chapter 5

Another proof of the convexity of $S(A)$

Remark 5.1. Alike Davis' treatment [1] for $W(A)$ when $A \in \mathbb{C}_{2\times 2}$, there is a simpler and more 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

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

Since $\{x \times y : x, y \in \mathbb{R}^3 \text{ are orthonormal}\} = \mathbb{S}^2$, $S(A)$ is the image under the linear map $z \in \mathbb{S}^2 \mapsto (c, -b, a) \cdot z \in \mathbb{C}$. Thus $S(A) \subset \mathbb{C}$ is an elliptical disk using Davis' idea. We just establish the convexity of $S(A)$ when $n = 3$.

Lemma 5.2. Let $A \in \mathfrak{so}_n(\mathbb{C})$ and $n \geq 3$.

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 $y_2^T Ay_1 = \pm x_2^T Ax_1$.

2. If $\xi \in S(A)$, then $t\xi \in S(A)$ if $|t| \leq 1$ and $t \in \mathbb{R}$.

Proof. (1) Notice that $[y_1\ y_2] = [x_1\ x_2]O$ where O is a 2×2 orthogonal matrix, 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 5.1 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, $y_1 := (\operatorname{Re} A)y \in \mathbb{R}^n$ 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 $t\xi = tx^T A y = w^T A y \in S(A)$. $\square$

A proof of the convexity of $S(A)$:

We now provide a proof of the convexity of $S(A)$ which is different from [14] and is based on Theorem 4.4. The cases $n = 3, 4$ are proved in Theorem 4.4. 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 $\hat{A} : C \to C$ be the compression of A onto $C := \operatorname{span}\{x_1, x_2, y_1, y_2\}$. It is easy to see that the matrix of $\hat{A}$ is complex skew symmetric and $S(\hat{A})$ contains w_1 and w_2. Since w_1, w_2 are distinct, $2 \leq \dim C \leq 4$.

Case 1: $3 \leq \dim C \leq 4$. By Theorem 4.4, $S(\hat{A})$ 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: $\dim C = 2$. Then $\operatorname{span}\{x_1, y_1\} = \operatorname{span}\{x_2, y_2\}$ so that by Lemma 5.2(1) $w_1 = -w_2 \neq 0$. Pick $x_3 \in \mathbb{R}^n$ such that $x_3 \notin \operatorname{span}\{x_1, y_1\}$ since $n \geq 4$. Apply the previous argument on $C' := \operatorname{span}\{x_1, x_3, y_1\}$ to have the desired result.

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

Example 5.3. If $A = \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}$, then

$$S(A) = \{-\cos\theta(\sin\theta - i\cos\theta) : \theta \in \mathbb{R}\}$$

by direct computation. So $S(A)$ contains the points $\pm\frac{1}{2} + \frac{i}{2}$ but not their midpoint $\frac{i}{2}$.

Bibliography

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

[2] D.Z. Djoković and T.Y. Tam, Some questions about semisimple Lie groups originating in matrix theory, Canad. Math. Bull., **46** (2003), 332–343.

[3] K.E. Gustafson, The Toeplitz-Hausdorff theorem for linear operators, Proc. Amer. Math. Soc., **25** (1970) 203–204.

[4] P.R. Halmos, Numerical ranges and normal dilations, Acta Szeged, **25** (1964), 1–5.

[5] P.R. Halmos, *A Hilbert Space Problem Book*, Springer-Verlag, New York, 1978.

[6] F. Hausdorff, Der Wertvorrat einer Bilinearform, Math Z. **3** (1919), 314–316.

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

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

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

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

[11] R. Raghavendran, Toeplitz-Hausdorff theorem on numerical ranges, Proc. Amer. Math. Soc. **20** (1969), 284–285.

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

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

[14] T.Y. Tam, Convexity of generalized numerical range associated with a compact Lie group, J. Austral. Math. Soc. **71** (2001), 1–10.

[15] T.Y. Tam, On the shape of numerical range associated with Lie groups, Taiwanese J. Math., **5** (2001), 497–506.

[16] O. Toeplitz, Das algebraishe Analogon zu einem Satze von Fejér, Math. Z. **2** (1918), 187–197.

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