Pacific Journal of Mathematics

THE K-ORBIT OF A NORMAL ELEMENT IN A COMPLEX SEMISIMPLE LIE ALGEBRA

TIN-YAU TAM AND WAI-SHUN CHEUNG

Volume 238 No. 2 December 2008

PACIFIC JOURNAL OF MATHEMATICS
Vol. 238, No. 2, 2008

THE K-ORBIT OF A NORMAL ELEMENT IN A COMPLEX SEMISIMPLE LIE ALGEBRA

TIN-YAU TAM AND WAI-SHUN CHEUNG

Given a complex semisimple Lie algebra $\mathfrak{g} = \mathfrak{k} + i\mathfrak{k}$, we consider the converse question of Kostant's convexity theorem for a normal $x \in \mathfrak{g}$. Let $\pi : \mathfrak{g} \to \mathfrak{h}$ be the orthogonal projection under the Killing form onto the Cartan subalgebra $\mathfrak{h} := \mathfrak{t} + i\mathfrak{t}$ where $\mathfrak{t}$ is a maximal abelian subalgebra of $\mathfrak{k}$. If $\pi(\mathrm{Ad}(K)x)$ is convex, then there is $k \in K$ such that each simple component of $\mathrm{Ad}(k)x$ can be rotated into the corresponding component of $\mathfrak{t}$. The result also extends a theorem of Au-Yeung and Tsing on the generalized numerical range.

1. Introduction

Let $A \in \mathbb{C}_{n\times n}$. Consider the set

$$\mathcal{W}(A) := \{\mathrm{diag}\,(UAU^{-1}) : U \in \mathrm{U}(n)\},$$

where $\mathrm{U}(n)$ denotes the unitary group. It is the image of the projection of the orbit

$$O(A) := \{UAU^{-1} : U \in \mathrm{U}(n)\}$$

onto the set of diagonal matrices. The following two results concern the geometric shape of $\mathcal{W}(A)$.

Theorem 1.1 (Schur–Horn [Schur 1923; Horn 1954]). *If $A \in \mathbb{C}_{n\times n}$ is Hermitian with eigenvalues $\lambda := (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$, then*

$$\mathcal{W}(A) = \mathrm{conv}\, S_n\lambda,$$

where $\mathrm{conv}\, S_n\lambda$ *is the convex hull of the orbit of λ under the action of the full symmetric group S_n.*

For general $A \in \mathbb{C}_{n\times n}$, $\mathcal{W}(A)$ is not convex. Indeed Tsing [1981] proved that $\mathcal{W}(A)$ is star-shaped with respect to the star center $\frac{1}{n}(\mathrm{tr}\, A)(1, \dots, 1)$.

Theorem 1.2 (Au-Yeung and Sing [1977]). *Let $A \in \mathbb{C}_{n\times n}$ be normal. If $\mathcal{W}(A)$ is convex*, *then the eigenvalues of A are collinear*, *that is*, *there exist $\alpha, \beta \in \mathbb{C}$ such that $\alpha A + \beta I$ is Hermitian.*

MSC2000: primary 22E10; secondary 17B20.
Keywords: K-orbit, convex, normal element, complex semisimple Lie algebra.

So Theorem 1.2 may be viewed as the converse to Theorem 1.1 as one restricts the attention on normal matrices. We remark that if $A \in \mathbb{C}_{n\times n}$ has zero trace, then $\alpha A + \beta I$ being Hermitian means that $e^{i\gamma} A$ is Hermitian for some $\gamma \in \mathbb{R}$. The following result of Au-Yeung and Tsing is stronger than Theorem 1.2. It affirmatively answers the conjecture of Marcus [1979] about the (stronger) converse of the result of Westwick [1975] on the convexity of c-numerical range. Bebiano and Da Providência [1996] gave another proof of Theorem 1.3.

Theorem 1.3 (Au-Yeung and Tsing [1983]). *Let $A \in \mathbb{C}_{n\times n}$ be normal. If*

$$W_{A^*}(A) := \{\operatorname{tr} A^* U A U^{-1} : U \in \mathrm{U}(n)\}$$

is convex, then A has collinear eigenvalues.

The above results can be reduced to the case $\operatorname{tr} A = 0$, that is, the simple Lie algebra $\mathfrak{sl}_n(\mathbb{C})$. We may write $A = \hat{A} + \frac{1}{n}(\operatorname{tr} A) I_n$, where $\hat{A} := A - \frac{1}{n}(\operatorname{tr} A) I_n$ has zero trace. Then

$$\mathcal{W}(A) = \mathcal{W}(\hat{A}) + \frac{\operatorname{tr} A}{n}(1, \dots, 1),$$

$$W_{A^*}(A) = W_{\hat{A}^*}(\hat{A}) + \frac{|\operatorname{tr} A|^2}{n^2}.$$

We will extend Theorems 1.2 and 1.3 in the context of semisimple Lie algebras.

2. Main results

Let $\mathfrak{g}$ be a complex semisimple Lie algebra and let $\mathfrak{k}$ be a real compact form of $\mathfrak{g}$. Let G be a complex Lie group with Lie algebra $\mathfrak{g}$. It has a finite center so K (the analytic group of $\mathfrak{k}$) is compact. As a real K-module, $\mathfrak{g}$ is just the direct sum of two copies of the adjoint module $\mathfrak{k}$ of K: $\mathfrak{g} = \mathfrak{k} + i\mathfrak{k}$ (direct sum), that is, Cartan decomposition of $\mathfrak{g}$. Denote by $\mathfrak{g}^*$ the dual space of $\mathfrak{g}$. Given $x \in \mathfrak{g}$, consider the orbit of x under the adjoint action of K

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

The orbit $K \cdot x$ depends on $\mathrm{Ad}_G K$ which is the analytic subgroup of the adjoint group $\mathrm{Int}(\mathfrak{g}) \subset \mathrm{Aut}(\mathfrak{g})$ corresponding to $\mathrm{ad}_{\mathfrak{g}}(\mathfrak{k})$. Thus $K \cdot x$ is independent of the choice of G. Let $\mathfrak{t}$ be a maximal abelian subalgebra of $\mathfrak{k}$. The complexification $\mathfrak{h} := \mathfrak{t} + i\mathfrak{t}$ (direct sum) is a Cartan subalgebra of $\mathfrak{g}$. The rank of $\mathfrak{g}$ is $\dim_{\mathbb{C}} \mathfrak{h}$, denoted by rank $\mathfrak{g}$. Let

$$\mathfrak{g} = \mathfrak{h} + \sum_{\alpha\in\Delta} \mathfrak{g}_\alpha$$

be the root space decomposition of $\mathfrak{g}$ with respect to $\mathfrak{h}$, where Δ denotes the set of all nonzero roots. Denote by $B(\cdot,\cdot)$ the Killing form of $\mathfrak{g}$. As $B(\cdot,\cdot)$ is a nondegenerate bilinear form, it induces a vector space isomorphism $\mathfrak{g} \to \mathfrak{g}^*$ sending

$x \to \varphi_x$, where $\varphi_x(y) = B(x, y)$ for all $y \in \mathfrak{g}$. Denote the inverse by $\varphi \to H_\varphi \in \mathfrak{g}$ ($\varphi \in \mathfrak{g}^*$), where $B(H_\varphi, y) = \varphi(y)$ for all $y \in \mathfrak{g}$. Let

$$\mathfrak{h}_\mathbb{R} := \sum_{\alpha \in \Delta} \mathbb{R} H_\alpha$$

so that $B(\cdot\,,\cdot)$ is a real inner product on $\mathfrak{h}_\mathbb{R}$ and $\mathfrak{h} = \mathfrak{h}_\mathbb{R} + i\mathfrak{h}_\mathbb{R}$ (direct sum). Hence rank $\mathfrak{g} = \dim_\mathbb{R} \mathfrak{h}_\mathbb{R}$. Moreover $\mathfrak{h}_\mathbb{R} = i\mathfrak{t}$ [Helgason 1978, p. 259]. Notice that $B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ [Helgason 1978, p. 166] whenever $\alpha + \beta \neq 0$ ($\mathfrak{g}_0 = \mathfrak{h}$) so the sum

$$\mathfrak{g} = \mathfrak{h} + \sum_{\alpha \in \Delta^+} (\mathfrak{g}_\alpha + \mathfrak{g}_{-\alpha})$$

is orthogonal under the Killing form. Thus we have the orthogonal projection $\pi : \mathfrak{g} \to \mathfrak{h}$ under $B(\cdot\,,\cdot)$. For $x \in \mathfrak{g}$, we consider $\pi(K \cdot x)$, that is, the projection of $K \cdot x$ onto $\mathfrak{h}$. When $x \in \mathfrak{k}$, $K \cdot x \subset \mathfrak{k}$ so $\pi(K \cdot x) \subset \mathfrak{t}$.

Kostant [1973] generalized Theorem 1.1 in the context of real semisimple Lie algebras. The following statement is for complex semisimple case. When $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$, it is reduced to Theorem 1.1.

Theorem 2.1 (Kostant [1973]). *If $x \in \mathfrak{k}$, then $\pi(K \cdot x) \subset \mathfrak{t}$ is convex and equals to* conv $W x_\mathfrak{t}$, *where $x_\mathfrak{t} \in K \cdot x \cap \mathfrak{t}$ and W is the Weyl group, that is, $W = N(T)/T$, the normalizer of T modulo T.*

Let θ be the Cartan involution of $\mathfrak{g}$ if $\mathfrak{g}$ is viewed as a real Lie algebra, that is, $\theta : \mathfrak{g} \to \mathfrak{g}$ such that $x + y \mapsto x - y$ if $x \in \mathfrak{k}$ and $y \in i\mathfrak{k}$. In other words, $\mathfrak{k}$ is the $+1$ eigenspace of θ and $i\mathfrak{k}$ is the -1 eigenspace of θ. Though θ is not an automorphism of $\mathfrak{g}$ over $\mathbb{C}$ (since $\theta(cx) = \bar{c}\,\theta x$ for $c \in \mathbb{C}$ and $x \in \mathfrak{g}$), it respects the bracket, that is,

$$\theta[x, y] = [\theta x, \theta y], \qquad x, y \in \mathfrak{g}.$$

Moreover Ad(k) and θ commute for all $k \in K$. Since $\mathfrak{g} = \mathfrak{k} + i\mathfrak{k}$ and $\mathfrak{k}$ is compact,

$$B_\theta(x, y) := -B(x, \theta y)$$

is an inner product on $\mathfrak{g}$ over $\mathbb{C}$. Let

$$\|x\|_\theta := B_\theta^{1/2}(x, x)$$

be the induced norm on $\mathfrak{g}$. The projection $\pi : \mathfrak{g} \to \mathfrak{h}$ under $B(\cdot\,,\cdot)$ coincides with that under $B_\theta(\cdot\,,\cdot)$ since $\theta\mathfrak{g}_\alpha = \mathfrak{g}_{-\alpha}$, for $\alpha \in \Delta$.

An element $x \in \mathfrak{g}$ is said to be *normal* if $[x, \theta x] = 0$, where θ is the Cartan involution. When $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$, the Cartan decomposition is the usual Hermitian decomposition, $K = \mathrm{SU}(n)$ and $\theta(z) = -z^*$, $z \in \mathfrak{sl}_n(\mathbb{C})$. When $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ and $\mathfrak{k} = \mathfrak{su}(n)$, normality reduces to the usual notion of normality of a matrix.

We want to know when $\pi(K \cdot x)$ is convex, that is, the converse question of Theorem 2.1 when we restrict ourselves to normal $x \in \mathfrak{g}$. Djoković and Tam [2003] proved that $B_\theta(K \cdot x, y) \subset \mathbb{C}$ is star shaped with respect the origin for each $y \in \mathfrak{g}$, if $x \in \mathfrak{g}$ is normal. In particular $B_\theta(K \cdot x, x)$ is star shaped. We also want to know when $B_\theta(K \cdot x, x)$ is convex. It turns out their answers coincide as suggested by Theorems 1.2 and 1.3. Indeed it is equivalent to say that $B_\theta(K \cdot x, y)$ is convex for all $y \in \mathfrak{g}$ in the following theorem.

Theorem 2.2. *Let $\mathfrak{g} = \mathfrak{g}_1 + \cdots + \mathfrak{g}_\ell$ be a complex semisimple Lie algebra with simple components $\mathfrak{g}_1, \ldots, \mathfrak{g}_\ell$. Let $x = x_1 + \cdots + x_\ell \in \mathfrak{g}$ be normal, where $x_i \in \mathfrak{g}_i$, $i = 1, \ldots, \ell$. The following statements are equivalent*:

(1) *$\pi(K \cdot x)$ is convex.*

(2) *$B_\theta(K \cdot x, x)$ is convex.*

(3) *$B_\theta(K \cdot x, x)$ is a closed line segment in $\mathbb{R}$.*

(4) *$K_j \cdot e^{i\theta_j} x_j \cap \mathfrak{t}_j$ is nonempty for some $\theta_j \in [0, 2\pi]$, $j = 1, \ldots, \ell$.*

(5) *$B_\theta(K \cdot x, y)$ is convex for all $y \in \mathfrak{g}$.*

Remark 2.3. Normality of $x \in \mathfrak{g}$ is necessary. When $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C})$ and $K = \mathrm{SU}(n)$, it is known that $B_\theta(K \cdot x, y)$ is convex for all $y \in \mathfrak{sl}_n(\mathbb{C})$ if $x \in \mathfrak{sl}_n(\mathbb{C})$ and the matrix rank of x is 1 (not necessarily normal), according to a result of Tsing [1984]. For example, if

$$x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \oplus 0_{n-2},$$

then $B_\theta(K \cdot x, y)$ is convex for all $y \in \mathfrak{sl}_n(\mathbb{C})$. However statement (3) in Theorem 2.2 does not hold.

We first establish some results in order to prove Theorem 2.2.

A line L is called a *support* of $B_\theta(K \cdot x, x) \subset \mathbb{C}$ at $\xi \in \partial B_\theta(K \cdot x, x)$ if $B_\theta(K \cdot x, x)$ lies in one of the closed half planes determined by L. A point $\xi \in B_\theta(K \cdot x, x)$ is called an *extreme point* of $B_\theta(K \cdot x, x)$ if ξ does not belong to any open line segment lying in $B_\theta(K \cdot x, x)$. It is clear that extreme points belong to $\partial B_\theta(K \cdot x, x)$. An extreme point $\xi \in B_\theta(K \cdot x, x)$ is called a *sharp point* if $B_\theta(K \cdot x, x)$ has more than one support line at ξ. Clearly a sharp point ξ of $B_\theta(K \cdot x, x)$ is an extreme point. The definitions are valid for convex sets in $\mathbb{C}$. The notions of extreme point and sharp point of a convex polygon in $\mathbb{C}$ coincide. We remark that $B_\theta(K \cdot x, x)$ is not necessarily a convex polygon.

Proposition 2.4. *Let $x \in \mathfrak{g}$ be normal.*

(a) *$B_\theta(K \cdot x, x) \subset \mathbb{C}$ is symmetric about the real axis.*

(b) *$B_\theta(K\cdot x, x)\subset\mathbb{C}$ is contained in the convex polygon*

$$B_\theta(\operatorname{conv} Wx_1 + i\operatorname{conv} Wx_2, x),$$

where $x = x_1 + ix_2$, $x_1, x_2 \in \mathfrak{k}$. Both sets contain the point $B_\theta(x, x)\geq 0$ which has the largest magnitude. Thus $B_\theta(x, x)$ is a sharp point of both $B_\theta(K\cdot x, x)$ and $B_\theta(\operatorname{conv} Wx_1 + i\operatorname{conv} Wx_2, x)$.

Proof. Since θ and $\operatorname{Ad}(k)$ ($k\in K$) commute, for $x, y\in\mathfrak{g}$,

$$B_\theta(\operatorname{Ad}(k)x, \operatorname{Ad}(k)y) = -B(\operatorname{Ad}(k)x, \operatorname{Ad}(k)\theta y) = B_\theta(x, y)$$

and hence $\operatorname{Ad}(k):\mathfrak{g}\to\mathfrak{g}$ is an isometry with respect to $B_\theta(\cdot,\cdot)$.

(a) Let $x\in\mathfrak{g}$ be normal. Clearly

$$\overline{B_\theta(\operatorname{Ad}(k)x, x)} = B_\theta(x, \operatorname{Ad}(k)x) = B_\theta(\operatorname{Ad}(k^{-1})x, x).$$

Hence (a) is established.

(b) Since $x = x_1 + ix_2\in\mathfrak{g}$ ($x_1, x_2\in\mathfrak{k}$) is normal, $K\cdot x$ intersects $\mathfrak{h}$ [Djoković and Tam 2003, Lemma 3.3.14]. So we may assume that $x_1, x_2\in\mathfrak{t}$. By Theorem 2.1

$$\begin{aligned}\pi(K\cdot x) &= \pi(K\cdot(x_1+ix_2))\\ &\subset \pi(K\cdot x_1 + iK\cdot x_2)\\ &= \pi(K\cdot x_1) + i\pi(K\cdot x_2)\\ &= \operatorname{conv} Wx_1 + i\operatorname{conv} Wx_2,\end{aligned}$$

where the sum $\operatorname{conv} Wx_1 + i\operatorname{conv} Wx_2$ is a convex polytope in $\mathfrak{h}$. Since $\pi:\mathfrak{g}\to\mathfrak{h}$ is also an orthogonal projection with respect to $B_\theta(\cdot,\cdot)$,

$$B_\theta(K\cdot x, x) = B_\theta(\pi(K\cdot x), x)$$

is contained in the convex polygon $B_\theta(\operatorname{conv} Wx_1 + i\operatorname{conv} Wx_2, x)$. Let

$$y\in\operatorname{conv} Wx_1\subset\mathfrak{t}\quad\text{and}\quad z\in\operatorname{conv} Wx_2\subset\mathfrak{t}.$$

Since $\mathfrak{h}_{\mathbb{R}} := \sum_{\alpha\in\Delta}\mathbb{R}H_\alpha = i\mathfrak{t}$ [Helgason 1978, p. 259] and $\alpha(H)\in\mathbb{R}$ for each $H\in\mathfrak{h}_{\mathbb{R}}$, $\alpha\in\Delta$, $\alpha(y), \alpha(x_1)\in i\mathbb{R}$ and $\alpha(iz), \alpha(ix_2)\in\mathbb{R}$. Hence

$$\alpha(\theta x) = -\overline{\alpha(x)}$$

so

$$\|x\|_\theta^2 = B_\theta(x, x) = \sum_{\alpha\in\Delta}|\alpha(x)|^2.$$

Moreover

$$\|y+iz\|_\theta^2 = \sum_{\alpha\in\Delta}\alpha(y+iz)\,\overline{\alpha(y+iz)} = \sum_{\alpha\in\Delta}(|\alpha(y)|^2 + |\alpha(iz)|^2) = \|y\|_\theta^2 + \|iz\|_\theta^2.$$

By Cauchy–Schwarz's inequality

$$|B_\theta(y+iz,x)|^2 \le \|y+iz\|_\theta^2\|x\|_\theta^2 = (\|y\|_\theta^2+\|iz\|_\theta^2)\|x\|_\theta^2. \tag{2-1}$$

Using triangle inequality, we have

$$\|y\|_\theta^2 \le \|x_1\|_\theta^2, \quad \|iz\|_\theta^2 \le \|ix_2\|_\theta^2, \tag{2-2}$$

since the elements in W are isometries. By (2-1) and (2-2)

$$|B_\theta(y+iz,x)|^2 \le B_\theta^2(x,x). \qquad \square$$

Remark 2.5. Given $x\in\mathfrak{h}$, $Wx\subset K\cdot x$ and thus $Wx\subset\pi(K\cdot x)$. We do not know whether $\pi(K\cdot x)\subset \operatorname{conv} Wx$ or not though it is true when $\mathfrak{g}=\mathfrak{sl}_n(\mathbb{C})$.

Lemma 2.6. *Let $x\in\mathfrak{g}$ be normal. Then $B_\theta(K\cdot x,x)$ is convex if and only if it is a closed interval in $\mathbb{R}$.*

Proof. One implication is trivial. Suppose $B_\theta(K\cdot x,x)$ is convex and we may assume $x\ne 0$. By Proposition 2.4

$$\xi := B_\theta(x,x) = \|x\|_\theta$$

is a sharp point of $B_\theta(K\cdot x,x)$. There are two supporting lines passing through ξ and one is the reflection of the other by Proposition 2.4 (a). Clearly $B_\theta(K\cdot x,x)$ is inside the cone determined by the two lines. Let L be the upper supporting line for definiteness. So $B_\theta(K\cdot x,x)$ is in the lower half plane determined by L.

By [Djoković and Tam 2003, Lemma 3.14] we may assume that $x=x_1+ix_2\in\mathfrak{h}$, $x_1,x_2\in\mathfrak{t}$. Let $\xi_j := B_\theta(\mathrm{Ad}(k_j)x,x)$ ($k_j\in K$) be on the upper boundary of $B_\theta(K\cdot x,x)$ so that $|\xi-\xi_j|<1/j$ but $\xi_j\ne\xi$, $j=1,2,\dots$. Since K is compact, there is a convergent subsequence $\{k_{j_m}\}_{m=1}^\infty$ of $\{k_j\}_{j=1}^\infty$. Let $\lim_{m\to\infty}k_{j_m}=k_0\in K$. So

$$B_\theta(\mathrm{Ad}(k_0)x,x)=\xi=B_\theta(x,x)=\|\mathrm{Ad}(k_0)x\|_\theta\|x\|_\theta$$

since $\mathrm{Ad}(k_0)$ is an isometry. By the equality case of Cauchy–Schwarz's inequality, $\mathrm{Ad}(k_0)x=x$. Thus

$$B_\theta(\mathrm{Ad}(k_j)x,x)=B_\theta(\mathrm{Ad}(k_j)x,\mathrm{Ad}(k_0)x)=B_\theta(\mathrm{Ad}(k_0^{-1}k_j)x,x).$$

We may replace k_{j_m} by $k_0^{-1}k_{j_m}\to e$ (the identity) or simply assume that $k_0=e$. The exponential map is an analytic diffeomorphism between an open neighborhood of $0\in\mathfrak{k}$ and an open neighborhood of $e\in K$. So for each sufficiently large m, there is $s_{j_m}\in\mathfrak{k}$ such that

$$\exp s_{j_m}=k_{j_m}\to e.$$

Since $x \in \mathfrak{h}$,

$$\begin{aligned}\xi_{j_m} &= B_\theta(\mathrm{Ad}(e^{s_{j_m}})x, x) = B_\theta(e^{\mathrm{ad}\, s_{j_m}} x, x) \\ &= B_\theta(x, x) + B_\theta(\mathrm{ad}\, s_{j_m} x, x) + \frac{1}{2} B_\theta((\mathrm{ad}\, s_{j_m})^2 x, x) \\ &\quad + \sum_{k=3}^{\infty} \frac{1}{k!} B_\theta((\mathrm{ad}\, s_{j_m})^k x, x).\end{aligned} \tag{2-3}$$

The first term of (2-3) is just ξ. The second term is

$$-B(\mathrm{ad}\,(s_{j_m})x, \theta x) = -B([s_{j_m}, x], \theta x) = -B(s_{j_m}, [x, \theta x]) = 0,$$

because $[x, \theta x] = 0$. Since the elements in $\mathrm{ad}\,\mathfrak{k}$ are skew Hermitian with respect to $B_\theta(\,\cdot\,,\,\cdot\,)$, the third term is

$$B_\theta((\mathrm{ad}\, s_{j_m})^2 x, x) = -B_\theta(\mathrm{ad}\, s_{j_m} x, \mathrm{ad}\,(s_{j_m})x) = -\|\mathrm{ad}\,(s_{j_m})x\|_\theta^2.$$

Taking the absolute value of the last term of (2-3), we have

$$\begin{aligned}\left|\sum_{k=3}^{\infty} \frac{1}{k!} B_\theta((\mathrm{ad}\, s_{j_m})^k x, x)\right| &= \left|\sum_{k=3}^{\infty} \frac{1}{k!} B_\theta(\mathrm{ad}\, s_{j_m} \circ (\mathrm{ad}\, s_{j_m})^{k-2} \circ \mathrm{ad}\,(s_{j_m})x, x)\right| \\ &= \sum_{k=3}^{\infty} \frac{1}{k!} |B_\theta((\mathrm{ad}\, s_{j_m})^{k-2} \circ \mathrm{ad}\,(s_{j_m})x, \mathrm{ad}\,(s_{j_m})x)| \\ &\le \sum_{k=3}^{\infty} \frac{1}{k!} \|(\mathrm{ad}\, s_{j_m})^{k-2} \mathrm{ad}\,(s_{j_m})x\|_\theta \|\mathrm{ad}\,(s_{j_m})x\|_\theta \\ &\le \sum_{k=3}^{\infty} \frac{1}{k!} \|(\mathrm{ad}\, s_{j_m})^{k-2}\|_\theta \|\mathrm{ad}\,(s_{j_m})x\|_\theta^2 \\ &\le (e^{\|\mathrm{ad}\, s_{j_m}\|} - 1) \|\mathrm{ad}\,(s_{j_m})x\|_\theta^2\,,\end{aligned}$$

where

$$\|\mathrm{ad}\, s_{j_m}\| := \max_{y \in \mathfrak{g}, \|y\|_\theta = 1} \|\mathrm{ad}\,(s_{j_m})y\|_\theta$$

is the operator norm of $\mathrm{ad}\, s_{j_m} : \mathfrak{g} \to \mathfrak{g}$ with respect to $\|\cdot\|_\theta$. Notice that $\mathrm{ad}\,(s_{j_m})x \neq 0$ otherwise $\xi = \xi_{j_m}$ from (2-3). Since $s_{j_m} \to 0$ ($x \neq 0$, $s_{j_m} \neq 0$),

$$\lim_{m \to \infty} \frac{\left|\sum_{k=3}^{\infty} \frac{1}{k!} B_\theta((\mathrm{ad}\, s_{j_m})^k x, x)\right|}{\|\mathrm{ad}\,(s_{j_m})x\|_\theta^2} = 0.$$

Consequently we have

$$\lim_{m \to \infty} \frac{\xi - \xi_{j_m}}{\|\mathrm{ad}\,(s_{j_m})x\|_\theta^2} = \frac{1}{2}. \tag{2-4}$$

Since $B_\theta(K\cdot x, x)$ is convex, there is $\xi' \in L\cap\partial B_\theta(K\cdot x, x)$ so that the line segment $[\xi, \xi'] \subset \partial B_\theta(K\cdot x, x)$. For sufficiently large m, $\xi_{j_m} \in [\xi, \xi']$, thus the limit on the left-hand side of (2-4) must be a positive multiple of $\xi - \xi'$. So $\xi' \in \mathbb{R}$ and thus $L \subset \mathbb{R}$. Therefore the compact connected set $B_\theta(K\cdot x, x)$ is a closed interval in $\mathbb{R}$. □

Proposition 2.7. *Let $\mathfrak{g} = \mathfrak{g}_1 + \cdots + \mathfrak{g}_\ell$ be a complex semisimple Lie algebra with simple components $\mathfrak{g}_1, \dots, \mathfrak{g}_\ell$. Let $x, y \in \mathfrak{h} := \mathfrak{t} + i\mathfrak{t}$. Write $x = x_1 + \cdots + x_\ell$ and $y = y_1 + \cdots + y_\ell$, where $x_i, y_i \in \mathfrak{h}_i$, $i = 1, \dots, \ell$. Suppose that x_i, y_i are nonzero for all $i = 1, \dots, \ell$. Then the following statements are equivalent.*

(1) *$B_\theta(K\cdot x, y)$ is a (closed) line segment in $\mathbb{C}$.*

(2) *$B_\theta(W\cdot x, y)$ is on a line segment in $\mathbb{C}$, where W is the Weyl group.*

(3) *$K_j \cdot e^{i\theta_j} x_j \cap \mathfrak{t}_j$ and $K_j \cdot e^{i\rho_j} y_j \cap \mathfrak{t}_j$ are nonempty for some $\theta_j, \rho_j \in [0, 2\pi]$, $j = 1, \dots, \ell$, and $\kappa := \theta_j - \rho_j$ is a constant for all $j = 1, \dots, \ell$.*

Proof. (1) ⇒ (2) is trivial.

(3) ⇒ (1): We may assume that $e^{i\theta_j} x_j \in \mathfrak{t}_j$ and $e^{i\rho_j} x_j \in \mathfrak{t}_j$ since

$$B_\theta(K\cdot x, y) = B_\theta(K\cdot x, K\cdot y).$$

Now

$$\begin{aligned} B_\theta(K\cdot x, y) &= B_\theta(K_1\cdot x_1, y_1) + \cdots + B_\theta(K_\ell\cdot x_\ell, y_\ell) \\ &= e^{-i\kappa} \sum_{j=1}^{\ell} B_\theta(K_1\cdot e^{i\theta_j} x_j, e^{i\rho_j} y_j) \end{aligned}$$

and each summand $B_\theta(K_1\cdot e^{i\theta_j} x_j, e^{i\rho_j} y_j) \subset \mathbb{R}$.

(2) ⇒ (3): Suppose $B_\theta(Wx, y)$ is a (closed) line segment. By rotation on x or y we may assume that $B_\theta(Wx, y) \subset \mathbb{R}$. Since

$$B_\theta(Wx, y) = B_\theta(W_1 x_1, y_1) + \cdots + B_\theta(W_\ell x_\ell, y_\ell),$$

each $B_\theta(W_j x_j, y_j)$ is a real line segment, $j = 1, \dots, \ell$. So it suffices to consider simple $\mathfrak{g}_j$. To simplify notations, from now on we drop the index j from $\mathfrak{g}_j, \mathfrak{k}_j, \mathfrak{t}_j, \mathfrak{h}_j, x_j, r_j$ and so on, or simply assume that $\mathfrak{g}$ is simple.

Notice that

$$\tau_{H_\beta}(H_\alpha) = H_\alpha - \frac{2B(H_\alpha, H_\beta)}{B(H_\beta, H_\beta)} H_\beta, \quad \alpha, \beta \in \Delta.$$

As a finite reflection group, the Weyl group W is generated by the reflections τ_{H_β}, $\beta \in \Delta$, and

$$B_\theta(Wx, \tau_{H_\beta} y) = B_\theta(Wx, y) \subset \mathbb{R}$$

so for all $\omega \in W$ and $\beta \in \Delta$,

$$B_\theta(\omega x, \tau_{H_\beta} y) = B_\theta\Big(\omega x, y - \frac{2\beta(y)}{\|\beta\|_\beta^2} H_\beta\Big)$$
$$= B_\theta(\omega x, y) - \frac{2\overline{\beta(y)}}{\|\beta\|_\beta^2} B_\theta(\omega x, H_\beta).$$

Hence for all $\beta \in \Delta$,

$$\frac{2\overline{\beta(y)}}{\|\beta\|_\theta^2} B_\theta(Wx, H_\beta) \subset \mathbb{R}$$

so either (a) $B_\theta(H_\beta, y) = \beta(y) = 0$ for all $\beta \in \Delta$, or (b) for some $\beta \in \Delta$ (depends on y), $\beta(y) \neq 0$, that is, $e^{i\gamma} B_\theta(Wx, H_\beta) \subset \mathbb{R}$ for some $\gamma \in \mathbb{R}$.

Since $\mathfrak{h} = \sum_{\beta\in\Delta} \mathbb{C} H_\beta$ and B is nondegenerate on $\mathfrak{h}$, (a) would not occur because we assume that $y \neq 0$. So (b) occurs, that is, $B_\theta(We^{i\gamma} x, H_\beta) \subset \mathbb{R}$. But then

$$B_\theta(WH_\beta, e^{i\gamma} x) = \overline{B_\theta(We^{i\gamma} x, H_\beta)} \subset \mathbb{R}.$$

Similarly for all $\alpha \in \Delta$,

$$\frac{2\overline{\alpha(e^{i\gamma} x)}}{\|H_\alpha\|_\theta^2} B_\theta(WH_\beta, H_\alpha) \subset \mathbb{R}.$$

Now $B_\theta(WH_\beta, H_\alpha) \subset \mathbb{R}$ since $H_\alpha, H_\beta \in \mathfrak{h}_\mathbb{R} = i\mathfrak{t}$. By contragradience the Weyl group permutes the roots. If $\omega \in W$ then $\omega H_\beta = H_{\omega\cdot\beta}$. We claim that

$$B_\theta(\omega H_\beta, H_\alpha) \neq 0, \text{ for some } \omega \in W.$$

It is because that the Weyl group acts simply transitively on each subset of roots of the same length [Helgason 1978, p. 523]. If H_α and H_β are of the same length, then $\omega H_\beta = H_\alpha$ for some $\omega \in W$ and $B_\theta(\omega H_\beta, H_\alpha) = \|H_\alpha\|_\theta^2 > 0$. Hence the claim follows immediately. When $\mathfrak{g} = \mathfrak{a}_n, \mathfrak{d}_n, \mathfrak{e}_6, \mathfrak{e}_7, \mathfrak{e}_8$, all the roots are of the same length [Helgason 1978, p. 462–474]. Notice that

$$\mathfrak{b}_n : \Delta = \{\pm e_i \pm e_j : 1 \le i \neq j \le n\} \cup \{\pm e_i : 1 \le i \le n\},$$
$$\mathfrak{c}_n : \Delta = \{\pm e_i \pm e_j : 1 \le i \neq j \le n\} \cup \{\pm 2e_i : 1 \le i \le n\}, \text{ and}$$
$$\mathfrak{f}_4 : \Delta = \pm\{e_i \ (i = 1, \dots, 4);\ e_i \pm e_j \ (1 \le i < j \le 4);\ \tfrac{1}{2}(e_1 \pm e_2 \pm e_3 \pm e_4)\}.$$

For each case, the root length squares are either 1 or 2 and the claim is clearly true for them. Finally when $\mathfrak{g} = \mathfrak{g}_2$, the root length squares are either 2 or 6,

$$\Delta = \pm\{e_1 - e_2, e_2 - e_3, e_1 - e_3, 2e_1 - e_2 - e_3, 2e_2 - e_1 - e_3, 2e_3 - e_1 - e_2\}$$

and the claim is also true. As a result $\alpha(e^{i\gamma}x) \in \mathbb{R}$ for all $\alpha \in \Delta$ so $e^{i\gamma}x \in \mathfrak{h}_{\mathbb{R}} = i\mathfrak{t}$. Similarly we have the same conclusion for y. Then clearly $\theta_j - \rho_j$ is a constant, $j = 1, \dots, \ell$. □

Proof of Theorem 2.2. We first show that the first four statements are equivalent.

(1) ⇒ (2): We may assume that $x \in \mathfrak{h}$. We have $B_\theta(K \cdot x, x) = B_\theta(\pi(K \cdot x), x)$ and it is convex since $\pi(K \cdot x)$ is convex.

(2) ⇔ (3): Lemma 2.6.

(3) ⇒ (4): The case $x = 0$ is trivial. For $x \neq 0$, we may assume that each component $x_j \neq 0$ in the expression $x = x_1 + \cdots + x_\ell \in \mathfrak{g}$. Then apply Proposition 2.7.

(4) ⇒ (1): By Theorem 2.1.

(5) ⇒ (2): obvious.

(4) ⇒ (5): Let $y = y_1 + \cdots + y_\ell \in \mathfrak{g}_1 + \cdots + \mathfrak{g}_\ell$. Then

$$B_\theta(K \cdot x, y) = B_\theta(K_1 \cdot x_1, y_1) + \cdots + B_\theta(K_1 \cdot x_\ell, y_\ell).$$

By (4) there exist $k_j \in K_j$ and $\theta_j \in \mathbb{R}$ so that $t_j := e^{i\theta_j}\mathrm{Ad}(k_j)x_j \in \mathfrak{t}_j$ for each $j = 1, \dots, \ell$. Write

$$y_j = y_j^{(1)} + i y_j^{(2)},$$

for $y_j^{(1)}, y_j^{(2)} \in \mathfrak{k}$. So

$$\begin{aligned} B_\theta(K_j \cdot x_j, y_j) &= e^{-i\theta_j} B_\theta(K \cdot t_j, y_j) \\ &= e^{-i\theta_j} \left\{ B\big(\mathrm{Ad}(k_j)t_j, y_j^{(1)}\big) + i B\big(\mathrm{Ad}(k_j)t_j, y_j^{(2)}\big) : k_j \in K_j \right\} \end{aligned}$$

which is convex by a result of Tam [2002]. Hence $B_\theta(K \cdot x, y)$ is a sum of convex sets and thus convex. □

Remark 2.8. The second author conjectured (see [Tam 2001, Conjecture 4.1]) that for a normal $x \in \mathfrak{g}$ (semisimple), if $B_\theta(K \cdot x, x)$ is convex, then there is $\gamma \in \mathbb{R}$ such that $e^{i\gamma}x \in \mathfrak{t}$. It is not true in view of Theorem 2.2. Consider the semisimple $\mathfrak{g} := \mathfrak{a}_1 \times \mathfrak{a}_1$. To be concrete, let $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C}) \times \mathfrak{sl}_2(\mathbb{C})$ with $K = \mathrm{SU}(2) \oplus \mathrm{SU}(2)$. Consider the normal $x = \mathrm{diag}\,(x_1, -x_1) \oplus \mathrm{diag}\,(x_2, -x_2)$, where $x_1, x_2 \in \mathbb{C}$. Then for $k = k_1 \oplus k_2 \in K$,

$$\begin{aligned} \mathrm{tr}\, kxk^{-1}x^* &= \mathrm{tr}\, k_1 \mathrm{diag}\,(x_1, -x_1)k_1^{-1}\mathrm{diag}\,(\bar{x}_1, -\bar{x}_1) \\ &\quad + \mathrm{tr}\, k_2 \mathrm{diag}\,(x_2, -x_2)k_2^{-1}\mathrm{diag}\,(\bar{x}_2, -\bar{x}_2). \end{aligned}$$

By Theorem 1.3 the set

$$\left\{ \mathrm{tr}\, k_i \mathrm{diag}\,(x_i, -x_i)k_1^{-1}\mathrm{diag}\,(\bar{x}_i, -\bar{x}_i) : k_i \in \mathrm{SU}(n) \right\}$$

is convex, $i = 1, 2$, so $\{\mathrm{tr}\, kxk^{-1}x^* : k \in K\}$ is the sum of two convex sets and thus is convex. However, x_1, x_2 need not be collinear with 0.

By Proposition 2.4 (a) $B_\theta(K \cdot x, x)$ is symmetric about the real axis. For some simple Lie algebras, more symmetry occurs for $B_\theta(K \cdot x, x)$ if $x \in \mathfrak{g}$ is normal. Indeed the symmetry is also true for $B_\theta(K \cdot x, y)$ for each $y \in \mathfrak{g}$.

Proposition 2.9. *Let $\mathfrak{g}$ be simple and of type $\mathfrak{b}_\ell$, $\mathfrak{c}_\ell$, $\mathfrak{d}_\ell$ (ℓ even), $\mathfrak{g}_2$, $\mathfrak{f}_4$, $\mathfrak{e}_7$ and $\mathfrak{e}_8$. Let $x \in \mathfrak{g}$ be normal. The sets $\pi(K \cdot x) \subset \mathfrak{h}$ and $B_\theta(K \cdot x, y) \subset \mathbb{C}$ are symmetric about the origin for each $y \in \mathfrak{g}$.*

Proof. We may assume that $x \in \mathfrak{h}$. The Weyl group W contains -1 [Helgason 1978, p. 523] so the desired result follows. □

It is known [Djoković and Tam 2003] that if $x \in \mathfrak{g}$ is normal, then $B_\theta(K \cdot x, y)$ is star-shaped with respect to the center 0 for each $y \in \mathfrak{g}$.

We do not know whether $\pi(K \cdot x)$ is star shaped or not and the following conjectures [Tam 2001] are still open.

Conjecture 2.10. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra. If $x, y \in \mathfrak{g}$, then $B_\theta(K \cdot x, y)$ is star-shaped with respect to the star center* 0.

Conjecture 2.11. *Let $\mathfrak{g}$ be a complex semisimple Lie algebra. If $x \in \mathfrak{g}$, then $\pi(K \cdot x)$ is star-shaped with respect to the star center* 0.

We remark that these conjectures can be reduced to the simple cases. The cases $\mathfrak{a}_\ell$ ($\ell \geq 1$), $\mathfrak{d}_\ell$ ($\ell \geq 2$), $\mathfrak{e}_6$, $\mathfrak{e}_7$ for Conjecture 2.10 are true [Cheung and Tsing 1996; Djoković and Tam 2003].

Added in proof

The authors very recently proved Conjecture 2.11 affirmatively.

References

[Au-Yeung and Sing 1977] Y. H. Au-Yeung and F. Y. Sing, "A remark on the generalized numerical range of a normal matrix", *Glasgow Math. J.* **18**:2 (1977), 179–180. MR 56 #8594 Zbl 0359.47001

[Au-Yeung and Tsing 1983] Y. H. Au-Yeung and N.-K. Tsing, "A conjecture of Marcus on the generalized numerical range", *Linear and Multilinear Algebra* **14**:3 (1983), 235–239. MR 85e:15029 Zbl 0521.15016

[Bebiano and Da Providência 1996] N. Bebiano and J. Da Providência, "Another proof of a conjecture of Marcus on the c-numerical range", *Linear and Multilinear Algebra* **41**:1 (1996), 35–40. MR 97g:15029 Zbl 0871.15028

[Cheung and Tsing 1996] W.-S. Cheung and N.-K. Tsing, "The C-numerical range of matrices is star-shaped", *Linear and Multilinear Algebra* **41**:3 (1996), 245–250. MR 97k:15060 Zbl 0876.15021

[Djoković and Tam 2003] D. Z. Djoković and T.-Y. Tam, "Some questions about semisimple Lie groups originating in matrix theory", *Canad. Math. Bull.* **46**:3 (2003), 332–343. MR 2004g:22008 Zbl 1047.22013

[Helgason 1978] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Pure and Applied Mathematics **80**, Academic Press, New York, 1978. MR 80k:53081 Zbl 0451.53038

[Horn 1954] A. Horn, "Doubly stochastic matrices and the diagonal of a rotation matrix", *Amer. J. Math.* **76** (1954), 620–630. MR 16,105c Zbl 0055.24601

[Kostant 1973] B. Kostant, "On convexity, the Weyl group and the Iwasawa decomposition", *Ann. Sci. École Norm. Sup.* (4) **6** (1973), 413–455 (1974). MR 51 #806 Zbl 0293.22019

[Marcus 1979] M. Marcus, "Some combinatorial aspects of numerical range", *Ann. New York Acad. Sci.* **319** (1979), 368–376. MR 81g:15028 Zbl 0483.15017

[Schur 1923] I. Schur, "Über eine Klasse von Mittelbildungen mit Anwendungen auf der Determinantentheorie", *Sitzungsber. Berl. Math. Ges.* **22** (1923), 9–20. JFM 49.0054.01

[Tam 2001] T.-Y. Tam, "On the shape of numerical ranges associated with Lie groups", *Taiwanese J. Math.* **5**:3 (2001), 497–506. MR 2002f:15043 Zbl 0986.15022

[Tam 2002] T.-Y. Tam, "Convexity of generalized numerical range associated with a compact Lie group", *J. Aust. Math. Soc.* **72**:1 (2002), 57–66. MR 2002g:15056 Zbl 1007.15021

[Tsing 1981] N. K. Tsing, "On the shape of the generalized numerical ranges", *Linear and Multilinear Algebra* **10**:3 (1981), 173–182. MR 82m:15026 Zbl 0471.47002

[Tsing 1984] N.-K. Tsing, "The constrained bilinear form and the C-numerical range", *Linear Algebra Appl.* **56** (1984), 195–206. MR 85b:15033 Zbl 0528.15007

[Westwick 1975] R. Westwick, "A theorem on numerical range", *Linear and Multilinear Algebra* **2** (1975), 311–315. MR 51 #11132 Zbl 0303.47001

Received December 15, 2007. Revised February 29, 2008.

TIN-YAU TAM
DEPARTMENT OF MATHEMATICS AND STATISTICS
221 PARKER HALL
AUBURN UNIVERSITY
AUBURN, AL 36849-5310
UNITED STATES
tamtiny@auburn.edu

WAI-SHUN CHEUNG
DEPARTMENT OF MATHEMATICS AND STATISTICS
221 PARKER HALL
AUBURN UNIVERSITY
AUBURN, AL 36849-5310
UNITED STATES
cheungwaishun@gmail.com