J. Austral. Math. Soc. **72** (2002), 57–66

CONVEXITY OF GENERALIZED NUMERICAL RANGE ASSOCIATED WITH A COMPACT LIE GROUP

TIN-YAU TAM

Dedicated to Professor Y. H. Au-Yeung

(Received 5 March 1999; revised 16 January 2001)

Communicated by A. H. Dooley

Abstract

Westwick's convexity theorem on the numerical range is generalized in the context of compact connected Lie groups.

2000 *Mathematics subject classification*: primary 15A60, 22E99.
Keywords and phrases: Numerical range, convexity, compact connected Lie group, co-adjoint orbit.

1. Introduction

The celebrated Toeplitz-Hausdorff theorem [21, 13] asserts that the numerical range of an $n \times n$ complex matrix A,

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

is a compact convex set in $\mathbb{C}$. Toeplitz [21] proved that $W(A)$ has a convex outer boundary and Hausdorff [13] showed that the intersection of every line with $W(A)$ is connected or empty. It is remarkable for it states that the image of the unit sphere in $\mathbb{C}^n$ (a hollow object) is a compact convex set in $\mathbb{C}$ under the nonlinear map, $x \mapsto x^*Ax$. Since then various generalizations have been considered ranging from

The author thanks Professor R. Westwick for providing the manuscript [19] shortly after they met in the Second Mini-Matrix Conference held in Hong Kong, December 1993. Thanks are given to Professor S. Cheung and Jennifer Rude for translating [19] and Professor M. Liao for helpful discussion.

finite dimensional linear and multilinear maps [17] to operators on normed spaces [8]. The volume of literature on the subject has been growing rapidly in the last decades [12]. Halmos introduced the k-numerical range of A: $W_k(A) = \{\sum_{i=1}^k x_i^* A x_i : x_1, \ldots, x_k$ are orthonormal vectors in $\mathbb{C}^n\}$, $k = 1, \ldots, n$. He conjectured and Berger [7] proved that $W_k(A)$ is always convex. Then Westwick [22] considered the c-numerical range of A, where $c \in \mathbb{C}^n$:

$$W_c(A) := \left\{ \sum_{i=1}^n c_i x_i^* A x_i : x_1, \ldots, x_n \text{ are orthonormal vectors in } \mathbb{C}^n \right\}.$$

It can be formulated as $W_C(A) := \{\operatorname{tr} CUAU^* : U \in U(n)\}$. Here $U(n)$ denotes the unitary group and C is normal with eigenvalues $c \in \mathbb{C}^n$. Notice that $W_C(A) = \{\operatorname{tr} CUAU^* : [U] \in U(n)/\triangle(n)\}$, where $\triangle(n) \subset U(n)$ is the subgroup of diagonal matrices and $U \mapsto [U]$ is the natural projection from $U(n)$ onto the homogenous space $U(n)/\triangle(n)$. Westwick proved that $W_C(A)$ is always convex for real c, that is, C is Hermitian (this is known as Westwick's convexity theorem) but fails to be convex for complex c. The main idea of Westwick's proof is the application of Morse theory on $U(n)/\triangle(n)$. Poon [18] was the first to give an elementary proof to Westwick's result. The result was later rediscovered by Ginsburg [6, page 8].

If $A = A_1 + iA_2$ is the Hermitian decomposition of A, then $W_C(A)$ may be identified as the subset of $\mathbb{R}^2$,

$$W_C(A_1, A_2) := \left\{(\operatorname{tr} CUA_1U^*, \operatorname{tr} CUA_2U^*) : U \in U(n)\right\}. \tag{1}$$

Westwick considered the map $f_B : U(n)/\triangle(n) \to \mathbb{R}$ defined by $[U] \mapsto \operatorname{tr} CUBU^*$, where B is a given Hermitian matrix. If the level surface $f_B^{-1}(a)$ is connected (or empty) in $U(n)/\triangle(n)$ for any $a \in \mathbb{R}$, then convexity follows by Hausdorff's argument. He examined the critical points of the function f_B and evaluated the Hessians at those points, assuming that B and C are both regular, that is, the Hermitian matrices B and C have distinct eigenvalues. The critical points have even indices. Then by the handlebody decomposition theorem, the level surface $f_B^{-1}(a)$ is connected. Westwick also affirmed that the connectedness is valid even for nonregular B and C. But Raïs [19] pointed out that this is not obvious.

It is well known that $U(n)$ is a compact connected Lie group whose Lie algebra $\mathfrak{u}(n)$ is the set of skew Hermitian matrices. Notice that $\operatorname{tr} CUBU^* = \operatorname{tr} BUCU^* = -\operatorname{tr}(iB)U(iC)U^*$ and thus (1) can be written as $W_C(A_1, A_2) = \{(\operatorname{tr} A_1L, \operatorname{tr} A_2L) : L \in O(C)\}$, where $O(C) := \{UCU^* : U \in U(n)\}$ is the adjoint orbit of C in $\mathfrak{u}(n)$ which is identified with the set of Hermitian matrices. Moreover, $O(C)$ and $U(n)/\triangle(n)$ can be identified. So the following consideration of Raïs [19] is natural: Let G be a compact Lie group with Lie algebra $\mathfrak{g}$ which is equipped with a G-invariant inner product $\langle \cdot, \cdot \rangle$, that is, $\langle \operatorname{Ad}(g)X, \operatorname{Ad}(g)Y \rangle = \langle X, Y \rangle$, $X, Y \in \mathfrak{g}$, $g \in G$.

For $X_1, X_2, Y \in \mathfrak{g}$, the Y-numerical range of (X_1, X_2) is defined to be the following subset of $\mathbb{R}^2$:

$$W_Y(X_1, X_2) :=\{(\langle X_1, \mathrm{Ad}(g)Y\rangle, \langle X_2, \mathrm{Ad}(g)Y\rangle) : g \in G\}. \tag{2}$$

Note that (2) can be rewritten as

$$W_Y(X_1, X_2) =\{(\langle X_1, L\rangle, \langle X_2, L\rangle) : L \in O(Y)\}, \tag{3}$$

where $O(Y) := \{\mathrm{Ad}(g)Y : g \in G\}$ is the adjoint orbit of Y in $\mathfrak{g}$. If $G(Y) := \{g \in G : \mathrm{Ad}(g)Y = Y\}$ denotes the centralizer of $Y \in \mathfrak{g}$ in G, then

$$W_Y(X_1, X_2) = \{(\langle X_1, \mathrm{Ad}(g)Y\rangle, \langle X_2, \mathrm{Ad}(g)Y\rangle) : [g] \in G/G(Y)\},$$

where $g \mapsto [g]$ is the natural projection from G onto $G/G(Y)$. Indeed, $O(Y)$ and $G/G(Y)$ can be identified.

We will use the fact that $O(Y) \cap \mathfrak{t}$ is a nonempty finite set, where $Y \in \mathfrak{g}$ and $\mathfrak{t}$ is the Lie algebra of a maximal torus T of G when G is compact and connected [16].

In Section 2, we will prove the convexity of $W_Y(X_1, X_2)$ via Atiyah's lemma on compact connected symplectic manifolds and the Kirillov-Kostant-Souriau symplectic structure of the co-adjoint orbits of a Lie group. The statements for classical groups, namely, $SO(n)$, $SU(n)$ and $Sp(n)$ are explicitly worked out. Convexity fails to be true when $G = O(2n)$ but remains valid when $G = O(2n + 1)$. It demonstrates that the connectedness is necessary. In Section 3, we suggest an approach for the convexity via Bott-Samelson-Raïs' result, without symplectic technique.

2. Convexity of the generalized numerical ranges

We now identify $\mathfrak{g}^*$ with $\mathfrak{g}$ via the isomorphism $\varphi : X \mapsto \langle X, \cdot\rangle$, $X \in \mathfrak{g}$, that is, $z(X) = \langle X, \varphi^{-1}(z)\rangle$, $z \in \mathfrak{g}^*$, and $\mathfrak{g}^*$ has an induced inner product $\langle\cdot, \cdot\rangle$ (abuse of notation) such that $\langle x, y\rangle := \langle\varphi^{-1}(x), \varphi^{-1}(y)\rangle$, $x, y \in \mathfrak{g}$. Notice that

$$\varphi(\mathrm{Ad}(g)Y) = \langle\mathrm{Ad}(g)Y, \cdot\rangle = \varphi(Y, \mathrm{Ad}(g^{-1})(\cdot)) = \mathrm{Ad}^*(g)(\varphi(Y)). \tag{4}$$

Here the co-adjoint representation $\mathrm{Ad}^* : G \to \mathrm{Aut}(\mathfrak{g}^*)$ of G in $\mathfrak{g}^*$ is defined by $g \mapsto \mathrm{Ad}^*(g)$ such that $\mathrm{Ad}^*(g)(y)Y = y(\mathrm{Ad}(g^{-1})Y)$, where $y \in \mathfrak{g}^*$, $Y \in \mathfrak{g}$. The differential of Ad^* yields the co-adjoint representation of $\mathfrak{g}$ on $\mathfrak{g}^*$, namely, $\mathrm{ad}^* : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}^*)$ such that

$$\mathrm{ad}^*(X)y(Y) = -y(\mathrm{ad}(X)Y) = y([Y, X]), \qquad X, Y \in \mathfrak{g}, \quad y \in \mathfrak{g}^*.$$

Similarly as in (3), given a compact Lie group G, we define

$$W_y(x_1, x_2) := \{(\langle x_1, \ell\rangle, \langle x_2, \ell\rangle) : \ell \in O_y\},$$

where $O_y := \{\mathrm{Ad}^*(g)y : g \in G\}$ is the co-adjoint orbit of $y \in \mathfrak{g}^*$. From (4) $\varphi(O(Y)) = O_{\varphi(y)}$. Thus

$$W_y(x_1, x_2) = W_{\varphi^{-1}(y)}(\varphi^{-1}(x_1), \varphi^{-1}(x_2)) = \{(\ell(\varphi^{-1}(x_1)), \ell(\varphi^{-1}(x_2))) : \ell \in O_y\}. \tag{5}$$

If $G_y := \{g \in G : \mathrm{Ad}^*(g)y = y\}$ denotes the stabilizer of $y \in \mathfrak{g}^*$, whose Lie algebra is $\mathfrak{g}_y = \{X \in \mathfrak{g} : \mathrm{ad}^*(X)(y) = 0\} = \{X \in \mathfrak{g} : y([Y, X]) = 0, \text{ for all } Y \in \mathfrak{g}\}$, then we have

$$W_y(x_1, x_2) = \{(\langle x_1, \mathrm{Ad}^*(g)y\rangle, \langle x_2, \mathrm{Ad}^*(g)y\rangle) : [g] \in G/G_y\},$$

where $g \mapsto [g]$ is the natural projection from G onto G/G_y. The tangent space of the co-adjoint orbit O_y and $\mathfrak{g}/\mathfrak{g}_y$ can be identified.

Atiyah [1, Lemma 1.3] obtained the following result (also see [10, 11, 15]).

LEMMA 2.1. *Let M be a compact connected symplectic manifold and $f : M \to \mathbb{R}$ a smooth function whose Hamiltonian vector field generates a torus action. Then for any $a \in \mathbb{R}$, the level surface $f^{-1}(a)$ is connected* (*or empty*).

A symplectic manifold M is a differentiable manifold of even dimension with an exterior differential 2-form ω satisfying (1) $d\omega = 0$, that is, ω is closed, and (2) ω is of maximal rank. A real-valued smooth function f on M defines a Hamiltonian vector field ξ_f which corresponds to the 1-form df using the duality defined by ω, that is, $\iota(\xi_f)\omega + df = 0$ [14, page 232].

LEMMA 2.2. *Let G be a compact Lie group. If X_1, X_2 and Y are in $\mathfrak{g}$, $x_1, x_2, y \in \mathfrak{g}^*$, then*

(1) $W_Y(X_1, X_2) = W_{\mathrm{Ad}(g_1)Y}(\mathrm{Ad}(g_2)X_1, \mathrm{Ad}(g_2)X_2)$ *for any $g_1, g_2 \in G$. Hence if G is connected and $\mathfrak{t}$ is the Lie algebra of a maximal torus T of G, then Y and one of the X's can be taken as elements of $\mathfrak{t}$;*
(2) $W_y(x_1, x_2) = W_{\mathrm{Ad}^*(g_1)y}(\mathrm{Ad}(g_2)x_1, \mathrm{Ad}(g_2)x_2)$ *for any $g_1, g_2 \in G$;*
(3) *rotating $W_Y(X_1, X_2)$ ($W_y(x_1, x_2)$) by an angle θ yields $W_Y(X_1', X_2')$ ($W_y(x_1', x_2')$) where $(X_1', X_2') = (X_1 \cos\theta - X_2 \sin\theta,\ X_1 \sin\theta + X_2 \cos\theta)$ and $(x_1', x_2') = (x_1 \cos\theta - x_2 \sin\theta,\ x_1 \sin\theta + x_2 \cos\theta)$.*

PROOF. (1) and (2). For any $g_1, g_2 \in G$,

$$\langle \mathrm{Ad}(g_2)X, \mathrm{Ad}(g)\,\mathrm{Ad}(g_1)Y\rangle = \langle X, \mathrm{Ad}(g_2^{-1}gg_1)Y\rangle.$$

As g runs through the group G, so does $g_2^{-1}gg_1$. Statement (3) follows from direct computation. □

THEOREM 2.3. *Let G be a compact connected Lie group. For $x_1, x_2, y \in \mathfrak{g}^*$ and $Y \in \mathfrak{g}$, $W_y(x_1, x_2)$ is a compact convex set in $\mathbb{R}^2$. Thus for $X_1, X_2, Y \in \mathfrak{g}$, $W_Y(X_1, X_2)$ is a compact convex set.*

PROOF. For any Lie group G, the co-adjoint orbit $\Omega := O_y$ has a natural symplectic structure, known as the Kirillov-Kostant-Souriau structure [14, pages 230–234]. Let $T_z\Omega$ be the tangent space of Ω at the point $z \in \Omega$. The symplectic form is given by $\omega_z(\alpha, \beta) = z([A, B])$, $\alpha, \beta \in T_z\Omega$, $z \in \Omega$, and α and β are corresponding to the elements A and $B \in \mathfrak{g}$, respectively (under the identification $T_z\Omega$ with $\mathfrak{g}/\mathfrak{g}_z$), that is, $\beta = \mathrm{ad}^*(B)(z) = d/dt|_{t=0}\, \mathrm{Ad}^*(e^{-tB})z$.

In view of (5), it is sufficient to consider the smooth function $f : \Omega \to \mathbb{R}$ defined by $f(z) = z(X)$, where $z \in \Omega$ for any given $X \in \mathfrak{g}$, that is, f is the restriction on Ω of the linear functional of $\mathfrak{g}^*$ corresponding to $X \in \mathfrak{g}$, and show that $f^{-1}(a)$ is connected (or empty) for any $a \in \mathbb{R}$. This implies that the intersection of $W_y(x_1, x_2)$ with every vertical (horizontal as well) straight line is connected (or empty). By Lemma 2.2 (3), the intersection of $W_y(x_1, x_2)$ with every straight line is connected (or empty). Now

$$\begin{aligned} df_z(\beta) &= \frac{d}{dt}\bigg|_{t=0} f(\mathrm{Ad}^*(e^{-tB})z) = \frac{d}{dt}\bigg|_{t=0} \mathrm{Ad}^*(e^{-tB})z(X) \\ &= \frac{d}{dt}\bigg|_{t=0} z(\mathrm{Ad}(e^{tB})X) = z([B, X]). \end{aligned}$$

So $\iota(\xi_f)\omega + df = 0$ means that $\omega_z(\xi_f(z), \beta) + df_z(\beta) = 0$ for all $\beta \in T\Omega$ and $z \in \Omega$. It amounts to $z([Z, B]) + z([B, X]) = 0$ for all $B \in \mathfrak{g}$ and $z \in \Omega$, where $Z \in \mathfrak{g}$ corrresponds to $\xi_f(z)$. So $z([X - Z, B]) = 0$ for all $B \in \mathfrak{g}$, that is, $Z = X \bmod \mathfrak{g}_z$. In other words, the corresponding Hamiltonian vector field associated with f is just the natural action of X on Ω. If G is compact connected, so is Ω. If, in addition, X is in $\mathfrak{t}$, the Lie algebra of a torus $T \subset G$, then the conditions of Lemma 2.1 are satisfied [1, page 2]. By Lemma 2.2 (a), the level set, $f^{-1}(a)$ is connected (or empty) for any $a \in \mathbb{R}$. □

We now work out the explicit statements for some classical groups, namely, the unitary group, the special unitary group, the orthogonal group $O(2n+1)$, the special orthogonal group $SO(n)$ and the symplectic group $Sp(n)$. The symplectic group $Sp(n) \subset U(2n)$ consists of

$$\begin{bmatrix} A & -\overline{B} \\ B & \overline{A} \end{bmatrix} \in U(2n).$$

COROLLARY 2.4. (1) (Westwick [22]) *Let $G = U(n)$ or $SU(n)$. The C-numerical range $W_C(A_1, A_2) = \{(\mathrm{tr}\, A_1UCU^*, \mathrm{tr}\, A_2UCU^*) : U \in G\}$ is convex, where A_1, A_2 and C are Hermitian matrices.*

(2) *The set* $W_C(A_1, A_2) = \{(\mathrm{tr}\, A_1OCO^T, \mathrm{tr}\, A_2OCO^T) : O \in SO(n)\}$ *is convex, where* A_1, A_2, *and* C *are real skew symmetric matrices.*
(3) *The set* $W_C(A_1, A_2) = \{(\mathrm{tr}\, A_1OCO^T, \mathrm{tr}\, A_2OCO^T) : O \in O(2n+1)\}$ *is convex and is equal to* $\{(\mathrm{tr}\, A_1OCO^T, \mathrm{tr}\, A_2OCO^T) : O \in SO(2n+1)\}$, *where* A_1, A_2, *and* C *are real skew symmetric matrices.*
(4) *The set* $W_C(A_1, A_2) = \{(\mathrm{tr}\, A_1UCU^*, \mathrm{tr}\, A_2UCU^*) : U \in Sp(n)\}$ *is convex, where* $A_1, A_2, C \in \mathfrak{sp}(n)$.

PROOF. (1) Notice that $W_C(A_1, A_2)$ is the reflection of the convex set $W_{iC}(iA_1, iA_2)$ about the line $x = y$ on the xy plane. When $G = SU(n)$, the Lie algebra is the set of traceless skew Hermitian matrices. Then for any $U \in SU(n)$,

$$(\mathrm{tr}\, A_1UCU^*, \mathrm{tr}\, A_2UCU^*) = (\mathrm{tr}\, \hat{A}_1U\hat{C}U^*, \mathrm{tr}\, \hat{A}_2U\hat{C}U^*) + \frac{1}{n}(\mathrm{tr}\, C\, \mathrm{tr}\, A_1, \mathrm{tr}\, C\, \mathrm{tr}\, A_2),$$

where $\hat{C} = C - (\mathrm{tr}\, C/n)I$ and $\hat{A}_1$ and $\hat{A}_2$ are similarly defined. They are traceless skew Hermitian matrices. So $W_C(A_1, A_2)$ is just a translation of the convex set $W_{\hat{C}}(\hat{A}_1, \hat{A}_2)$.

(2) and (4) are obvious.

(3) The orthogonal group $O(k) = SO(k) \cup DSO(k)$ has two connected components $SO(k)$ and $DSO(k) = \{DO : O \in \mathrm{SO}(k)\}$, where D is the diagonal matrix with $\mathrm{diag}(1, \ldots, 1, -1)$. So we have $W_C(A_1, A_2) = \{(\mathrm{tr}\, A_1OCO^T, \mathrm{tr}\, A_2OCO^T) : O \in O(k)\} = W_1 \cup W_2$, where

$$W_1 := \left\{(\mathrm{tr}\, A_1OCO^T, \mathrm{tr}\, A_2OCO^T) : O \in SO(k)\right\}$$

and

$$\begin{aligned} W_2 &:= \left\{(\mathrm{tr}\, A_1OCO^T, \mathrm{tr}\, A_2OCO^T) : O \in DSO(k)\right\} \\ &= \left\{(\mathrm{tr}\, A_1OC'O^T, \mathrm{tr}\, A_2OC'O^T) : O \in SO(k)\right\} \end{aligned}$$

are convex by (2) with $C' = D^TCD$.

When $k = 2n+1$, $W_1 = W_2$ since $\{OCO^T : O \in SO(2n+1)\} = \{OC'O^T : O \in DSO(2n+1)\}$. Hence $W_C(A_1, A_2)$ is convex. □

We remark that (2) and (3) are valid for general real C since $W_C(A_1, A_2) = W_{\hat{C}}(A_1, A_2)$, where $\hat{C} = (C - C^T)/2$. We also remark that the connectedness of G in Theorem 2.3 is necessary when we consider $O(2n)$. Let

$$C = \begin{bmatrix} 0 & c \\ -c & 0 \end{bmatrix}, \qquad A_1 = \begin{bmatrix} 0 & a_1 \\ -a_1 & 0 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0 & a_2 \\ -a_2 & 0 \end{bmatrix}.$$

Then $W_C(A_1, A_2) = \{\pm c(a_1, a_2)\}$ which is not convex if $c \neq 0$ and a_1 and a_2 are not both zero, because $W_1 = \{c(a_1, a_2)\}$ and $W_2 = \{-c(a_1, a_2)\}$. The argument extends to $2n$. Consider

$$C = \begin{bmatrix} 0 & c_1 \\ -c_1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & c_n \\ -c_n & 0 \end{bmatrix},$$

$$A_1 = \begin{bmatrix} 0 & a_1 \\ -a_1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & a_n \\ -a_n & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & b_1 \\ -b_1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & b_n \\ -b_n & 0 \end{bmatrix}.$$

Recall that $W_C(A_1, A_2) = W_1 \cup W_2$ and denote by $\mathscr{C}_1$ ($\mathscr{C}_2$) the convex hull of the elements $(\pm c_{\theta(1)}, \ldots, \pm c_{\theta(n)})$, $\theta \in S_n$ and for even (odd) number of negative signs. By a result in [20], W_1 (W_2) is the set of $-2\big(\sum_i a_i \xi_i, \sum_i b_i \xi_i\big)$, where $\xi = (\xi_1, \ldots, \xi_n)$ are in $\mathscr{C}_1$ ($\mathscr{C}_2$). So the set W_1 (W_2) is the convex hull of the points $\big(\sum_i \pm a_i c_{\theta(i)}, \sum_i \pm b_i c_{\theta(i)}\big)$, where $\theta \in S_n$ and for even (odd) number of negative signs. Now if we choose a's, b's and c's positive and set them in decreasing order, respectively, then $\big(\sum_i a_i c_i, \sum_i b_i c_i\big) \in W_1$ but not in W_2.

The statement of Theorem 2.3 is best possible in the sense that $W_Y(X_1, \ldots, X_p)$ may fail to be true if $p \geq 3$. Indeed, when $G = U(n)$ and $Y = \operatorname{diag}(1, 0, \ldots, 0)$, $W_Y(X_1, \ldots, X_p)$ fails to be convex [3] for some choice of X's when $p \geq 3$ or $n = 2$ while $p = 3$. But it is convex when $p = 3$ and $n > 2$ (also see [4]).

3. Remarks

Since the map $G \to \mathbb{R}$ defined by $g \mapsto \langle X, \operatorname{Ad}(g)Y \rangle$ (or $O(Y) \to \mathbb{R}$ defined by $L \mapsto \langle X, L \rangle$) is clearly continuous, $W_Y(X_1, X_2)$ is compact in $\mathbb{R}^2$ if G is a compact Lie group, where X's and Y are in $\mathfrak{g}$. The following result deals with the continuity of the map $\prod^3 \mathfrak{g} \to \mathscr{C}(\mathbb{R}^2)$, where $\mathscr{C}(\mathbb{R}^2)$ is the set of compact sets in $\mathbb{R}^2$, equipped with Hausdorff topology, such that $(X_1, X_2, Y) \mapsto W_Y(X_1, X_2)$. We will then discuss a possible approach to Theorem 2.3.

PROPOSITION 3.1. *Let G be a compact Lie group and let $\mathscr{C}(\mathbb{R}^2)$ be the set of compact subsets of $\mathbb{R}^2$ equipped with Hausdorff metric. Let $\|\cdot\|$ be the norm induced by the G-invariant inner product on $\mathfrak{g}$. Let $|||\cdot|||$ be the norm of $\prod^3 \mathfrak{g}$ induced by the norm of $\mathfrak{g}$, that is, $|||(Z_1, Z_2, Z_3)||| = \max_{i=1,2,3} \|Z_i\|$.*

(1) *The function $\mathscr{W} : \prod^3 \mathfrak{g} \to \mathscr{C}(\mathbb{R}^2)$ defined by $\mathscr{W}(X_1, X_2, Y) = W_Y(X_1, X_2)$ is continuous.*

(2) *If $Y \in \mathfrak{g}$, then the function $\mathscr{W}_Y : \prod^2 \mathfrak{g} \to \mathscr{C}(\mathbb{R}^2)$ defined by $\mathscr{W}_Y(X_1, X_2) = W_Y(X_1, X_2)$ is uniformly continuous.*

(3) *Similar results are true for $W_y(x_1, x_2)$.*

PROOF. (1) Recall the Hausdorff metric for $\mathscr{C}(\mathbb{R}^2)$: write $M + (\epsilon) = \{z + \alpha : z \in M, \|\alpha\|_2 < \epsilon\}$ for each $M \in \mathscr{C}(\mathbb{R}^2)$ and $\epsilon > 0$, where $\|\cdot\|_2$ denotes the Euclidean norm on $\mathbb{R}^2$. If $M, N \in \mathscr{C}(\mathbb{R}^2)$, then the Hausdorff metric $d(M, N)$ is defined to be the infimum of all positive numbers ϵ such that both $M \subset N + (\epsilon)$ and $N \subset M + (\epsilon)$ hold. Now by the triangle inequality and the Cauchy-Schwarz inequality,

$$\begin{aligned}
&\|(\langle X_1, \mathrm{Ad}(g)Y\rangle, \langle X_2, \mathrm{Ad}(g)Y\rangle) - (\langle X_1', \mathrm{Ad}(g)Y'\rangle, \langle X_2', \mathrm{Ad}(g)Y'\rangle)\|_2 \\
&\quad = \|(\langle X_1 - X_1', \mathrm{Ad}(g)Y\rangle, \langle X_2 - X_2', \mathrm{Ad}(g)Y\rangle) \\
&\qquad + (\langle X_1', \mathrm{Ad}(g)(Y - Y')\rangle, \langle X_2', \mathrm{Ad}(g)(Y - Y')\rangle)\|_2 \\
&\quad \le \|(\langle X_1 - X_1', \mathrm{Ad}(g)Y\rangle, \langle X_2 - X_2', \mathrm{Ad}(g)Y\rangle)\|_2 \\
&\qquad + \|(\langle X_1', \mathrm{Ad}(g)(Y - Y')\rangle, \langle X_2', \mathrm{Ad}(g)(Y - Y')\rangle)\|_2 \\
&\quad \le \left(\sum_{i=1}^{2} \|X_i - X_i'\|^2 \|\mathrm{Ad}(g)Y\|^2\right)^{1/2} + \left(\sum_{i=1}^{2} \|X_i'\|^2 \|\mathrm{Ad}(g)(Y - Y')\|^2\right)^{1/2} \\
&\quad = \left(\sum_{i=1}^{2} \|X_i - X_i'\|^2\right)^{1/2} \|Y\| + \left(\sum_{i=1}^{2} \|X_i'\|^2\right)^{1/2} \|Y - Y'\|.
\end{aligned}$$

So

(6)
$$\begin{aligned}
&d\big(W_Y(X_1, X_2), W_{Y'}(X_1', X_2')\big) \\
&\quad \le \left(\sum_{i=1}^{2} \|X_i - X_i'\|^2\right)^{1/2} \|Y\| + \left(\sum_{i=1}^{2} \|X_i'\|^2\right)^{1/2} \|Y - Y'\| \\
&\quad \le \sqrt{2} \max_{i=1,2} \|X_i - X_i'\| \|Y\| + \sqrt{2} \max_{i=1,2} \|X_i'\| \|Y - Y'\|.
\end{aligned}$$

For $\epsilon > 0$, we choose

$$0 < \delta < \min\left\{1, \frac{\epsilon}{2\sqrt{2}(\|Y\| + \max_{i=1,2}\|X_i\| + 1)}\right\}.$$

Then $\|(\langle X_1, \mathrm{Ad}(g)Y\rangle, \langle X_2, \mathrm{Ad}(g)Y\rangle) - (\langle X_1', \mathrm{Ad}(g)Y'\rangle, \langle X_2', \mathrm{Ad}(g)Y'\rangle)\|_2 < \epsilon$, whenever $|||(X_1, X_2, Y) - (X_1', X_2', Y')||| = \max_{i=1,2}\{\|X_i - X_i'\|, \|Y - Y'\|\} < \delta$. In other words, $d(W_Y(X_1, X_2), W_{Y'}(X_1', X_2')) < \epsilon$, whenever $|||(X_1, X_2, Y) - (X_1', X_2', Y')||| < \delta$.

(2) When $Y = Y'$, (6) becomes

$$d\big(W_Y(X_1, X_2), W_Y(X_1', X_2')\big) \le \sqrt{2} \max_{i=1,2} \|X_i - X_i'\| \|Y\|.$$

So $\mathscr{W}_Y$ is uniformly continuous. □

We remark that Proposition 3.1 is true for $W_Y(X_1, \dots, X_p)$ as well.

Without symplectic technique Raïs [19] showed that if X is a *regular* element of $\mathfrak{g}$, then the critical points of the function $F : O(Y) \to \mathbb{R}$ defined by $F(Z) = \langle X, Z\rangle$

are all nondegenerate, that is, F is nondegenerate, and the indices of F on the critical points are always even. So the level surface $F^{-1}(a)$ is connected (or empty) for $a \in \mathbb{R}$. Indeed, Bott and Samelson [9] (see [2, page 76]) had proved a stronger result: F is nondegenerate and an index of a critical point is equal to twice the number of hyperplanes crossed by a line joining X to the critical point. But this does not yield the convexity of $W_Y(X_1, X_2)$ yet, where $X_1, X_2, Y \in \mathfrak{g}$, since X is assumed to be regular. However, if one can show that for any given $X_1, X_2 \in \mathfrak{g}$, there exist sequences of regular elements $X_1^{(n)}, X_2^{(n)} \in \mathfrak{g}$ such that $X_1^{(n)} \to X_1$ and $X_2^{(n)} \to X_2$ as $n \to \infty$ and $X_1'(n) = X_1^{(n)} \cos\theta - X_2^{(n)} \sin\theta$ and $X_2'(n) = X_1^{(n)} \sin\theta + X_2^{(n)} \cos\theta$ are both regular for all $\theta \in [0, \pi/2]$, then the convexity of $W_Y(X_1, X_2)$ follows. The reason is that by Proposition 3.1 (2), $W_Y(X_1^{(n)}, X_2^{(n)}) \to W_Y(X_1, X_2)$ with respect to Hausdorff topology. The sets $W_Y(X_1^{(n)}, X_2^{(n)})$ are convex by Lemma 2.2 (3), Bott-Samelson-Raïs' result, and the Hausdorff-Westwick argument. Since the space of compact convex subsets of $\mathbb{R}^2$ is closed, $W_Y(X_1, X_2)$ is convex.

References

[1] M. F. Atiyah, 'Convexity and commuting Hamiltonians', *Bull. London Math. Soc.* **308** (1982), 1–15.

[2] M. F. Atiyah and R. Bott, 'The Yang-Mills equations over Riemann surfaces', *Philos. Trans. Roy. Soc. London Ser. A* **14** (1982), 523–615.

[3] Y. H. Au-Yeung and Y. T. Poon, 'A remark on the convexity and positive definiteness concerning Hermitian matrices', *Southeast Asian Bull. Math.* **3** (1979), 85–92.

[4] Y. H. Au-Yeung and N. K. Tsing, 'An extension of the Hausdorff-Toeplitz theorem on the numerical range', *Proc. Amer. Math. Soc.* **89** (1983), 215–218.

[5] ———, 'Some theorems on the numerical range', *Linear and Multilinear Algebra* **15** (1984), 3–11.

[6] M. Audin, *The topology of torus actions on symplectic manifolds*, Progress in Mathematics 93 (Birkäuser, Boston, 1991).

[7] C. A. Berger, *Normal dilations* (Ph.D. Thesis, Cornell University, 1963).

[8] F. F. Bonsall and J. Duncan, *Numerical ranges of operators on normed spaces and elements of normed algebras*, London Math. Soc. Lecture Notes 2 (Cambridge University Press, 1971).

[9] R. Bott and H. Samelson, 'Applications of the theory of Morse to symmetric spaces', *Amer. J. Math.* **80** (1958), 964–1029.

[10] V. Guillemain and S. Sternberg, 'Convexity properties of the moment mapping', *Invent. Math.* **67** (1982), 491–513.

[11] ———, 'Convexity properties of the moment mapping II', *Invent. Math.* **77** (1984), 533–546.

[12] K. E. Gustafson and D. K. M. Rao, *Numerical range: the field of values of linear operators and matrices* (Springer, New York, 1997).

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

[14] A. A. Kirillov, *Elements of the theory of representations* (Springer, Berlin, 1976).

[15] F. Kirwan, 'Convexity properties of the moment mapping III', *Invent. Math.* **77** (1984), 547–552.

[16] A. Knapp, *Lie groups beyond an introduction*, Progress in Mathematics 140 (Birkhäuser, Boston, 1996).

[17] C. K. Li, 'C-numerical ranges and c-numerical radii', *Linear and Multilinear Algebra* **37** (1994), 51–82.
[18] Y. T. Poon, 'Another proof of a result of Westwick', *Linear and Multilinear Algebra* **9** (1980), 35–37.
[19] M. Raïs, 'Remarques sur un theoreme de R. Westwick', unpublished manuscript.
[20] T. Y. Tam, 'Kostant's convexity theorem and classical compact groups', *Linear and Multilinear Algebra* **43** (1997), 87–113.
[21] O. Toeplitz, 'Das algebraische Analogon zu einem Satze von Fejér', *Math. Z.* **2** (1918), 187–197.
[22] R. Westwick, 'A theorem on numerical range', *Linear and Multilinear Algebra* **2** (1975), 311–315.

Department of Mathematics
218 Parker Hall
Auburn University
AL 36849-5310
USA
e-mail: tamtiny@auburn.edu