Lie groups, convexity and symplectic structure

E.P. van den Ban

Orientation seminar, January 18, 2016

1 Lie groups

Definition 1.1 A Lie group is a group $(G, \cdot, e)$ equipped with the structure of a smooth (i.e., C^∞) manifold, such that the maps

(a) $\mu : G \times G \to G, (x, y) \mapsto xy,$

(b) $\iota : G \to G, x \mapsto x^{-1}$

are smooth maps between the indicated manifolds.

If $a \in G$ we define left multiplication by a to be the map

$$l_a : G \to G,\ x \mapsto ax.$$

Since $l_a = \mu(a, \cdot)$, we see that $l_a : G \to G$ is a smooth map. Furthermore, it is readily verified that $l_e = I_G$ and that $l_{a^{-1}}$ is a two-sided inverse to l_a. In particular, l_a is a smooth bijection with smooth inverse, i.e., a diffeomorphism from G onto itself. This means that the smooth structure of G is determined by any given local coordinate system on a neighborhood U of a. Indeed, assume that (U, χ) is a chart at e. By this we mean that U is an open neighborhood of e in G, and χ is a diffeomorphism from U onto an open subset of $\mathbb{R}^n$, where $n = \dim G$. For a given $a \in G$, the set $l_a(U) = aU$ is an open neighborhood of a in G and $\chi \circ l_{a^{-1}} : aU \to \chi(U)$ is a diffeomorphism, hence a chart at a.

Likewise, if $a \in G$ then $r_a : G \to G, x \mapsto xa$ is a diffeomorphism from G onto itself.

Definition 1.2 A homomorphism of Lie groups is a map $\varphi : G \to H$ with G and H Lie groups and such that

(a) φ is a group homomorphism;

(b) φ is a smooth map of manifolds.

Example 1.3

(a) $(\mathbb{R}^n, +, 0)$ is a Lie group.

(b) $(\mathbb{R}^*, \cdot, 1)$ is Lie group (here $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$).

(c) The circle group $\mathbb{R}/Z$ is a Lie group, and so is any product of the form $(\mathbb{R}/\mathbb{Z})^r$ (the r-dimensional torus).

(d) Let $\mathrm{M}(n, \mathbb{R})$ denote the linear space of $n \times n$ matrices with real entries. Then the set $\mathrm{GL}(m, \mathbb{R})$ of invertible elements of $\mathrm{M}(n, \mathbb{R})$, equipped with matrix multiplication, is a group (called the general linear group). We note that

$$\mathrm{GL}(n, \mathbb{R}) = \{A \in \mathrm{M}(n, \mathbb{R}) \mid \det A \neq 0\}.$$

Since $\det : \mathrm{M}(n, \mathbb{R}) \to \mathbb{R}$ is smooth, hence in particular continuous, whereas $\mathbb{R} \setminus \{0\}$ is an open subset of $\mathbb{R}$, it follows that $\mathrm{GL}(n, \mathbb{R})$ is an open subset of $\mathrm{M}(n, \mathbb{R})$. On the other hand, $\mathrm{M}(n, \mathbb{R})$ is a n^2 dimensional linear space, hence a manifold of dimension n^2, and we see that $\mathrm{GL}(n, \mathbb{R})$ is an n^2-dimensional manifold as well. Furthermore, since the matrix multiplication map $m : \mathrm{M}(n, \mathbb{R}) \times \mathrm{M}(n, \mathbb{R}) \to \mathrm{M}(n, \mathbb{R})$ is bilinear, it is smooth, and it follows that its restriction μ to $\mathrm{GL}(n, \mathbb{R}) \times \mathrm{GL}(n, \mathbb{R})$ is a smooth map $\mathrm{GL}(n, \mathbb{R}) \times \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$. By application of Cramer's rule, one sees that the inversion map $\iota : \mathrm{GL}(n, \mathbb{R}) \to \mathrm{GL}(n, \mathbb{R})$ is smooth. Accordinlgy, $\mathrm{GL}(n, \mathbb{R})$ is a Lie group.

(e) Since the determinant map $\mathrm{M}(n, \mathbb{R}) \to \mathbb{R}$ is polynomial in the matrix entries, it follows that $\det : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}^*$ is a homomorphism of Lie groups.

A given Lie group G possesses a particular coordinate system at its identity element e, given by the exponentional map, which we shall characterize below.

Definition 1.4 A one-parameter subgroup of G is defined to be a Lie group homomorphism $\alpha : (\mathbb{R}, +, 0) \to G$.

The condition on α means that $\alpha : \mathbb{R} \to G$ is a smooth map such that $\alpha(0) = e$ and $\alpha(s + t) = \alpha(s)\alpha(t)$ for all $s, t \in \mathbb{R}$.

The following result can be proved by using the local existence and uniqueness theorem for ordinary differential equations in $\mathbb{R}^n$. For details we refer to a regular course on Lie groups, see [4], [5].

Proposition 1.5 *Let G be a Lie group. For every tangent vector $X \in T_eG$ there exists a unique one parameter subgroup $\alpha : \mathbb{R} \to G$ such that $\alpha'(0) = X$.*

We write $\alpha_X = \alpha_X^G$ for the unique one parameter subgroup of G specified in Proposition 1.5

Example 1.6 Since $\mathrm{GL}(n, \mathbb{R})$ is an open subset of $\mathrm{M}(n, \mathbb{R})$, the tangent space of $T_I\mathrm{GL}(n, \mathbb{R})$ may be canonically identified with $\mathrm{M}(n, \mathbb{R})$. In fact, if $X \in \mathrm{M}(n, \mathbb{R})$ then the corresponding tangent vector $\underline{X}$ of $\mathrm{GL}(n, \mathbb{R})$ is given by

$$\underline{X} = \left.\frac{d}{dt}\right|_{t=0} (I + tX).$$

From now on, we shall identify X with $\underline{X}$ and accordingly write $T_I\mathrm{GL}(n,\mathbb{R}) = \mathrm{M}(n,\mathbb{R})$. It is now a straightforward exercise to show that for $X \in \mathrm{M}(n,\mathbb{R})$ we have

$$\alpha_X(t) = e^{tX} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k X^k.$$

Definition 1.7 For a Lie group G, the exponential map $\exp : T_eG \to G$ is defined by

$$\exp(X) = \alpha_X(1).$$

Lemma 1.8 *The exponential map* $\exp : T_eG \to G$ *is smooth and maps* 0 *to* e. *In accordance with the identification* $T_0T_eG = T_eG$, *its tangent map at* 0 *is given by* $T_0 \exp = I : T_eG \to T_eG$. *In particular,* $\exp$ *is a local diffeomorphism at* 0.

The final assertion of the above lemma means that there exists an open neighborhood Ω of 0 in T_eG such that $\exp$ maps Ω diffeomorphically onto an open neighborhood U of e in G. Accordinly, the inverse $(\exp|_\Omega)^{-1} : U \to T_eG$ defines a chart of G at e.

Sketch of proof: By the theory of ordinary differential equations, the curve α_X depends smoothly on the parameter X. This implies that $T_eG \times \mathbb{R} \to G$, $(X,t) \mapsto \alpha_X(t)$ is smooth. Restricting to $t = 1$ we see that $\exp$ is smooth. To prepare for the assertion about the tangent map, we note that for $X \in T_eG$ and $s \in \mathbb{R}$ the map $t \mapsto \alpha_X(st)$ is a one-parameter subgroup. Furthermore, by the chain rule,

$$\left.\frac{d}{dt}\right|_{t=0} \alpha_X(st) = s\alpha_X'(0) = sX,$$

hence $\alpha_X(st) = \alpha_{sX}(t)$ by the uniqueness part of Proposition 1.5. It follows that

$$\alpha_X(t) = \alpha_{tX}(1) = \exp tX,$$

for all $X \in T_eG$ and $t \in \mathbb{R}$. Therefore,

$$T_0(\exp)(X) = \left.\frac{d}{dt}\right|_{t=0} \exp(tX) = \left.\frac{d}{dx}\right|_{t=0} \alpha_X(t) = X.$$

This implies $T_0 \exp = I$. The final assertion follows by application of the inverse function theorem. $\square$

The exponential map behaves well with respect to Lie group homomorphisms.

Lemma 1.9 *Let* $\mathfrak{g} : G \to H$ *be a homomorphism of Lie groups. Then the following diagram commutes:*

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & H \\ {\scriptstyle \exp_G} \uparrow & & \uparrow {\scriptstyle \exp_H} \\ T_eG & \xrightarrow{T_e\varphi} & T_eH. \end{array}$$

Proof. The proof is natural. Let $X \in T_eG$. Then $\varphi \circ \alpha_X^G$ is a one parameter subgroup of H, hence of the form α_Y^H for $Y \in T_eH$ given by

$$Y = \left.\frac{d}{dt}\right|_{t=0} \varphi \circ \alpha_X^G(tX) = T_e\varphi \left.\frac{d}{dt}\right|_{t=0} \alpha_X^G(tX) = T_e(\varphi)X.$$

Evaluation at $t = 1$ leads to $\varphi(\exp tX) = \exp Y = \exp(T_e\varphi X)$. □

With the notions introduced above we can now define the so-called adjoint action of G on T_eG.

Definition 1.10 Given an element $g \in G$ we define the map $\mathrm{Ad}(g) : T_eG \to T_eG$ by

$$\mathrm{Ad}(g)X = \left.\frac{d}{dt}\right|_{t=0} g \exp tXg^{-1}.$$

Example 1.11 Let $G = \mathrm{GL}(n, \mathbb{R})$; then $T_eG = \mathrm{M}(n, \mathbb{R})$. It is now readily seen that

$$\mathrm{Ad}(g)X = \left.\frac{d}{dt}\right|_{t=0} g \exp tXg^{-1} = \left.\frac{d}{dt}\right|_{t=0} \exp(tgXg^{-1}) = gXg^{-1}.$$

In general, $\mathrm{Ad}(g) : T_eG \to T_eG$ may be viewed as conjugation by $g \in G$ in the following sense.

Lemma 1.12 *Let $g \in G$ and $X \in T_eG$. Then*

$$g(\exp X)g^{-1} = \exp(\mathrm{Ad}(g)X).$$

Proof. Since $C_g := l_g \circ r_{g^{-1}} : G \to G,\ x \mapsto gxg^{-1}$ is a Lie group automorphism, it follows that $C_g \circ \alpha_X$ is a one-parameter subgroup of G. Hence,

$$C_g \circ \alpha_X = \alpha_Y, \tag{1}$$

where

$$Y = \left.\frac{d}{dt}\right|_{t=0} C_g \circ \alpha_X(t) = \left.\frac{d}{dt}\right|_{t=0} C_g(\exp tX) = \mathrm{Ad}(g)X.$$

The result now follows by evaluation of (1) at the identity at 1. □

Lemma 1.13 *For every $g \in G$, the map $\mathrm{Ad}(g) : T_eG \to T_eG$ is an invertible linear map. Furthermore,*

$$\mathrm{Ad} : G \to \mathrm{GL}(T_eG)$$

defines a homomorphism of Lie groups.

Proof. For $g \in G$, let $C_g : G \to G$ be defined as above. Then, for $X \in T_eG$,

$$\mathrm{Ad}(g)X = \left.\frac{d}{dt}\right|_{t=0} C_g \circ \exp(tX) = T_e(C_g)X.$$

Hence, $\mathrm{Ad}(g) = T_e(C_g)$ and we see that $\mathrm{Ad}(g) : T_eG \to T_e(G)$ is linear. From $C_{gh} = C_g \circ C_h$ it follows by taking tangent maps at e and application of the chain rule that

$$\mathrm{Ad}(gh) = \mathrm{Ad}(g) \circ \mathrm{Ad}(h).$$

From $C_g = I_G$ it follows that $\mathrm{Ad}(e) = I_{T_eG}$. Hence, $Ad : G \to \mathrm{GL}(T_eG)$ is a group homomorphism. Finally, let $X \in T_eG$. Since $G \times \mathbb{R} \to G$, $(g, t) \mapsto g \exp tXg^{-1}$ is smooth, it follows by differentiation at $t = 0$ that $g \mapsto \mathrm{Ad}(g)X$ is smooth. This implies that $\mathrm{Ad} : G \to \mathrm{GL}(T_eG)$ is smooth. The result follows. $\square$

The map Ad behaves naturally with respect to Lie group homomoprhisms.

Lemma 1.14 *Let $\varphi : G \to H$ be a homomorphism of Lie groups. Then for all $g \in G$ the following diagram commutes*

$$\begin{array}{ccc} T_eG & \xrightarrow{T_e\varphi} & T_eH \\ {\scriptstyle \mathrm{Ad}_G(g)} \uparrow & & \uparrow {\scriptstyle \mathrm{Ad}_H(\varphi(g))} \\ T_eG & \xrightarrow{T_e\varphi} & T_eH. \end{array}$$

Proof. Again the proof is natural. We leave it to the reader. $\square$

We can now define a bracket structure on T_eG which turns this linear space into a Lie algebra.

Definition 1.15 A Lie algebra (over the field $\mathbb{R}$) is a real linear space L, equipped with a bilinear map $L \times L \to L, (X, Y) \mapsto [X, Y]$, such that, for all $X, Y, Z \in L$,

(a) $[X, Y] = -[X, Y]$ (anti-symmetry);

(b) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ (Jacobi-identity).

We define $[\,\cdot\,,\,\cdot\,] : T_eG \times T_eG \to T_eG$ by

$$[X, Y] := \left.\frac{d}{dt}\right|_{t=0} \mathrm{Ad}(\exp tX)Y. \tag{2}$$

Example 1.16 Consider the standard example $G = \mathrm{GL}(n, \mathbb{R})$. Then $T_eG = \mathrm{M}(n, \mathbb{R})$. For $X, Y \in T_eG$, we have

$$[X, Y] = \left.\frac{d}{dt}\right|_{t=0} e^{tX}Ye^{-tX} = \left.\frac{d}{dt}\right|_{t=0} e^{tX}Y + \left.\frac{d}{dt}\right|_{t=0} Ye^{-tX} = XY - YX,$$

by straightforward application of the Leibniz rule. Thus, in this example, $[\,\cdot\,,\,\cdot\,]$ is given by the so-called commutator bracket. One readily verifies that $\mathrm{M}_n(n)$, equipped with the commutator bracket, is a Lie algebra.

Example 1.17 More generally, if A is an associative algebra (over $\mathbb{R}$ or more generally over a field k), then A equipped with the bracket $[X, Y] = XY - YX$ is a Lie algebra (over k).

Lemma 1.18 *Let G be a Lie group. The bracket defined by (2) is bilinear. Furthermore, it behaves well with respect to Lie group homomorphisms in the following sense. If $\varphi : G \to H$ is a Lie group homomorphism, then*

$$T_e\varphi([X, Y]_G) = [T_e\varphi(X), T_e\varphi(Y)]_H. \tag{3}$$

Proof. We first address the bilinearity. The linearity in the second component is obvious. Consider the map $\Phi : T_eG \to T_eG$ given by

$$\Phi(Z) := \mathrm{Ad}(\exp Z)Y.$$

Then Φ is differentiable and $\Phi(0) = Y$. It follows that

$$[X, Y] = \frac{d}{dt}\bigg|_{t=0} \Phi(tX) = T_0\Phi(X),$$

which is linear in X. This establishes the bilinearity. The rest of the proof is natural and left to the reader. $\square$

Corollary 1.19 *Let G be a Lie group. Then for all $g \in G$ and $X, Y \in T_eG$ we have*

$$\mathrm{Ad}(g)[X, Y] = [\mathrm{Ad}(g)X, \mathrm{Ad}(g)Y].$$

Proof. Use that $\mathrm{Ad}(g) : T_eG \to T_eG$ is the tangent map of $C_g : G \to G$ at the point e. $\square$

Theorem 1.20 *Let G be a Lie group. Then the bracket $[\,\cdot\,,\,\cdot\,]$ turns T_eG into a Lie algebra.*

Proof. It remains to establish the Jacobi-identity. Let $Y, Z \in T_eG$. Then for all $X \in T_eG$ we have

$$\mathrm{Ad}(\exp tX)[Y, Z] = [\mathrm{Ad}(\exp tX)Y, \mathrm{Ad}(\exp tX)Z].$$

Differentiating at $t = 0$ and using the bilinearity of the bracket, we find that

$$[X, [Y, Z]] = [[X, Y], Z] + [Y, [X, Z]].$$

The Jacobi identity now follows by invoking the anti-symmetry of the bracket. $\square$

From now on, we shall write $\mathfrak{g}$ for T_eG, equipped with the bracket (2). Then, $\mathfrak{g}$ is called the Lie algebra of the Lie group G. Furthermore, we shall write ad for the linear map $\mathfrak{g} \to \mathrm{End}(\mathfrak{g})$ given by $\mathrm{ad}(X)Y = [X, Y]$, for $X, Y \in \mathfrak{g}$.

Let $\varphi : G \to H$ be a Lie group homomorphism. Then by using (3) we infer that $\varphi_* := T_e\varphi : \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism in the obvious sense.

The following result, for which we refer to a basic course of Lie groups is truly remarkable.

Lemma 1.21 *Let G be a Lie group, and let H be a subgroup of G which is closed in the sense of topology. Then H is a submanifold of G. This manifold structure turns H into a Lie group such that the inclusion map $i : H \to G$ is a Lie group homomorphism.*

In the setting of the above theorem, the tangent map $i_* := T_e i : \mathfrak{h} \to \mathfrak{g}$ is an injective homomorphism of Lie algebras, realizing $\mathfrak{h}$ as the Lie subalgebra $i_*(\mathfrak{h})$ of $\mathfrak{g}$. We shall identify theses Lie algebras via the map i_*.

Lemma 1.22 *In the above setting,*

$$\mathfrak{h} = \{X \in \mathfrak{g} \mid \exp \mathbb{R}X \subset H\}.$$

Exercise 1.23 We consider the following subsets of $\mathrm{GL}(n, \mathbb{R})$,

$$\begin{aligned} \mathrm{SL}(n, \mathbb{R}) &:= \{x \in \mathrm{GL}(n, \mathbb{R}) \mid \det x = 1\}, \\ \mathrm{O}(n) &:= \{x \in \mathrm{GL}(n, \mathbb{R}) \mid x^{\mathrm{T}} = x^{-1}\}, \\ \mathrm{SO}(n) &:= \mathrm{SL}(n, \mathbb{R}) \cap \mathrm{O}(n). \end{aligned}$$

Show that all these are closed subgroups of $\mathrm{GL}(n, \mathbb{R})$. Show that their Lie algebras, realized as Lie subalgebras of $\mathrm{M}_n(\mathbb{R})$, equipped with the commutator bracket are given by

$$\begin{aligned} \mathfrak{sl}(n, \mathbb{R}) &= \{X \in \mathrm{M}(n, \mathbb{R}) \mid \mathrm{tr} X = 0\}, \\ \mathfrak{o}(n) &= \{X \in \mathrm{M}(n, \mathbb{R}) \mid X^{\mathrm{T}} = -X\}, \\ \mathfrak{so}(n) &= \{X \in \mathrm{M}(n, \mathbb{R}) \mid X^{\mathrm{T}} = -X\}. \end{aligned}$$

For the first identity you may use that $\det e^X = e^{\mathrm{tr} X}$ for all $X \in \mathrm{M}(n, \mathbb{R})$.

Show that $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ are compact. Using the 2×2 block decomposition of element in $\mathrm{O}(n)$ it is fairly easy to show that $\mathrm{SO}(n)$ is connected.

Exercise 1.24 Observe the

$$\mathrm{GL}(n, \mathbb{C}) = \{x \in \mathrm{M}(n, \mathbb{C}) \mid \det_{\mathbb{C}} x \neq 0\}$$

may be viewed as a real Lie group. Show that its Lie algebra is $\mathrm{M}_n(n, \mathbb{C})$, equipped with the commutator bracket. Show that

$$\begin{aligned} \mathrm{SL}(n, \mathbb{C}) &:= \{x \in \mathrm{GL}(n, \mathbb{C}) \mid \det x = 1\}, \\ \mathrm{U}(n) &:= \{x \in \mathrm{GL}(n, \mathbb{C}) \mid x^* = x^{-1}\}, \\ \mathrm{SU}(n) &:= \mathrm{SL}(n, \mathbb{C}) \cap \mathrm{U}(n). \end{aligned}$$

are closed subgroups of $\mathrm{GL}(n, \mathbb{C})$, hence real Lie groups of their own right. Here $*$ indicates that the Hermitean conjugate has been taken. Show that their Lie algebras may be realized as the following Lie subalgebras of $\mathrm{M}(n, \mathbb{C})$,

$$\begin{aligned} \mathfrak{sl}(n, \mathbb{C}) &= \{X \in \mathrm{M}(n, \mathbb{C}) \mid \mathrm{tr}_{\mathbb{C}} X = 0\}, \\ \mathfrak{u}(n) &= \{X \in \mathrm{M}(n, \mathbb{C}) \mid X^* = -X\}, \\ \mathfrak{su}(n) &= \{X \in \mathrm{M}(n, \mathbb{R}) \mid X^* = -X \text{ and } \mathrm{tr}_{\mathbb{C}} X = 0\}. \end{aligned}$$

2 Kostant's convexity theorem

In this section, we assume that G is a (connected) compact Lie group.

Lemma 2.1 *The Lie algebra* $\mathfrak{g}$ *carries a positive definite inner product for which* $\mathrm{Ad}(G) \subset \mathrm{GL}(\mathfrak{g})$ *consists of orthogonal transformations.*

Sketch of proof. There exists a positive left invariant regular Borel measure dx on G. This so called left Haar measure is unique up to a positive scalar factor. The measure allows one to integrate continuous functions $f \in C(G)$ over G. The associated integral is denoted by

$$I(f) := \int_G f(x)\, dx.$$

Positivity of the measure means that $f \geq 0 \Rightarrow I(f) \geq 0$ and $f \geq 0, I(f) = 0 \Rightarrow f = 0$. Left invariance of the measure means that

$$\int_G f(gx)\, dx = \int_G f(x)\, dx,$$

for all $f \in C(G)$ and $g \in G$. We may fix the measure uniquely by requiring it to be normalized in the sense that $\int_G dx = 1$.

Let $(\,\cdot\,,\,\cdot\,)$ be a choice of positive definite inner product on G. Then by averaging over G one may define the following symmetric bilinear form $\langle\,\cdot\,,\,\cdot\,\rangle$ on $\mathfrak{g}$,

$$\langle v\,,\, w\rangle := \int_G (\mathrm{Ad}(x)^{-1}v, \mathrm{Ad}(x)^{-1}w)\, dx. \tag{4}$$

We leave it as an exercise to the reader to verify that $\langle\,\cdot\,,\,\cdot\,\rangle$ is a positive definite inner product, and that $\mathrm{Ad}(G) \subset \mathrm{O}(\mathfrak{g})$ with respect to this inner product. By connectedness of G we even have $\mathrm{Ad}(G) \subset \mathrm{SO}(\mathfrak{g})$. $\square$

Exercise 2.2 Complete the final part of the above proof.

We now fix a maximal torus $\mathfrak{t} \subset \mathfrak{g}$. By this we mean a subspace $\mathfrak{t} \subset \mathfrak{g}$ which is maximal subject to the condition that $[X, Y] = 0$ for all $X, Y \in \mathfrak{t}$. Furthermore, we denote the orthogonal projection $\mathfrak{g} \to \mathfrak{t}$ by $P_\mathfrak{t}$.

Given an element $X \in \mathfrak{g}$ we consider the adjoint orbit

$$\mathrm{Ad}(G)X := \{\mathrm{Ad}(g)X \mid g \in G\}.$$

Lemma 2.3 *The orbit* $\mathrm{Ad}(G)X$ *is a smooth compact submanifold of* $\mathfrak{g}$.

In the next section on actions, this result will be explained in more detail. We mention the following result without proof.

Lemma 2.4 *The intersection* $\mathrm{Ad}(G)X \cap \mathfrak{t}$ *is a non-empty finite subset of* $\mathfrak{t}$.

The following remarkable result, known as Kostant's convexity theorem, first appeared in [8].

Theorem 2.5 *Let* $X \in \mathfrak{g}$ *and let* $\mathrm{Ad}(G)X$ *be the adjoint orbit through* X. *Then the image of* $\mathrm{Ad}(G)X$ *under* P *is given by*

$$P(\mathrm{Ad}(G)X) = \mathrm{conv}\,(\mathrm{Ad}(G)X \cap \mathfrak{t}).$$

In particular, $P(\mathrm{Ad}(G))$ *is a convex polyhedral subset of* $\mathfrak{t}$.

The intersection $\mathrm{Ad}(G)X \cap \mathfrak{t}$ has an interesting description, which can be proven by using the detailed structure theory of connected compact Lie groups (in terms of root spaces). It can be shown that $T = \exp(\mathfrak{t})$ is a compact subgroup of G, hence a Lie group of its own right, which is is a compact torus, i.e., isomorphic to a finite product of circles, $(\mathbb{R}/\mathbb{Z})^r$, where $r = \dim \mathfrak{t}$.

Lemma 2.6 *The intersection* $\mathrm{Ad}(G)X \cap \mathfrak{t}$ *equals the set* $[\mathrm{Ad}(G)X]^T$ *of fixed points for the natural action of* T *on* $\mathrm{Ad}(G)X$. *In particular, this set of fixed points is finite.*

Proof. It suffices to show that $\mathfrak{t}$ equals the set $\mathfrak{g}^T$ of fixed points for $\mathrm{Ad}(T)$ in $\mathfrak{g}$. We will first prove that $\mathfrak{t} \subset \mathfrak{g}^T$. Let $X \in \mathfrak{t}$. Then it follows that $[Y, X] = 0$ for all $Y \in \mathfrak{t}$. Consider the function $c : \mathbb{R} \to \mathfrak{g}$ given by $c(t) := \mathrm{Ad}(\exp tY)X$. Then it follows that

$$c'(t) = \left.\frac{d}{ds}\right|_{s=0} c(t+s) = \mathrm{Ad}(tX)[Y, X] = 0,$$

hence $c(t) = c(0) = X$ for all t. We infer that $X \in \mathfrak{g}^T$.

For the converse inclusion, let $X \in \mathfrak{g}^T$. Then $\mathrm{Ad}(\exp tY)X = 0$ for all $Y \in \mathfrak{t}$ and $t \in \mathbb{R}$. Differentiating at $t = 0$ we see that $[\mathfrak{t}, X] = 0$. Since $\mathfrak{t}$ is maximal abelian, we conclude that $X \in \mathfrak{t}$. $\square$

Thus, Kostant's result may also be reformulated as

$$P(\mathrm{Ad}(G)X) = \mathrm{conv}\,([\mathrm{Ad}(G)X]^T).$$

This formulation has the virtue that it corresponds to a similar result in $\mathfrak{g}^* = \mathrm{Hom}_{\mathbb{R}}(\mathfrak{g}, \mathbb{R})$, where it can be given a natural interpretation in the context of symplectic geometry. This will be discussed in the remainder of these notes.

Finally, we mention that the set $[\mathrm{Ad}(G)X]^T$ can be described in terms of the so called Weyl group W of $\mathfrak{t}$ in G. This group may be defined by

$$W := N_G(\mathfrak{t})/Z_G(\mathfrak{t}),$$

where $N_G(\mathfrak{t})$ stands for $\{g \in G \mid \mathrm{Ad}(g)\mathfrak{t} \subset \mathfrak{t}\}$, the normalizer of $\mathfrak{t}$ in G and $Z_G(\mathfrak{t})$ for $\{g \in G \mid \mathrm{Ad}(g)|_{\mathfrak{t}} = I_{\mathfrak{t}}\}$, the centralizer of $\mathfrak{t}$ in G. The Weyl group is finite and naturally embeds into $\mathrm{GL}(\mathfrak{t})$ as the reflection group associated with the root system of $\mathfrak{t}$ in $\mathfrak{g}$. Let $X \in \mathfrak{g}$, then $\mathrm{Ad}(G)X \cap \mathfrak{t} \neq \emptyset$, hence $\mathrm{Ad}(G)X = \mathrm{Ad}(G)Y$ for a suitable element $Y \in \mathfrak{t}$. The set $[\mathrm{Ad}(G)X]^T$ can now be described as the image of the map $W \to \mathrm{Ad}(G)Y$, induced by $N_G(\mathfrak{t}) \to \mathrm{Ad}(G)Y$, $g \mapsto \mathrm{Ad}(g)Y$. Since obviously this image is contained in $\mathfrak{t}$, we obtain the following version of Kostant's convexity theorem, for $Y \in \mathfrak{t}$,

$$P(\mathrm{Ad}(G)Y) = \mathrm{conv}\,(WY).$$

Actually, this is the theorem Kostant originally formulated in [8].

Exercise 2.7 We consider the group $G = \mathrm{U}(n)$. From exercise 1.24 we recall that $\mathfrak{u}(n)$ consists of the anti-Hermitian matrices in $\mathrm{M}(n, \mathbb{C})$.

(a) Show that $\mathrm{U}(n)$ is compact.

(b) By using diagonalization, show that $\mathrm{U}(n)$ is connected.

(c) Given $x \in \mathbb{R}^n$ we write H_x for the diagonal matrix with diagonal entries $ix_1, \dots, ix_n$. Here i denotes the imaginary unit. Show that $\mathfrak{t} := \{H_x \mid x \in \mathbb{R}^n\}$ is a maximal torus in $\mathfrak{u}(n)$.

We view $\mathfrak{u}(n)$, equipped with the commutator bracket, as a real Lie algebra, hence in particular as a real linear space. For $X, Y \in \mathfrak{u}(n)$ we define $\langle X \,,\, Y\rangle = \mathrm{tr}(Y^* X)$.

(d) Show that $\langle \cdot \,,\, \cdot \rangle$ is a symmetric real bilinear form, which is positive definite. Show that for all $g \in \mathrm{U}(n)$ and $X, Y \in \mathfrak{u}(n)$ we have $\langle \mathrm{Ad}(g)X \,,\, \mathrm{Ad}(g)Y\rangle = \langle X \,,\, Y\rangle$, i.e., $\mathrm{Ad}(g)$ is orthogonal with respect to the given inner product.

(e) Describe the orthogonal projection $P : \mathfrak{u}(n) \to \mathfrak{t}$ with respect to this inner product.

(f) Show that there exists a natural action of the permutation group S_n on $\mathfrak{t}$ such that for all $H \in \mathfrak{t}$,

$$[\mathrm{Ad}(\mathrm{U}(n))H] \cap \mathfrak{t} = S_n \cdot H.$$

(g) Let A be a Hermitian matrix in $\mathrm{M}(n, \mathbb{C})$ with eigenvalues $\lambda = (\lambda_1, \dots, \lambda_n)$, viewed as a vector in $\mathbb{R}^n$. Let $a := (a_1, \dots, a_n) \in \mathbb{R}^n$ be the diagonal part of the matrix A. Show that $a \in \mathrm{conv}\,(S_n \cdot \lambda)$. This result is due to I. Schur. Show the converse result, due to A. Horn, that that every point a in the convex hull of $S_n \cdot \lambda$ is the diagonal part of a Hermitian matrix with eigenvalues λ.

(h) Show that $\mathfrak{su}(n)$ is a codimension one real subspace of $\mathfrak{u}(n)$ which is invariant under the adjoint action of $\mathrm{U}(n)$. Determine the orthocomplement of $\mathfrak{su}(n)$ on $\mathfrak{u}(n)$. Show that this orthocomplement is invariant under the adjoint action of $\mathrm{U}(n)$ as well.

(i) Show that $\mathfrak{t}_0 := \mathfrak{t} \cap \mathfrak{su}(n)$ and that its orthocomplement in $\mathfrak{t}$ is invariant under the action of S_n.

(j) Take $n = 2$ and show that the orbits for the adjoint action of $\mathrm{U}(2)$ on $\mathfrak{su}(2)$ are two dimensional spheres and the point 0. In this case, verify Kostant's convexity theorem.

(k) We now consider the case $n = 3$. By using the cosine rule for the inner product, determine the angle between the elements $a := H_{(1,-1,0)}$ and $b := H_{(0,1,-1)}$. Using a, b as a basis for $\mathfrak{t}_0$, express the orbit $S_3 \cdot a$ in terms of linear combinations of a and b.

(l) Define an orthogonal isomorphism φ from $\mathfrak{t}_0$ onto $\mathbb{R}^2$ (equipped with the standard inner product) such that $\varphi(a) = (1, 0)$, and draw the image $S_3 \cdot a$. Show that under φ the action of S_3 corresponds to the action of a finite subgroup of $\mathrm{GL}(2, \mathbb{R})$, generated by reflections.

(m) Make a sketch which represents the image under φ of the generic S_3 orbit in $\mathfrak{t}$.

3 Actions and orbits

Kostant's convexity theorem was later put in a symplectic context. The natural setting turns out to be the co-adjoint action of G on the dual space $\mathfrak{g}^* = \mathrm{Hom}(\mathfrak{g}, \mathbb{R})$, rather than the adjoint action of the compact group G.

We recall that a (left) action of a group G on a set M is a map $\alpha : G \times M \to M$, $(g, m) \mapsto gm$ such that the following conditions are fulfilled, for all $g_1, g_2 \in G$ and $m \in M$,

(a) $em = m$,

(b) $g_1(g_2 m) = (g_1 g_2)m$.

If we write $\alpha(g) = \alpha(g, \cdot\,)$, then the above conditions mean that $g \mapsto \alpha(g)$ is a group homomorphism to the group of bijections of M. If G is a Lie group and M a smooth manifold, then the action α is said to be smooth if the map $\alpha : G \times M \to M$ is smooth.

Assume that $\alpha : G \times M \to M$ is a smooth action and a a point of M. Then the stabilizer

$$G_a := \{g \in G \mid ga = a\}$$

is readily verified to be a closed subgroup of G, hence a Lie group, by Lemma 1.21. It is clear that the map $\alpha_a : G \to M, g \mapsto ga$ factors through an injective map $\bar{\alpha}_a : G/G_a \to M$, with image equal to the orbit Ga.

For the following lemma, we refer the reader to [4].

Lemma 3.1 *If H is a closed subgroup of a Lie group G, then G/H carries a unique structure of manifold such that the natural projection $\pi : G \to G/H$ is smooth and submersive.*

The condition of submersiveness means that for every $g \in G$ the tangent map $T_g\pi : T_gG \to T_{\pi(g)}(G/H)$ is surjective. By the local structure of submersions it follows that $T_e\pi : \mathfrak{g} \to T_{\pi(e)}(G/H)$ has kernel equal to $T_e\pi^{[e]} = T_eH = \mathfrak{h}$, hence induces a linear isomorphism

$$\overline{T_e\pi} : \mathfrak{g}/\mathfrak{h} \xrightarrow{\simeq} T_{[e]}(G/H),$$

where we have written $[e] = \pi(e)$. We will always use this linear isomorphism to identify $T_{[e]}(G/H)$ with $\mathfrak{g}/\mathfrak{h}$.

We return to the setting of a smooth action $\alpha : G \times M \to M$ of a Lie group G on a manifold M. The action of G on M induces a linear map $X \mapsto X_M$ from $\mathfrak{g}$ to the space $\mathfrak{X}(M)$ of vector fields on M given by

$$X_M(m) = \frac{d}{dt}\bigg|_{t=0} \exp tX \cdot m.$$

It can be shown that the map $X \mapsto X_M$ is an anti-homomorphism of Lie algebras, i.e. $[X, Y]_M = -[X_M, Y_M]$, for $X, Y \in \mathfrak{g}$.

Let $a \in M$ and consider the map $\alpha_a : G \to M,\ g \mapsto ga$.

Lemma 3.2 *The tangent map $T_e\alpha_a : \mathfrak{g} \to T_aM$ is given by $X \mapsto X_M(a)$. Its kernel equals $\mathfrak{g}_a$.*

Proof. Let $X \in \mathfrak{g}$. Then for $t \in \mathbb{R}$ we have

$$\alpha_a(\exp tX) = \exp tX \cdot a.$$

Differentiating with respect to t at $t = 0$ and applying the chain rule, we find

$$T_e\alpha_a(X) = X_M(a).$$

This establishes the first assertion. We turn to the second.

By Lemma 1.22 the Lie algebra $\mathfrak{g}_a$ equals the space of $X \in \mathfrak{g}$ for which $\exp \mathbb{R}X \subset G_a$. Let $X \in \mathfrak{g}$. Then $X \in \mathfrak{g}_a$ implies $\exp tX \cdot a = a$ for all $t \in \mathbb{R}$. By differentiation at $t = 0$, the latter condition implies $X_M(a) = 0$. Conversely, assume $X_M(a) = 0$. Then the integral curve of X_M with initial point a is stationary with image $\{a\}$. On the other hand, it is readily verified that $t \mapsto \exp tXa$ describes the integral curve. By uniqueness of integral curves it follows that $\exp tX \cdot a = a$ for all t, hence $X \in \mathfrak{g}_a$. $\square$

Using Lemma 3.1, Lemma 3.2 and basic manifold theory one may prove the following lemma.

Lemma 3.3 *Let G be a compact Lie group, M a smooth manifold, and $\alpha : G \times M \to M$ a smooth action. Then for every $a \in M$ the orbit Ga is a compact smooth submanifold of M and the induced map $\bar{\alpha}_a : G/G_a \to Ga$ is a diffeomorphism. Finally, the map $T_e(\alpha_a) :\ \mathfrak{g} \to T_aM,\ X \mapsto X_M(a)$ has kernel $\mathfrak{g}_a$ and image $T_a(Ga)$.*

Proof. By application of the submersion theorem one sees that the smooth map $\alpha_a : G \to M$ factors trough a smooth map $\bar{\alpha}_a : G/G_a \to M$. Hence, $\alpha_a = \bar{\alpha}_a \circ \pi$ so that $T_e\alpha = T_{[e]}\bar{\alpha}_a \circ T_e\pi$ by the chain rule. Since $T_e\alpha_a$ and $T_e\pi$ both have kernel $\mathfrak{g}_a$, it follows that $T_{[e]}\bar{\alpha}_a$ is injective. By homogeneity it follows that the tangent map of $\bar{\alpha}_a$ is injective everywhere. Since $\bar{\alpha}_a$ is injective, and G compact, it follows by application of standard manifold theory that $\bar{\alpha}_a$ is an embedding of G/G_a into M. Therefore, its image Ga is a submanifold of M and $\bar{\alpha}_a : G/G_a \to M$ is a diffeomorphism onto Ga. It follows that $T_a(Ga)$ equals the image of $T_{[e]}\bar{\alpha}_a$, which is also the image of $T_e\alpha_a$. As we have seen, the latter map is given by $\mathfrak{g} \to T_aM,\ X \mapsto X_M(a)$. $\square$

Let now G be a compact connected Lie group. Then it follows from the above lemma that the adjoint orbits $\mathrm{Ad}(G)X$, for $X \in \mathfrak{g}$, are compact smooth submanifolds of $\mathfrak{g}$. From the perspective of symplectic geometry it turns out to be more natural to look at the co-adjoint action $\mathrm{Ad}^\vee$ of G on the dual space $\mathfrak{g}^* = \mathrm{Hom}(\mathfrak{g}, \mathbb{R})$. This action is given by

$$g \cdot \xi = \mathrm{Ad}^\vee(g)\xi = \xi \circ \mathrm{Ad}(g)^{-1} \qquad (g \in G, \xi \in \mathfrak{g}^*)$$

Here the inverse is needed for this to be a left action. It is readily seen that the given action is smooth. Its orbits, the co-adjoint orbits in $\mathfrak{g}^*$, are connected compact submanifolds.

Exercise 3.4 Let $\langle \cdot\, ,\, \cdot \rangle$ be an $\mathrm{Ad}(g)$-invariant inner product on $\mathfrak{g}$ defined as in (4).

(a) Show that the map $\varphi : \mathfrak{g} \to \mathfrak{g}^*$ given by $\varphi(X) = \langle X\, ,\, \cdot \rangle$ is a linear isomorphism which maps the adjoint orbit $\mathrm{Ad}(G)X$, for $X \in \mathfrak{g}$, diffeomorphically onto the co-adjoint orbit $\mathrm{Ad}^\vee(G)\varphi(X)$ in $\mathfrak{g}^*$.

(b) The inclusion map $\iota : \mathfrak{t} \to \mathfrak{g}$ naturally induces the surjective linear map $\iota^* : \mathfrak{g}^* \to \mathfrak{t}^*$ given by $\xi \mapsto \xi \circ \iota = \xi|_{\mathfrak{t}}$. Show that the following diagram commutes:

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\varphi} & \mathfrak{g}^* \\ {\scriptstyle P}\downarrow & & \downarrow \iota^* \\ \mathfrak{t} & \xrightarrow{\varphi_{\mathfrak{t}}} & \mathfrak{t}^* \end{array}$$

Here $\varphi_{\mathfrak{t}}$ denotes the linear isomorphism $\mathfrak{t} \to \mathfrak{t}^*$ induced by the restriction of $\langle \cdot\, ,\, \cdot \rangle$ to $\mathfrak{t}$.

(c) Show that Kostant's convexity theorem may be reformulated as

$$\iota^*(\mathrm{Ad}^\vee(G)\xi) = \mathrm{conv}\,([\mathrm{Ad}^\vee(G)\xi]^T)$$

4 Symplectic structure of co-adjoint orbits

We recall that a differential two form on M is a smooth assignment $M \ni x \to \omega_x$, where $\omega_x : T_xM \times T_xM \to \mathbb{R}$ is an anti-symmetric bilinear form. In local coordinates $x^1, \ldots, x^m$ on M (defined on an open set U), such a form has a local expression

$$\omega = \sum_{i<j} \omega_{ij} dx^i \wedge dx^j,$$

with ω_{ij} uniquely defined smooth functions on U. The form ω is said to be closed if its exterior derivative $d\omega$ vanishes. In local coordinates with U contractible, this is equivalent to the existence of a one form λ on U such that $d\lambda = \omega$. In turn, this is equivalent to the existence of functions $\lambda_1, \ldots, \lambda_m$ on U such that

$$\omega_{ij} = \partial_j \lambda_i - \partial_i \lambda_j \qquad (i < j).$$

The 2-form ω is said to be non-degenerate if for all $x \in M$ the bilinear form ω_x on T_xM is non-degenerate. The latter means that for every $v \in T_xM$ we have $\omega_x(v, \cdot\,) = 0 \Rightarrow v = 0$. In turn, this is equivalent to the assertion that $v \mapsto \omega_x(v, \cdot\,)$ defines a linear isomorphism $T_xM \to T_x^*M$.

Definition 4.1 A symplectic form on a manifold M is a two form ω which is closed and non-degenerate. A symplectic manifold is a manifold equipped with such a symplectic form.

We will now show that the co-adjoint orbits in $\mathfrak{g}^*$ can all be equipped with a natural symplectic structure. Let $\mathcal{O}$ be a co-adjoint orbit and let $\xi \in \mathcal{O}$ and $X \in \mathfrak{g}$. Then from the above discussion with $M = \mathfrak{g}^*$ we see that

$$X_M(\xi) = \left.\frac{d}{dt}\right|_{t=0} \exp tX \cdot \xi = \left.\frac{d}{dt}\right|_{t=0} \xi \circ \mathrm{Ad}(\exp(-tX)) = -\xi \circ \mathrm{ad}(X).$$

This implies that

$$\mathfrak{g}_\xi = \{X \in \mathfrak{g} \mid \xi \circ \mathrm{ad}(X) = 0\} = \{X \in \mathfrak{g} \mid \xi = 0 \text{ on } [X, \mathfrak{g}]\}.$$

From this it follows that the alternating bilinear form $\omega' = \omega'_\xi : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ given by

$$\omega'(X, Y) = \xi([X, Y])$$

factors through a bilinear form

$$\omega'' : \mathfrak{g}/\mathfrak{g}_\xi \times \mathfrak{g}/\mathfrak{g}_\xi \to \mathbb{R}.$$

It is readily checked that ω'' is non-degenerate.

Exercise 4.2 Show that ω'' is non-degenerate.

We recall that the map $X \mapsto X_M(\xi)$ induces a linear isomorphism from $\mathfrak{g}/\mathfrak{g}_\xi$ onto $T_\xi(cO)$ and define the two form ω_ξ on $T_\xi(\mathcal{O})$ by the requirement that

$$\omega_\xi(X_M(\xi), Y_M(\xi)) = \omega'_\xi(X, Y).$$

From this definition it naturally follows that $\omega : \xi \mapsto \omega_\xi$ defines a smooth 2-form on $\mathcal{O}$ which is everywhere non-degenerate. We claim that ω is G-invariant in the sense that

$$\omega_\xi = \omega_{g\xi} \circ [T_\xi(l_g) \times T_\xi(l_g)],$$

for all $\xi \in \mathcal{O}$ and $g \in G$.

Exercise 4.3 Prove that the two form ω is G-invariant.

Proposition 4.4 *The two form ω on $\mathcal{O}$ is closed.*

Sketch of proof: We will show that $d\omega = 0$. Fix $\xi \in \mathcal{O}$ and consider the map $\varphi : G \to G\xi$ given by $g \mapsto g\xi$. Since this map is submersive, it suffices to show that $\varphi^* d\omega = 0$. Since $\varphi^* d\omega = d\varphi^*\omega$ it suffices to show that $\varphi^*\omega$ is closed. We will prove this by showing that the two form $\varphi^*\omega$ on G is exact, i.e., there exists a one form λ on G such that $d\lambda = \varphi^*\omega$.

From the G-invariance of ω and the fact that $\varphi : G \to G\xi$ intertwines the left actions of G, i.e., $l_g \circ \varphi = \varphi \circ l_g$ for all $g \in G$, it follows that $\varphi^*\omega$ is left G-invariant. We will in fact define λ

to be the left G-invariant one form on G given by $\lambda_e = \xi \in \mathfrak{g}^*$. By G-invariance it now suffices to show that

$$\varphi^*(\omega)_e = (d\lambda)_e. \tag{5}$$

For $X \in \mathfrak{g}$ we define the vector field v_X on G by

$$v_X(g) = \left.\frac{d}{dt}\right|_{t=0} g \exp tX.$$

This vector field is left G-invariant and has the value X at $g = e$. Let $X, Y \in \mathfrak{g}$; then

$$(d\lambda)_e(X, Y) = d\lambda(v_X, v_Y)_e = [v_X\lambda(v_Y) - v_Y(\lambda(v_X)) - \lambda([v_X, v_Y])]_e,$$

by a well known formula for the exterior derivative. We now use that $\lambda(v_X)$ and $\lambda(v_Y)$ are left-invariant, hence constant functions. Furthermore, it can be shown that $[v_X, v_Y] = v_{[X,Y]}$. It follows that

$$(d\lambda)_e(X, Y) = \lambda_e([X, Y]) = \xi([X, Y]) = (\varphi^*\omega)_e(X, Y).$$

Hence, (5). $\square$

5 Hamiltonian actions

We assume that (M, ω) is a symplectic manifold. Given a function $f : M \to \mathbb{R}$ we define the associated vector field X_f on M by

$$\omega(X_f, \cdot\,) = df.$$

This definition is justified by the non-degeneracy of ω. It can be seen that H_f us smooth. The above equation is also written as

$$\iota_{X_f}\omega = df.$$

Lemma 5.1 $\mathcal{L}_{X_f}\omega = 0$.

Equivalently, this means that ω is invariant under the flow of X_f.

Proof. From the definition of X_f it follows that $d\iota_{X_f}\omega = ddf = 0$. On the other hand, $d\omega = 0$, and it follows by application of Cartan's formula that

$$\mathcal{L}_{X_f}\omega = d\iota_{X_f}\omega + \iota_{X_f}d\omega = 0.$$

$\square$

Conversely, if X is a vector field such that $\mathcal{L}_X\omega = 0$ then it follows by Cartan's formula that $d\iota_X\omega = 0$ so that locally at any given point one may find a function f such that $X_f = X$. If this is possible for a global choice of f, then X is said to be a Hamiltonian vector field. The function f is said to be a Hamiltonian for X.

We now assume that G has a smooth action on the symplectic manifold M preserving the symplectic form ω; such an action will be called symplectomorphic (or canonical).

If all vector fields X_M, for $X \in \mathfrak{g}$, are Hamiltonian, then there exists a linear map $J : \mathfrak{g} \to C^\infty(M)$ such that

$$X_{J(Z)} = Z_M$$

for all $Z \in \mathfrak{g}$ (use a basis of $\mathfrak{g}$ to see this). Such a map J gives rise to a map $\mathbf{J} : M \to \mathfrak{g}^*$ defined by $\langle \mathbf{J}(m)\, ,\, Z\rangle = J(Z)(m)$. The assignment $J \mapsto \mathbf{J}$ defines a linear isomorphism $\mathrm{Hom}_{\mathbb{R}}(\mathfrak{g}, C^\infty(M)) \simeq C^\infty(M, \mathfrak{g}^*)$ via which we shall identify. Accordingly, we shall not distinguish between J and $\mathbf{J}$ anymore. A map of this type will be called a momentum map for the symplectomorphic G-action.

Definition 5.2 Let G act smoothly by symplectomorphisms on the symplectic manifold G. A momentum map for this action is defined to be a smooth map $J : M \to \mathfrak{g}^*$ such that for every $Z \in \mathfrak{g}$ the function $J_Z : m \to J(m)(Z)$ is a Hamiltonian for the vector field Z_M, i.e., $X_{J_Z} = Z_M$.

Definition 5.3 A Hamiltonian action of G on the symplectic manifold M is an action by symplectomorphisms for which there exists a momentum map $J : M \to \mathfrak{g}^*$ which is equivariant in the sense that

$$J(gm) = \mathrm{Ad}^\vee(g)J(m) \qquad (g \in G, m \in M).$$

If G is compact, then from a given momentum map one can always obtain an equivariant momentum map by the process of averaging.

We now have the following beautiful convexity theorem for momentum maps of torus actions, discovered by Atiyah [1] and Guillemin-Sternberg [6], independently.

We start with an observation about the momentum map associated to the G-action on a co-adjoint orbit $\mathcal{O}$.

Theorem 5.4 *Let G be a Lie group and $\mathcal{O} \subset \mathfrak{g}^*$ a co-adjoint orbit, equipped with its natural symplectic form. Then the co-adjoint action of G on $\mathcal{O}$ is Hamiltonian, with associated equivariant momentum map equal to the inclusion map $j : \mathcal{O} \to \mathfrak{g}^*$.*

Proof. The action of G is by symplectomorphisms, and j is equivariant. Let $Z \in \mathfrak{g}$ and let $j_Z : \mathcal{O} \to \mathbb{R}$ be defined by $j_Z(\xi) = \langle \xi\, ,\, Z\rangle$. Then it suffices to show that $X_{j_Z} = Z_{\mathcal{O}}$. We will do this at a given point $\xi \in \mathcal{O}$. We have

$$dj_Z(\xi) = \langle\, \cdot\, ,\, Z\rangle \quad \text{on} \quad T_\xi\mathcal{O},$$

where the latter space is canonically identified with a linear subspace of $\mathfrak{g}^*$. The map $\mathfrak{g} \to T_\xi\mathcal{O}$, $Y \mapsto Y_{\mathcal{O}}(\xi)$ is surjective. Let $Y \in \mathfrak{g}$; then it follows that

$$dj_Z(\xi)Y_{\mathcal{O}}(\xi) = -\langle \xi \circ \mathrm{ad}(Y)\, ,\, Z\rangle = -\xi([Y, Z]) = \omega_\xi(Z_{\mathcal{O}}(\xi), Y_{\mathcal{O}}(\xi)).$$

We thus see that

$$dj_Z(\xi) = \omega_\xi(Z_{\mathcal{O}}(\xi), \cdot).$$

Hence, $X_{j_Z} = Z_{\mathcal{O}}$, as required. $\square$

We now assume that in the above setting $H \subset G$ is a closed subgroup, and write ι for the associated inclusion map $\mathfrak{h} \to \mathfrak{g}$. Then ι interwines the adjoint action of H on $\mathfrak{h}$ with the action $\mathrm{Ad}^G|_H$ on $\mathfrak{g}$. It follows that $\iota^* : \mathfrak{g}^* \to \mathfrak{h}^*$ intertwines the co-adjoint actions of H on $\mathfrak{h}^*$ and $\mathfrak{g}^*$.

Lemma 5.5 *The restricted action of H on $\mathcal{O}$ is Hamiltonian with equivariant momentum map $J : \mathcal{O} \to \mathfrak{h}^*$ given by $J = \iota^* \circ j$.*

Proof. Equivariance is obvious. If $Z \in \mathfrak{h}$, then $J_Z(\xi) = \langle J(\xi)\,,\, Z\rangle = \langle j(\xi)\,,\, \iota(Z)\rangle = j_{\iota(Z)}(\xi)$. It follows that

$$X_{J_Z} = X_{j_{\iota(Z)}} = \iota(Z)_M = Z_M.$$

$\square$

The following result is due to M. Atiyah [1] and independently, V. Guillemin and S. Sternberg [6].

Theorem 5.6 *Let T be a torus, i.e., a Lie group isomorphic to $(\mathbb{R}/Z)^r$ for some $r \in \mathbb{N}$. Assume that the symplectic manifold M is equipped with a symplectomorphic action of T, with associated momentum map $J : M \to \mathfrak{t}^*$. If the action of T has a finite set M^T of fixed points, then*

$$J(M) = \mathrm{conv}\,(M^T).$$

In particular, $J(M)$ is a convex polyhedron.

We will now show that this result generalizes Kostant's convexity theorem and in fact, gives a natural explanation of it from the viewpoint of symplectic geometry.

Let G be a connected compact Lie group. Let $\mathfrak{t}$ be a maximal (infinitesimal) torus in $\mathfrak{g}$ and $T = \exp(\mathfrak{t})$ the associated torus in G. Let $\mathcal{O} \subset \mathfrak{g}^*$ be a coadjoint orbit. Then it follows from the previous lemmas that the action of T on $\mathcal{O}$ is Hamiltonian with momentum map $J : \mathcal{O} \to \mathfrak{t}^*$ equal to $\iota^*|_{\mathcal{O}} : \mathcal{O} \to \mathfrak{t}^*$. Thus, $J(\mathcal{O}) = \iota^*(\mathcal{O})$. It follows from Lemma and Exercise 3.4 that the action of T on $\mathcal{O}$ has finitely many fixed points. The set of these is given by

$$\mathcal{O}^T = \varphi_{\mathfrak{t}}([\mathrm{Ad}(G)X]^T) = \varphi_{\mathfrak{t}}(\mathrm{Ad}(X) \cap \mathfrak{t}).$$

By Theorem 5.6 it follows that

$$\iota^*(\mathcal{O}) = J(\mathcal{O}) = \mathrm{conv}\,(J(\mathcal{O}^T)) = \mathrm{conv}\,\iota^*(\mathcal{O}^T).$$

This is the assertion of Kostant's theorem.

6 Further convexity results

For further convexity results in Lie theory, we refer the reader to [7], [3], [2].

Appendix: tangent map and chain rule

If M is a smooth n-dimensional manifold, and $a \in M$, we denote by T_aM the tangent space of M at a. This is real linear space of dimension $\dim M$ defined in such a way that the following rules are valid.

If $c : [0,1] \to M$ is a differentiable curve, then its velocity vector $c'(t)$ at $t \in [0,1]$ is a vector in the tangent space $T_{c(t)}M$. Furthermore, if differentiable curves $c, d : [0,1] \to M$ have the same initial point $a = c(0) = d(0)$, then c and d have the same velocity vector in any coordinate system at a if and only if $c'(0) = d'(0)$ (in T_aM).

If $f : M \to N$ is a smooth map of manifolds and $a \in M$, then there exists a uniquely defined linear map $T_af : T_aM \to T_{f(a)}N$, which generalizes that notion of total derivative. This map is called the tangent map of f at a.

If $c : [0,1] \to M$ is a differentiable curve with initial point $a \in M$ then $f \circ c$ is a differentiable curve in N with initial point $f(a)$, and we have the following version of the chain rule:

$$(f \circ c)'(0) = T_af(c'(0)). \tag{6}$$

By using this rule one may derive the following chain rule.

Lemma 6.1 (Chain rule) *Let $f : M \to \mathbb{N}$ and $g : N \to R$ be smooth maps, $a \in M$, then*

$$T_a(g \circ f) = T_{f(a)}(g) \circ T_af.$$

Remark 6.2 With notation as in the above lemma, let $b = f(a)$ and $c = g(b)$. Then we may view f and g as smooth maps of pointed manifolds, $f : (M,a) \to (N,b)$ and $g : (N,b) \to (R,c)$. To a pointed space (M,a) one may assign the tangent space T_aM and to a morphism $f : (M,a) \to (N,b)$ the tangent map $T_af : T_aM \to T_bN$. The chain rule now amounts to the assertion that taking tangent spaces and maps defines a functor from the category of pointed manifolds to the category of finite dimensional linear spaces.

References

[1] M.F. Atiyah. Convexity and commuting Hamiltonians. *Bull. London Math. Soc.*, 14(1):1–15, 1982.

[2] D. Balibanu and E.P. van den Ban. Convexity theorem for semisimple symmetric spaces. *Forum Math.*, pages 1 – 41, 2016.

[3] E.P. van den Ban. A convexity theorem for semisimple symmetric spaces. *Pacific J. Math.*, 124(1):21–55, 1986.

[4] E.P. van den Ban. Lie groups, Lecture Notes. pages 1 – 152, 2010.

[5] T. Bröcker and T. tom Dieck. *Representations of compact Lie groups*, volume 98 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. Translated from the German manuscript, Corrected reprint of the 1985 translation.

[6] V. Guillemin and S. Sternberg. Convexity properties of the moment mapping. *Invent. Math.*, 67(3):491–513, 1982.

[7] G.J. Heckman. Projections of orbits and asymptotic behaviour of multiplicities for compact lie groups. *Thesis Rijksuniversiteit Leiden*, 1980.

[8] B. Kostant. On convexity, the Weyl group and the Iwasawa decomposition. *Ann. Sci. École Norm. Sup. (4)*, 6:413–455 (1974), 1973.