HOFER'S METRIC IN COMPACT LIE GROUPS

GABRIEL LAROTONDA AND MARTÍN MIGLIOLI

ABSTRACT. In this article we study the Hofer geometry of a compact Lie group K which acts by Hamiltonian diffeomorphisms on a symplectic manifold M. Generalized Hofer norms on the Lie algebra of K are introduced and analyzed with tools from group invariant convex geometry, functional and matrix analysis. Several global results on the existence of geodesics and their characterization in finite dimensional Lie groups K endowed with bi-invariant Finsler metrics are proved. We relate the conditions for being a geodesic in the group K and in the group of Hamiltonian diffeomorphisms. These results are applied to obtain necessary and sufficient conditions on the moment polytope of the momentum map, for the commutativity of the Hamiltonians of geodesics. Particular cases are studied, where a generalized non-crossing of eigenvalues property of the Hamiltonians hold.

(G. Larotonda) DEPARTAMENTO DE MATEMÁTICA, FCEYN-UBA, AND INSTITUTO ARGENTINO DE MATEMÁTICA, CONICET. BUENOS AIRES, ARGENTINA.

(M. Miglioli) INSTITUTO ARGENTINO DE MATEMÁTICA, CONICET. BUENOS AIRES, ARGENTINA.

E-mail addresses: glaroton@dm.uba.ar, martin.miglioli@gmail.com.

2010 *Mathematics Subject Classification.* 58B20, 53D20 (primary), and 53C22, 58D05 (secondary).

Key words and phrases. Hofer's metric, compact Lie group, Hamiltonian action, moment map, moment polytope, Finsler length structure, geodesic, coadjoint orbit, commuting Hamiltonians, symplectic energy.

The authors were supported by grants PIP 2010-0757 (CONICET), 2010-2478 (ANPCyT).

CONTENTS

1. Introduction 2
2. The generalized Hofer norm and its convex geometry 4
3. Hofer's metric on compact Lie groups 13
4. Convex geometry of Hofer's norm on Cartan algebras 19
5. The geometry of groups with bi-invariant Finsler metrics 25
6. Spheres with abelian faces 41
Acknowledgments 52
References 52

1. INTRODUCTION

The Finsler length structure on the group $\mathrm{Ham}(M,\omega)$ of Hamiltonian diffeomorphisms of a symplectic manifold M with symplectic form ω, was introduced by Hofer in the paper [Hof90]. The Lie algebra of this group is the set of Hamiltonian vector fields, which can be identified (by means of the symplectic gradient) with the set of Hamiltonian functions in M -modulo constant functions-. The norm of a vector field is then the quotient $L^\infty(M)$-norm of the generating Hamiltonian function (modulo constant functions). A natural problem of current interest in the literature is the study of geodesics in this Finsler manifold. There has been a significant amount of progress, and fairly deep work on the properties of this metric, mostly from the point of view of symplectic topology, see [BiaPol94, LaMcD95a, LaMcD95b, Pol01, PolShe15] and also the textbook [McDSa17] and the references therein.

In this article we consider (almost) effective Hamiltonian actions of compact semi-simple Lie groups K on a symplectic manifold (M,ω). This action defines an inclusion (modulo the discrete kernel of the action) $K \hookrightarrow \mathrm{Ham}(M,\omega)$, which allows us to introduce a Hofer metric on K by means of

$$\|x\|_{\mu(M)} = \max_{y\in\mu(M)} \langle y,x\rangle - \min_{y\in\mu(M)} \langle y,x\rangle.$$

Here $x \in \mathfrak{k} = T_1K$ and μ is the momentum map of the Hamiltonian action. For this norm, the intersection of its unit ball with a maximal abelian subalgebra $\mathfrak{h} \subseteq \mathfrak{k}$ is the polar dual of a certain polytope P, which is derived from the moment polytope of the action μ. The rectifiable distance in K is defined as the infima of the lenghts of paths joining given endpoints, and we define a *geodesic* in K as a distance minimising path.

We first show that one-parameter groups in K are always geodesics for this distance, provided the speed is in the domain of injectivity of the exponential map of K. These paths are mapped to paths in $\mathrm{Ham}(M,\omega)$ with autonomous Hamiltonian. Since the norm is not strictly convex, it is natural to expect other geodesics in K, and this is the case for all these norms; we show that these geodesics correspond to quasi-autonomous Hamiltonianans. The extreme points of the polytope $P \subseteq \mathfrak{k}$ encode the privileged directions that the speeds of geodesics can take. Relevant geometrical properties of the geodesics can be expressed in terms of these extreme points. Of particular relevance are the polytopes with only

regular extreme points in $\mathfrak{k}$, which are fully characterized both in terms of the Lie algebra (by polar duality) and in terms of the geometry of geodesics in K. We obtain similar results for the pull-back of the one-sided Hofer norm, which is usually only positively homogeneous:

$$\|x\|^{+}_{\mu(M)} = \max_{y\in\mu(M)} \langle y, x\rangle.$$

In fact, the general results on geodesics in this article are stated for any Finsler length structure given by (left or right) translation of an Ad-invariant Finsler norm in $\mathfrak{k}$, therefore we include non-symmetric distances in our discussion. It is worthile mentioning here that the notion of majorization of real vectors $\overrightarrow{v} \prec \overrightarrow{w}$ (identified with the eigenvalues of the operators $\operatorname{ad} v, \operatorname{ad} w$, where $v, w \in \mathfrak{k}$) plays a significant role in the proofs concerning minimality of geodesics, and it is related to the condition $v \in \operatorname{conv}(O_w)$, where the later set is the convex closure of the coadjoint orbit of w in $\mathfrak{k}$.

There are several relevant applications related to this setting of actions of compact Lie groups: for instance, as shown in [Entov01], the geometry of the canonical Hamiltonian action $\mathrm{SU}(n) \to \mathrm{Ham}(Gr_{r,n}, \omega)$ can be used as a tool to study the eigenvalue inequalities in the quantum version of Horn's problem [Bel08]. Here $Gr_{r,n}$ is the Grassmannian of r-dimensional planes in $\mathbb{C}^n$, and ω is the canonical Kirilov-Kostant-Souriau symplectic form. We plan to extend some of the results in this article to the case of infinite dimensional groups using ideas connected to the results in [BoLeZh99] and [Lar19]. Some of the techniques developed in this article might also be relevant to the study of Finsler length structures derived from mechanics with non-smooth energy.

The article is organized as follows: in Section 2 we define generalized Hofer norms. We study the faces and norming functionals of the unit balls of these norms, based on two different theories. We first analyse the structure of these balls with functional analytic techniques via an embedding in certain function spaces, an approach that will be useful in the study of the stability under geodesy at the end of Section 5.4. We also study these norms using results from convex geometry.

In Section 3 we recall basic results on Hamiltonian actions and Hofer's metric on groups of Hamiltonian diffeomorphisms, and we pull-back these metrics to compact groups using the homomorphism $K \to \mathrm{Ham}(M, \omega)$. These are the motivations and main examples for the norms and length structures on groups studied in this article. Nevertheless, this part of symplectic geometry is not necessary for the understanding of several results in the article which are solely based on convexity and Finsler length structures.

In Section 4 a characterization of the intersection of the unit balls of the Hofer norms with maximal abelian algebras is given based on group invariant convex analysis and symplectic convexity theorems.

In Section 5 we first recall several results obtained in [Lar19] for groups endowed with Finsler length structures obtained from Ad-invariant norms which are valid for the groups studied in this article. Then we prove several global results on geodesics in the case of finite dimensional groups endowed with continuous Finsler metrics. We show that geodesics in groups K with Hofer's metric are quasi-autonomous, which provides a link between the conditions for length minimization in K and the corresponding conditions in $\mathrm{Ham}(M, \omega)$.

Finally, in Section 6 we study actions of groups with commuting Hamiltonians. We start with the important special case of actions on regular coadjoint orbits and related groups: the Hamiltonians of length minimizing curves have the interesting feature of "non-crossing of eigenvalues". We characterize the Ad-invariant norms such that in groups with Finsler structures defined from these norms, all geodesics have commuting speeds. Based on this result we characterize the compact groups of Hamiltonian diffeomorphisms such that length minimizing curves have commuting Hamiltonians. We then show how conditions on Kirwan's polytope can characterize this property. The paper ends with a study of how these properties behave when we consider the direct product of Hamiltonaian actions. It is proved then that it suffices to have geodesics with commuting Hamiltonians for one of the actions, to obtain the same property for the geometry induced in K by the direct product of actions.

2. The generalized Hofer norm and its convex geometry

In this section we define the generalized Hofer norms and we study them with two approaches. In the first we embed the normed space in a quotient of a space of continuous functions and use functional analytic techniques. In the second we use the polar duality from convex geometry.

Definition 2.1. A subset $E \subseteq V$ of a vector space V is called *full* if it affinely generates the space. A set $B \subseteq V$ is a *convex body* if B is a compact convex set with non empty interior. If additionally, B is centrally symmetric ($v \in B \Rightarrow -v \in B$), then it is called a *symmetric convex body*. Equivalently, a symmetric convex body is a convex balanced absorbing set in V.

Definition 2.2. Let $(V, \langle \cdot, \cdot \rangle)$ be a finite dimensional inner product space and let $E \subseteq V$ be a compact full subset. Consider the norm $\| \cdot \|_E$ on V given by the embedding

$$\iota : V \hookrightarrow C(E)/\mathbb{R}\mathbf{1}, \quad x \mapsto [\varphi_x] := \varphi_x + \mathbb{R}\mathbf{1},$$

where $\varphi_x(y) = \langle x, y \rangle$ for $y \in E$. Then

$$\|x\|_E = \max_{y \in E} \varphi_x(y) - \min_{y \in E} \varphi_x(y) = 2\|[\varphi_x]\|_\infty, \tag{2.1}$$

where $\|[\varphi_x]\|_\infty = \inf\{\|\varphi_x - \lambda\mathbf{1}\|_\infty : \lambda \in \mathbb{R}\}$ is the quotient norm. We call $\| \cdot \|_E$ a *generalized Hofer norm*.

This is a norm since E is full. If there is a group acting isometrically on V and leaving the set E invariant then the action is also isometric for the norm $\| \cdot \|_E$.

Remark 2.3. Another (semi)-norm which is invariant for isometric actions on V that we will consider is the *generalized second Hofer norm* given by the supremum norm

$$\|x\|'_E = \max\{\max_{y \in E} \varphi_x(y), -\min_{y \in E} \varphi_x(y)\} = \|\varphi_x\|_\infty. \tag{2.2}$$

It is also of interest to consider a third invariant norm which is only a Finsler norm (i.e. only positively homogeneous), given by

$$\|x\|^+_E = \max_{y \in E} \varphi_x(y).$$

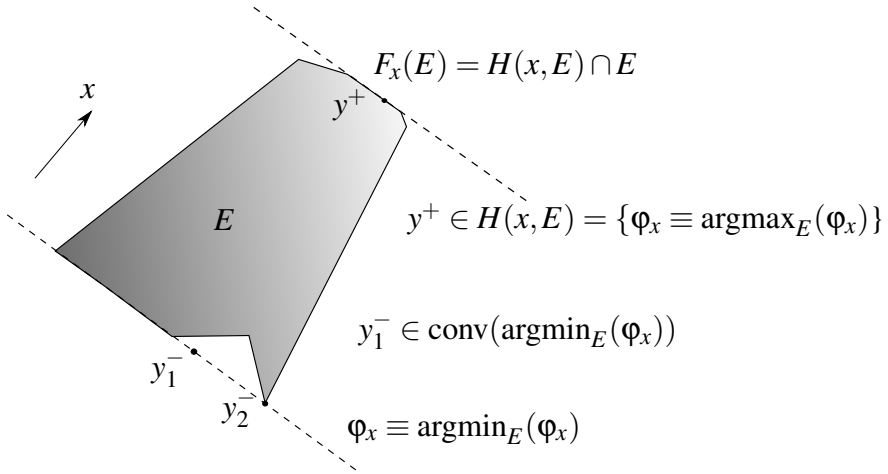

FIGURE 1. Norming functionals, faces and supporting hyperplanes

We will call this the *one-sided Hofer norm*. This defines a Finsler norm for a bounded set $E \subseteq V$ such that its convex hull contains 0 in its interior, or equivalently, such that the cone generated by E is all of V. It also possible to consider

$$\|x\|_E^- = -\min_{y \in E} \varphi_x(y).$$

but note that this one can be obtained from the previous by replacing E with $-E$. Clearly $\|x\|_E' = \max\{\|x\|_E^+, \|x\|_E^-\}$ is the second Hofer norm and $\|x\|_E = \|x\|_E^+ + \|x\|_E^-$ is the first Hofer norm.

2.1. **Maximal faces and norming functionals.** We begin studying the faces of the sphere and the norming functionals of Hofer's norm by relating this norm on the space V to the norm on the much larger space $C(E)/\mathbb{R}\mathbf{1}$ (where E is a compact full set) by means of the embedding $\iota : V \hookrightarrow C(E)/\mathbb{R}\mathbf{1}$ given by $x \mapsto \varphi_x + \mathbb{R}\mathbf{1}$, where $\varphi_x = \langle x, \cdot \rangle$.

Definition 2.4. Let V be a normed space and denote $\|\varphi\| = \sup\{\varphi(v) : \|v\| = 1\}$ for $\varphi \in V^*$. The dual space V^* with this norm is a Banach space. We say that $\varphi \in V^*$ is a *norming functional* of $v \in V$ if $\varphi(v) = \|v\|$ and $\|\varphi\| = 1$. A functional φ is *extremal* if φ is an extreme point of the unit ball B_{V^*}. If $\|\cdot\|$ is only positively homogeneous (to remark it we say that it is a *Finsler norm*) the same definitions apply, and the norm given to V^* is only positively homogeneous. In any case we refer to it as the *dual norm*.

§ Note that the difference bewtween the ball of a norm and that of a Finsler norm is that the last one might not be balanced (i.e. symmetric). In both cases, it is an absorbing, open convex set containing $0 \in V$.

Definition 2.5. A *face* F of the unit ball B_V of a normed space V is the intersection of the unit ball B_V with the hyperplane determined by a unit norm functional $\varphi \in V^*, \|\varphi\| = 1$, i.e.

$$F_\varphi = B_V \cap \{v \in V : \varphi(v) = 1\}.$$

We say that the face is *maximal* if φ is extremal. Every face is contained in a maximal face: if φ is a unit norm functional then $\|\varphi\| = 1$ and since B_{V^*} is compact and convex

there exists by the Krein-Milman theorem extremal functionals $\{\varphi_i\}_{i=1,\dots,n} \subseteq B_{V^*}$ such that φ is a convex combination of the φ_i:

$$\varphi = \sum_i \lambda_i \varphi_i, \quad \lambda_i \geq 0, \quad \sum_i \lambda_i = 1.$$

It is then easy to check that if $\varphi(v) = \|v\|$ then $\varphi_i(v) = \|v\|$ for all i. Therefore if $v \in F_\varphi$, $v \in F_{\varphi_i}$ for all i and in fact F_φ is the intersection of all the maximal faces that contain it.

The *cone generated* by a face F_φ is $\mathbb{R}_+ F_\varphi$. Note that this cone consists exactly of those $v \in V$ such that $\varphi(v) = \|v\|$.

The following elementary characterization will be useful:

Lemma 2.6. *In a vector space V, $\|v_1 + \cdots + v_n\| = \|v_1\| + \cdots + \|v_n\|$ holds if and only if $v_1, \dots, v_n$ belong to the cone generated by a face.*

Proof. If the v_i are in the cone of φ, then

$$\|v_1\| + \cdots + \|v_n\| \geq \|v_1 + \cdots + v_n\| \geq \varphi(v_1 + \cdots + v_n) = \|v_1\| + \cdots + \|v_n\|.$$

On the other hand, if $\|v_1 + \cdots + v_n\| = \|v_1\| + \cdots + \|v_n\|$ holds, by means of Hahn-Banach's theorem pick a unit norm functional that norms the sum of the v_i, that is $\varphi(\sum_i v_i) = \|\sum_i v_i\|$. Then

$$\|\sum_i v_i\| = \varphi(\sum_i v_i) = \sum_i \varphi(v_i) \leq \sum_i \|v_i\| = \|\sum_i v_i\|$$

since $\varphi(v_i) \leq \|v_i\|$ for all i; this is only possible if equality holds for each i. Therefore φ is a norming functional for all the v_i. □

2.1.1. *Norming functionals as Borel measures.* Let X be a compact Hausdorff topological space and let $C(X)$ be the continuous real valued functions on X. Endow $C(X)/\mathbb{R}\mathbf{1}$ with (twice) the quotient L^∞ norm (the factor 2 is there to be consistent with (2.1)). By Riesz-Markov's theorem, its dual space can be identified with the regular finite Borel signed measures in X such that $\mu(X) = 0$ (that is because the identification $\mu \mapsto \varphi_\mu$ is given by integration $\varphi_\mu(f) = \int_X f d\mu$, and we require that $\varphi_\mu(\mathbf{1}) = 0$). The norm of φ_μ is given by the total variation of μ, therefore unit norm functionals are characterized by having total variation equal to *two*. From here it is immediate the following characterization:

Remark 2.7. For $f \in C(X)/\mathbb{R}\mathbf{1}$ different from zero, its norming functionals are given by $\varphi = \mu^+ - \mu^-$, where μ^+ and μ^- are probability measures in X, supported in $\operatorname{argmax}(f)$ and $\operatorname{argmin}(f)$ respectively. Since the extreme points of the probability measures are the Dirac measures, the maximal faces are given by norming functionals $\varphi = \delta^+ - \delta^-$, with delta measures δ^+, δ^- supported in $x^+, x^- \in X$ respectively.

Proposition 2.8. *A set $\{f_i\}_{i \in I} \subseteq C(X)/\mathbb{R}\mathbf{1}$ is a subset of a cone generated by a face if and only if $\cap_{i \in I} \operatorname{argmin}(f_i) \neq \emptyset$ and $\cap_{i \in I} \operatorname{argmax}(f_i) \neq \emptyset$.*

Proof. If both intersections are non empty, pick x^-, x^+ respectively in each of them and consider φ in the dual given by $\varphi(f) = f(x^+) - f(x^-)$. Then φ is a unit norm functional and

$$\sum_i \|f_i\| \geq \|\sum_i f_i\| \geq \varphi(\sum_i f_i) = \sum_i f_i(x^+) - f_i(x^-) = \sum_i \|f_i\|,$$

therefore by Lemma 2.6 the f_i are in the cone generated by the face given by φ. Reciprocally, if there are, say, $f = f_k$ and $g = f_l$ such that the maximal argument of f does not intersect the maximal argument of g, then $\max(f+g) < \max f + \max g$, therefore

$$\|f+g\| = \max(f+g) - \min(f+g) < \max f + \max g - \min f - \min g = \|f\| + \|g\|$$

and the conclusion follows by Lemma 2.6. □

Corollary 2.9. *For $f, g \in C(X)/\mathbb{R}\mathbf{1}$ we have $\|f+g\| = \|f\| + \|g\|$ if and only if* $\operatorname{argmin}(f) \cap \operatorname{argmin}(g) \neq \emptyset$ *and* $\operatorname{argmax}(f) \cap \operatorname{argmax}(g) \neq \emptyset$.

Putting together the previous characterizations (and recalling that the norm we are considering is twice the quotient norm), it is clear that

Corollary 2.10. *The maximal faces of the ball of $C(X)/\mathbb{R}\mathbf{1}$ are given by the sets*

$$F_{x^-,x^+} = \{[f] : \|[f]\|_\infty = 2, x^- \in \operatorname{argmin}(f), x^+ \in \operatorname{argmax}(f)\}$$

for a choice of points $x^-, x^+ \in X$.

Since a face of the ball of V is contained in a face of $C(E)/\mathbb{R}\mathbf{1}$ by means of the map ι of Definition 2.2, we have the following result (see Figure 2):

Corollary 2.11. *Let V be a vector space with the norm defined by the map ι and the compact full set $E \subseteq V$. A set $S \subseteq V$ is a subset of a cone generated by a face if and only if*

$$\bigcap_{x \in S} \operatorname{argmin}_E(\varphi_x) \neq \emptyset \quad \textit{and} \quad \bigcap_{x \in S} \operatorname{argmax}_E(\varphi_x) \neq \emptyset.$$

Definition 2.12. Given a compact $E \subseteq V$ and $x^-, x^+ \in E$ we define the cone

$$C_{x^-,x^+}(E) := \{x \in V : x^- \in \operatorname{argmin}_E(\varphi_x) \text{ and } x^+ \in \operatorname{argmax}_E(\varphi_x)\}.$$

Proposition 2.13. *Each cone generated by a maximal face is equal to $C_{x^-,x^+}(E)$ for some $x^-, x^+ \in E$.*

Proof. Let $\mathbb{R}_+ F_{max}$ be the cone generated by a maximal face. By Corollary 2.11 it is contained in $C_{x^-,x^+}(E)$ for some $x^-, x^+ \in E$. Since $C_{x^-,x^+}(E)$ satisfies the condition of Corollary 2.11 it is contained in the cone $\mathbb{R}_+ F$ generated by a face F. By maximality of F_{max} we get $F_{max} = F$ and the conclusion follows. □

The cones of Definition 2.12 have good properties with respect to sum of sets.

Proposition 2.14. *Let $E_1, \dots, E_n$ be compact sets in V and let $x_i^-, x_i^+ \in E_i$ for $i = 1, \dots, n$. If we define $E = E_1 + \cdots + E_n$, $x^- = x_1^- + \cdots + x_n^-$ and $x^+ = x_1^+ + \cdots + x_n^+$ then*

$$C_{x^-,x^+}(E) = \bigcap_{i=1,\dots,n} C_{x_i^-,x_i^+}(E_i).$$

Proof. It is easy to verify that that for $x \in V$ the functional φ_x has a maximum at $x_i^-, x_i^+ \in E_i$ in E_i for $i = 1, \dots, n$ if and only if it has a maximum at $x_1^- + \cdots + x_n^-$ in $E_1 + \cdots + E_n$. The same holds for the minimizers and the proof follows. □

Remark 2.15. Similar results can be obtained for the one-sided Hofer norm $\|\cdot\|_E^+$ (Remark 2.3), for a compact set E such that its connvex hull contains 0 in its interior. A set $S \subseteq V$ is a subset of a cone generated by a face if and only if

$$\bigcap_{x\in S} \operatorname{argmax}_E(\varphi_x) \neq \emptyset.$$

Given a compact $E \subseteq V$ and $x^+ \in E$ we define the cone

$$C_{x^+}(E) := \{x \in V : x^+ \in \operatorname{argmax}_E(\varphi_x)\}.$$

Each cone generated by a maximal face is equal to $C_{x^+}(E)$ for some $x^+ \in E$. These cones have also good properties with respect to sum of sets. Let $E_1, \dots, E_n$ be compact sets in V and let $x_i^+ \in E_i$ for $i = 1, \dots, n$. If we define $E = E_1 + \cdots + E_n$ and $x^+ = x_1^+ + \cdots + x_n^+$ then

$$C_{x^+}(E) = \bigcap_{i=1,\dots,n} C_{x_i^+}(E_i).$$

We now characterize norming functionals for the Hofer norm, see Figure 1 and Figure 2. The convex hull of a set X is denoted by $\operatorname{conv}(X)$.

Theorem 2.16. *The norming functionals of $x \in (V, \|\cdot\|_E)$ are $\varphi_{y^+-y^-}$, with*

$$y^+ \in \operatorname{conv}(\operatorname{argmax}_E \varphi_x) \quad \textit{and} \quad y^- \in \operatorname{conv}(\operatorname{argmin}_E \varphi_x).$$

Proof. Suppose first that $y^+ \in \operatorname{conv}(\operatorname{argmax}_E \varphi_x)$ and $y^- \in \operatorname{conv}(\operatorname{argmin}_E \varphi_x)$, then we have $\varphi_x(y^+) = \max_E(\varphi_x)$ and likewise with y^-. It is immediate from the definitions that $\|\varphi_{y^+-y^-}\| \leq 1$, and on the other hand

$$\varphi_{y^+-y^-}(x) = \langle y^+ - y^-, x\rangle = \max_E(\varphi_x) - \min_E(\varphi_x) = \|x\|_E,$$

thus $\varphi_{y^+-y^-}$ has unit norm and it is norming for x. Suppose now φ is a norming functional of x in $V \simeq \iota(V) \subseteq C(E)/\mathbb{R}\mathbf{1}$. We can extend it by the Hahn-Banach theorem to all $C(E)/\mathbb{R}\mathbf{1}$, so that it is given by integration with $\mu^+ - \mu^-$ for probability measures μ^+, μ^- supported in $\operatorname{argmax}(\varphi_x)$ and $\operatorname{argmin}(\varphi_x)$ respectively. Hence

$$\begin{aligned}\varphi(z) &= \int_E \langle w, z\rangle \, d(\mu_1 - \mu_2)(w) \\ &= \langle \int_E w \, d\mu_1(w) - \int_E w \, d\mu_2(w), z\rangle \\ &= \langle \operatorname{cent}(\mu_1) - \operatorname{cent}(\mu_2), z\rangle,\end{aligned}$$

where $\operatorname{cent}(\mu)$ denotes the center of mass of the probability measure μ. The result follows if we take $y^+ = \operatorname{cent}(\mu_1)$ and $y^- = \operatorname{cent}(\mu_2)$. □

In Theorem 2.26 we will give another proof of the previous result, based on polar duality.

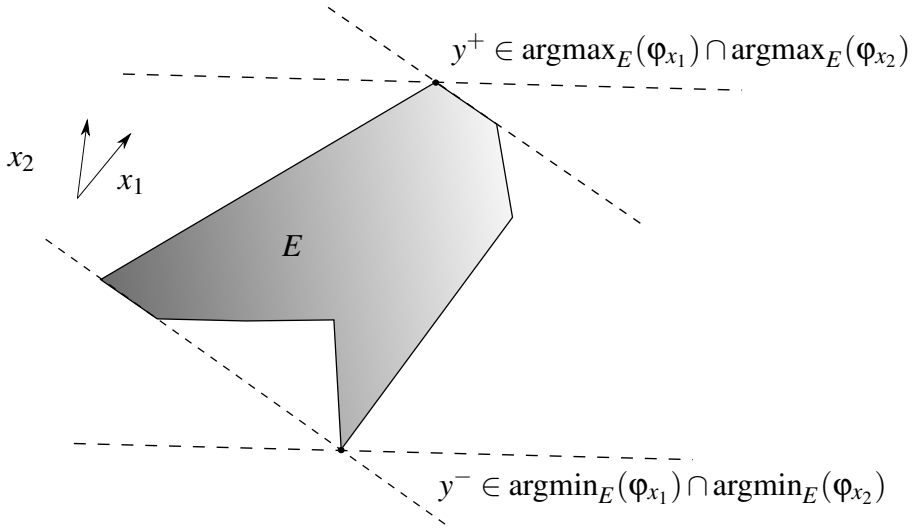

FIGURE 2. $\varphi = \varphi_{y^+ - y^-}$ is a norming functional of x_1 and x_2

2.2. **Convex sets, polar duality and Minkowski norms.** We present some basic results on convex geometry and polar duality which will be used in this section to characterize the Hofer norms. We refer to Chapter I of [Bron83] for basic results on convex sets and Chapter II of the same book for basic results on convex polytopes. Another reference is [Bar02]. Let V be a finite dimensional inner product space. For a non-zero vector $x \in V$ and a scalar $a \in \mathbb{R}$ we define the hyperplane

$$H_{x,a} = \{y \in H : \langle x, y\rangle = a\}.$$

For $a = 1$ we set $H_x := H_{x,1}$. We define the negative halfspace as

$$H^-_{x,a} = \{y \in H : \langle x, y\rangle \leq a\}.$$

The *polar duality* is given by the following bijection between non-zero points in V and hyperplanes in V not containing zero: $x \mapsto H_x$.

Definition 2.17 (Supporting hyperplanes)**.** The *support function* of a bounded subset $E \subseteq V$ is the function

$$h_E : V \to \mathbb{R}, \quad h_E(u) = \sup_{x \in E} \langle x, u\rangle.$$

Note that $h_E = h_{\mathrm{conv}(E)}$, and that if $\mathrm{conv}(E)$ contains 0 in its interior, then h_E is a Finsler norm, our one-sided Hofer norm (Remark 2.3).

If $0 \neq x \in V$, the hyperplane given by

$$H(E,x) := \{v \in V : \langle v, x\rangle = h_E(x)\}$$

is the *supporting hyperplane* of E for x (See Figure 1). For $x \in E$, the set

$$F_x(E) := E \cap H(E,x) = \mathrm{argmax}_E(\varphi_x)$$

is called the *face* of E defined by x, or also the *support set* of E for x. For $v \in F_x(E)$ we say that $H(E,x)$ supports E at v.

In the literature the faces defined above are usually called exposed faces.

Definition 2.18 (Minkowski gauge)**.** The *Minkowski gauge* or *gauge* of a bounded convex set $B \subseteq V$ which contains the origin in its interior is the function

$$g_B : V \to \mathbb{R}, \quad g_B(x) = \inf\{t > 0 : x \in tB\}.$$

The set B is a symmetric convex body (Definition 2.1) if and only if the gauge function g_B is a norm on V whose unit ball is B. Otherwise it is a Finsler norm, i.e. only positively homogeneous.

Remark 2.19. If $x \in V$ is such that $g_B(x) = 1$ then $H_y = \varphi_y^{-1}(1) = \{z \in V : \langle z, y\rangle = 1\}$ is a supporting hyperplane of B at x if and only if φ_y is a norming functional of x.

Definition 2.20. The *polar* of a nonempty bounded set $E \subseteq V$ is

$$E^\circ = \{x \in V : \langle x, y\rangle \leq 1 \text{ for all } y \in E\}.$$

Note that if E is invariant by an isometric action so is its polar E°. A *polytope* is the convex hull of a finite set of points.

These are the result of applying the polar operation to some standard sets

- If B_p is the unit ball of the ℓ^p space for $1 \leq p \leq \infty$ then $B_p^\circ = B_q$ where q is conjugate to p.
- The polar of an ellipsoid $E \subseteq \mathbb{R}^n$ with axes of length $a_1, \ldots, a_n$ is the ellipsoid with axes of length $1/a_1, \ldots, 1/a_n$.
- The polar of a polytope $P \subseteq V$ containing 0 in its interior and which has n faces and k vertices is a polytope with k faces and n vertices.

Remark 2.21. These are some standard properties that will be used later; let $E, F \subseteq V$ be compact convex sets containing $0 \in V$ in the interior, then

(1) $E^{\circ\circ} = E$, this is the bipolar property.
(2) $(\lambda E)^\circ = \lambda^{-1} E^\circ$ for $\lambda > 0$.
(3) If $E \subseteq F$, then $F^\circ \subseteq E^\circ$.
(4) $(E \cup F)^\circ = E^\circ \cap F^\circ$.
(5) $(E \cap F)^\circ = \operatorname{conv}(E^\circ \cup F^\circ)$.
(6) For a polytope $E = \operatorname{conv}\{x_1, \ldots, x_n\}$ we have

$$E^\circ = \{y \in V : \langle x_i, y\rangle \leq 1 \text{ for } i = 1, \ldots, n\}.$$

(7) For $E = \{y \in V : \langle x_i, y\rangle \leq 1 \text{ for } i = 1, \ldots, n\}$ its polar is the polytope

$$E^\circ = \operatorname{conv}\{0, x_1, \ldots, x_n\}.$$

The next result can be found in Theorem 6.4 and Corollary 6.5 of [Bron83], and Theorem 14.5 of [Rock70].

Theorem 2.22. *Let $E \subseteq V$ be a compact convex set containing* 0 *in its interior. Then*

(1) $h_E = g_{E^\circ}$ *and* $g_E = h_{E^\circ}$.
(2) *The supporting hyperplanes of E are the hyperplanes determined by the points of the boundary* $\operatorname{bd} E^\circ$ *of E°, i.e. H_x with $x \in \operatorname{bd} E^\circ$.*

The following are equivalent, see Figure 3

- *the hyperplane H_y supports E at $x \in E$.*
- *the hyperplane H_x supports E° at $y \in E^\circ$.*
- $x \in \mathrm{bd}\,E$, $y \in \mathrm{bd}\,E^\circ$ *and* $\langle x, y\rangle = 1$.

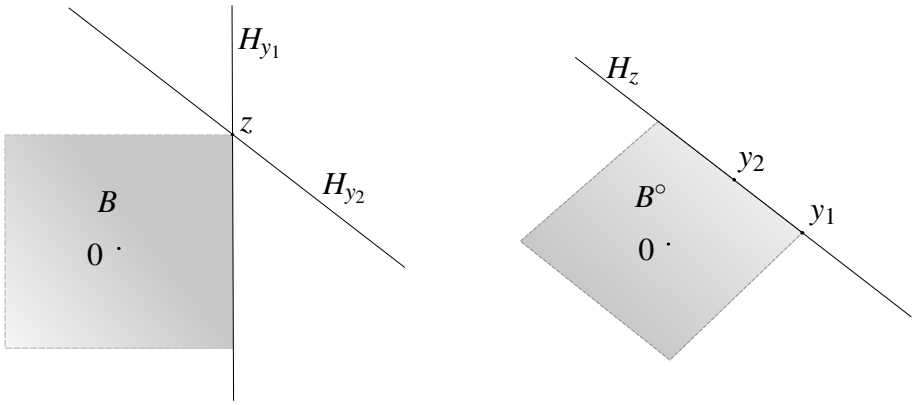

FIGURE 3. Polar duality and supporting hyperplanes.

The following, which relates the polar operation to orthogonal projections and sections with subspaces will also be used. Its proof is elementary therefore omitted.

Remark 2.23. If B is a convex body in the inner product space V, W is a subspace of V, and p_W is the orthogonal projection onto W, then the following holds:

i) $(B \cap W)^\circ = p_W(B^\circ)$.
ii) $B^\circ \cap W = p_W(B)^\circ$.

Item ii) actually holds for any subset B of V and item i) follows from polar duality.

2.3. **Convex structure of the unit ball of Hofer norms.** In this section, we characterize the generalized Hofer norms in terms of the supporting and gauge functions h_E, g_E.

Proposition 2.24. *Let $E \subseteq V$ be a compact full set. Then we have*

$$\|y\|_E = h_{E-E}(y) = g_{(E-E)^\circ}(y).$$
$$\|y\|'_E = h_{E\cup -E}(y) = g_{(E\cup -E)^\circ}(y).$$

If in addition 0 *is in the interior of* $\mathrm{conv}(E)$

$$\|y\|^+_E = h_E(y) = g_{E^\circ}(y).$$

Proof. For the first Hofer norm observe that

$$\begin{aligned}
\|y\|_E &= \max_{x\in E}\langle x, y\rangle - \min_{x\in E}\langle x, y\rangle = \max_{x\in E}\langle x, y\rangle - (-\max_{x\in E}\langle x, -y\rangle) \\
&= h_E(y) - (-h_E(-y)) = h_E(y) + h_E(-y) && \text{by Definition 2.17} \\
&= h_E(y) + h_{-E}(y) = h_{E-E}(y) \\
&= g_{(E-E)^\circ}(y) && \text{by Theorem 2.22.}
\end{aligned}$$

For the second Hofer norm

$$\begin{aligned}\|y\|'_E &= \max\{\max_{x\in E}\langle x,y\rangle, -\min_{x\in E}\langle x,y\rangle\}\\ &= \max\{\max_{x\in E}\langle x,y\rangle, \max_{x\in E} -\langle x,y\rangle\}\\ &= \max\{\max_{x\in E}\langle x,y\rangle, \max_{x\in -E}\langle x,y\rangle\} = \max_{x\in E\cup -E}\langle x,y\rangle = h_{E\cup -E}(y)\\ &= g_{(E\cup -E)^\circ}(y) \qquad \text{by Theorem 2.22.}\end{aligned}$$

For the one-sided Hofer norm

$$\|y\|^+_E = \max_{x\in E}\langle x,y\rangle = h_E(y) = g_{E^\circ}(y).$$

□

Remark 2.25. Note that the unit ball of Hofer's norm is exactly

$$(E-E)^\circ = (\mathrm{conv}(E-E))^\circ = (\mathrm{conv}(E)-\mathrm{conv}(E))^\circ.$$

Therefore all Ad-invariant norms are Hofer norms if we take $E = \frac{1}{2}B^\circ$ where B is the unit ball of the norm.

With these tools, we give a second proof of the characterization of the norming functionals of vectors in the unit sphere (Theorem 2.16), or equivalently, the supporting hyperplanes of the unit ball for points in its sphere.

Theorem 2.26. *The norming functionals of $x \in V$ are given by $\varphi_{y^+-y^-}$, with*

$$y^+ \in \mathrm{conv}(\mathrm{argmax}_E\,\varphi_x) \quad \textit{and} \quad y^- \in \mathrm{conv}(\mathrm{argmin}_E\,\varphi_x).$$

Proof. Rescaling, we can assume that $1 = \|x\|_E = g_{\mathrm{conv}(E-E)^\circ}(x)$, so that that $x \in \mathrm{bd}(E-E)^\circ$. By Remark 2.19 a functional φ_y is a norming functional of x if and only if the hyperplane H_y supports $(E-E)^\circ$ at x. Theorem 2.22 implies that H_y supports $(E-E)^\circ$ at x if and only if H_x supports $(E-E)^{\circ\circ} = \mathrm{conv}(E-E) = \mathrm{conv}(E)-\mathrm{conv}(E)$ at $y \in \mathrm{conv}(E-E)$. Moreover, H_x supports $\mathrm{conv}(E-E)$ at $y \in \mathrm{conv}(E-E)$ if and only if $y \in \mathrm{argmax}_{\mathrm{conv}(E)-\mathrm{conv}(E)}(\varphi_x)$, and this happens if and only if

$$y \in \mathrm{argmax}_{\mathrm{conv}(E)}(\varphi_x) - \mathrm{argmin}_{\mathrm{conv}(E)}(\varphi_x).$$

That is, $y = y^+ - y^-$ with

$$y^+ \in \mathrm{argmax}_{\mathrm{conv}(E)}(\varphi_x) = \mathrm{conv}(\mathrm{argmax}_E\,\varphi_x)$$

and likewise with y^-. This finishes the proof. □

Remark 2.27. For the one-sided Hofer norm the norming functionals of $x \in V$ are given by φ_{y^+}, with

$$y^+ \in \mathrm{conv}(\mathrm{argmax}_E\,\varphi_x).$$

3. Hofer's metric on compact Lie groups

In this section, we present actions of compact semi-simple Lie groups K on compact connected manifolds M with symplectic form ω, as a nice setting for the convex geometry that was discussed in the previous sections. It should serve as motivation and also as a source of examples. We refer to [PolShe15] for general background on the geometry of Hamiltonian actions.

3.1. **Hamiltonian diffeomorphisms.** Let (M,ω) be a connected closed symplectic manifold and let $H : [0,1]\times M \to \mathbb{R}$ be a smooth function. We denote $H_t(m) := H(t,m)$. This function H induces a time dependent Hamiltonian vector field X_{H_t} by Hamilton's equations

$$dH_t = \omega(\cdot, X_{H_t}) = -\iota_{X_{H_t}}\omega, \tag{3.1}$$

and hence an isotopy $\phi_t^H : M \to M$, $t \in [0,1]$ by the prescription that

$$\phi_0^H = \phi \text{ and } \frac{d}{dt}\phi_t^H(m) = X_{H_t}(\phi_t^H(m)).$$

For a symplectic map φ. The Hamiltonian diffeomorphism group $\mathrm{Ham}(M,\omega)$ is by definition the set of diffeomorphisms $\phi : M \to M$ which can be written as $\phi = \phi_1^H$ for some H and $\phi_0^H = \mathrm{id}$ as above.

The set $\mathrm{Ham}(M,\omega)$ is an infinite dimensional group under composition, all elements of which are symplectomorphisms of (M,ω). Its Lie algebra are the Hamiltonian vector fields in $\mathcal{X}(M)$, which can be identified with the smooth functions in M modulo constant functions, via (3.1), hence

$$T_{\mathrm{id}}\,\mathrm{Ham}(M,\omega) \simeq \mathcal{X}_{\mathrm{Ham}}(M) \simeq C^\infty(M)/\mathbb{R}\mathbf{1}.$$

Since $\psi^* X_H = X_{H\circ\psi}$ the adjoint action in this group is given by $\mathrm{Ad}_\psi[H] = [H\circ\psi]$, where $[H]$ will denote the class of H modulo constant functions. If the Hamiltonian is time independent, i.e. $H_t = H$ for $t \in [0,1]$, it is called *autonomous*. Note that the flow of such H is ruled by the equation

$$\frac{d}{dt}\phi_t(m) = X_H(\phi_t(m)) = D(R_{\phi_t})_1(X_H)$$

when we interpret the differential of the right translation R_g in the group of diffeomorphisms, as composition from the right. Therefore ϕ_t is the flow of the right invariant field $X_g = D(R_g)_1(X_H)$ and with the initial condition $\phi_0 = \mathrm{id} = 1_{\mathrm{Ham}(M,\omega)}$ it is clear that $\phi_1 = \exp(X_H)$, where exp is the exponential map of the group of Hamiltonian diffeomorphism. However, this exponential map is not well-suited as a chart for the group, since it is not a local diffeomorphism in any reasonable neighbourhood of the 0 vector field, see [PolShe15].

3.1.1. *Hofer's norm.* The Ad-invariant L^∞ norm

$$\|[H]\| = \max_M H - \min_M H \tag{3.2}$$

on the Lie algebra $T_{\mathrm{id}}\operatorname{Ham}(M,\omega) \simeq C^\infty(M)/\mathbb{R}\mathbf{1}$ of the group $\operatorname{Ham}(M,\omega)$ is *Hofer's norm*. It induces a Finsler length structure on curves $(\phi_t)_{t\in[0,1]}$ in $\operatorname{Ham}(M,\omega)$ by means of

$$\begin{aligned}\operatorname{length}(\phi_t^H) &= \int_0^1 \|\frac{d}{dt}\phi_t^H\|dt = \int_0^1 \|H_t \circ \phi_t^H\|dt \\ &= \int_0^1 \left(\max_M H_t - \min_M H_t\right) dt,\end{aligned}$$

and hence a bi-invariant distance

$$\operatorname{dist}(\phi_0,\phi_1) = \inf\left\{\operatorname{length}(\phi_H^t) : \phi_H^0 = \phi_0, \quad \phi_H^1 = \phi_1\right\}.$$

As was shown for $\mathbb{R}^{2n}$ in [Hof90] and for general symplectic manifolds in [LaMcD95b], dist is a non degenerate, bi-invariant metric on $\operatorname{Ham}(M,\omega)$.

Hofer proved that the path of any autonomous Hamiltonian on $\mathbb{C}^n$ is length minimizing (among homotopic paths with fixed endpoints) as long as the corresponding Hamilton's equation has no non-constant time-one periodic orbit. This result was generalized in [LaMcD95b] to general symplectic manifolds.

Remark 3.1. The *symplectic energy* on $\operatorname{Ham}(M,\omega)$ is $\mathrm{E}(\psi) = \operatorname{dist}(\mathrm{id},\psi)$. It is invariant under conjugation by symplectomorphisms. This function is also referred to as the Hofer energy.

Definition 3.2. A path $(\phi_t)_{t\in[0,1]} \subseteq \operatorname{Ham}(M,\omega)$ is a geodesic of the Hofer metric if each $t \in [0,1]$ has a neighbourhood I such that $\phi|_I$ is minimal, i.e. shorter than any other path joining its endpoints.

A Hamiltonian H_t is called *quasi-autonomous* if there exists two points $x^-, x^+ \in M$ such that

$$H_t(x^-) = \min_M H_t, \quad H_t(x^+) = \max_M H_t$$

for all $t \in [0,1]$. The Hamiltonian is called *slow* if it has no contractible periodic orbits of period less than or equal to one.

It is worth mentioning here the following results concerning quasi-autonomous Hamiltonians and the Hofer metric (but we won't be making use of them). See [McDSa17, Section 12.3] and the references therein for the proofs:

Theorem 3.3. *Let $(\phi_t)_{t\in[0,1]} \subseteq \operatorname{Ham}(M,\omega)$ be a regular path (C^1 and with non-vanishing derivative). If ϕ is short for the Hofer metric, then the corresponding Hamiltonian H_t is quasi-autonomous. If the Hamiltonian is quasi-autonomous, then ϕ is locally short (locally here refers to the time-variable). There exists a C^2 neighbourhood V of the identity of $Ham(M,\omega)$ such that every point of V can be joined to the identity by a minimizing geodesic.*

A very similar norm used in the literature is defined without taking the quotient of the Hamiltonian functions by the constant functions. It is defined via the normalization $\int_M H\omega^n = 0$, where ω^n is the Liouville measure on M. The norm is the Ad-invariant L^∞ norm

$$\|H\|' = \max\{\max H, -\min H\} = \|H\|_\infty$$

which we are going to relate to the second Hofer norm in Remark 3.9.

3.2. **Hamiltonian actions.** Let K be compact semi-simple Lie group with Lie algebra $\mathfrak{k} = T_1 K$ and dual space $\mathfrak{k}^*$, which we identify with $\mathfrak{k}$ through the duality pairing given by the (opposite of) the Ad-invariant Killing form $\langle \cdot, \cdot \rangle$. This is an inner product in $\mathfrak{k}$ because $\mathfrak{k}$ is compact; since it is also Ad-invariant, we can identify the coadjoint action with the adjoint action. The Lie algebra $\mathfrak{k}$ acts by skew-symmetric transformations, that is $\operatorname{ad} x$ is skew-adjoint for this inner product, for any $x \in \mathfrak{k}$.

Definition 3.4. Assume that the action of K on M is a *symplectic action*, that is, there exists a smooth map $\Phi : K \times M \to M$ that we denote $\Phi(g,m) = g \cdot m$ such that for each fixed $g \in K$, the automorphism $\Phi_g = \Phi(g, \cdot)$ is a symplectomorphism of M. We also assume that the action is *almost effective* ($g \mapsto \Phi_g$ has discrete kernel). For fixed $m \in M$ we denote by $\pi_m : K \to M$ the map $\pi_m(g) = \Phi(g,m)$.

For $x \in \mathfrak{k}$, let $x_M(m) = \frac{d}{dt}\,|_{t=0} \exp(tx) \cdot m \in T_m M$ denote the infinitesimal action on M, the assumption that the action is almost effective implies that $\mathfrak{k} \ni x \mapsto x_M \in \mathfrak{X}(M)$ is injective; the assumption that the action is symplectic implies that the field x_M is symplectic, i.e. $\mathcal{L}_{x_M}\omega = 0$.

A *moment map* for a symplectic K action on (M,ω) is a map $\mu : M \to \mathfrak{k}^*$ defined by

$$\mu^x(m) = \langle \mu(m), x \rangle, \quad \mu^x : M \to \mathbb{R}$$

such that μ intertwines the K-action on M and the coadjoint action on $\mathfrak{k}^*$, i.e. $\mu(g \cdot m) = \operatorname{Ad}_g \mu(m)$ for all $g \in K, m \in M$, and such that μ satisfies Hamilton's equation

$$d\mu^x = -\iota(x_M)\omega = -\omega(x_M, \cdot) \text{ for all } x \in \mathfrak{k}. \tag{3.3}$$

A symplectic K action is called *Hamiltonian* if it admits a moment map.

Remark 3.5. A symplectic K action on a symplectic manifold (M,ω) is the same thing as a homomorphism $K \to \operatorname{Symp}(M,\omega)$ to the group of symplectomorphisms of (M,ω) such that the map $K \times M \to M$, $(g,m) \to g \cdot m$ is smooth. The action has a moment map if and only if the image of the identity component of K is contained in $\operatorname{Ham}(M,\omega)$, i.e. there is a map $\mu : M \to \mathfrak{k}^*$ which satisfies Hamilton's equations (3.3). Averaging with the Haar probability measure dk on K yields an equivariant moment map

$$\widetilde{\mu}(x) = \int_{k \in K} \operatorname{Ad}_k \mu(k^{-1} \cdot x)\, dk.$$

Note that if $x \in \mathfrak{k}$ then x_M is a Hamiltonian vector field and $H = \mu^x$ is a Hamiltonian; from (3.3) it is apparent that $X_{\mu^x} = x_M$. Also note that the Poisson bracket of the Hamiltonian functions μ^x and μ^y is given by $\{\mu^x, \mu^y\} = \mu^{[x,y]}$ for $x, y \in \mathfrak{k}$.

The action defines a homomorphism with discrete kernel

$$K \to \operatorname{Ham}(M,\omega).$$

At the level of Lie algebras this morphism is injective

$$\begin{aligned} T_1(K) = \mathfrak{k} \hookrightarrow C^\infty(M)/\mathbb{R}\mathbf{1} &\simeq T_{\mathrm{id}} \operatorname{Ham}(M,\omega) \\ x \mapsto [\mu^x] = \mu^x + \mathbb{R}\mathbf{1} &\mapsto x_M. \end{aligned} \tag{3.4}$$

Given $u \in K$ and a path $x : [0,1] \to \mathfrak{k}$, $t \mapsto x_t$, we calculate the isotopy $\phi_t \in \operatorname{Ham}(M,\omega)$ given by the initial condition $\phi_0 = \Phi_u$ and the time dependent Hamiltonian $H_t = \mu^{x_t}$ (as noted, the Hamiltonian vector fields are $(x_t)_M$).

Proposition 3.6. *If $\gamma\colon [0,1] \to K$ is the smooth solution to $\dot\gamma_t\gamma_t^{-1} = x_t$ and $\gamma_0 = u$, then the isotopy is given by $\phi_t = \Phi_{\gamma_t}$.*

Proof. We have to check that $\dot\phi_t\phi_t^{-1} = (x_t)_M$: for each $m \in M$,

$$\begin{aligned}\dot\phi_t(m) &= (\gamma_t \cdot m)^{\cdot} = D(\pi_m)_{\gamma_t}(\dot\gamma_t) = D(\pi_m)_{\gamma_t}(x_t\gamma_t) = D(\pi_m)_{\gamma_t}(\frac{d}{dr}\Big|_{r=0}\exp(rx_t)\gamma_t)\\ &= \frac{d}{dr}\Big|_{r=0}\pi_m(e^{rx_t}\gamma_t) = \frac{d}{dr}\Big|_{r=0}(e^{rx_t}\gamma_t)\cdot m = \frac{d}{dr}\Big|_{r=0}e^{rx_t}\cdot(\gamma_t\cdot m)\\ &= (x_t)_M(\gamma_t\cdot m) = (x_t)_M(\phi_t(m)) = ((x_t)_M\circ\phi_t)(m),\end{aligned}$$

and therefore $\dot\phi_t = (x_t)_M \circ \phi_t$ as claimed. □

Remark 3.7. The image of the moment map $\mu(M)$ is a union of (co)adjoint orbits in $\mathfrak{k}$ which we denote by $O_\lambda = \mathrm{Ad}_K(\lambda)$ for $\lambda \in \mathfrak{k}$. Let T be a maximal torus of K and $\mathfrak{h}$ its Lie algebra, a Cartan subalgebra. The image of the moment map can therefore be parametrized by a subset $A = \mu(M) \cap \mathfrak{h}_+$ of a closed positive Weyl chamber $\mathfrak{h}_+$, corresponding to a choice positive simple roots. Hence we have

$$\mu(M) = \bigsqcup_{\lambda\in A}\mathrm{Ad}_K(\lambda) = \bigsqcup_{\lambda\in A} O_\lambda.$$

3.2.1. *Hofer's metric given by Hamitonian actions.* We use the inclusion of Lie algebras (3.4) to pull-back the Finsler length structure given by Hofer's norm. The L^∞ norm on $C^\infty(M)/\mathbb{R}\mathbf{1}$ restricted to the image $T_1(K) = \mathfrak{k}$ is therefore by definition of the Hofer norm (3.2) given by

$$\begin{aligned}\|\mu^x\| &= \max_{m\in M}\mu^x(m) - \min_{m\in M}\mu^x(m) = \max_{m\in M}\langle\mu(m),x\rangle - \min_{m\in M}\langle\mu(m),x\rangle\\ &= \max_{y\in\mu(M)}\langle y,x\rangle - \min_{y\in\mu(M)}\langle y,x\rangle = \|x\|_{\mu(M)},\end{aligned} \tag{3.5}$$

where the last equality is by definition of the generalized Hofer norm (Definition 2.1) for the Ad-invariant set $E = \mu(M) \subseteq \mathfrak{k}$.

Remark 3.8 ($\|\cdot\|_{\mu(M)}$ is a norm)**.** Since M is compact the image of the moment map is bounded. Also, because the image of the moment map is full in $\mathfrak{k}$ if and only if the action is almost effective: The image of the moment map is not full if and only if its image is contained in a hyperplane, which in turn happens if and only if there exists $x \in \mathfrak{k}$ such that φ_x is constant on $\mu(M)$. This is equivalent to the map $\mu^x : M \to \mathbb{R}$ being a constant Hamiltonian which generates the zero vector field. Finally, this is also equivalent to the one-parameter group $\exp(tx)$ acting trivially, i.e. the action not being almost effective.

Therefore, the Hofer norm just described is the norm $\|\cdot\|_{\mu(M)}$ for a moment map $\mu : M \to \mathfrak{k}^* \simeq \mathfrak{k}$, given by equation (3.5) above. For a general Ad-invariant set E in $\mathfrak{k}$ we obtain an Ad-invariant norm $\|\cdot\|_E$ in $\mathfrak{k}$. This norm is Ad-invariant and induces a bi-invariant Finsler metric on K by left (or right) translation given by

$$\mathrm{length}(\gamma) = \int_0^1 \|\dot\gamma_t\|_{\gamma_t}dt = \int_0^1 \|\dot\gamma_t\gamma_t^{-1}\|_{\mu(M)}dt = \int_0^1 \|\gamma_t^{-1}\dot\gamma_t\|_{\mu(M)}dt,$$

$$\mathrm{dist}(u,v) := \inf\{\mathrm{length}(\gamma_t) : \gamma_0 = u, \quad \gamma_1 = v\}. \tag{3.6}$$

Remark 3.9 (Barycenter)**.** In the case of actions by semi-simple Lie groups, all Hamiltonians are normalized in the following sense:

$$\int_M \mu^x \omega^n = 0.$$

This is because

$$\int_M \langle \mu, x\rangle \omega^n = \int_{\mu(M)} \langle y, x\rangle d\nu(y),$$

where $\nu = \mu_*(\omega^n)$ is the pushforward of the normalized Liouville measure ω^n by the moment map. The measure ν is Ad invariant hence its center of mass is also invariant. The center of mass is zero since the group is semi-simple, therefore $\int_{\mu(M)} \langle y, x\rangle d\nu(y) = 0$. The resulting L^∞ norm is

$$\|\mu^x\|_\infty = \max\{\sup_{m\in M} \mu^x(m), -\inf_{m\in M} \mu^x(m)\} = \|x\|_{\mu(M)}, \tag{3.7}$$

by the definition of the second generalized Hofer norm in 2.2 for the Ad-invariant set $E = \mu(M) \subseteq \mathfrak{k}$. Therefore, this second Hofer norm just described is the norm defined in equation (2.2) as $\|\cdot\|'_{\mu(M)}$ for a momentum map $\mu : M \to \mathfrak{k}^* \simeq \mathfrak{k}$.

Remark 3.10. We can define also an L^p-norm for $1 \leq p < \infty$ via the embedding

$$\iota : V \hookrightarrow L^p(E, \nu)/\mathbb{R}\mathbf{1}, \quad x \mapsto [\varphi_x] = \varphi_x + \mathbb{R}\mathbf{1},$$

where ν is an Ad-invariant measure. Consider the case of an L^p-norm defined by the inclusion

$$\iota : \mathfrak{k} \hookrightarrow L^p(M, \nu)/\mathbb{R}\mathbf{1}.$$

Since $L^p(M, \nu)$ is uniformly convex when $1 < p < \infty$, the quotient $L^p(M, \nu)/\mathbb{R}\mathbf{1}$ is uniformly convex. Therefore the L^p-norm is uniformly convex and this implies by [Lar19, Theorem 4.15] that K with the induced metric is uniquely geodesic. We won't be pursuing the geometry of these norms in this paper.

Remark 3.11. Recall that for a compact semisimple group K, each element $\lambda \in K$ is semi-simple. A *regular* element λ is such that its commutant $\mathfrak{h} = \mathfrak{z}(\lambda)$ is a maximal abelian subalgebra; equivalently in this setting, $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{k}$.

In the case of $K = \mathrm{SU}(n)$, an element of $\mathfrak{k} = \mathfrak{su}(n)$ is regular if and only if all its eigenvalues are distinct.

Example 3.12 (Coadjoint orbits)**.** An important special case is that of a compact connected semi-simple Lie group K acting via Ad in a Hamiltonian way on a (co)adjoint orbit $\mathcal{O}_\lambda$ containing $\lambda \in \mathfrak{k}^* \simeq \mathfrak{k}$.

For the action to be almost effective, it is necessary and sufficient that λ is non-zero in each simple summand of $\mathfrak{k}$ (because in that case each simple factor of K acts nontrivially on the orbit, and this is equivalent to the fact that the (co)adjoint orbit is full in $\mathfrak{k}$ (see [BiGhiHe14, Lemma 17] and Remark 3.8). In particular this occurs if λ is regular.

The adjoint orbit is endowed with the Kostant-Kirillov-Souriau symplectic form ω. The kernel of this homomorphism is the center $Z(K)$ of K which is a discrete subgroup. We have therefore an inclusion

$$K/Z(K) \hookrightarrow \mathrm{Ham}(\mathcal{O}_\lambda, \omega).$$

At the level of Lie algebras this inclusion is $x \mapsto [\mu^x]$ where $\mu^x : O_\lambda \to \mathbb{R}$ is the Hamiltonian map given by the $\mu^x(y) = \langle x, y\rangle$. The Hamiltonian is the component of the moment map $\iota : O_\lambda \hookrightarrow \mathfrak{k}^* \simeq \mathfrak{k}$ along x, that is $\mu^x = \varphi_x \circ \iota$. Moreover, since the action of K is given by $g \mapsto gxg^{-1}$, it is apparent that if $x \in \mathfrak{k}$ then the induced Hamiltonian vector field is given by

$$x_{O_\lambda}(m) = \frac{d}{dt}\Big|_{t=0} \exp(tx)\cdot m = \frac{d}{dt}\Big|_{t=0} \mathrm{Ad}_{e^{tx}}(m) = (\mathrm{ad}\, x)(m) = [x, m].$$

The symplectic form is

$$\omega_x(y_{O_\lambda}(x), z_{O_\lambda}(x)) = \langle x, [y, z]\rangle$$

where the pairing is given by the opposite Killing form of $\mathfrak{k}$ as before. The induced Hofer norm on $T_1(K) = \mathfrak{k}$ is therefore

$$\begin{aligned} \|x\|_{O_\lambda} &= \max_{y\in O_\lambda} \mu^x(y) - \min_{y\in O_\lambda} \mu^x(y) = \max_{k\in K} \mu^x(\mathrm{Ad}_k \lambda) - \min_{k\in K} \mu^x(\mathrm{Ad}_k \lambda) \\ &= \max_{k\in K} \langle x, \mathrm{Ad}_k \lambda\rangle - \min_{k\in K} \langle x, \mathrm{Ad}_k \lambda\rangle. \end{aligned} \tag{3.8}$$

Since the action is almost effective $\|\cdot\|_{O_\lambda}$ is a norm. Since it is also Ad-invariant it induces a bi-invariant Finsler metric on K.

Example 3.13. In the case of the special unitary group $\mathrm{SU}(n)$ we have the inclusion

$$\mathrm{SU}(n)/\mathbb{Z}_n \simeq \mathrm{PU}_n \hookrightarrow \mathrm{Ham}(O_\lambda, \omega)$$

where O_λ is a (co)adjoint orbit of $\mathrm{SU}(n)$ passing through the given diagonal matrix $\lambda = i\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n) \in \mathfrak{su}(n)$ with $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$. At the Lie algebra level this inclusion is

$$T_1(\mathrm{SU}(n)) = \mathfrak{su}(n) \hookrightarrow C^\infty(O_\lambda)/\mathbb{R}\mathbf{1} \simeq T_{\mathrm{id}}\,\mathrm{Ham}(O_\lambda, \omega)$$

and $\mu^x : O_\lambda \to \mathbb{R}$ is the linear Hamiltonian given by the opposite Killing form

$$\mu^x(y) = -2n\,\mathrm{tr}(xy^*) = 2n\,\mathrm{tr}(xy).$$

The Hofer norm (3.8) is given by

$$\|v\|_{O_\lambda} = 2n \max\{\mathrm{tr}(\lambda uvu^{-1}) : u \in \mathrm{SU}(n)\} - 2n \min\{\mathrm{tr}(\lambda wvw^{-1}) : w \in \mathrm{SU}(n)\},$$

which is exactly ($2n$-times) the λ-numerical diameter of v, see the next remark.

Example 3.14. The group $\mathrm{SU}(2)/\{\mathrm{id}, -\mathrm{id}\} \simeq \mathrm{SO}(3)$ acts on a nontrivial (co)adjoint orbit $O_\lambda \simeq S^2$ which is the two dimensional sphere endowed with the area form ω. The action is given by rotations, see Example 1.4.H in [Pol01].

Remark 3.15. Another bi-invariant norm related to the Hofer norm for skew-symmetric matrices $v \in \mathfrak{su}(n)$, is the *c-numerical radius*

$$\omega(v)_c = \max\{|\,\mathrm{tr}(cuvu^{-1})| : u \in U_n\},$$

given by a fixed matrix c such that $\mathrm{tr}(c) \neq 0$ and c is not a scalar multiple of the identity (see [LiTsi89]). This is our second Hofer norm applied to non semi-simple groups. The usual numerical radius of v can be recovered by picking the one-dimensional projection $c = e_1 \otimes e_1$.

Remark 3.16 (One-sided norms)**.** For fixed $0 \neq \lambda \in \mathfrak{k}$, and $v \in \mathfrak{k}$, consider the function

$$\|v\|^+_{\mathcal{O}_\lambda} = \max_{k\in K} \langle v, \mathrm{Ad}_k \lambda\rangle.$$

Note that by the Ad-invariance of the norm, this can be also computed as

$$\|v\|^+_{\mathcal{O}_\lambda} = \max_{k\in K} \langle \mathrm{Ad}_k v, \lambda\rangle = \max\{\langle y, \lambda\rangle : y \in \mathrm{conv}(\mathcal{O}_v)\}.$$

When λ is non-zero on each simple-summand of $\mathfrak{k}$ (in particular if λ is regular), then 0 is in the interior of its closed convex hull (see Remark 5.10 below). Therefore this is a Finlser norm, in fact it is our one-sided Hofer norm of Remark 2.3. In the context of the group of Hamiltonian symplectomorphisms, these norms were studied in several papers; in particular see [McD02] where the name *one-sided* was used. This norm is positively homogeneous, but if there is a Weyl group element w such that $w.\lambda = -\lambda$ then it is homogeneous. Note by replacing λ with minus λ, we can also obtain $\|v\|^-_{\mathcal{O}_\lambda} = -\min_{k\in K}\langle \mathrm{Ad}_k v, \lambda\rangle$.

4. Convex geometry of Hofer's norm on Cartan algebras

In this section we study generalized Hofer norms on the Lie algebra of a compact semisimple Lie group. We use several convexity theorems and the convex analysis of Ad-invariant convex functions. This convex analysis expresses properties of Ad-invariant functions in terms of properties of its restrictions to Cartan algebras and positive Weyl chambers, which are fundamental domains for the adjoint action.

We first recall Kirwan's nonabelian convexity theorem (see [GuSja05]), which will be useful for our purposes of characterizing the Hofer norm restricted to a Cartan subalgebra $\mathfrak{h} \subseteq \mathfrak{k}$. As before $\mathfrak{h}_+ \subseteq \mathfrak{h}$ is a choice of closed positive Weyl chamber.

Theorem 4.1 (Kirwan)**.** *If $K \curvearrowright M$ is a Hamiltonian action with K a compact connected Lie group and (M,ω) compact and connected, then $\mu(M) \cap \mathfrak{h}_+$ is a convex polytope, i.e. the convex hull of a finite set of points $x_1,\dots,x_n$ in $\mathfrak{h}_+$, that is*

$$\mu(M) \cap \mathfrak{h}_+ = \mathrm{conv}\{x_1,\dots,x_n\}.$$

This theorem is a generalization of the Atiyah-Guillemin-Sternberg theorem, which states that for a Hamiltonian action of a torus on a compact connected manifold the image of the moment map is the convex hull of the image of the fixed point set of the action. The Atiyah-Guillemin-Sternberg theorem is a generalization of Kostant's convexity theorem which will be useful here:

Theorem 4.2 (Kostant)**.** *If K is a compact Lie group and T is a maximal torus in K and $p : \mathfrak{k} \to \mathfrak{h}$ is the projection of the Lie algebra of K onto the Lie algebra of T, then*

$$p(\mathcal{O}_\lambda) = \mathrm{conv}(\mathcal{W}.\lambda),$$

where $\mathcal{W}.\lambda$ is the Weyl group orbit of $\lambda \in \mathfrak{k}$.

The projection p is taken along the orthogonal direction given by (minus) the Killing form of $\mathfrak{k}$.

4.1. **The Hofer polytope.** With these tools we characterize the intersection with Cartan algebras of unit balls (of Hofer norms given by compact, full and Ad-invariant sets $E \subseteq \mathfrak{k}$).

Lemma 4.3. *If E is* Ad*-invariant, and $p : \mathfrak{k} \to \mathfrak{h}$ is the orthogonal projection, then*

$$p(\mathrm{conv}(E)) = \mathrm{conv}\{w.x : x \in E \cap \mathfrak{h}_+,\, w \in \mathcal{W}\},$$

and if $E \cap \mathfrak{h}_+ = \mathrm{conv}\{x_1, \dots, x_n\}$, then

$$p(\mathrm{conv}(E)) = \mathrm{conv}\{w.x_i : i = 1, \dots, n,\, w \in \mathcal{W}\}.$$

Proof. This follows from the identities

$$\begin{aligned}
p(\mathrm{conv}(E)) &= p(\mathrm{conv}(\cup_{\lambda \in E \cap \mathfrak{h}_+} O_\lambda)) \\
&= \mathrm{conv}(p(\cup_{\lambda \in E \cap \mathfrak{h}_+} O_\lambda)) && \text{since } p \text{ is affine} \\
&= \mathrm{conv}(\cup_{\lambda \in E \cap \mathfrak{h}_+} p(O_\lambda)) \\
&= \mathrm{conv}(\cup_{\lambda \in E \cap \mathfrak{h}_+} \{w.\lambda : w \in \mathcal{W}\}) && \text{by Theorem 4.2} \\
&= \mathrm{conv}\{w.\lambda : \lambda \in E \cap \mathfrak{h}_+,\, w \in \mathcal{W}\}.
\end{aligned}$$

Now observe that if $E \cap \mathfrak{h}_+ = \mathrm{conv}\{x_1, \dots, x_n\}$

$$\begin{aligned}
p(\mathrm{conv}(E)) &= \mathrm{conv}(\cup_{\lambda \in E \cap \mathfrak{h}_+} \{w.\lambda : w \in \mathcal{W}\}) \\
&= \mathrm{conv}(\cup_{w \in \mathcal{W}} w.(E \cap \mathfrak{h}_+)) = \mathrm{conv}(\cup_{w \in \mathcal{W}} \mathrm{conv}(w.(E \cap \mathfrak{h}_+))) \\
&= \mathrm{conv}(\cup_{w \in \mathcal{W}} \mathrm{conv}\{w.x_i : i = 1, \dots, n\}) \\
&= \mathrm{conv}\{w.x_i : i = 1, \dots, n,\, w \in \mathcal{W}\}
\end{aligned}$$

where the third and last equalities follow from basic properties of the convex hull operation. □

Remark 4.4. Let $B \subseteq \mathfrak{k}$ be an Ad-invariant convex body, and $p : \mathfrak{k} \to \mathfrak{h}$ the orthogonal projection to a Cartan sub-algebra, then Kostant's Theorem 4.2 implies that $B \cap \mathfrak{h} = p(B)$. This follows from $p(b) \in p(O_b) = \mathrm{conv}(O_b \cap \mathfrak{h}) \subseteq \mathrm{conv}(B \cap \mathfrak{h}) = B \cap \mathfrak{h}$ for $b \in B$.

Proposition 4.5. *Set $A = E \cap \mathfrak{h}_+$, then*

$$\begin{aligned}
(\mathrm{conv}(E) - \mathrm{conv}(E)) \cap \mathfrak{h} &= \mathrm{conv}\{w_1.x_1 - w_2.x_2 : x_1, x_2 \in A,\, w_1, w_2 \in \mathcal{W}\}. \\
(\mathrm{conv}(E \cup -E)) \cap \mathfrak{h} &= \mathrm{conv}\{w.x : x \in A \cup -A, w \in \mathcal{W}\}. \\
\mathrm{conv}(E) \cap \mathfrak{h} &= \mathrm{conv}\{w.x : x \in A, w \in \mathcal{W}\}.
\end{aligned}$$

If furthermore $A = \mathrm{conv}\{x_1, \dots, x_n\}$ is a polytope, then

$$\begin{aligned}
(\mathrm{conv}(E) - \mathrm{conv}(E)) \cap \mathfrak{h} &= \mathrm{conv}\{w.x - w'.x' : x, x' \in \{x_1, \dots, x_n\},\, w, w' \in \mathcal{W}\}. \\
(\mathrm{conv}(E) \cup \mathrm{conv}(E)) \cap \mathfrak{h} &= \mathrm{conv}\{w.x : x \in \{x_1, \dots, x_n, -x_1, \dots, -x_n\}, w \in \mathcal{W}\}. \\
\mathrm{conv}(E) \cap \mathfrak{h} &= \mathrm{conv}\{w.x : x \in \{x_1, \dots, x_n\}, w \in \mathcal{W}\}.
\end{aligned}$$

Proof. For the first equality note that by the previous remark,

$$\begin{aligned}
(\mathrm{conv}(E) - \mathrm{conv}(E)) \cap \mathfrak{h} &= p(\mathrm{conv}(E) - \mathrm{conv}(E)) \\
&= p(\mathrm{conv}(E)) - p(\mathrm{conv}(E)) \\
&= \mathrm{conv}\{w_1.x_1 - w_2.x_2 : x_1, x_2 \in A,\, w_1, w_2 \in \mathcal{W}\}
\end{aligned}$$

where the last equality is due to Lemma 4.3. For the fourth assertion note that

$$\begin{aligned}(\mathrm{conv}(E)-\mathrm{conv}(E))\cap\mathfrak{h} &= p(\mathrm{conv}(E)-\mathrm{conv}(E)) = p(\mathrm{conv}(E))-p(\mathrm{conv}(E))\\ &= \mathrm{conv}\{w.x-w'.x' : x,x'\in\{x_1,\dots,x_n\},\quad w,w'\in\mathcal{W}\}\end{aligned}$$

holds, where we used the second assertion of Lemma 4.3 in the last equality.

For the second equality note that by Remark 4.4,

$$\begin{aligned}(\mathrm{conv}(E\cup -E)\cap\mathfrak{h} &= p(\mathrm{conv}(E\cup -E))\\ &= p(\mathrm{conv}(\mathrm{conv}(E)\cup\mathrm{conv}(-E)))\\ &= \mathrm{conv}(p(\mathrm{conv}(E)\cup\mathrm{conv}(-E)))\\ &= \mathrm{conv}(p(\mathrm{conv}(E))\cup p(\mathrm{conv}(-E)))\\ &= \mathrm{conv}(p(\mathrm{conv}(E))\cup -p(\mathrm{conv}(E)))\\ &= \mathrm{conv}(\{w.x : x\in A,\, w\in\mathcal{W}\}\cup-\{w.x : x\in A,\, w\in\mathcal{W}\})\\ &= \mathrm{conv}(\{w.x : x\in A\cup -A,\, w\in\mathcal{W}\})\end{aligned}$$

where the penultimate equality is due to Lemma 4.3. For the fifth assertion note that

$$\begin{aligned}(\mathrm{conv}(E\cup -E)\cap\mathfrak{h} &= p(\mathrm{conv}(E\cup -E))\\ &= \mathrm{conv}\{\pm w.x : x\in\mathrm{conv}\{x_1,\dots,x_n\},\, w\in\mathcal{W}\}\\ &= \mathrm{conv}\{w.x : x\in\{x_1,\dots,x_n,-x_1,\dots,-x_n\},\, w\in\mathcal{W}\}\end{aligned}$$

holds, where we used the previous identity and properties of the convex hull operation.

The proof of the third and fourth equalities is simpler and we omit it. □

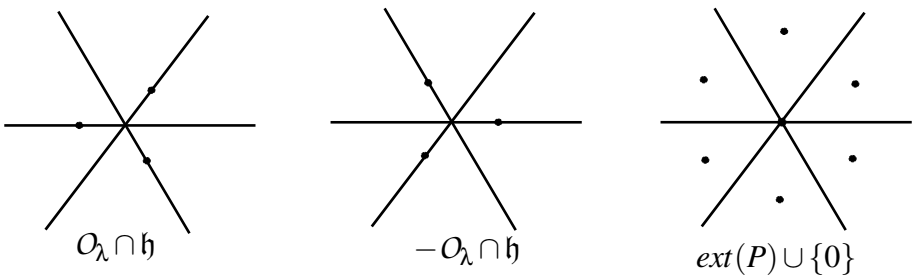

FIGURE 4. Extremal points of the Hofer norm polytope.

In Figure 4 we illustrate the first couple of equalities of Proposition 4.5 in the case $E = O_\lambda$ where $\lambda\in\mathfrak{su}(3)$ is singular and non zero.

Definition 4.6. For $E\cap\mathfrak{h}_+ = \mathrm{conv}\{x_1,\dots,x_n\}$, we call the polytope

$$P = \mathrm{conv}\{w.x-w'.x' : x,x'\in\{x_1,\dots,x_n\},\quad w,w'\in\mathcal{W}\}$$

of Proposition 4.5 the *Hofer norm polytope*. We call

$$P' = \mathrm{conv}\{w.x : x\in\{x_1,\dots,x_n,-x_1,\dots,-x_n\},\, w\in\mathcal{W}\}$$

the *second Hofer norm polytope*. Finally, we call

$$P^+ = \mathrm{conv}\{w.x : x\in\{x_1,\dots,x_n\},\, w\in\mathcal{W}\}$$

the *one-sided Hofer norm polytope*.

Remark 4.7 (Polytopes, norms, unit balls)**.** 1 If $\mathrm{ext}(P) = \{y_1, \dots, y_m\}$ are the extreme points of the Hofer norm polytope P then this set is Weyl group invariant. If

$$B = (\mathrm{conv}(E) - \mathrm{conv}(E))^\circ$$

is the unit ball of the Hofer norm (Remark 2.25), from Remarks 2.23 and 4.4, we obtain

$$B \cap \mathfrak{h} = (\mathrm{conv}(E) - \mathrm{conv}(E))^\circ \cap \mathfrak{h} = P^\circ. \tag{4.1}$$

The same holds for the other Hofer norms with $B' = (\mathrm{conv}(E) \cap -\mathrm{conv}(E))^\circ$ and $B^+ = \mathrm{conv}(E)^\circ$ and the polytopes P' and P^+ respectively (for the second Hofer norm and the one-sided Hofer norm, see Remark 2.3). The polytopes P and P' are centrally symmetric, and in general P^+ is not.

From (4.1) and Proposition 2.24 we obtain the following

Corollary 4.8. *Let $\|\cdot\|_{\mu(M)}$ be the Hofer norm derived from the* Ad*-invariant set $E = \mu(M)$, and let P be the corresponding Hofer norm polytope. Then Hofer norm restricted to the Cartan subalgebra $\mathfrak{h}$ is given by the Minkowski gauge g_{P°, that is*

$$\|x\|_{\mu(M)} = g_{P^\circ}(x) = \inf\{t > 0 : x \in tP^\circ\} \qquad \forall x \in \mathfrak{h}.$$

Remark 4.9. From Remark 2.21 it follows that

$$P^\circ = \{x \in V : \langle y_i, x\rangle \leq 1 \text{ for } i = 1, \dots, m\}.$$

4.1.1. *Direct sums of manifolds and polytopes.* If $M_1, \dots, M_n$ are symplectic manifolds equipped with Hamiltonian actions of K, with moment maps $\mu_1, \dots, \mu_n$, then the induced action on $M_1 \times \cdots \times M_n$ is also Hamiltonian with moment map $\mu : M_1 \times \cdots \times M_n \to \mathfrak{k}$ given by

$$\mu(m_1, \dots, m_n) = \mu_1(m_1) + \cdots + \mu_n(m_n).$$

Therefore the image of μ is given by

$$\mu(M_1 \times \cdots \times M_n) = \mu_1(M_1) + \cdots + \mu_n(M_n). \tag{4.2}$$

To find the Hofer norm polytope of this action we need the following

Lemma 4.10. *If $E_1, \dots, \mathrm{E}_n$ are* Ad*-invariant and compact subsets of $\mathfrak{k}$, and $p : \mathfrak{k} \to \mathfrak{h}$ is the orthogonal projection, then*

$$\begin{aligned}(\mathrm{conv}(E_1 + \cdots + \mathrm{E}_n) &- \mathrm{conv}(E_1 + \cdots + \mathrm{E}_n)) \cap \mathfrak{h} \\ &= (\mathrm{conv}(E_1) - \mathrm{conv}(E_1)) \cap \mathfrak{h} + \cdots + (\mathrm{conv}(E_n) - \mathrm{conv}(E_n)) \cap \mathfrak{h}.\end{aligned}$$

Proof. Again, by Remark 4.4

$$\begin{aligned}(\mathrm{conv}(E_1 + \cdots + \mathrm{E}_n) &- \mathrm{conv}(E_1 + \cdots + \mathrm{E}_n)) \cap \mathfrak{h} \\ &= p((\mathrm{conv}(E_1 + \cdots + \mathrm{E}_n) - \mathrm{conv}(E_1 + \cdots + \mathrm{E}_n))) \\ &= p((\mathrm{conv}(E_1) - \mathrm{conv}(E_1)) + \cdots + (\mathrm{conv}(E_n) - \mathrm{conv}(E_n))) \\ &= p(\mathrm{conv}(E_1) - \mathrm{conv}(E_1)) + \cdots + p(\mathrm{conv}(E_n) - \mathrm{conv}(E_n)) \\ &= (\mathrm{conv}(E_1) - \mathrm{conv}(E_1)) \cap \mathfrak{h} + \cdots + (\mathrm{conv}(E_n) - \mathrm{conv}(E_n)) \cap \mathfrak{h}.\end{aligned}$$

□

The next proposition is now a straightforward consequence of Lemma 4.10.

Proposition 4.11. *If $M_1,\dots,M_n$ are symplectic manifolds equipped with Hamiltonian almost effective actions of a compact semi-simple group K such that the Hofer norm polytopes are $P_1,\dots,P_n$. Then the induced action on $M_1\times\cdots\times M_n$ has Hofer norm polytope $P_1+\cdots+P_n$.*

Remark 4.12. Proposition 4.11 is also valid in the case of the one-sided Hofer Finsler norm (Remark 3.16), with polytopes $P_1^+,\dots,P_n^+$ and $P_1^++\cdots+P_n^+$.

4.1.2. *Faces of balls of* Ad-*invariant norms.* In this section we will use Lewis' results on convex analysis of Ad-invariant functions and their restriction to Cartan subalgebras [Lew00] to describe the faces of the ball of Ad-invariant norms.

Since Lewis' results are stated in terms of subgradients of gauge function we provide next the correspondence between the faces of the balls and the gradients of the gauge functions: *the subdifferential of a gauge g_B at a unit norm x corresponds via $y\mapsto\varphi_y$ to the supporting functionals of the polar of the ball at x.*

Recall that the subdifferential of a convex function $f:V\to\mathbb{R}$ at $x_0\in V$ is defined as

$$\partial f(x_0)=\{y\in V: f(x)-f(x_0)\geq\langle x-x_0,y\rangle \qquad \forall x\in V\}.$$

Each such y which satisfies the condition stated above is called a *subgradient*. It defines a supporting hyperplane to the graph of f at $(x_0,f(x_0))$.

Proposition 4.13. *Let $B\subseteq V$ be a convex body containing $0\in V$ and let g_B be the gauge function associated to B. For a non-zero $x\in V$*

$$\partial g_B(x)=F_x(B^\circ) \quad \textit{and} \quad \partial g_{B^\circ}(x)=F_x(B)$$

i.e. $\partial g_B(x)$ is the exposed face of B° defined by x.

Proof. If $x\neq 0$, we first claim that

$$\partial g_B(x)=\{y\in V:\langle y,x\rangle=g_B(x), g_{B^\circ}(y)=1\}.$$

This is Corollary 23.5.3 in [Rock70], we give a direct proof for the convenience of the reader. Note that since $g_B=h_{B^\circ}$ (Proposition 2.24), $\langle y,x\rangle=g_B(x)$ and $g_{B^\circ}(y)=1$ imply $\langle x,y\rangle=\sup_{w\in B^\circ}\langle x,w\rangle=g_B(x)$ and $\langle y,z\rangle\leq g_B(z)\quad\forall z\in V$ respectively. Thus for such $y\in V$ and any $z\in V$, we have $\langle z,y\rangle-\langle x,y\rangle\leq g_B(z)-g_B(x)$ which shows that $y\in\partial g_B(x)$. Reciprocally, if $y\in\partial g_B(x)$, then replacing with $z=0$ and $z=2x$ in

$$\sup_{w\in B^\circ}\langle w,z\rangle-\sup_{w\in B^\circ}\langle w,x\rangle=g_B(z)-g_B(x)\geq\langle z,y\rangle-\langle x,y\rangle,$$

we obtain $g_B(x)=\langle x,y\rangle$; in particular $g_{B^\circ}(y)\geq 1$. But then from $g_B(z)\geq\langle z,y\rangle$ for all z it also follow that $g_{B^\circ}(y)\leq 1$, thus proving the claim.

Therefore, noting that $x/g_B(x) \in \operatorname{bd} B$

$$\begin{aligned}
\partial g_B(x) &= \{y \in V : \langle y,x\rangle = g_B(x), g_{B^\circ}(y) = 1\} \\
&= \{y \in V : \langle y,x\rangle = g_B(x), y \in \operatorname{bd} B^\circ\} \\
&= \{y \in V : \langle y,x/g_B(x)\rangle = 1, y \in \operatorname{bd} B^\circ\} \\
&= \{y \in V : H_{x/g_B(x)} \text{ supports } B^\circ \text{ at } y\} && \text{by Theorem 2.22} \\
&= F_{x/g_B(x)}(B^\circ) = F_x(B^\circ). && \text{by Definition 2.17}
\end{aligned}$$

The other identity is immediate from the polar duality. □

The next theorem relates the subdifferential of the gauges to the subdifferentials of the gauges restricted to Cartan subalgebras $\mathfrak{h}$. It's proof can be found in [Lew00, Theorem 3.5]. We define the map $\delta : \mathfrak{k} \to \mathfrak{h}^+$ by $O_x \cap \mathfrak{h}^+ = \{\delta(x)\}$, that is, $\delta(x)$ is the unique element in the positive Weyl chamber $\mathfrak{h}^+$ that is Ad-conjugated to x.

Theorem 4.14 (Lewis)**.** *Let K be a semi-simple compact group and let $g_{\mathfrak{k}} : \mathfrak{k} \to \mathbb{R}_+$ be an* Ad*-invariant gauge function. Then $y \in \partial g_{\mathfrak{k}}(x)$ if and only if $\delta(y) \in \partial g_{\mathfrak{h}}(\delta(x))$ and there is $u \in K$ such that $\operatorname{Ad}_u(x), \operatorname{Ad}_u(y) \in \mathfrak{h}_+$.*

If $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is the Cartan decomposition of a semi-simple Lie algebra $\mathfrak{g}$, then Lewis' theorem was stated for an Ad invariant gauge on $\mathfrak{p}$. We can take the complexified Lie algebra $\mathfrak{g} = \mathfrak{k} \oplus i\mathfrak{k}$ and apply the theorem to $\mathfrak{p} = i\mathfrak{k}$. In the case $K = \mathrm{SU}(n)$, the matrices x and y such that there is $u \in K$ with $\operatorname{Ad}_u(x), \operatorname{Ad}_u(y) \in \mathfrak{h}_+$ are said to be *simultaneously ordered diagonalizable*.

We now restate Lewis' theorem in a form convenient to its application in Section 6; there is a related formulation due to Lewis in [Lew98] for the orthogonal group of matrices.

Theorem 4.15. *Let K be a semi-simple compact group, let B be an* Ad*-invariant convex body in $\mathfrak{k}$ containing* 0*, and let $\mathfrak{h}$ be a Cartan algebra in $\mathfrak{k}$ and $\mathfrak{h}_+$ a positive Weyl chamber. Then for non-zero $x \in \mathfrak{h}_+$*

$$F_x(B) = \operatorname{Ad}_{Z(x)}(F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+),$$

and for general $x' \in \mathfrak{k}$, if $v \in K$ is such that $\operatorname{Ad}_v(x') \in \mathfrak{h}_+$, then

$$F_{x'}(B) = \operatorname{Ad}_{v^{-1}} \operatorname{Ad}_{Z(\operatorname{Ad}_v(x'))}(F_{\operatorname{Ad}_v(x')}(B \cap \mathfrak{h}) \cap \mathfrak{h}_+).$$

Proof. We restate Theorem 4.14 in the case that $x \in \mathfrak{h}_+$. Let $g = g_B$ be the gauge function, see Definition 2.18. The theorem states that $y \in \partial g(x)$ if and only if $\gamma(y) \in \partial g_{\mathfrak{h}}(x)$ and for some $u \in K$ we have $\operatorname{Ad}_u(x), \operatorname{Ad}_u(y) \in \mathfrak{h}_+$. Observe that $\operatorname{Ad}_u(x) \in \mathfrak{h}_+$ for $u \in K$ if and only if $u \in Z(x)$, so the condition $\operatorname{Ad}_u(y) \in \mathfrak{h}_+$ reduces to $y \in \operatorname{Ad}_{Z(x)} \mathfrak{h}_+$. The conditions $\gamma(y) \in \partial g_{\mathfrak{h}}(x)$ and $y \in \operatorname{Ad}_{Z(x)} \mathfrak{h}_+$ are equivalent to $y \in \operatorname{Ad}_{Z(x)}(\partial g_{\mathfrak{h}}(x) \cap \mathfrak{h}_+)$. Hence

$$\partial g(x) = \operatorname{Ad}_{Z(x)}(\partial g_{\mathfrak{h}}(x) \cap \mathfrak{h}_+).$$

For a general $x' \in \mathfrak{k}$ take $v \in K$ such that $\operatorname{Ad}_v(x') \in \mathfrak{h}_+$. Then, since the adjoint action is isometric for the norm g we get

$$\partial g(x') = \operatorname{Ad}_{v^{-1}} \operatorname{Ad}_{Z(\operatorname{Ad}_v(x'))}(\partial g_{\mathfrak{h}}(\operatorname{Ad}_v(x')) \cap \mathfrak{h}_+).$$

In the case that $g = g_B$ and $g_{\mathfrak{h}} = g_{B\cap\mathfrak{h}}$ we have

$$F_x(B^\circ) = \partial_B(x) = \mathrm{Ad}_{Z(x)}(\partial g_{B\cap\mathfrak{h}}(x)\cap\mathfrak{h}_+) = \mathrm{Ad}_{Z(x)}(F_x(B^\circ\cap\mathfrak{h})\cap\mathfrak{h}_+),$$

where we used $F_x(B^\circ) = \partial_B(x)$ and $F_x(B^\circ\cap\mathfrak{h}) = \partial g_{B\cap\mathfrak{h}}$ (this follows from Proposition 4.13). If we take polars, by Remark 2.23 and the bipolar property, we obtain the required statement. □

If $B = \mathrm{conv}(E-E)^\circ$, then $B\cap\mathfrak{h} = \mathrm{conv}(E-E)^\circ\cap\mathfrak{h} = (\mathrm{conv}(E-E)\cap\mathfrak{h})^\circ = P^\circ$, so we get the following:

Corollary 4.16. *If $B = \mathrm{conv}(E-E)^\circ$ is the unit ball of Hofer's norm and $P = \mathrm{conv}(E-E)\cap\mathfrak{h}$ is Hofer's norm polytope then for non-zero $x\in\mathfrak{h}_+$*

$$F_x(B) = \mathrm{Ad}_{Z(x)}(F_x(P^\circ)\cap\mathfrak{h}_+).$$

The same characterization holds for the one-sided Hofer norm and its polytope P^+.

5. The geometry of groups with bi-invariant Finsler metrics

In this section we focus our study of geodesics of Lie groups with Finsler metrics obtained by (left or right) translation of Ad-invariant Finsler norms. We want to emphasize that in this section, we only require that norms are positively homogeneous, i.e. $\|\lambda x\| = \lambda\|x\|$ for $\lambda\geq 0$, thus including our one-sided Hofer norms. Therefore the distance obtained is not necessarily symmetric.

We need to begin this section with some remarks concerning compact semi-simple Lie algebras and their root decomposition.

Remark 5.1. A connected finite dimensional Lie group K admits a bi-invariant *Riemannian* metric if and only if it is isomorphic to the cartesian product of a compact group and an additive vector group (see [Mil76, Lemma 7.5]). Therefore in that case $\mathfrak{k}\simeq\mathfrak{k}_0\oplus\mathfrak{a}$ where a is an abelian sub-algebra and $\mathfrak{k}_0$ is semi-simple. To simplify the discussion, we will only consider here compact groups K with semisimple Lie algebra $\mathfrak{k}$. Thus K is always unimodular, in particular $\mathrm{tr}\,\mathrm{ad}\,v = 0$ for any $v\in\mathfrak{k}$. The Ad-invariant inner product is given by (minus) the Killing form of $\mathfrak{k}$, as before, and for this Ad-invariant inner product, the operator $\mathrm{ad}\,v$ is skew-adjoint for any $v\in\mathfrak{k}$ ([Mil76, Lemmas 6.3 & 7.2]).

What follows, in the form of a Remark, is in fact a recollection of useful facts that we will use about the real root decomposition of a real semi-simple Lie algebra. It will also help to fix and clarify the notation.

Remark 5.2 (Real root decomposition)**.** Let Δ be the set of (real) roots of $\mathfrak{k}$ with respect to a fixed Cartan subalgebra $\mathfrak{h}$, and denote Δ_+ the positive roots. There is an orthonormal set (the real root vectors)

$$\{u_\alpha, v_\alpha : \alpha\in\Delta_+\}\subseteq\mathfrak{k}$$

with respect to this inner product, such that for each $h\in\mathfrak{h}$

$$[h,u_\alpha] = -\alpha(h)v_\alpha \qquad [h,v_\alpha] = \alpha(h)u_\alpha \qquad [u_\alpha,v_\alpha] = h_\alpha, \tag{5.1}$$

where $h_\alpha \in \mathfrak{h}$ is the unique element such that $\langle h_\alpha, \cdot\rangle = -\alpha(\cdot)$ (see for instance [Kna02, Chapter 6]). Set

$$Z_\alpha = \mathbb{R}u_\alpha \oplus \mathbb{R}v_\alpha.$$

Then $\mathfrak{k} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Delta_+} Z_\alpha$, and moreover for each $h \in \mathfrak{h}$, we have

$$\operatorname{ad} h = i \sum_{\alpha\in\Delta_+} \alpha(h)(u_\alpha \otimes v_\alpha - v_\alpha \otimes u_\alpha) = i \sum_{\alpha\in\Delta_+} \alpha(h) T_\alpha \tag{5.2}$$

were we write $T_\alpha = (u_\alpha \otimes v_\alpha - v_\alpha \otimes u_\alpha)$ for short. Note that $T_\alpha T_\beta = 0$ when $\alpha \neq \beta$.

Remark 5.3. The Weyl group of K acts transitively on the roots (see the remark in [Serre01, p. 47]). Thus, if σ is a permutation of the roots, there exists $k_\sigma \in K$ such that $\operatorname{Ad}_{k_\sigma} u_\alpha = u_{\sigma\alpha}$ and likewise with v_α. Let $w \in \mathfrak{h}$, and assumme that T_σ is the linear transform in $\mathfrak{k}$ obtained by permuting the eigenvalues of $\operatorname{ad} w$, then

$$\begin{aligned} T_\sigma &= i \sum_{\alpha\in\Delta_+} \sigma\alpha(w)(u_\alpha \otimes v_\alpha - v_\alpha \otimes u_\alpha) = i \sum_{\alpha\in\Delta_+} \alpha(w)(u_{\sigma\alpha} \otimes v_{\sigma\alpha} - v_{\sigma\alpha} \otimes u_{\sigma\alpha}) \\ &= \operatorname{Ad}_{k_\sigma} i \sum_{\alpha\in\Delta_+} \alpha(w)(u_\alpha \otimes v_\alpha - v_\alpha \otimes u_\alpha) \operatorname{Ad}_{k\sigma}^{-1} \\ &= \operatorname{Ad}_{k_\sigma} \operatorname{ad} w \operatorname{Ad}_{k_\sigma}^{-1} = \operatorname{ad}(\operatorname{Ad}_{k_\sigma} w). \end{aligned}$$

Then $T_\sigma = \operatorname{ad}(\operatorname{Ad}_{k_\sigma} w)$. It is clear that T_σ commutes with $\operatorname{ad} w$, and since $\mathfrak{k}$ is semi-simple, $\operatorname{Ad}_{k_\sigma} w$ commutes with w and in particular $w_\sigma = \operatorname{Ad}_{k_\sigma} w \in \mathfrak{h}$.

5.1. **Geodesics of groups with bi-invariant Finsler metrics.** First we recall here some useful facts regarding norming functionals in $\mathfrak{k}$, their proofs are quite elementary and can be found in [Lar19, Section 4].

Remark 5.4 (Gauss' Lemma)**.** Let $\mathfrak{k}$ be the Lie algebra of K equipped with an Ad-invariant norm (or Finsler norm). As mentioned in the introduction, it is relevant to remark here that the norm does not need to be homogeneous, it is only necessary that it is positively homogeneous. For $v \in \mathfrak{k}$ let φ be a norming functional of v or $-v$. Then

(1) We have $\varphi([w,v]) = 0$ for all $w \in \mathfrak{k}$. Equivalently, $\varphi \circ \operatorname{ad} v \equiv 0$ on $\mathfrak{k}$.
(2) $\varphi(e^{\lambda \operatorname{ad} v} w) = \varphi(w)$ for all $w \in \mathfrak{k}$ and $\lambda \in \mathbb{R}$.
(3) $\varphi(e^{-v} D\exp_v w) = \varphi(w)$ for all $w \in \mathfrak{k}$.

When the norm $\|\cdot\|$ is a norm derived from an inner product, the third item above is in fact Gauss' Lemma of Riemannian geometry (it is well-known that the Riemannian exponential map, for a bi-invariant Riemannian metric in a Lie group, coincides with the group exponential). We next show how to describe the boundary of a ball in K. We first recall a definition.

Definition 5.5. Let B be a convex body in a vector space V and let $v \in \operatorname{bd} B$. The *solid tangent cone* at v is

$$TC_v := \bigcap \{H_v^- : H_v \text{ is a supporting hyperplane of } B \text{ at } v\}.$$

If $v \in (V, \|\cdot\|)$ and $\|v\| = r > 0$, then taking $B = B_r(0)$ it is apparent that

$$TC_v := \bigcap \{\varphi^{-1}(-\infty, r] : \varphi \text{ is a norming functional of } v\}.$$

Therefore, by Gauss' Lemma above, for each $0 \neq v \in \mathfrak{k}$, if $w \in TC_v$ then $e^{-v} D\exp_v w \in TC_v$. Moreover, if v is such that the differential of the exponential map is invertible (for instance, if v is smaller that the injectivity radius of K, see Definition 5.11 below), then it is clear that

$$D\exp_v(TC_v) = e^v TC_v.$$

Geometrically, the differential of the exponential map at v acts on the tangent cone at v as the left translation.

Remark 5.6. Any functional φ can be described as $\varphi(v) = \langle v, a\rangle$ for some $a \in \mathfrak{k}$, via the inner product in $\mathfrak{k}$ (Remark 5.1). That is $\varphi = -\operatorname{tr}(\operatorname{ad} a \cdot\ )$. If $\varphi(z) = \|z\|$ for some $z \in \mathfrak{k}$, then $\varphi([v,z]) = 0$ for all $v \in \mathfrak{k}$ (Remark 5.4). Thus for any $v \in \mathfrak{k}$,

$$0 = \varphi([v,z]) = \langle [v,z], a\rangle = \langle v, [z,a]\rangle$$

and then $[z,a] = 0$. Therefore $\operatorname{ad} a$ and $\operatorname{ad} z$ commute, and they are simultaneously diagonalizable. Let $\{e_i\}_{i=1,\dots,N}$ be an orthonormal basis of $\mathfrak{k}^{\mathbb{C}}$ that diagonalizes both simultaneously, let $q_{i,j} = e_i \otimes e_j$ and $q_i = q_{i,i}$ the corresponding rank-one orthogonal projections. If $\operatorname{ad} z = i\sum_{j\in J} z_j q_j$ is the spectral decomposition of the skew-adjoint operator $\operatorname{ad} z$, we can write

$$A = \operatorname{ad} a = i\sum_{j\in J} a_j q_j + iB$$

for certain $a_j \in \mathbb{R}$ and $B^* = B = \sum_{k,l\notin J} B_{kl} q_{kl}$, with $Bq_j = 0$ for all $j \in J$.

Lemma 5.7 (Supporting norming functionals)**.** *If φ norms $z \in \mathfrak{k}$ as in the previous remark, and $\|\varphi\| = 1$, drop the term B and all the $a_j = 0$, and consider $\psi = -\operatorname{tr}(\widetilde{A}\ \ \cdot)$, where $\widetilde{A} = i\sum_j a_j q_j$. Then ψ is still norming for z and has unit norm. If we drop some of the non-zero a_j, its norm shrinks, and moreover $a_j z_j > 0$ for each non-zero z_j.*

Proof. Note first that ψ is still norming for z:

$$\psi(z) = \sum_j a_j z_j \operatorname{tr}(q_j) = \operatorname{tr}(A \circ \operatorname{ad} z) = \varphi(z) = \|z\|.$$

Now for any $x \in \mathfrak{k}$, write $\operatorname{ad} x$ as a block matrix in terms of the q_j and its orthogonal complement,

$$A = \operatorname{ad} a = \begin{pmatrix} ia_j & 0 \\ 0 & B \end{pmatrix} \qquad \widetilde{A} = \begin{pmatrix} ia_j & 0 \\ 0 & 0 \end{pmatrix} \qquad \operatorname{ad} x = \begin{pmatrix} x_{ij} & * \\ * & x_0 \end{pmatrix}$$

where $x_0 q_j = 0$ for all j. Then

$$\begin{aligned} 1 = \|\varphi\| &= \sup_{\|x\|=1} \varphi(x) = \sup_{\|x\|=1} -\operatorname{tr}(\operatorname{ad} a \circ \operatorname{ad} x) = \sup_{\|x\|=1} \sum_j a_j x_{jj} + \operatorname{tr}(Bx_0) \\ &\geq \sup_{\|x\|=1, x_0=0} \sum_j a_j x_{jj} + \operatorname{tr}(Bx_0) = \sup_{\|x\|=1, x_0=0} \sum_j a_j x_{jj} \\ &= \sup_{\|x\|=1} \sum_j a_j x_{jj} = \sup_{\|x\|=1} -\operatorname{tr}(\widetilde{A} \circ \operatorname{ad} x) = \sup_{\|x\|=1} \psi(x) = \|\psi\|. \end{aligned}$$

This proves that $\|\psi\| \leq 1$, but since $\psi(z) = \|z\|$, it must be $\|\psi\| = 1$. Now consider $\widetilde{\psi}$ which is exactly as ψ but omitting a_d with $d \in J$. Then if $x \in \mathfrak{k}$ and $\operatorname{ad} x = \sum_{k,l} x_{kl} q_{kl}$ with

the q_{kl} as in the previous Remark, we have

$$\begin{aligned}\|\widetilde{\psi}\| &= \sup_{\|x\|=1} \widetilde{\psi}(x) = \sup_{\|x\|=1} \sum_{j\in J, j\neq d} a_j x_{jj} = \sup_{\|x\|=1, x_{dd}=0} \sum_{j\in J} a_j x_{jj} \\ &\leq \sup_{\|x\|=1} \sum_{j\in J} a_j x_{jj} = \sup_{\|x\|=1} \psi(x) = \|\psi\| = 1.\end{aligned}$$

Finally, we claimed that since $\psi(z) = \|z\|$ it must be $a_j z_j > 0$ for each non-zero z_j. Otherwise, suppose that $a_d z_d < 0$ then

$$\|z\| = \psi(z) = \sum_{j\in J} a_j z_j < \sum_{j\neq d} a_j z_j = \sum_{j\neq d} \widetilde{a}_j z_j = \widetilde{\psi}(z) \leq \|\widetilde{\psi}\|\|z\| \leq \|z\|,$$

a contradiction. □

We now charachterize the linear order given by Finsler Ad-invariant norms; results in the same vein can be found in [AtBo83, Section 12], [BhaHo89, Proposition 6] and [HoTa99, Proposition 2.8]. For $z, w \in \mathfrak{k}$ we denote $\overrightarrow{w} = (w_1, w_2, \ldots, w_N)$ the string of real numbers such that iw_j are the eigenvalues of $\operatorname{ad} w$ in the complexification $\mathfrak{k}^{\mathbb{C}}$ of $\mathfrak{k}$, and likewise with $\overrightarrow{z}$.

Proposition 5.8 (Majorization and norms)**.** *Let $z, w \in \mathfrak{k}$, let $N = \dim(\mathfrak{k})$. The following are equivalent:*

(1) *$z \in \operatorname{conv} O_w$, more precisely there exist (at most) $N+1$ points $k_i \in K$ and $N+1$ real numbers $\lambda_i \geq 0$ with $\sum_i \lambda_i = 1$ such that*

$$z = \sum_{i=1}^{N+1} \lambda_i \operatorname{Ad}_{k_i} w.$$

(2) *$\overrightarrow{z} \prec \overrightarrow{w}$ (strong majorization).*
(3) *$\|z\| \leq \|w\|$ for all* Ad*-invariant Finsler norms in $\mathfrak{k}$.*
(4) *$\max_{k\in K}\langle z, \operatorname{Ad}_k x\rangle \leq \max_{k\in K}\langle w, \operatorname{Ad}_k x\rangle$ for all $x \in \mathfrak{k}$.*

If moreover equality holds for some Finsler norm, then z and all the $\operatorname{Ad}_{k_i} w$ lie in the same face of the ball for that norm (and in fact lie in the intersection of all the faces that z lies in). If that norm is strictly convex then $z = \operatorname{Ad}_k w$ for some $k \in K$.

Proof. The fact that any element in the orbit can be written with a prescribed number of combinations $(N+1)$ is a consequence of Caratheodory's theorem. Let us establish first the equivalence $(1) \Leftrightarrow (2)$. Assume (1), passing to the adjoint representation we have

$$\operatorname{ad} z = \sum_{i=1}^{N} \lambda_i \operatorname{Ad}_{k_i} \operatorname{ad} w \operatorname{Ad}_{k_i}^{-1}.$$

Each Ad_{k_i} is a unitary operator acting in $\mathfrak{k}^{\mathbb{C}}$ therefore $\operatorname{ad} z$ belongs to the convex hull of the coadjoint orbit of $\operatorname{ad} w$, and this in turn (and by Schur-Horn's theorem) implies that strong majorization $\overrightarrow{z} \prec \overrightarrow{w}$ holds. Now assume (2) holds, let $\mathfrak{h}$ be a Cartan subalgebra containing w, let Δ^+ be the positive simple roots and let $k \in K$ such that $\operatorname{Ad}_k z \in \mathfrak{h}$. The spectrum of

$$\operatorname{ad}(\operatorname{Ad}_k z) = \operatorname{Ad}_k \operatorname{ad} z \operatorname{Ad}_k^{-1}$$

is also the string $\overrightarrow{z}$, and the assumption implies $\overrightarrow{z}$ is in the convex hull of the permutations of the eigenvalues of $\operatorname{ad} w$ (see [Bhatia97, Theorem II.1.10]). Equivalently (and again invoking Caratheodory's theorem), there are $N+1$ such elements with

$$\operatorname{ad}(\operatorname{Ad}_k z) = \sum_{i=1}^{N+1} \lambda_i T_{\sigma_i}$$

for a certain string of non-negative numbers $(\lambda_i)_{i=1,\dots,N+1}$ such that $\sum_i \lambda_i = 1$. Now each T_{σ_i} is obtained permuting the eigenvalues of $\operatorname{ad} w$ or equivalently, permuting the roots $\alpha \in \Delta_+$. By Remark 5.3, any such permutation is obtained by an inner automorphism, therefore

$$\operatorname{ad}(\operatorname{Ad}_k z) = \sum_{i=1}^{N+1} \lambda_i \operatorname{ad}(\operatorname{Ad}_{h_i} w) = \operatorname{ad}(\sum_{i=1}^{N+1} \lambda_i \operatorname{Ad}_{h_i} w),$$

and by the semi-simplicity of $\mathfrak{k}$, we obtain $z = \sum_{i=1}^{N+1} \lambda_i \operatorname{Ad}_{k_i} w \in \operatorname{conv} O_w$, where $k_i = k^{-1} h_i$.

Clearly (1) $\Rightarrow$ (3), and (3) $\Rightarrow$ (4) when $x \in \mathfrak{k}$ is a regular element (so we obtain a non-degenerate Finsler norm, see Remarks 3.16 and 5.10). Since regular elements are dense in $\mathfrak{k}$, it is straightforward to see that (4) must then hold for any $x \in \mathfrak{k}$, if (3) holds. Now assume that (1) does not hold, then by Hahn-Banach's theorem there exists a linear functional φ separating z and the convex capsule of the orbit, i.e.

$$\varphi(y) \leq r < \varphi(z) \qquad \forall y \in \operatorname{conv}(O_w).$$

Let $x \in \mathfrak{k}$ such that $\varphi = \langle x, \cdot \rangle$, then by Remark 3.16

$$\max_{k \in K} \langle w, \operatorname{Ad}_k x \rangle = \max_{k \in K} \langle \operatorname{Ad}_k w, x \rangle = \max\{\langle y, x \rangle : y \in \operatorname{conv}(O_w)\} \leq r < \langle z, x \rangle.$$

This shows that (4) cannot hold, finishing the proof of the equivalences. Now assumme that any of the conditions hold, and we have equality of norms $\|z\| = \|w\|$ holds for some norm, then

$$\|w\| = \|z\| = \|\sum \lambda_i \operatorname{Ad}_{k_i} w\| \leq \sum \lambda_i \|\operatorname{Ad}_{k_i} w\| = \sum \lambda_i \|w\| \leq \|w\|.$$

We note that there is a common norming functional φ for all the $\lambda_i \operatorname{Ad}_{k_i} w$ by Lemma 2.6; since $\lambda_i \geq 0$ this amounts for the $\operatorname{Ad}_{k_i} w$ in the same face of a sphere of the norm. Then also

$$\varphi(z) = \varphi(\sum_i \lambda_i \operatorname{Ad}_{k_i} w) = \sum \lambda_i \varphi(\operatorname{Ad}_{k_i} w) = \sum \lambda_i \|w\| = \|w\| = \|z\|.$$

If φ is any functional norming z, then

$$\|z\| = \varphi(z) = \sum \lambda_i \varphi(\operatorname{Ad}_{k_i} w) \leq \sum \lambda_i \|\operatorname{Ad}_{k_i} w\| = \|w\| = \|z\|$$

shows that it must be $\varphi(\operatorname{Ad}_{k_i} w) = \|\operatorname{Ad}_{k_i} w\|$ for all i, thus these vectors are in fact in the intersection of all the faces where z lives. If the norm is strictly convex all the $\operatorname{Ad}_{k_i} w$ are aligned, but being normed by the same functional they must be equal therefore $z = \operatorname{Ad}_k w$ as claimed. □

Remark 5.9. It suffices to check condition (3) above only for *strictly convex norms* to obtain the equivalences. This is because any Ad-invariant Finsler norm $\|\cdot\|$ can be approximated explicitly with a strictly convex (Ad-invariant, Finsler) norm by means of

$$\|x\|_\varepsilon = \|x\| + \varepsilon \|x\|_2.$$

Here $\|x\|_2^2 = -\operatorname{tr}(\operatorname{ad} x \circ \operatorname{ad} x)$ is the norm derived from the Killing form of $\mathfrak{k}$. Likewise, it suffices to check (4) for regular $x \in \mathfrak{k}$.

Remark 5.10 (One-sided norms)**.** Since K is compact it is unimodular and then $0 = \sum_j w_j = i\operatorname{tr}(\operatorname{ad} w)$ (Remark 5.1). This easily implies that all the partial sums $\sum_{k=1}^m w_k$, with the w_k rearranged in decreasing order, must be non-negative. Thus the vector $\overrightarrow{w}$ strongly majorizes the zero vector in $\mathbb{R}^N$, i.e. $\overrightarrow{0} \prec \overrightarrow{w}$, and by the previous proposition, $0 \in \operatorname{conv}(O_w)$ for any $w \in \mathfrak{k}$. In particular the one-sided Hofer norms (Remark 3.16) are in fact-nonnegative, regardless their degeneracy or non-degeneracy. Assumming that the orbit O_w is full, then 0 must be an interior point of $\operatorname{conv}(O_w)$, and then we obtain a true Finlser norm. This can be seen using the argument in [BiGhiHe14, Lemma 6]: if 0 is in the boundary of $\operatorname{conv}(O_w)$, by Hahn-Banach's separation theorem, there exists $0 \neq x \in \mathfrak{k}$ such that $\varphi_x(0) = 0$ and $\varphi_x(\operatorname{conv}(O_w)) \geq 0$. But then it must be $\varphi_x(O_w) = 0$ because otherwise

$$\int_{k\in K} \varphi_x(\operatorname{Ad}_k(w))dk = \varphi_x(\int_{k\in K} \operatorname{Ad}_k(w)dk) > 0$$

contradicting that $\int_{k\in K} \operatorname{Ad}_k(w)dk$ is a fixed point of the adjoint action, therefore it is 0 because $\mathfrak{k}$ is semi-simple. Since $\varphi_x(O_w) = 0$, the orbit is not full.

5.1.1. *Domain of injectivity of the exponential.* Since K is a finite-dimensional Lie group, the exponential map of K is a local diffeomorphism for some open ball of the norm $\|\cdot\|$. More precisely, let $\mathbf{D} \subseteq \mathfrak{k}$ be a maximal open convex Ad-invariant set, such that $V = \exp(\mathbf{D})$ is open in $\mathfrak{k}$ and $\exp : \mathbf{D} \to V$ is a diffeomorphism. It will be convenient to denote $\|\cdot\|_\infty$ to the Minkowski gauge of the set $\mathbf{D}$; it is an Ad-invariant Finsler norm in $\mathfrak{k}$ and we will refer to it as the *uniform norm*. Then we also define

Definition 5.11. Let K be a Lie group, $\|\cdot\|$ an Ad-invariant Finsler norm on $\mathfrak{k}$, and for $R > 0$ let $B_R = \{v \in \mathfrak{k} : \|v\| < R\}$, $V_R = \exp(B_R)$. If $\exp : B_R \to V_R$ is a diffeomorphism between open sets and R is maximal, we call R the *radius of injectivity* for the given norm.

Note that if $B_R \subseteq \mathbf{D}$ then $\exp : B_R \to V_R$ is a diffeomorphism between open sets, thus one looks for balls of the given norm that fit inside the domain of injectivity of the exponential map.

Remark 5.12. The condition $B_R \subseteq \mathbf{D}$ is equivalent to $B_R \cap \mathfrak{h} \subseteq \mathbf{D} \cap \mathfrak{h}$. In the case of Hofer norms $B_R \cap \mathfrak{h} = R\{y_1, \ldots, y_m\}^\circ$, where $\{y_1, \ldots, y_m\}$ are the extreme points of the Hofer norm polytope. Note also that $\mathbf{D} \cap \mathfrak{h}$ can be taken as the interior of $\cup_{w\in\mathcal{W}} w.C$, where C is a Weyl alcove.

Example 5.13. For the group $\mathrm{SU}(n)$ we take $\mathbf{D} = \{z \in \mathfrak{su}_n : \|z\|_\infty < \pi\}$ where $\|\cdot\|_\infty$ now is exactly the usual spectral norm. Equivalently, $\mathbf{D} \cap \mathfrak{h}$ is

$$\{\operatorname{diag}(x_1, \ldots, x_n) \in \mathfrak{h} : |x_i| < \pi, \text{ for } i = 1, \ldots, n\}.$$

For the group $\mathrm{SU}(n)/\mathbb{Z}_n$ we take

$$\mathbf{D} \cap \mathfrak{h} = \{\operatorname{diag}(x_1, \ldots, x_n) \in \mathfrak{h} : |x_i| < \pi/n, \text{ for } i = 1, \ldots, n\}.$$

We now show that the convex body $\mathbf{D}$ (which depends only on K) is optimal in terms of lengths of segments (one-parameter subgroups), for any bi-invariant distance:

Theorem 5.14 (Exponential map and Finsler norms)**.** *Let $z, w \in \mathfrak{k}$ such that $e^z = e^w$. Assumme that $z \in \mathbf{D}/2$, Then w commutes with z and $\|z\| \leq \|w\|$ for any* Ad*-invariant Finsler norm $\|\cdot\|$ in $\mathfrak{k}$. If equality of norms holds for a strictly convex norm, then $w = z$.*

Proof. Assumme first that w is regular, let $\mathfrak{z}(w) = \mathfrak{h}$ denote the Cartan subalgebra. Note that

$$\exp(e^{t\,\mathrm{ad}\,w} z) = \exp(\mathrm{Ad}_{e^{tw}} z) = e^{tw} e^z e^{-tw} = e^{tw} e^w e^{-tw} = e^w.$$

Therefore, differentiating at $t = 0$ we obtain $D\exp_z([w,z]) = 0$, and since $z \in \mathbf{D}$, we conclude that $[w,z] = 0$, and $z \in \mathfrak{h}$. Since $e^{\mathrm{ad}\,w} = e^{\mathrm{ad}\,z}$, then $\exp(\mathrm{ad}(w-z)) = \exp(\mathrm{ad}\,w - \mathrm{ad}\,z) = 1$, implying that $\sigma(\mathrm{ad}(w-z)) \subseteq 2\pi i\mathbb{Z}$. This implies, using equation (5.2), that we can write

$$i \sum_{\alpha\in\Delta_+} w_\alpha T_\alpha = \mathrm{ad}\,w = i \sum_{\alpha\in\Delta_+} (z_\alpha + 2\pi n_\alpha) T_\alpha$$

wiht $w_\alpha, z_\alpha \in \mathbb{R}$ and $n_\alpha \in \mathbb{Z}$. It will be convenient to number the roots, so we let $J = card(\Delta_+)$ and we have $w_j = z_j + 2\pi n_j$ for all $j \in J$, where some of the z_j might be zero. Note that since $z \in \mathbf{D}/2$, then $2z \in \mathbf{D}$ therefore the exponential map is injective and a diffeomorphism in $t2z$ for $t \in [-1,1]$, and in particular it must be $\sigma(\mathrm{ad}\,2z) \subseteq (-2\pi i, 2\pi i)$ (by inspection of the formula of the differential of the exponential map, see Remark 5.4 below). Thus we have $|2z_j| < 2\pi$ for all $j \in J$ or equivalently $-\pi < z_j < \pi$. We can asumme that the z_j are given in order $z_1 \geq z_2 \geq \cdots \geq z_J$ (recall also $\sum z_j = 0$). Let us reorder the w_j in decreasing order also, so there is a permutation σ of $\{1,\ldots,J\}$ such that if $v_j = w_{\sigma j}$ then $v_1 \geq v_2 \geq \cdots \geq v_J$ (and we also have $\sum v_j = 0$). From here it is also clear that $\sum n_j = 0$. Let's spare for a moment those j such that $n_j = 0$, and for the others, note that if $n_j > 0$ and $n_r < 0$ then

$$z_j + 2\pi n_j > -\pi + 2\pi = \pi > -\pi = -2\pi + \pi > 2\pi n_r + z_r.$$

This shows that the $v_j = z_{\sigma j} + 2\pi n_{\sigma j}$, with positive $n_{\sigma j}$, are always bigger than those with negative $n_{\sigma j}$. We split the indices in two sets: let j_0 be such that if $j \in 1,\ldots,j_0$ then $n_{\sigma j} > 0$ and otherwise $n_{\sigma j} < 0$ when $j_0 + 1 \leq j \leq J$. We compute the sum of the z_j up to any such $1 \leq j \leq j_0$, we have

$$\sum_{k=1}^{j} z_k \leq \sum_{k=1}^{j} \pi \leq j\pi.$$

On the other hand, it is clear that the sum of the first j bigger v_k, for $k \leq j$, must be of those v_k with $n_{\sigma k} > 0$, therefore

$$\sum_{k=1}^{j} v_j = \sum_{k=1}^{j} z_{\sigma k} + 2\pi n_{\sigma k} \geq -j\pi + 2\pi j = j\pi.$$

Thus if $j \leq j_0$,

$$z_1 + z_2 + \cdots + z_j \leq v_1 + v_2 + \cdots + v_j. \tag{5.3}$$

Now assumme that $j \geq j_0 + 1$, and note that

$$\sum_{k=1}^{j} z_k = -\sum_{k=j+1}^{N} z_k < \pi(N - (j+1)).$$

Likewise

$$\sum_{k=1}^{j} v_k = -\sum_{k=j+1}^{N} v_k = -\sum_{k=j+1}^{N} (z_{\sigma k} + 2\pi n_{\sigma k})$$

and note that now all the $n_{\sigma k} \leq -1$, since $k \geq j+1 \geq j_0+1 > j_0$. Therefore

$$\sum_{k=1}^{j} v_k \geq -\pi(N-(j+1)) + 2\pi(N-(j+1)) = \pi(N-(j+1)),$$

and equation (5.3) is also valid for $j \geq j_0$. Let $\overrightarrow{v} = (v_1, v_2, \ldots, v_N)$ and likewise $\overrightarrow{z} = (z_1, z_2, \ldots, z_N)$. Then equation (5.3) together with $\sum z_j = \sum v_j = 0$ tells us that $\overrightarrow{z} \prec \overrightarrow{v}$, that is $\overrightarrow{v}$ majorizes $\overrightarrow{z}$. Then by [Bhatia97, Theorem II.1.10], $\overrightarrow{z}$ is in the convex hull of all vectors obtained by permutating the coordinates of $\overrightarrow{v}$. Clearly, we can add those w_j such that $w_j = z_j$ ($n_j = 0$) and this still holds true. Since the v_j are just a permutation of the w_j, then $\overrightarrow{z}$ is in fact in the convex hull of all vectors obtained permutating the coordinates of $\overrightarrow{w} = (w_1, w_2, \ldots, w_J)$. By Remark 5.3, we have

$$\operatorname{ad} z = \sum_{\sigma} \lambda_\sigma \operatorname{ad}(\operatorname{Ad}_{k_\sigma} w) = \operatorname{ad}(\sum_{\sigma} \lambda_\sigma \operatorname{Ad}_{k_\sigma} w),$$

and since $\mathfrak{k}$ is semi-simple, it must be $z = \sum \lambda_\sigma \operatorname{Ad}_{k_\sigma} w$. Then

$$\|z\| \leq \sum \lambda_\sigma \|\operatorname{Ad}_{k_\sigma} w\| = \sum \lambda_\sigma \|w\| = \|w\|$$

proving the claim for regular w. If w is not regular, fix the bi-invariant distance in K given by the uniform norm (the Minkowski Finsler norm of the convex set $\mathbf{D}$). For each $\varepsilon > 0$ pick $w_\varepsilon \in \mathfrak{k}$ such that $\|w_\varepsilon\|_\infty < \varepsilon$ and $w + w_\varepsilon$ is regular (regular elements are dense). Observe that

$$\operatorname{dist}(e^w, e^{w+w_\varepsilon}) \leq \|w_\varepsilon\|_\infty < \varepsilon$$

by Theorem 5.18. Therefore there exists $y_\varepsilon \in \mathfrak{k}$ such that $e^w e^{y_\varepsilon} = e^{w+w_\varepsilon}$. Again, note that

$$\|y_\varepsilon\|_\infty = \operatorname{dist}(1, e^{y_\varepsilon}) = \operatorname{dist}(1, e^{-w} e^{w+w_\varepsilon}) = \operatorname{dist}(e^w, e^{w+w_\varepsilon}) \leq \|w\|_\infty < \varepsilon.$$

Now consider the map $f : v \mapsto e^{z+v}$. Since $f(0) = e^z$ and $Df_0 = D\exp_z$, the hypotesis $z \in \mathbf{D}/2$ guarantees that f is a local diffeomorphism from a 0-neighbourhood to a neighbourhood of e^z. Therefore there exists a unique $z_\varepsilon \in \mathfrak{k}$ in that neighbourhood, such that $e^{z+z_\varepsilon} = e^z e^{y_\varepsilon}$. Note also that when $y_\varepsilon \to 0$, then also $z_\varepsilon \to 0$. In particular, for small $\varepsilon > 0$, $z + z_\varepsilon \in \mathbf{D}/2$ just like z. Then from

$$e^{z+z_\varepsilon} = e^z e^{y_\varepsilon} = e^w e^{y_\varepsilon} = e^{w+w_\varepsilon}.$$

and the previous proof, we can conclude that $\|z + z_\varepsilon\| \leq \|w + w_\varepsilon\|$ for any Ad-invariant Finsler norm in $\mathfrak{k}$. Letting $\varepsilon \to 0$ gives us the desired inequality $\|z\| \leq \|w\|$.

Now assumme that there is an equality of norms for a strictly convex norm, then by Proposition 5.8, $z = \operatorname{Ad}_k w$ for some $k \in K$, and in particular $w \in \mathbf{D}/2$ also. But $e^w = e^z$ and the injectivity of the exponential map implies $z = w$. □

Remark 5.15. If z, w are as in the previous theorem, then by Proposition 5.8, we have

$$z = \sum_{i=1}^{N+1} \lambda_i \operatorname{Ad}_{k_i} w,$$

for some $k_i \in K$, $\lambda_i \geq 0$ with $\sum_i \lambda_i = 1$. We also mention here that the proof of the previous theorem shows that when w is regular, the k_i are in the Weyl group of K.

For linear Lie groups such as $K = \mathrm{SU}(n)$, the passage to the adjoint representation is unnecessary. Thus from $e^z = e^w$ with $z \in \mathbf{D}$ we can conclude that $[w,z] = 0$ and therefore

$$w = z + 2\pi i \sum_j n_j q_j.$$

With the same proof as the previous theorem, we now obtain the same result for $z \in \mathbf{D}$, i.e. $\|z\|_\infty < \pi$ (and not just $z \in \mathbf{D}/2$):

Corollary 5.16 (Linear groups)**.** *Let $K \subseteq M_n(\mathbb{C})$ be compact semi-simple linear Lie group. Let $z, w \in \mathfrak{k}$ such that $e^z = e^w$, and assumme that $z \in \mathbf{D}$. Then $\|z\| \leq \|w\|$ for any* Ad*-invariant Finsler norm $\|\cdot\|$ in $\mathfrak{k}$, and if equality of norms holds for a strictly convex norm, then $w = z$.*

5.1.2. *Local Hopf-Rinow theorem and the characterization of geodesics.* Let us take a look at geodesics of a Lie group with a bi-invariant Finsler metric. We recall here the fundamental results about geodesics, for proofs see [Lar19, Section 4]. For a given Finsler norm, let $R > 0$ be an injectivity radius. The results are stated in terms of the left logarithmic derivatives $\gamma_t^{-1}\dot{\gamma}_t$ and hold also if stated in terms of the right logarithmic derivative since the norm is Ad-invariant.

Definition 5.17 (Short paths)**.** We call a curve $\gamma : [0,1] \to K$ *short* or we say that γ is a *geodesic* if it minimizes the length functional

$$L(\gamma) = \mathrm{length}(\gamma) = \int_a^b \|\dot{\gamma}_t\|_{\gamma_t} dt = \int_a^b \|\dot{\gamma}_t \gamma_t^{-1}\| dt = \int_a^b \|\gamma_t^{-1}\dot{\gamma}_t\| dt$$

among all curves in K with the same endpoints.

Theorem 5.18 (Geodesics)**.** *Let $u_0, u_1 = u_0 e^z \in K$ with $\|z\| < R$.*

(1) *If $\delta(t) = u_0 e^{tz}$, $t \in [0,1]$, then δ is shorter than any other piecewise C^1 path γ in K joining u_0, u_1 and* $\mathrm{dist}(u_0, u_1) = \|z\|$.

(2) *If $v, w \in \mathfrak{k}$ then*

$$\mathrm{dist}(e^v, e^w) \leq \|w - v\|$$

and if w, v commute and $\|w - v\| \leq R$, then equality holds (this is known as the exponential metric decreasing property).

(3) *Let $\Gamma : [a,b] \to \mathfrak{k}$ be a piecewise C^1 short path joining $0, z$, let $\gamma = e^\Gamma$. Then $\|\gamma_t^{-1}\dot{\gamma}_t\| = \|\dot{\Gamma}_t\|$ for all t, and γ is short in K with the same length than Γ. Moreover if φ is norming functional for z, then*

$$\varphi(\gamma_t^{-1}\dot{\gamma}_t) = \|\gamma_t^{-1}\dot{\gamma}_t\| = \|\dot{\Gamma}_t\| = \varphi(\dot{\Gamma}_t) \quad \forall t \in [a,b],$$

thus $\gamma^{-1}\dot{\gamma}$ (normalized) sits inside a face of the unit sphere of the norm.

(4) *$\gamma : [a,b] \to K$ is a piecewise C^1 short path joining $1, e^z$ in K if and only if $\gamma = e^\Gamma$ for a piecewise C^1 path $\Gamma : [a,b] \to \mathfrak{k}$ joining $0, z$ (with $\|\Gamma_t\| \leq R$) and*

$$\varphi(\dot{\Gamma}_t) = \|\gamma_t^{-1}\dot{\gamma}_t\| = \varphi(\gamma_t^{-1}\dot{\gamma}_t)$$

for some unit norm functional φ *and all* $t \in [a,b]$ *(and then this holds for any norming functional of* z*).*

(5) *If* $z/\|z\|$ *is an extremal point of the unit sphere of* $\mathfrak{k}$*, then the only short piecewise* C^1 *path joining* $1, e^z$ *in* K *is (a reparametrization of) the segment* $\delta(t) = e^{tz}$*.*

These results were established in [Lar19] with some generality; for finite dimensional groups we can improve the existence of short paths invoking the metric version of Hopf-Rinow's theorem. As usual, here K denotes a connected compact Lie group with semi-simple Lie algebra $\mathfrak{k}$. The last item of this theorem extends significanly (to this family of Lie groups) the results obtained in [AnLVa14] for the group $U(n)$.

Theorem 5.19. *Let* dist *be a bi-invariant metric in* K *(i.e. from an* Ad*-invariant Finsler norm* $\|\cdot\|$ *in* $\mathfrak{k}$*). Then*

(1) *For each* $u_1, u_2 \in K$ *there exists a short polygonal path* δ *joining them, i.e. a concatenation of segments* $t \mapsto u_i e^{tz_i}$ *such that* $\|z_i\| < R$ *and*

$$L(\delta) = \sum \|z_i\| = \operatorname{dist}(u_1, u_2).$$

(2) *If the norm is strictly convex, there exists* $w \in \mathfrak{k}$ *such that* $\|w\| = \operatorname{dist}(u_1, u_2)$ *and the segment* $t \mapsto u_1 e^{tw}$ *is a short path joining them. Any short path is a reparametrization of a segment, and if* $\operatorname{dist}(u_1, u_2) < R$*, there is exactly one short segment joining them.*

(3) *If* $z \in \mathbf{D}/2$ *(*$z \in \mathbf{D}$ *for linear Lie groups), then* $\operatorname{dist}(1, e^z) = \|z\|$*. If the norm is strictly convex,* $t \mapsto e^{tz}$ *is the unique short path joining them.*

Proof. Let dist_g denote the bi-invariant distance induced by the Ad-invariant metric given by the Killing form in $\mathfrak{k}$. It is well-known that such dist_g metric has one-parameter groups as Riemannian geodesics, therefore it is geodesically complete. By Hopf-Rinow's theorem $(K, \operatorname{dist}_g)$ is metrically complete, and therefore (K, dist) is metrically complete since both metrics are uniformly equivalent (since both are bi-invariant and the tangent norms are uniformly equivalent, $\mathfrak{k}$ being finite dimensional). Since any of these metrics induce the original topology of K, and since K is a finite dimensional manifold, it is locally compact and we can apply Cohn-Vossen's theorem [BuBuIv01, Theorem 2.5.28] to the metric space (K, dist). This theorem tells us that any approximating sequence of paths in K has a limit point $\gamma\colon [0,1] \to K$ such that γ is rectifiable and $L(\gamma) = \operatorname{dist}(u_1, u_2)$. Partition γ into finite pieces, in points denoted $\gamma_i = \gamma_{t_i}$, such that $\operatorname{dist}(\gamma_i, \gamma_{i+1}) < R$. Write $\gamma_{i+1} = \gamma_i e^{z_i}$ using Theorem 5.18(1), hence if δ is the concatenation of these paths,

$$L(\delta) = \sum \|z_i\| = \sum \operatorname{dist}(\gamma_i, \gamma_{i+1}) = \operatorname{dist}(u_1, u_2)$$

which shows that δ is minimizing. If the norm is strictly convex, it can be shown that the z_i commute, hence we can replace the concatenation of segments with a segment; the proof of this and the local uniqueness can be found in [Lar19, Theorem 4.15]. Now assumme that z is in (half of) the domain of injectivity of the exponential map. Assumme first that the norm is strictly convex. By the previous item of this theorem, there exists $w \in \mathfrak{k}$ such that $t \mapsto e^{tw}$ joins $1, e^z$ and such that $\|w\| = \operatorname{dist}(1, e^z) \leq \|z\|$. But Theorem 5.14 also tells us that $\|z\| \leq \|w\|$, therefore $\|z\| = \|w\| = \operatorname{dist}(1, e^z)$. Now let $\|\cdot\|$ be any

Ad-invariant norm, let $\varepsilon > 0$ and let $\|v\|_g = \sqrt{\langle v, v\rangle}$ be an Ad-invariant Riemannian metric in $\mathfrak{k}$. Consider

$$|v|_\varepsilon = \|v\| + \varepsilon\|v\|_g,$$

and note that $|\cdot|_\varepsilon$ is Ad-invariant and strictly convex. Therefore by what we just proved, if $z \in \mathbf{D}/2$ then

$$\|z\| \leq |z|_\varepsilon = \mathrm{dist}_\varepsilon(1, e^z) \leq L_\varepsilon(\gamma) = L(\gamma) + \varepsilon L_g(\gamma)$$

for any piecewise smooth path γ joining $1, e^z$ in K. Letting $\varepsilon \to 0^+$ first, and taking the infimum over the paths γ, shows that $\|z\| \leq \mathrm{dist}(1, e^z)$ as claimed. If the norm is strictly convex, any other short path is also a segment $t \mapsto e^{tw}$ by the second item of this theorem. Thus $e^w = e^z$ and $\|w\| = \|z\|$, and by Theorem 5.14, we conclude that $w = z$. For linear Lie groups we can replace half of $\mathbf{D}$ with the full set $\mathbf{D}$ by Corollary 5.16. □

From the last assertion of the theorem we can give a nice characterization of the product of exponentials. This is connected with the noncommutative Horn inequalities as studied by Belkale et al, see [Bel08, Entov01] and the references therein.

Corollary 5.20 (Product of exponentials)**.** *Let $x, y, z \in \mathfrak{k}$ with $z \in \overline{\mathbf{D}}/2$ such that $e^x e^y = e^z$. Then*

(1) $\|z\| \leq \|x+y\|$ *for any* Ad*-invariant Finsler norm in* $\mathfrak{k}$.
(2) *Let $N = \dim(\mathfrak{k})$, then there exist (at most) $N+1$ points $k_i \in K$ and $N+1$ real numbers $\lambda_i \geq 0$ with $\sum_i \lambda_i = 1$ such that*

$$z = \sum_{i=1}^{N+1} \lambda_i \,\mathrm{Ad}_{k_i}(x+y).$$

(3) *If equality holds for some Finsler norm, then z and all the $\mathrm{Ad}_{k_i}(x+y)$ lie in the same face of the ball for that norm (in fact, in the intersection of all the faces where z sits).*
(4) *If equality holds for a strictly convex norm then x, y commute, $z = \mathrm{Ad}_k(x+y)$ for some $k \in K$ (thus $x+y \in \overline{\mathbf{D}}/2$) and k commutes with e^z. If moreover $z \in \mathbf{D}/2$ then x, y commute and $z = x+y$.*

Proof. Let $\beta(t) = e^{tx}e^{ty}$, which joins $1, e^z$ in K. Note that $\beta_t^{-1}\dot{\beta}_t = e^{-t\,\mathrm{ad}\,y}(x+y)$ therefore $L(\beta) = \|x+y\|$ for any Ad-invariant Finsler norm in $\mathfrak{k}$. By the previous theorem, the third assertion also holds for the closure of $\mathbf{D}/2$ therefore we must have

$$\|z\| \leq L(\beta) = \|x+y\|$$

for any such Finsler norm. Assertions 2 and 3 follow from Proposition 5.8. If the norm is strictly convex then $z = \mathrm{Ad}_k(x+y)$ by Proposition 5.8, but also x, y commute by [Lar19, Theorem 4.17]. This implies $e^z = e^{\mathrm{Ad}_k(x+y)} = ke^{x+y}k^{-1} = ke^x e^y k^{-1} = ke^z k^{-1}$ thus k commutes with e^z. When $z \in \mathbf{D}/2$, the condition $e^{\mathrm{Ad}_k z} = e^z$ is only possible if $\mathrm{Ad}_k z = z = \mathrm{Ad}_k(x+y)$ therefore $z = x+y$. □

Remark 5.21. For linear Lie groups such as $K = \mathrm{SU}(n)$, and by Corollary 5.16, the same results stated in the previous corollary hold for $z \in \overline{\mathbf{D}}$, i.e. $\|z\|_\infty \leq \pi$ (and not just $z \in \mathbf{D}/2$).

5.2. **Characterization of all short paths.** Before we proceed with the characterization of other short paths (for the case of non-strictly convex norms), we recall two results concerning the exponential map and the adjoint representation of the group K and its differentials:

Remark 5.22. If g_t is a smooth path in K,

$$\frac{d}{dt}\operatorname{Ad}_{g_t} v = \operatorname{Ad}_{g_t}[g_t^{-1}\dot{g}_t, v] \qquad \forall v \in \mathfrak{k}. \tag{5.4}$$

This follows by noting that to compute the derivative we can take $g_t = g_0 e^{t g_0^{-1}\dot{g}_0}$, and differentiate at $t = 0$, thus

$$\operatorname{Ad}_{g_t} v = \operatorname{Ad}_{g_0}\operatorname{Ad}_{e^{t g_0^{-1}\dot{g}_0}} v = \operatorname{Ad}_{g_0} e^{t \operatorname{ad} g_0^{-1}\dot{g}_0} v = \operatorname{Ad}_{g_0} v + t\operatorname{Ad}_{g_0}\operatorname{ad}(g_0^{-1}\dot{g}_0)v + o(t^2).$$

On the other hand if $v, w \in \mathfrak{k}$ then it is well-known that

$$e^{-v} D\exp_v w = \int_0^1 e^{-\lambda \operatorname{ad} v} w\, d\lambda = F(\operatorname{ad} v)w \tag{5.5}$$

where $F(\lambda) = \frac{1-e^{-\lambda}}{\lambda}$ is extended by $F(0) = 1$ to be an holomorphic function in $\mathbb{C}$. In particular, note that if $v \in \mathbf{D}$ (the domain of injectivity of exp), then $F(\operatorname{ad} v)$ must be nonsingular and in particular $\pm 2\pi i \notin \sigma(\operatorname{ad} v)$; otherwise we would have $0 \in \sigma(F(\operatorname{ad} v))$. Moreover, it must be $\sigma(\operatorname{ad} v) \subseteq i(-2\pi, 2\pi)$, otherwise replacing v with tv for some $t \in (0,1)$ we would obtain a contradiction (recall that $\mathbf{D}$ is convex).

With these tools at hand we now characterize short paths γ without the restriction of having length less that the radius of inyectivity of the group. In the next theorem K is a connected compact semi-simple Lie group with the metric induced by an Ad-invariant Finsler norm in $\mathfrak{k}$.

Theorem 5.23 (Characterization of geodesics)**.** *Let $\gamma\colon [a,b] \to K$ be a piecewise C^1 path in K. If γ is short for the bi-invariant metric, then for (almost) all t*

$$\varphi(\gamma_t^{-1}\dot{\gamma}_t) = \|\gamma_t^{-1}\dot{\gamma}_t\| \tag{5.6}$$

for some unit norm functional φ. Reciprocally, if the equality holds for some φ and (almost) all $t \in [t_0, t_1]$, and $L(\gamma)_{t_0}^{t_1} \le R$, then γ is short in $[t_0, t_1] \subseteq [a,b]$.

Proof. By the invariance of the metric, it suffices in all cases to consider paths starting at $u = 1$. If γ is short, assume first that its length is smaller than R, then for any $t \in [a,b]$ we have $\operatorname{dist}(\gamma(a), \gamma(t)) = L(\gamma)_a^t < R$. Therefore we can lift $\gamma_t = e^{\Gamma_t}$ for a rectifiable path $\Gamma : [a,b] \to \mathfrak{k}$, which does not leave the ball B_R. Then by Theorem 5.18, for any norming functional of z and any t we have

$$\varphi(\Gamma_t) = \int_a^t \varphi(\dot{\Gamma}) = \int_a^t \|\gamma^{-1}\dot{\gamma}\| = L(\gamma)_a^t = \operatorname{dist}(\gamma_a, \gamma_a e^{\Gamma_t}) = \|\Gamma_t\|.$$

Therefore again by the previous theorem and by Gauss' Lemma 5.4,

$$\begin{aligned}\|z\| = \varphi(z) &= \int_a^b \varphi(\dot{\Gamma}_s)ds = \int_a^b \varphi(e^{-\Gamma_s} D\exp_{\Gamma_s}\dot{\Gamma}_s)ds = \int_a^b \varphi(\gamma_s^{-1}\dot{\gamma}_s)ds \\ &\le \int_a^b \|\gamma_s^{-1}\dot{\gamma}_s\|ds = L(\gamma)_a^b = \operatorname{dist}(1, e^z) = \|z\|,\end{aligned}$$

and this is only possible if $\varphi(\gamma_t^{-1}\dot\gamma_t) = \|\gamma_t^{-1}\dot\gamma_t\|$ for (almost) all t. If γ is short but its longer than R, we can still partition γ in pieces of length smaller than R and its still short on each piece, in any of these intervals $[a_i, a_{i+1}]$ $(i = 1, \dots, k)$. There we have $\operatorname{dist}(\gamma(a_i), \gamma(t)) \leq L(\gamma)_{a_i}^t < R$. Write $\gamma_{a_{i+1}} = \gamma_{a_i} e^{z_i}$, thus we can lift $\gamma_t = \gamma_{a_i} e^{\Gamma_t^i}$ for a rectifiable $\Gamma^i : [a_i, a_{i+1}] \to \mathfrak{k}$, which does not leave the ball B_R. Using the translation invariance of the metric (from γ_{a_i} to 1), and repeating the argument above, we have for each i

$$\varphi_i(\gamma_t^{-1}\dot\gamma_t) = \|\gamma_t^{-1}\dot\gamma_t\| \qquad \text{if } \varphi_i \text{ norms } z_i \text{ and } \|\varphi_i\| = 1, \tag{5.7}$$

for (almost) all $t \in [a_i, a_{i+1}]$. Now let $\beta(s) = e^{sz_1} e^{sz_2} \dots e^{sz_k}$, then $\beta(0) = 1 = \gamma(0)$, $\beta(1) = \gamma(1)$ and

$$\operatorname{dist}(1, \gamma_1) \leq L(\beta) \leq \sum_i \|z_i\| = \sum_i \operatorname{dist}(\gamma_i, \gamma_{i+1}) = \sum_i L(\gamma|_{[a_i, b_i]}) = L(\gamma)_a^b = \operatorname{dist}(1, \gamma_1),$$

therefore β is short also. Since β is smooth, there exists $c \in (0,1)$ such that

$$\sum_i \|z_i\| = \operatorname{dist}(1, \gamma_1) = L(\beta) = \int_0^1 \|\beta_s^{-1}\dot\beta_s\| ds = \|\beta_c^{-1}\dot\beta_c\|.$$

Now let $w_0 = z_k$, $w_1 = e^{-c \operatorname{ad} z_k} z_{k-1}$, $w_2 = e^{-c \operatorname{ad} z_k} e^{-c \operatorname{ad} z_{k-1}} z_{k-2}$ and in general

$$w_j = e^{-c \operatorname{ad} z_k} e^{-c \operatorname{ad} z_{k-1}} \dots e^{-c \operatorname{ad} z_{k-j+1}} z_{k-j}$$

for $j = 0, \dots, k-1$. Note that $\|w_j\| = \|z_{k-j}\|$ for each j. A straitghtforward computation shows that

$$\beta_c^{-1}\dot\beta_c = \sum_{j=0}^{k-1} w_j,$$

therefore

$$\sum_j \|w_j\| = \sum_i \|z_i\| = L(\beta) = \|\beta_c^{-1}\dot\beta_c\| = \|\sum_j w_j\|,$$

and by Lemmma 2.6, there exists a unit norm functional such that $\varphi(w_j) = \|w_j\|$ for all $j = 0, \dots, k-1$. In particular $\varphi(z_k) = \varphi(w_0) = \|w_0\| = \|z_k\|$, thus also by Gauss' Lemma (Remark 5.4) we have

$$\varphi(z_{k-1}) = \varphi(e^{-c \operatorname{ad} z_k} z_{k-1}) = \varphi(w_1) = \|w_1\| = \|z_{k-1}\|.$$

Proceedings backwards in this fashion, we can conclude that $\varphi(z_i) = \|z_i\|$ for all $i = 1, \dots, k$. Thus by (5.7) we have that (5.6) holds for this φ, for (almost) all $t \in [a, b]$.

Now assume that (5.6) holds for some φ, described as $\varphi(v) = \langle v, a\rangle = -\operatorname{tr}(\operatorname{ad} v \circ \operatorname{ad} a)$ for some $a \in \mathfrak{k}$ (Remark 5.6), and let $t_0 \leq t \leq t_1$ then $\operatorname{dist}(\gamma_{t_0}, \gamma_t) \leq L(\gamma)_{t_0}^t \leq L(\gamma)_{t_0}^{t_1} = R$. Using the invariance of the metric, we can assumme that $\gamma_{t_0} = 1$. Then there exists a rectifiable lift Γ of γ such that $\Gamma \subseteq B_R$, $\Gamma_{t_0} = 0$, $e^{\Gamma_{t_1}} = \gamma_{t_1}$. Note that $\|\Gamma_{t_1}\| = R = \operatorname{dist}(1, \gamma_{t_1})$. By Remark 5.6, since φ norms $\gamma_t^{-1}\dot\gamma_t$, we have that $\gamma_t^{-1}\dot\gamma_t$ and a commute for all t. By formula (5.4) of Remark 5.22

$$\frac{d}{dt} \operatorname{Ad}_{\gamma_t} a = \operatorname{Ad}_{\gamma_t}[\gamma_t^{-1}\dot\gamma_t, a] = 0,$$

which shows that $\mathrm{Ad}_{\gamma_t} a = a$ for all t (since $\gamma_{t_0} = 1$). Then by the formula (5.5) for the differential of the exponential map of K,

$$e^{-\Gamma_t} D\exp_{\Gamma_t}[\Gamma_t, a] = (\int_0^1 e^{-\lambda \operatorname{ad}\Gamma_t} \operatorname{ad}\Gamma_t \, d\lambda)(a) = -e^{-\lambda \operatorname{ad}\Gamma_t}\Big|_0^1 (a)$$
$$= -e^{-\operatorname{ad}\Gamma_t} a + a = -\mathrm{Ad}_{\gamma_t^{-1}} a + a = 0.$$

Since Γ_t is inside the injectivity radius of the exponential map, it follows that $[\Gamma_t, a] = 0$ for all $t \in [t_0, t_1]$. Then for fixed t, the operators $\operatorname{ad}\Gamma_t$ and $\operatorname{ad} a$ commute, and since they are both skew-adjoint operators acting on $\mathfrak{k}$, they can be simultaneously diagonalized in an orthonormal basis, say $\{e_1, \dots, e_k\}$ of $\mathfrak{k}$. Let $p_i = e_i \otimes e_i$ denote the rank-one orthogonal projection with range e_i, then $\operatorname{ad} a = \sum_{i=1}^k a_i p_i$ and $\operatorname{ad}\Gamma_t = \sum_i \lambda_i p_i$. Note that $\mathrm{Ad}_{\gamma_t} = e^{\operatorname{ad}\Gamma_t} = \sum_i e^{\lambda_i} p_i$. Now we apply the formula (5.5) for the differential of the exponential map to the path $t \mapsto e^{\operatorname{ad}\Gamma_t}$, and noting that $\operatorname{ad}\dot\Gamma_t = (\operatorname{ad}\Gamma_t)^{\cdot}$ we obtain

$$e^{-\operatorname{ad}\Gamma_t} D\exp_{\operatorname{ad}\Gamma_t} \operatorname{ad}\dot\Gamma_t = \int_0^1 e^{-s\operatorname{ad}(\operatorname{ad}\Gamma_t)} \operatorname{ad}\dot\Gamma_t ds = \int_0^1 e^{-s\operatorname{ad}\Gamma_t} \operatorname{ad}\dot\Gamma_t e^{s\operatorname{ad}\Gamma_t} ds.$$

Due to (5.4) we also have

$$\operatorname{ad}\gamma_t^{-1}\dot\gamma_t = \mathrm{Ad}_{\gamma_t^{-1}} \frac{d}{dt} \mathrm{Ad}_{\gamma_t} = e^{-\operatorname{ad}\Gamma_t} D\exp_{\operatorname{ad}\Gamma_t}(\operatorname{ad}\dot\Gamma_t). \tag{5.8}$$

Then since $\operatorname{ad} a$ and $e^{\operatorname{ad}\Gamma_t}$ are diagonal in the orthonormal basis $\{e_i\}_{i=1,\dots,k}$ (for this particular t), we have

$$p_i(e^{-s\operatorname{ad}\Gamma_t} \operatorname{ad}\dot\Gamma_t e^{s\operatorname{ad}\Gamma_t} \circ \operatorname{ad} a) p_i = e^{-s\lambda_i} p_i \operatorname{ad}\dot\Gamma_t e^{s\lambda_i} a_i p_i = p_i \operatorname{ad}\dot\Gamma_t a_i p_i$$
$$= a_i (\operatorname{ad}\dot\Gamma_t)_{ii} p_i = p_i (\operatorname{ad} a \circ \operatorname{ad}\dot\Gamma_t) p_i.$$

Integrating s in $[0,1]$ it follows that

$$(\operatorname{ad} a \circ e^{-\operatorname{ad}\Gamma_t} D\exp_{\operatorname{ad}\Gamma_t} \operatorname{ad}\dot\Gamma_t)_{ii} = (\operatorname{ad} a \circ \operatorname{ad}\dot\Gamma_t)_{ii}$$

and using (5.8) shows that

$$(\operatorname{ad} a \circ (\operatorname{ad}\gamma_t^{-1}\dot\gamma_t))_{ii} = (\operatorname{ad} a \circ \operatorname{ad}\dot\Gamma_t)_{ii}$$

for all i. Thus

$$\varphi(\gamma_t^{-1}\dot\gamma_t) = \langle a, \gamma_t^{-1}\dot\gamma_t \rangle = -\operatorname{tr}(\operatorname{ad}\gamma_t^{-1}\dot\gamma_t \circ \operatorname{ad} a) = -\sum_i (\operatorname{ad}\gamma_t^{-1}\dot\gamma_t \circ \operatorname{ad} a)_{ii}$$
$$= -\sum_i (\operatorname{ad} a \circ \operatorname{ad}\dot\Gamma_t)_{ii} = -\operatorname{tr}(\operatorname{ad} a \circ \operatorname{ad}\dot\Gamma_t) = \langle \dot\Gamma_t, a \rangle = \varphi(\dot\Gamma_t).$$

Since this holds for any $t \in [t_0, t_1]$, we have $\varphi(\dot\Gamma_t) = \|\gamma_t^{-1}\dot\gamma_t\|$ there, and then

$$\operatorname{dist}(\gamma_{t_0}, \gamma_{t_1}) = \|\Gamma_{t_1}\| \geq \varphi(\Gamma_{t_1}) = \varphi(\int_{t_0}^{t_1} \dot\Gamma_t \, dt) = \int_{t_0}^{t_1} \varphi(\dot\Gamma_t) \, dt = \int_{t_0}^{t_1} \|\gamma_t^{-1}\dot\gamma_t\| dt = L_{t_0}^{t_1}(\gamma)$$

which proves that γ is short in that interval. □

Remark 5.24. It is clear from the proof of the previous theorem, that

i) If $\gamma = e^{\Gamma}$ is short, then there exists $a \in \mathfrak{k}$ such that $\operatorname{ad}\Gamma_t$ (and then also $\operatorname{ad}\dot{\Gamma}_t$, $\operatorname{Ad}_{\gamma_t}$, $\operatorname{ad}\gamma_t^{-1}\dot{\gamma}_t$) commute with $\operatorname{ad}a$ for all t. From the faithfulness of the adjoint representation, this is equivalent to $\Gamma_t, \dot{\Gamma}_t, \gamma_t, \gamma_t^{-1}\dot{\gamma}_t$ commuting with a for each t. For linear Lie groups it then follows that a and Γ_t can be simultaneously diagonalized for each t, and this basis also diagonalizes $\gamma_t = e^{\Gamma_t}$. However, a word of caution: the orthonormal basis depends on t.

ii) Once γ is short in a certain interval $[a,b]$, and assuming that we first write $\gamma_b = \gamma_a e^{z_1} e^{z_2} \dots e^{z_k}$, with $\gamma(t_{i+1}) = \gamma(t_i)e^{z_i}$, and the z_i giving the distance among the endpoints and $\|z_i\| \leq R$, then the equality (5.6) if full-filled *for any* norming functional of $z_1 + \dots + z_n$. If $\gamma_b = \gamma_a e^z$ with $z \in \mathbf{D}$, when can we ensure that z and the z_i are in the same face of the sphere? Equivalently, z and $\gamma^{-1}\dot{\gamma}$ are in the same face of the sphere? If $\|z\| < R$ this follows from Theorem 5.18, see also the next remark.

iii) If $\gamma_0 = 1, \gamma_1 = e^z$ with $z \in \mathbf{D}$, and γ is short for a given norm, does it follow that $\gamma = e^{\Gamma}$ with $\Gamma \subset \mathbf{D}$? The argument used in repeated occasions is that if $\|z\| < R$ and γ is short for that norm, then there is a lift $\Gamma \subset B_R$; but this radius depends on the norm.

Remark 5.25 (Non-optimal lifts)**.** By Theorem 5.18, we know that if Γ is short in $\mathfrak{k}$ for a given norm, then its exponential $\gamma = e^{\Gamma}$ is short in K for the bi-invariant metric of that norm. Is there any concrete example of Lie group K with Ad-invariant Finsler norm such that: there exists a path $\gamma = e^{\Gamma} : [a,b] \to K$, with γ short in K joining $1, e^z$ but Γ (which joins $0, z$ in $\mathfrak{k}$) not short in $\mathfrak{k}$? For strictly convex norms this is not possible since the only short paths are segments; as we will see in the next section it is also impossible if all the faces of the unit sphere are abelian (Corollary 5.27). Another relevant example, studied by Antezana, Ghighlioni and Stojanoff in [AnGhiSt19] is the full unitary group (or in our setting, $K = \mathrm{SU}_n$) with the spectral norm: examining [AnGhiSt19, Theorem 2.1] it is apparent that lifted short paths are also short there.

5.3. **Norms with abelian faces.** In Theorem 5.19 we showed that when the faces of the unit ball are singletons, the short paths are one-parameter groups. We will show here that the situation is somewhat similar if we allow the faces of the unit ball to be inside *abelian* subalgebras of $\mathfrak{k}$. In Section 6 we will give a full characterization of those norms with this important structural property.

Proposition 5.26. *Let B the unit ball of the bi-invariant norm in $\mathfrak{k}$. Then all the faces of B are abelian if and only if for all piecewise C^1 short curves $\gamma \subseteq K$, the logarithmic derivatives $x_t = \gamma_t^{-1}\dot{\gamma}_t$ of γ commute for all t.*

Proof. Assume first that each face of the unit ball is abelian. By Theorem 5.23, if γ is short, it has logarithmic derivative (after normalizing) inside a face of the ball, and these commute for all t.

Now assume that for some $w \in \mathfrak{k}$ the face $F_w(B)$ is non-abelian. Let $x, y \in F_w(B)$ with $[x,y] \neq 0$. Let $R > 0$ be the radius of injectivity and for $0 < b < R$ let $\gamma\colon [0,b] \to K$ be the path which solves

$$\gamma_t^{-1}\dot{\gamma}_t = tx + (b-t)y$$

for $t \in [0,b]$. Then by Theorem 5.23 γ is a geodesics, furthermore the logarithmic derivatives $\gamma_0^{-1}\dot\gamma_0 = x$ and $\gamma_b^{-1}\dot\gamma_b = y$ do not commute. □

Corollary 5.27. *Let B be the unit ball of the bi-invariant norm in $\mathfrak{k}$ and assume that the faces of B are all abelian. Let $\gamma\colon [a,b] \to K$ be a short piecewise C^1 path. Then*

i) *There exists $z \in \mathfrak{k}$ such that $\delta_t = \gamma_a e^{tz}$ is also short with the same endpoints.*
ii) *If $L(\gamma) \le R$ (thus $\gamma_b = \gamma_a e^z$ for some $\|z\| \le R$), then $\gamma = \gamma_a e^\Gamma$ where $\Gamma\colon [a,b] \to \mathfrak{k}$ and $[\Gamma_t, \Gamma_s] = 0$ for all $s,t \in [a,b]$, thus $\gamma_t^{-1}\dot\gamma_t = \dot\Gamma_t$. In particular the logarithmic derivatives of γ commute and Γ is short in $\mathfrak{k}$, with the same length than γ.*

Proof. As always we can assume that $\gamma_a = 1$, partition γ in small pieces such that $\gamma(t_{i+1}) = \gamma(t_i)e^{z_i}$ with $\|z_i\| \le R$ as in the proof of Theorem 5.23. Let $\beta(s) = e^{sz_1}e^{sz_2}$, then β joins $1, e^{z_1}e^{z_2} = \gamma(t_2)$ and $L(\beta) = \|z_1 + z_2\|$, thus

$$\operatorname{dist}(1,\gamma_{t_2}) \le L(\beta) \le \|z_1\| + \|z_2\| = \operatorname{dist}(1,\gamma_{t_1}) + \operatorname{dist}(\gamma_{t_1},\gamma_{t_2}) = \operatorname{dist}(1,\gamma_{t_2}),$$

which shows that β is short among its endpoints. For small s, let $\beta_s = e^{B_s}$ with $B_s = s(z_1 + z_2) + \frac{s^2}{2}[z_1,z_2] + o(s^3)$, then by Theorem 5.18(1), $\|B_s\| = \operatorname{dist}(1,\beta_s) = L_0^s(\beta) = s\|z_1+z_2\|$. On the other hand, by Theorem 5.18(4), we also have $\varphi(\dot B_s) = \|\beta_s^{-1}\dot\beta_s\| = \|z_1+z_2\|$, then integrating we have $\varphi(B_s) = s\|z_1+z_2\| = \|B_s\|$, and this shows that B_s (normalized) is inside a face of the ball, which implies that the B_s commute for all (small) s (therefore they also commute with $\dot B_s$). Then using the formula (5.5) we have

$$e^{-s\operatorname{ad} z_2}(z_1+z_2) = \beta_s^{-1}\dot\beta_s = e^{-B_s}D\exp_{B_s}\dot B_s = \dot B_s = z_1 + z_2 + s[z_1,z_2] + o(s^2).$$

Differentiating at $s = 0$ shows that $[z_1,z_2] = 0$. Thus $\gamma(t_2) = e^{z_1}e^{z_2} = e^{z_1+z_2}$ and $\gamma(t_3) = e^{z_1}e^{z_2}e^{z_3} = e^{z_1+z_2}e^{z_3}$. We repeat the argument now using $\beta_s = e^{s(z_1+z_2)}e^{sz_3}$, this shows that $[z_3, z_1+z_2] = 0$. Thus $\gamma_b = e^{z_1}e^{z_2}e^{z_3}\dots e^{z_k} = e^{\sum z_i}$, and if we let $z = \sum z_i$, then $\|z\| = \sum\|z_i\| = \operatorname{dist}(\gamma_a,\gamma_b)$ and the first claim follows.

For the second assertion let Γ be the smooth lift of γ, $\Gamma_a = 0$, then by Theorem 5.18(4) we have $\varphi(\dot\Gamma_t) = \|\gamma_t^{-1}\dot\gamma_t\|$ for some unit norm φ, and again integrating $\varphi(\Gamma_t) = L_0^t(\gamma) = \operatorname{dist}(1,e^{\Gamma_t}) = \|\Gamma_t\|$ shows that Γ_t (normalized) is inside a face of the sphere, which by hypothesis is abelian. But then Γ_t and $\dot\Gamma_t$ also commute for all t and it is apparent that

$$L(\Gamma) = \int_a^b \|\dot\Gamma\|dt = \int_a^b \|e^{-\Gamma}D\exp_\Gamma\dot\Gamma\|dt = \int_a^b \|\gamma^{-1}\dot\gamma\|dt = L(\gamma)$$

thus if $\Gamma_b = z$ then $\|z\| \le L(\Gamma) = L(\gamma) = \operatorname{dist}(1,e^z) = \|z\|$ and Γ is short. □

5.4. **Geodesic are quasi-autonomous.** Throughout, the action of the semi-simple compact group K on the symplectic manifold M complies the hypothesis of Section 3.2. Recall (Definition 3.2) that a Hamiltonian H_t is called *quasi-autonomous* if there exists $x^-, x^+ \in M$ such that $H_t(x^-) = \min_M H_t$, $H_t(x^+) = \max_M H_t$ for all $t \in [a,b]$. As before, we use R to indicate the injectivity radius of the exponential map of K, for the given norm. In the next theorem, the geometry of the group K is the one given by the generalized Hofer norm (3.5) obtained by the almost effective action.

Theorem 5.28. *Let $K \curvearrowright (M,\omega)$ be a Hamiltonian almost effective action. Let $\gamma\colon [a,b] \to K$ be piecewise C^1, and denote its right logarithmic derivative by $x_t = \dot\gamma_t\gamma_t^{-1}$. Then if γ is*

short, $(\mu^{x_t})_{t\in[0,1]}$ is a quasi-autonomous Hamiltonian, and if μ is quasi-autonomous, γ is locally short (in each interval of length $\leq R$).

Proof. By Theorem 5.23, γ is locally short if and only if there is a common norming functional for $x_t = \dot{\gamma}_t\gamma_t^{-1}$, for all t. By Corollary 2.11 the set of logarithmic derivatives $\{x_t\}_{t\in[0,1]} \subseteq \mathfrak{k}$ is contained in a cone generated by a face if and only if there exist x^- and x^+ such that $x^- \in \cap_{t\in[0,1]} \operatorname{argmin}(\varphi_{x_t})$ and $x^+ \in \cap_{t\in[0,1]} \operatorname{argmax}(\varphi_{x_t})$, where $\varphi_{x_t} : \mu(M) \to \mathbb{R}$. We can choose $m^-, m^+ \in M$ such that $x^- = \mu(m^-)$ and $x^+ = \mu(m^+)$. The result follows since $\mu^{x_t} = \varphi_{x_t} \circ \mu$ for all $t \in [0,1]$. □

Remark 5.29. A similar characterization of geodesics (in fact, of their logarithmic derivative $\gamma_t^{-1}\dot{\gamma}_t = x_t$) can be obtained for the one-sided norm induced by the action of K in M, if one replaces the condition of quasi-autonomous for the Hamiltonian $H_t = \mu^{x_t}$, with the condition that there exists a point $x^+ \in M$ such that $H_t(x^+) = \max_M H_t$ for all $t \in [a,b]$.

6. Spheres with abelian faces

In this section we characterize those Ad-invariant norms which have unit balls with abelian faces in terms of conditions on its intersection with a Cartan subalgebra. When this intersection is a polytope the condition reduces to the regularity of the extreme points of its polar dual. Based on this result and Kirwan's Theorem 4.1 we characterize the compact semi-simple groups of Hamiltonian diffeomorphims such that geodesics have commuting Hamiltonians. We start with the important special case of Hofer norms derived from coadjoint actions on regular coadjoint orbits. Groups with length structures derived from these norms have the property that the logarithmic speed of its geodesics lie in a Weyl chamber.

6.1. Regular coadjoint actions and non-crossing of eigenvalues.

In this section we follow the setting of Example 3.12 on (co)adjoint orbits in $O \subseteq \mathfrak{k} \simeq \mathfrak{k}^*$.

Definition 6.1. The (co)adjoint orbit $O = O_\lambda \subseteq \mathfrak{k}$ of the action is *regular* if $\lambda \in \mathfrak{k}$ is a regular element (Remark 3.11).

In order to characterize the faces of the unit ball in $\mathfrak{k}$ via Corollary 2.11 we need to study, for a nonzero x in $\mathfrak{k}$, the sets $\operatorname{argmax}_O(\varphi_x)$ and $\operatorname{argmin}_O(\varphi_x)$ of $\varphi_x = \mu^x : O \to \mathbb{R}$. Let $m_t = \operatorname{Ad}_{e^{tv}} m$ be a path through $m \in O$ with $\dot{m}_0 = [v,m] \in T_m O$, differentiating $\varphi_x(m_t) = \langle x, m_t\rangle$ at $t = 0$ we obtain

$$D(\varphi_x)_m(\dot{m}_0) = \langle x, [v,m]\rangle = -\langle x, [m,v]\rangle = \langle [m,x], v\rangle,$$

since $\operatorname{ad} m$ is skew-adjoint for the Killing form. Since $v \in \mathfrak{k}$ is arbitrary, this implies that $m \in O$ is a critical point of φ_x if and only if $m \in \mathfrak{z}(x)$, where $\mathfrak{z}(x)$ is the centralizer of x, i.e.

$$Crit(\varphi_x) = O \cap \mathfrak{z}(x). \tag{6.1}$$

The proof of following proposition can be found in [BiGhiHe14, Lemma 23].

Proposition 6.2. *Fix a maximal torus $T \subseteq K$, a nonzero vector $x \in \mathfrak{h} = Lie(T)$ and a point $m \in O \cap \mathfrak{h}$; thus $m \in Crit(\varphi_x)$. Then m is a maximum point of φ_x if and only if there is a Weyl chamber in $\mathfrak{h}$ whose closure contains both x and m.*

With these tools at hand, we next characterize the faces of the unit ball of $\|\cdot\|_O$.

Proposition 6.3. *Let O be a regular (co)adjoint orbit and let $\|\cdot\|_O$ be the associated* Ad*-invariant Hofer norm. A set of vectors has a common norming functional if and only if it is contained in a Weyl chamber (given by a choice of torus and positive simple roots). Hence, the maximal cones generated by faces of the unit ball are Weyl chambers.*

Proof. By Corollary 2.11 a set of elements $S \subseteq \mathfrak{k}$ have the same norming functional if and only if the set $\{\varphi_u : u \in S\}$ has a common maximizer $x^+ \in O$ and a common minimizer $x^- \in O$.

We denote by $\mathfrak{h}_+$ the Weyl chamber that contains x^+, it is unique since x^+ is regular. If $u \in S$, then the functional φ_u has a maximum at x^+, hence by equation (6.1) u commutes with x and therefore $u \in \mathfrak{h}$. By Proposition 6.2 we conclude that $u \in \mathfrak{h}_+$ if and only if the functional φ_u has a maximum at x^+. Let us denote by $x' \in \mathfrak{k}$ the element defined by $\{x'\} = O \cap -\mathfrak{h}_+$, that is, the element of O in the opposite Weyl chamber. A functional φ_u has a minimum at x' if and only if φ_{-u} has a maximum at x'. This holds if and only if $-u$ and x' belong to the same Weyl chamber, which is equivalent to $u \in \mathfrak{h}_+$. Hence, φ_u has a maximum at x^+ and φ_u has a minimum at x', are both equivalent to $u \in \mathfrak{h}_+$. If we take $x^- = x'$ as the common minimizer we get a maximal face which is $\mathfrak{h}_+$.

If $\mathfrak{h}_+$ is a Weyl chamber we denote by x^+ and x^- the elements defined by $\{x^+\} = O \cap \mathfrak{h}_+$ and $\{x^-\} = O \cap -\mathfrak{h}_+$. Then we can reverse the argument in the previous paragraph to conclude that the functionals φ_u with $u \in \mathfrak{h}_+$ have $x^+ \in O$ as maximizer and $x^- \in O$ as minimizer. □

From this characterization of the faces of the unit ball in $\mathfrak{k}$ we can determine the geodesics in K. Let $\gamma\colon [a,b] \to K$ be a curve in a group K endowed with the length structure obtained from Hofer's norm $\|\cdot\|_O$ for a regular (co)adjoint orbit O. As a combination of the previous proposition and Theorem 5.23, we obtain

Theorem 6.4. *If γ is short then all its logarithmic derivatives are contained in the same Weyl chamber $\mathfrak{h}_+$ (given by a choice of torus and positive simple roots). If its derivatives are contained in a Weyl chamber, then γ is locally short (in each interval of* length $\leq R$*, where R is the injectivity radius of the norm).*

Example 6.5 ($K = \mathrm{SU}(n)$ and the non-crossing of eigenvalues)**.** Consider the case of $\mathrm{SU}(n)$ acting on a regular (co)adjoint orbit O_λ containing $\lambda = i\,\mathrm{diag}(\lambda_1, \dots, \lambda_n)$, with $\lambda_1 < \cdots < \lambda_n$. Recall that here we obtain in $\mathfrak{k}$ the λ-numerical radius as Hofer norm (Example 3.13). The local condition for the logarithmic derivatives to be the speeds of geodesics is that there exist a fixed orthonormal basis such that the speeds are simultaneously diagonal in this basis for all t (Corollary 5.27) and if $x_i(t)$ are the eigenvalues of $x_t = \dot{\gamma}_t \gamma_t^{-1}$, then

$$x_1(t) \leq \cdots \leq x_n(t) \qquad \forall t \in [a,b].$$

This is because in $\mathrm{SU}(n)$ the condition of being in the same Weyl chamber is given by the non-crossing of the eigenvalues.

Remark 6.6. The Hofer norm associated to *singular* (co)adjoint orbits can be studied using the following result on maximizers and minimizers of linear functionals restricted to the orbits, see [BiGhiHe14, Lemma 22]. Let $Z(x)$ be the centraliser of x in K and let $F_x(O)$ be the face of O defined by x. Then

- $\operatorname{argmax}(\varphi_x)$ is a $Z(x)$-orbit,
- $\operatorname{ext}(F_x(\operatorname{conv}(O)) = \operatorname{argmax}(\varphi_x)$, so $\operatorname{ext}(F_x(\operatorname{conv}(O))$ is a $Z(x)$-orbit.
- $F_x(\operatorname{conv}(O)) \subseteq \mathfrak{z}(x)$.

6.2. **Invariant norms with abelian faces.** Let K be a compact connected semi-simple Lie group and let $\langle \cdot, \cdot \rangle$ be as usual the opposite Killing form of $\mathfrak{k}^{\mathbb{C}}$ restricted to $\mathfrak{k}$. Let Δ be the set of (real) roots of $\mathfrak{k}$ with respect to a fixed Cartan subalgebra $\mathfrak{h}$, as in Remark 5.2. For $x \in \mathfrak{h}$ we have

$$\mathfrak{z}(x) = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta_+, \alpha(x)=0} Z_\alpha, \tag{6.2}$$

here $\mathfrak{z}(x)$ is the Lie algebra of $Z(x)$ as usual (the centralizer of $x \in \mathfrak{k}$). We define the *smallest Weyl chamber wall containing* x as the linear space

$$W_x = \bigcap_{\{\alpha \in \Delta_+ : \alpha(x)=0\}} \ker(\alpha).$$

Theorem 6.7. *Let K be a semi-simple compact connected group and let B be an* Ad*-invariant convex body in $\mathfrak{k}$ containing* 0*. Let $\mathfrak{h}$ be a Cartan algebra in $\mathfrak{k}$ and $\mathfrak{h}_+$ a positive Weyl chamber. Then all faces of B are abelian if and only if for all $x \in \mathfrak{h}^+$ we have*

$$F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+ \subseteq W_x,$$

and in this case $F_x(B) = F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+$.

Proof. We assume for simplicity that $x \in \mathfrak{h}^+$. Theorem 4.15 states that

$$F_x(B) = \operatorname{Ad}_{Z(x)}(F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+).$$

Next we show that this set is abelian when the inclusion stated in the theorem holds. If x is in the interior of the Weyl chamber then $Z(x)$ is trivial and $F_x(B) = F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+$, hence the face is abelian. If x is not regular, and hence $Z(x)$ is not trivial, then by (5.1) and (6.2) the centralizer $Z(x)$ acts trivially on W_x. Therefore, the inclusion $F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+ \subseteq W_x$ implies that $F_x(B) = \operatorname{Ad}_{Z(x)}(F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+) = F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+$, so the face is abelian. Note also that the last assertion of the theorem follows from this argument.

If the inclusion in the statement of the theorem is not satisfied there exists $y \in F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+$ such that $y \notin W_x$; it follows that there exists α such that $\alpha(y) \neq 0$ and $\alpha(x) = 0$. Consider the orbit $\operatorname{Ad}_{Z(x)}(y) \subseteq \mathfrak{k}$ as a submanifold with its differentiable structure, and observe that

$$[\mathfrak{z}(x), y] \subseteq T_y(\operatorname{Ad}_{Z(x)}(y)).$$

By (6.2), we have $v_\alpha, u_\alpha \in \mathfrak{z}(x)$; the equations $[y, u_\alpha] = -\alpha(y) v_\alpha$ and $[y, v_\alpha] = \alpha(y) u_\alpha$ hold by (5.1), therefore we conclude that

$$\{u_\alpha, v_\alpha\} \subseteq T_y(\operatorname{Ad}_{Z(x)}(y)).$$

Since $[u_\alpha, v_\alpha] = h_\alpha \neq 0$ the tangent $T_y(\mathrm{Ad}_{Z(x)}(y))$ is not an abelian space, so that $\mathrm{Ad}_{Z(x)}(y)$ cannot be an abelian set. The inclusion $\mathrm{Ad}_{Z(x)}(y) \subseteq \mathrm{Ad}_{Z(x)}(F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+) = F_x(B)$ implies that $F_x(B)$ is not an abelian set. □

Remark 6.8. From the previous theorem we know that if the faces of B are abelian then

$$F_x(B) = F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+$$

for a Weyl chamber $\mathfrak{h}_+$ given by a choice of torus and positive simple roots such that $x \in \mathfrak{h}_+$. Therefore the balls of the norms studied in Section 6.1 have maximal commuting cones generated by faces.

6.3. **Polytopes with regular extreme points.** The aim of this section is to prove the following theorem. Note that an element $x \in \mathfrak{k}$ is regular if and only if $\alpha(x) \neq 0$ for all $\alpha \in \Delta_+$.

Theorem 6.9. *Let K be a semi-simple compact connected group, let B be an* Ad*-invariant convex body in $\mathfrak{k}$ containing* 0*, such that $B \cap \mathfrak{h} = P^\circ$ is a polytope. Then all faces of B are abelian if and only if all extreme points of P are regular.*

For the proof of this theorem we need a couple of preliminary lemmas on polytopes invariant under finite reflection groups. These lemmas will be later applied to the Hofer norm polytopes defined in Definition 4.6, which are a polytopes invariant under the Weyl group.

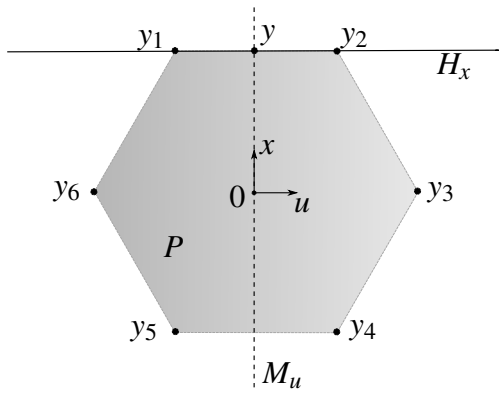

FIGURE 5. Mirror hyperplane and extreme points of a polytope P.

Definition 6.10 (Finite reflection groups)**.** Let $\mathcal{W}$ be a finite reflection group acting isometrically on an inner product vector space $(V, \langle\cdot,\cdot\rangle)$ and generated by reflections $\{r_u\}_{u\in\Phi}$. Here r_u is the refection which fixes the mirror hyperplane $M_u = \{u\}^\perp$ and Φ is a finite subset of the unit sphere that is called the positive root system. For a finite reflection group the space can be subdivided into chambers bounded by mirror hyperplanes. We can choose any (closed) chamber C and call it the fundamental chamber. Given a positive root system Φ, a fundamental chamber is given by

$$C = \{x \in V : \langle x, u\rangle \geq 0 \text{ for all } u \in \Phi\}.$$

Let P be a polytope with extreme points $\mathrm{ext}(P) = \{y_1, \ldots, y_m\}$ and assume that P is invariant under $\mathcal{W}$.

Lemma 6.11. *Let $\Pi \subseteq \Phi$ and assume that $\cup_{u\in\Pi} M_u \cap \{y_1,\dots,y_m\} = \emptyset$. If H_x supports P at y and $y \in \cap_{u\in\Pi} M_u$ it follows that $x \in \cap_{u\in\Pi} M_u$.*

Proof. Assume P is invariant under a reflection r_u and the mirror hyperplane does not contain any extreme point of P, i.e. $M_u \cap \{y_1,\dots,y_m\} = \emptyset$, see Figure 5. Then, if H_x is a hyperplane that supports P at an $y \in M_u$ we claim that $x \in M_u$. To prove this we write y as a convex combination of the $y_1,\dots,y_m$, i.e. $y = \lambda_1 y_1 + \dots \lambda_n y_n$, with $n \le m$, $\lambda_1,\dots,\lambda_n > 0$, and $\sum_{i=1}^n \lambda_i = 1$. If we write

$$y = \frac{1}{2}y + \frac{1}{2}r_u(y) = \frac{1}{2}(\lambda_1 y_1 + \dots \lambda_n y_n) + \frac{1}{2}r_u(\lambda_1 y_1 + \dots \lambda_n y_n),$$

we see that we can assume that $y = \lambda_1 y_1 + \dots \lambda_n y_n$, with $\lambda_1,\dots,\lambda_n > 0$, $\sum_{i=1}^n \lambda_i = 1$ and $r_u(y_j) \in \{y_1,\dots,y_n\}$ for $i = 1,\dots,n$. Since H_x supports P at $y \in M_u$ we get $\varphi_x(y) = 1$ and $\varphi_x(y_1),\dots\varphi_x(y_m) \le 1$. From $y = \lambda_1 y_1 + \dots + \lambda_n y_n$, with $\lambda_1,\dots,\lambda_n > 0$, $\sum_{i=1}^n \lambda_i = 1$, it follows that $\varphi_x(y_1),\dots,\varphi_x(y_n) = 1$. Since $y_1 \notin M_u$ we know that $r_u(y_1) \in \{y_2,\dots,y_n\}$ so that $y_1 - r_u(y_1)$ is non zero. Since $y_1 - r_u(y_1)$ is orthogonal to the mirror M_u and $\varphi_x(y_1 - r_u(y_1)) = 1 - 1 = 0$ we conclude that x is orthogonal to $y_1 - r_u(y_1)$ and is therefore contained in M_u. The Lemma follows by applying this result to each of the mirrors $\{M_u\}_{u\in\Pi}$. □

Lemma 6.12. *Assume that there is an extreme point y_1 of P such that $y_1 \in C$ and y_1 is contained in a mirror. Then there is a supporting hyperplane $H_{x'}$ of P at y_1 such that x' is in the interior of the fundamental chamber.*

Proof. By assumption $\Pi := \{u \in \Phi : y_1 \in M_u\} \neq \emptyset$. The set of reflections $\{r_u : u \in \Pi\}$ generates the stabilizer $\mathrm{Stab}(y_1)$ of y_1 (see e.g. [BoBo10, Theorem 12.6]). Since P is a polytope there exists x which determines a supporting hyperplane of P at y_1 such that $\varphi_x(y_1) = 1$ and $\varphi_x(y_j) < 1$ for $j = 2,\dots,m$. If we denote by x' the average

$$x' = \frac{1}{|\mathrm{Stab}(y_1)|}\sum_{w\in\mathrm{Stab}(y_1)} w.x,$$

then $\varphi_{x'}(y_1) = 1$, $\varphi_{x'}(y_j) < 1$ for $j = 2,\dots,m$, and $x' \in \cap_{u\in\Pi} M_u$. To show that $x' \in C$ we need to verify that $\langle x', u\rangle \ge 0$ for all $u \in \Phi$. If u is not in Π then y_1 is not in M_u, therefore $\varphi_{x'}(y_1) = 1$ and $\varphi_{x'}(r_u(y_1)) < 1$ which implies that $\varphi_{x'}(y_1 - r_u(y_1)) > 0$. On the other hand, since y_1 is in the positive chamber C it follows that $y_1 - r_u(y_1) = d.u$ for $d > 0$. Hence $\varphi_{x'}(d.u) > 0$ which is equivalent to $\langle x', u\rangle > 0$.

The x' can be perturbed so that $H_{x'}$ is a supporting hyperplane of P at y_1 and x' is in the interior of the fundamental chamber. □

Proof. (of Theorem 6.9). We apply the previous lemmas to the Weyl group action on $\mathfrak{h}$. It is easy to check that the condition of Theorem 6.7 can be stated using the more general notation of the previous lemmas as follows:

$$F_y(P^\circ) \cap C \subseteq W_y := \bigcap_{u\in\Phi : y\in M_u} M_u$$

for all nonzero y in $C = \mathfrak{h}_+$. Here, for example, $M_{h_\alpha} = \ker(\alpha)$, with h_α as in (5.1). We can multiply y by a positive scalar and assume that it is in $\mathrm{bd}\,P$. Then by Theorem 2.22

$$\begin{aligned} F_y(P^\circ) &= \{x \in V : H_y \text{ supports } P^\circ \text{ at } x\} \\ &= \{x \in V : H_x \text{ supports } P \text{ at } y\}. \end{aligned}$$

By Lemma 6.11, if the extreme points of P are all regular and $y \in W_y = \bigcap_{u\in\Phi: y\in M_u} M_u$, then $x \in W_y$ holds when H_x supports P at y, i.e. x is in $F_y(P^\circ)$.

If there is a non-regular extreme point of P then by invariance of P under the reflection group $\mathcal{W}$ we can choose a non-regular extreme point y_1 of P in C. Lemma 6.12 implies that there is a supporting hyperplane $H_{x'}$ of P at y_1 such that x' is in the interior of C. Then $x' \in F_y(P^\circ) \cap C$ and x' does not belong to W_y, so that the condition of Theorem 6.7 is not satisfied and there is a face of B which is not abelian. □

Let K be a semi-simple compact connected group, let B be an Ad-invariant convex body in $\mathfrak{k}$ containing 0, such that $B \cap \mathfrak{h} = P^\circ$ is a polytope. Let g_B be the associated Ad-invariant norm and endow K with the corresponding Finsler length structure.

Theorem 6.13. *The extreme points of P are regular if and only if all short curves γ in K have commuting logarithmic derivatives.*

Proof. By Theorem 6.9 the extreme points of P are regular if and only if B has abelian faces, and by Proposition 5.26 this holds if and only if all short curves have commuting logarithmic derivatives. □

We now specialize Theorem 6.13 to compact semi-simple groups of Hamiltonian diffeomorphisms.

Theorem 6.14. *Let $K \curvearrowright M$ be an almost effective Hamiltonian action with moment map $\mu : M \to \mathfrak{k}^* \simeq \mathfrak{k}$ and endow K with the pullback metric of Section 3.2.1. Let $\mu(M) \cap \mathfrak{h}^+ = \mathrm{conv}\{x_1, \dots, x_n\}$ be Kirwan's polytope given by Theorem 4.1, and let*

$$P = \mathrm{conv}\{w.x - w'.x' : x, x' \in \{x_1, \dots, x_n\}, \quad w, w' \in \mathcal{W}\}$$

be the Hofer norm polytope derived from it. Then all short curves in K have commuting Hamiltonians if and only if all the extreme points of P are regular.

Proof. By Proposition 3.6 a curve γ has Hamiltonians μ^{x_t}, where $x_t = \dot\gamma_t \gamma_t^{-1}$ is the right logarithmic derivative. Recalling that $\{\mu^{x_s}, \mu^{x_t}\} = \mu^{[x_s, x_t]}$ the theorem follows. □

Note that the same result Holds if we endow $K \to \mathrm{Ham}(M, \omega)$ with the second Hofer norm by taking the second Hofer norm polytope

$$P' = \mathrm{conv}(\{w.x : x \in \{x_1, \dots, x_n, -x_1, \dots, -x_n\}, \quad w \in \mathcal{W}\})$$

of Definition 4.6, and also if we consider the one-sided Finlser Hofer norm with its polytope

$$P^+ = \mathrm{conv}\{w.x : x \in \{x_1, \dots, x_n\}, w \in \mathcal{W}\}.$$

Example 6.15. An example of a group K with non-commuting Hamiltonians is $\mathrm{SU}(4)$ acting on the singular (co)adjoint orbit containing $x = i\,\mathrm{diag}(3, -1, -1, -1)$. The Hofer norm polytope is the convex hull of the permutations of the matrix $i\,\mathrm{diag}(4, 0, 0, -4)$. Its

extreme points are the permutations of the matrix given by $i\,\mathrm{diag}(4,0,0,-4)$, which are all singular. We can give a short informal explanation for this which is similar to the proof of Proposition 6.3. If the maximum of φ_x is at $x^+ = i\,\mathrm{diag}(3,-1,-1,-1)$ and the minimum is at $x^- = i\,\mathrm{diag}(-1,-1,-1,3)$ then x^+ block diagonalizes x and x^- block diagonalizes x, so that

$$x = i(\lambda_1 P_{\mathbb{C}e_1} \oplus A \oplus \lambda_4 P_{\mathbb{C}e_4}),$$

with A a selfadjoint operator on $\mathbb{C}e_2 \oplus \mathbb{C}e_3$ such that its spectrum $spec(A)$ satisfies $\lambda_1 \geq spec(A) \geq \lambda_4$.

6.4. **Properties determined by Kirwan's polytope and product actions.** In Section 6.3 we characterized Hofer norms with abelian faces using the regularity of the norm polytope (Theorem 6.9), and before that in Section 6.1 we studied the case of faces generating maximal abelian cones. In this section we show how conditions on Kirwan's polytope can characterize these properties, and then to finish the paper we study how these properties behave if we take products of Hamiltonian actions.

We are here in the context of finite reflection groups of Section 6.3, in particular see Definition 6.10. We recall from [EaPe77, Section 4] or [MOZZ03, Lemma 2.9] the following generalization of the rearrangement inequality

$$\sum_{i=1}^{n} x_{\pi(i)} y_i \leq \sum_{i=1}^{n} x_i y_i$$

for $x_1 \leq \cdots \leq x_n$, $y_1 \leq, \cdots \leq y_n$ and certain permutations π.

Lemma 6.16. *Suppose $x, y \in V$, then*

$$\sup_{w \in \mathcal{W}} \langle x, w.y \rangle = \langle x, y \rangle$$

if and only if x and y belong to the same Weyl chamber. In this case

$$\langle x, w.y \rangle = \langle x, y \rangle$$

if and only if there exists $w' \in \mathrm{Stab}(x)$ such that $w.y = w'.y$. That is, the set of maximizers of φ_x in $\mathcal{W}.y$ is exactly $\mathrm{Stab}(x).y$.

Definition 6.17. Let $E \subseteq V$ be a convex set with $x \in E$. The normal cone to E at x is defined as

$$N(x,E) := \{y \in V : \langle y, z - x \rangle \leq 0 \text{ for all } z \in E\}.$$

It is a closed convex cone.

We will also need the following simple fact about a polytope and its normal cones.

Lemma 6.18. *Let $P \subseteq V$ be a convex polytope. If $A \subseteq \mathrm{ext}(P)$ and*

$$\bigcup_{y \in A} N(y,P) = V$$

then $A = \mathrm{ext}(P)$.

Note that if P is a convex polytope that contains 0 in its interior, then for $y \in \mathrm{ext}(P)$ we have $\mathbb{R}_+ F_y(P^\circ) = N(y,P)$.

6.4.1. *Conditions on Kirwan's polytope.* If a Hamiltonian action has the same Hofer norm polytope as the action on a *regular* (co)adjoint orbit, then it has the same geodesics; therefore they are characterized by Theorem 6.4. We next characterize the Hofer norm polytopes which are derived from regular coadjoint orbits.

Definition 6.19. Given a Weyl chamber $\mathfrak{h}_+$ let $-\mathfrak{h}_+ = w^*.\mathfrak{h}_+$ be the opposite Weyl chamber (there is a unique such $w^* \in \mathcal{W}$). We say that $y \in \mathfrak{h}_+$ is *symmetric* if $w^*.y = -y$.

Proposition 6.20. *A $\mathcal{W}$-invariant symmetric polytope P is the Hofer norm polytope of $\|\cdot\|_O$ for a coadjoint orbit O if and only if* $\mathrm{ext}(P) = \mathcal{W}.y$ *for a symmetric $y \in \mathfrak{h}_+$.*

A $\mathcal{W}$-invariant symmetric polytope P is the Hofer norm polytope of $\|\cdot\|_O$ with O a regular coadjoint orbit if and only if $\mathrm{ext}(P) = \mathcal{W}.y$ *for a symmetric regular $y \in \mathfrak{h}_+$.*

Proof. For $x \in \mathfrak{h}_+$ let O_x be a coadjoint orbit. The Hofer norm polytope of $\|\cdot\|_{O_x}$ is given by

$$P = \mathrm{conv}\{w.x - w'.x : w, w' \in \mathcal{W}\}.$$

If we take $y = x - w^*x \in \mathfrak{h}_+$ we claim that P has an extreme point at $y \in \mathfrak{h}^+$ with normal cone at this point which equals $\cup_{w\in \mathrm{Stab}(y)} w.\mathfrak{h}_+$. To see this note that by Lemma 6.16 for $z \in \mathfrak{h}_+$ and $w \in \mathcal{W}$

$$\varphi_z(x) \geq \varphi_z(w.x),$$

and

$$\varphi_z(-w^*.x) = \varphi_{-z}(w^*.x) \geq \varphi_{-z}(ww^*.x) = \varphi_z(-ww^*.x).$$

Therefore

$$\varphi_z(x - w^*.x) \geq \varphi_z(w.x - w'.x),$$

for $w, w' \in \mathcal{W}$ and we conclude that the normal cone to P at y includes $\mathfrak{h}_+$.

Since this argument is Weyl group invariant, for $w \in \mathcal{W}$ the normal cone to P at $w.y$ includes $w.\mathfrak{h}_+$. In the case of singular y there can be repetitions: if $\mathrm{Stab}(y)$ is not trivial then the normal cone to P at y includes $\cup_{w\in \mathrm{Stab}(y)} w.\mathfrak{h}_+$. We have $\cup_{w\in\mathcal{W}} w.\mathfrak{h}_+ = \mathfrak{h}$, hence by Lemma 6.18 $\mathcal{W}.y$ are all the extreme points of P and the normal cone to P at $w'.y$ is $w'.\cup_{w\in \mathrm{Stab}(y)} w.\mathfrak{h}_+$.

If y is symmetric then we can take $O_{(1/2)y}$ and the Hofer norm polytope of $\|\cdot\|_{O_{(1/2)y}}$ satisfies $\mathrm{ext}(P) = \mathcal{W}.y$.

The proof of the second assertion is similar and we omit it. □

Theorem 6.21. *Let $E \subseteq \mathfrak{k}$ be an* Ad*-invariant set such that $E \cap \mathfrak{h}_+ = \mathrm{conv}\{x_1, \ldots, x_n\}$. The Hofer norm polytope derived from E has extreme points $\mathcal{W}.y$ for regular y if there is an x_i, say x_1, such that*

- *The point $x_1 - w^*.x_1 \in \mathfrak{h}_+$ is regular.*
- *For $x \in \mathfrak{h}_+$ we have $\varphi_x(x_1) \geq \varphi_x(x_j)$ for $j \in \{2, \ldots, n\}$, that is, $\mathfrak{h}_+$ is contained in the normal cone of* $\mathrm{conv}\{x_1, \ldots, x_n\} \subseteq \mathfrak{h}$ *at x_1.*

Proof. The Hofer norm polytope is given by

$$P = \mathrm{conv}\{w.x - w'.x' : x, x' \in \{x_1, \ldots, x_n\}, \quad w, w' \in \mathcal{W}\}.$$

We are going to prove that $\mathrm{ext}(P) = \mathcal{W}.(x_1 - w^*.x_1)$. Since $x_1 - w^*.x_1$ is regular $\mathcal{W}.(x_1 - w^*.x_1)$ has $|\mathcal{W}|$ points. We claim that P has an extreme point at $x_1 - w^*.x_1 \in \mathfrak{h}^+$ with a

normal cone at this point which equals $\mathfrak{h}_+$. To see this note that by the second assumption in the statement of the theorem and by Lemma 6.16 for $x \in \mathfrak{h}_+$, $j \in \{1,\dots,n\}$ and $w \in \mathcal{W}$

$$\varphi_x(x_1) \geq \varphi_x(x_j) \geq \varphi_x(w.x_j),$$

and

$$\varphi_x(-w^*.x_1) = \varphi_{-x}(w^*.x_1) \geq \varphi_{-x}(w^*.x_j) \geq \varphi_{-x}(ww^*.x_j) = \varphi_x(-ww^*.x_j).$$

Therefore

$$\varphi_x(x_1 - w^*.x_1) \geq \varphi_x(w.x - w'.x'),$$

for $x, x' \in \{x_1,\dots,x_n\}$ and $w, w' \in \mathcal{W}$ and we conclude that the normal cone to P at $x_1 - w^*.x_1$ includes $\mathfrak{h}_+$. Since this argument is Weyl group invariant, for $w \in \mathcal{W}$ the normal cone to P at $w.(x_1 - w^*.x_1)$ includes $w.\mathfrak{h}_+$. We have $\cup_{w\in\mathcal{W}} w.\mathfrak{h}_+ = \mathfrak{h}$, hence by Lemma 6.18 the orbit $\mathcal{W}.(x_1 - w^*.x_1)$ are all the extreme points of P and the normal cone to P at $w.(x_1 - w^*.x_1)$ is $w.\mathfrak{h}_+$. □

Example 6.22. An example of this is the action of SU(3) on the singular (co)adjoint orbit containing $\lambda = i\,\mathrm{diag}(2,-1,-1)$. The Hofer norm polytope is the convex hull of 0 and the permutations of $i\,\mathrm{diag}(3,0,-3)$. Its extreme points are the permutations of $i\,\mathrm{diag}(3,0,-3)$. This is the same Hofer norm polytope as the one derived from the (co)adjoint action on the regular (co)adjoint orbit containing $\frac{i}{2}\,\mathrm{diag}(-3,0,3)$. Several other examples can be computed from the cases of SU(3) acting on products of $\mathbb{P}^2$ studied in [MoSha18]. In fact, it is easy to see from all the figures in [MoSha18] that all the Kirwan polytopes listed there lead by Theorem 6.21 to Hofer norm polytopes with extreme points $\mathcal{W}.y$ for regular y.

Corollary 6.23. *Let O_x for $x \in \mathfrak{h}_+$ be a coadjoint orbit. The Hofer norm polytope derived from O has extreme points $\mathcal{W}.y$ for regular y if and only if $x - w^*.x \in \mathfrak{h}_+$ is regular.*

Example 6.24. An example of the previous corollary is the case of (co)adjoint orbits of SU(n). Let y be an extreme point of P. The set of x such that φ_x has a unique maximum at y is an open cone so we can assume that all eigenvalues of x are different, and without loss of generality we assume that they are ordered increasingly: $x_1 < \cdots < x_n$. Hence if the eigenvalues of the (co)adjoint orbit are $\lambda_1 \leq \cdots \leq \lambda_n$ the extreme point of the Hofer norm polytope which maximizes φ_x is

$$i\,\mathrm{diag}(\lambda_1,\dots,\lambda_n) - i\,\mathrm{diag}(\lambda_n,\dots,\lambda_1) = i\,\mathrm{diag}(\lambda_1-\lambda_n, \lambda_2-\lambda_{n-1},\dots,\lambda_n-\lambda_1).$$

This is the same maximum as the one that would be obtained from looking at the Hofer norm polytope derived from the (co)adjoint orbit with eigenvalues $x_1 = \frac{i}{2}(\lambda_1-\lambda_n) < \cdots < x_n = \frac{i}{2}(\lambda_n-\lambda_1)$. If these eigenvalues are all distinct then the Hofer norm polytope is equal to the Hofer norm polytope of a regular (co)adjoint orbit.

Remark 6.25. For the one-sided Hofer norm polytope the condition for being derived from a regular coadjoint orbit, is that there exists a regular $x_i \in P^+$, say x_1 such that for $x \in \mathfrak{h}_+$ we have $\varphi_x(x_1) \geq \varphi_x(x_j)$ for $j \in \{2,\dots,n\}$, where

$$P^+ = \mathrm{conv}\{w.x : x \in \{x_1,\dots,x_n\}, w \in \mathcal{W}\}.$$

The extreme points of the Hofer norm polytope are Weyl group invariant, so we can partition them into orbits

$$\mathrm{ext}(P) = \mathcal{W}.y_1 \sqcup \cdots \sqcup \mathcal{W}.y_m$$

for $y_1, \ldots, y_m \in \mathfrak{h}_+$. If $y_1, \ldots, y_m$ are regular then by Theorem 6.9 the unit ball B has abelian faces. We next give a sharper characterization of these faces.

Proposition 6.26. *If* $\mathrm{ext}(P) = \mathcal{W}.y_1 \sqcup \cdots \sqcup \mathcal{W}.y_m$ *for regular* $y_1, \ldots, y_m \in \mathfrak{h}_+$, *then for* $i \in \{1, \ldots, m\}$

$$F_{y_i}(B) = \{x \in \mathfrak{h}_+ : \varphi_x(y_i) = 1 \textit{ and } \varphi_x(y_i) \leq 1 \textit{ for } i \neq j\}.$$

Proof. From Theorem 6.7 we know that if the faces of B are abelian then

$$F_x(B) = F_x(B \cap \mathfrak{h}) \cap \mathfrak{h}_+$$

for a Weyl chamber $\mathfrak{h}_+$ given by a choice of torus and positive simple roots such that $x \in \mathfrak{h}_+$. We take for simplicity $y_i = y_1$, note that

$$\begin{aligned} F_{y_1}(B) &= F_{y_1}(B \cap \mathfrak{h}) \cap \mathfrak{h}_+ = F_{y_1}(P^\circ) \cap \mathfrak{h}_+ \\ &= \{x \in \mathrm{bd}\, P^\circ : H_{y_1} \text{ supports } P^\circ \text{ at } x\} \\ &= \{x \in \mathrm{bd}\, P^\circ : H_x \text{ supports } P \text{ at } y_1\} \\ &= \{x \in \mathfrak{h}_+ : \varphi_x(y_1) = 1 \text{ and } \varphi_x(y_j) \leq 1 \text{ for } j = 2, \ldots, n\}. \end{aligned}$$

Note that H_x supports P at y_1 is equivalent to $\varphi_x(y_1) = 1$ and $\varphi_x(w.y_j) \leq 1$ for $w \in \mathcal{W}$ and $j = 1, \ldots, n$. This in turn is equivalent by Lemma 6.16 to $\varphi_x(y_1) = 1$ and $\varphi_x(y_j) \leq 1$ for $w \in \mathcal{W}$ and $j = 2, \ldots, n$ since $x \in \mathfrak{h}_+$. This establishes the last equality. The third and fourth equalities follow from previous results on polar duality. □

Remark 6.27. We now have an alternative proof of Proposition 6.3: the extreme points of the Hofer norm polytope are $\mathcal{W}.y$ for a regular $y \in \mathfrak{h}_+$, hence $F_y(B) = \{x \in \mathfrak{h}_+ : \varphi_x(y) = 1\}$. The cone generated by this face is $\mathfrak{h}_+$.

Next we obtain conditions in terms of Kirwan's polytope which imply that $y_1, \ldots, y_m \in \mathfrak{h}_+$ are regular. This follows from the previous discussion therefore we omit the proof.

Theorem 6.28. *Let* $E \subseteq \mathfrak{k}$ *be an* Ad*-invariant set such that* $E \cap \mathfrak{h}_+ = \mathrm{conv}\{x_1, \ldots, x_n\}$. *The Hofer norm polytope derived from* E *has regular extreme points if the extreme points* y *of*

$$A = \mathrm{conv}\{x_i - w^*.x_j : i, j = 1, \ldots, n\}$$

such that the normal cone $N(y, A)$ *intersects the interior of* $\mathfrak{h}_+$ *are all regular.*

6.4.2. *Products of actions.* We now study products of Hamiltonian actions where the image of the moment map is given by (4.2). A property of regularity on one of the factors implies the same property in the product, in the following two cases:

Consider the case of the Hofer norm $\|\cdot\|_{O_1 + \cdots + O_n}$ for coadjoint orbits $O_1, \ldots, O_n$. This norm arises by (4.2) from the canonical symplectic action of K on $O_1 \times \cdots \times O_n$.

Proposition 6.29. *Let $P_1,\dots,P_n$ be $\mathcal{W}$-invariant convex polytopes in $\mathfrak{h}$ such that* $\operatorname{ext}(P_i) = \mathcal{W}.y_i$ *for $y_i \in \mathfrak{h}_+$ and $i = 1,\dots,n$. Then* $\operatorname{ext}(P_1+\cdots+P_n) = \mathcal{W}.(y_1+\cdots+y_n)$. *Hence, if there is a y_i that is regular, then the extreme points of $P_1+\cdots+P_n$ are the Weyl group orbit of a regular element.*

Proof. Let x be a point in the interior of $\mathfrak{h}_+$. Then, by Lemma 6.16 $\varphi_x(y_i) > \varphi_x(w.y_i)$ for $i = 1,\dots,n$ and $w \in \mathcal{W}\setminus \operatorname{Stab}(y_i)$. Hence φ_x has a unique maximum in P_i at y_i. This implies that φ_x has a unique maximum in $P_1+\cdots+P_n$ at $y_1+\cdots+y_n$, i.e. $y_1+\cdots+y_n$ is an extreme point of $P_1+\cdots+P_n$ and $x \in N(P_1+\cdots+P_n, y_1+\cdots+y_n)$. Since the normal cones are closed $\mathfrak{h}_+ \subseteq N(P_1+\cdots+P_n, y_1+\cdots+y_n)$ and since this argument is Weyl group invariant

$$w.\mathfrak{h}_+ \subseteq N(P_1+\cdots+P_n, w.(y_1+\cdots+y_n))$$

for $w \in \mathcal{W}$. By Lemma 6.18 all the extreme points are $\mathcal{W}.(y_1+\cdots+y_n)$. If there is a $j \in \{1,\dots,n\}$ such that y_j is regular, then this point is in the interior of the cone $\mathfrak{h}_+$ so that $y_1+\cdots+y_n$ is also in the interior, i.e. is regular. □

From Proposition 6.20 we get the following

Corollary 6.30. *Let $O_1,\dots,O_n$ be (co)adjoint orbits and let $\|\cdot\|_{O_1+\cdots+O_n}$ be the* Ad-*invariant Hofer norm defined by $O_1+\cdots+O_n$. If at least one coadjoint orbit is regular, then the Hofer norm polytope derived from $\|\cdot\|_{O_1+\cdots+O_n}$ has extreme points equal to $\mathcal{W}.y$ for a symmetric regular $y \in \mathfrak{h}_+$, i.e. it is the Hofer norm polytope derived from a regular coadjoint orbit.*

Remark 6.31. By Corollary 2.11 a cone $C \subseteq \mathfrak{k}$ generated by a face has the same norming functional, so the set $\{\varphi_x : x \in C\}$ has a common maximizer $x^+ \in O_1+\cdots+O_n$ and a common minimizer $x^- \in O_1+\cdots+O_n$. Let us write $x^- = x_1^- + \cdots + x_n^-$ and $x^+ = x_1^+ + \cdots + x_n^+$ with $x_i^-, x_i^+ \in O_i$ for $i = 1,\dots,n$. Then by Proposition 2.14

$$C_{x^-,x^+}(O_1+\cdots+O_n) = \bigcap_{i=1,\dots,n} C_{x_i^-,x_i^+}(O_i).$$

If O_1 is regular then by Theorem 6.3 $C_{x_1^-,x_1^+}(O_1)$ is contained in a Weyl chamber (given by a choice of torus and positive simple roots).

Conversely, if $\mathfrak{h}_+$ is a Weyl chamber (given by a choice of torus and positive simple roots) we choose x_i^+ to be the intersection of O_i with $\mathfrak{h}_+$ and x_i^- to be the intersection of O_i with $-\mathfrak{h}_+$. By Proposition 6.2 the Weyl chamber $\mathfrak{h}_+$ is contained in $C_{x_i^-,x_i^+}(O_i)$ for $i = 2,\dots,n$ and by Proposition 6.3 $C_{x_1^-,x_1^+}(O_1) = \mathfrak{h}_+$. Hence if we write $x^- = x_1^- + \cdots + x_n^-$ and $x^+ = x_1^+ + \cdots + x_n^+$ we get by Proposition 2.14

$$C_{x^-,x^+}(O_1+\cdots+O_n) = \bigcap_{i=1,\dots,n} C_{x_i^-,x_i^+}(O_i) = \mathfrak{h}_+,$$

which is a set with the same norming functionals. We conclude again that the Weyl chambers (given by a choice of torus and positive simple roots) are the cones generated by maximal faces.

We now turn to $\mathcal{W}$-invariant polytopes such that its extreme points are more than one $\mathcal{W}$-orbit.

Proposition 6.32. *Let $P_1,\dots,P_n$ be $\mathcal{W}$-invariant convex polytopes in $\mathfrak{h}$. If the extreme points of one of the polytopes are all regular then the extreme points of $P_1+\cdots+P_n$ are all regular.*

Proof. Let y be an extreme point of $P_1+\cdots+P_n$. There exists an $x\in\mathfrak{h}$ such that φ_x attains its unique maximum in $P_1+\cdots+P_n$ at y. Since the normal cone to the polytope $P_1+\cdots+P_n$ at y is open we can chose a regular x, therefore $\mathrm{Stab}(x)$ is trivial. This regular x is in a unique Weyl chamber which we denote by $\mathfrak{h}_+$. We have $y=y_1+\cdots+y_n$, where $y_1,\dots,y_n$ are points where φ_x attains its unique maximums in $P_1,\dots,P_n$ respectively. Since for $i=1,\dots,n$ we have $\mathcal{W}.y_i\subseteq P_i$ and φ_x attains a maximum at y_i Lemma 6.16 implies that $y_1,\dots,y_n$ are in $\mathfrak{h}_+$. Since one of the y_i is in the interior of $\mathfrak{h}_+$ so is their sum $y=y_1+\cdots+y_n$, hence y is regular. □

To finish this paper, from Proposition 6.32 and Theorem 6.14 we obtain the following structural property of geodesics (with different metrics) in the group K:

Theorem 6.33. *Let Let $K\curvearrowright M_1,\dots,K\curvearrowright M_n$ be Hamiltonian almost effective actions of a compact semi-simple group K and let $K\curvearrowright M_1\times\cdots\times M_n$ be the product action. Endow K with the pullback metrics of Section 3.2.1. Assume that each short curve in K with the metric derived from the action on one factor has commuting Hamiltonians: then all short curves in K with the metric derived from the product action have commuting Hamiltonians.*

Remark 6.34. Another proof of the previous theorem can be given as follows. If the extreme point of one of the Hofer norm polytopes, say P_1, are regular, then all the maximal faces of the sphere of the norm $\|\cdot\|_{\mu_1(M)}$ are abelian. The cones of the form $C_{x_1^-,x_1^+}(\mu_1(M))$ for $x_1^-,x_1^+\in\mu_1(M)$ are contained in cones generated by faces and are therefore abelian. The cone generated by faces of the sphere of the norm $\|\cdot\|_{\mu(M)}$ are contained in $C_{x^-,x^+}(\mu(M))$ for $x^-,x^+\in\mu(M)$. Also

$$C_{x^-,x^+}(\mu(M))=\bigcap_{i=1,\dots,n}C_{x_i^-,x_i^+}(\mu_i(M))$$

for $x_i^-,x_i^+\in\mu_i(M)$ and $i=1,\dots,n$ by Proposition 2.14. Since the first set of the intersection is abelian the result follows.

Acknowledgments

This research was supported by Instituto Argentino de Matemática (CONICET) and Universidad de Buenos Aires.

References

[AnLVa14] J. Antezana, G. Larotonda, A. Varela, *Optimal paths for symmetric actions in the unitary group*. Comm. Math. Phys. 328 (2014), no. 2, 481–497.

[AnGhiSt19] J. Antezana, E. Ghighlioni, D. Stojanoff, *Minimal curves in $U(n)$ and $Gl(n)^+$ with respect to the spectral and the trace norms*, J. Math. Anal. Appl. 483 (2020), no. 2, 123632.

[AtBo83] M. F. Atiyah, R. Bott, *The Yang-Mills equations over Riemann surfaces*. Philos. Trans. Roy. Soc. London Ser. A 308 (1983), no. 1505, 523–615.

[Bar02] A. Barvinok, *A course in convexity*. Graduate Studies in Mathematics, 54. American Mathematical Society, Providence, RI, 2002.

[Bel08] P. Belkale. *Quantum generalization of the Horn conjecture*. J. Amer. Math. Soc. 21 (2008), no. 2, 365–408.

[BhaHo89] R. Bhatia, J. A. R. Holbrook. *A softer, stronger Lidskiĭ theorem*. Proc. Indian Acad. Sci. Math. Sci. 99 (1989), no. 1, 75–83.

[Bhatia97] R. Bhatia, *Matrix analysis*. Graduate Texts in Mathematics, 169. Springer-Verlag, New York, 1997.

[BiaPol94] M. Bialy, L. Polterovich. *Geodesics of Hofer's metric on the group of Hamiltonian diffeomorphisms*. Duke Math. J. 76 (1994), no. 1, 273–292.

[BiGhiHe14] L. Biliotti, A. Ghigi, P. Heinzner. *Coadjoint orbitopes.* Osaka J. Math. 51 (2014), no. 4, 935–968.

[BoBo10] A. Borovik, A. Borovik. *Mirrors and reflections*. The geometry of finite reflection groups. Universitext. Springer, New York, 2010.

[BoLeZh99] J. M. Borwein, A. S. Lewis, Q. J. Zhu, *Convex spectral functions of compact operators. II. Lower semicontinuity and rearrangement invariance. Optimization and related topics* (Ballarat/Melbourne, 1999), 179–196, Appl. Optim., 47, Kluwer Acad. Publ., Dordrecht, 2001.

[Bron83] A. Brøndsted. *An introduction to convex polytopes*. Graduate Texts in Mathematics, 90. Springer-Verlag, New York-Berlin, 1983.

[BuBuIv01] D. Burago, Y. Burago, S. Ivanov. *A course in metric geometry*. Graduate Studies in Mathematics, 33. American Mathematical Society, Providence, RI, 2001.

[EaPe77] M. L. Eaton, M. D. Perlman, *Reflection groups, generalized Schur functions, and the geometry of majorization.* Ann. Probability 5 (1977), no. 6, 829–860.

[Entov01] Entov, Michael. *K-area, Hofer metric and geometry of conjugacy classes in Lie groups*. Invent. Math. 146 (2001), no. 1, 93–141.

[GuSja05] V. Guillemin, R. Sjamaar. *Convexity properties of Hamiltonian group actions*. CRM Monograph Series, 26. American Mathematical Society, Providence, RI, 2005.

[Hof90] Hofer, H. On the topological properties of symplectic maps. Proc. Roy. Soc. Edinburgh Sect. A 115 (1990), no. 1-2, 25–38.

[HoTa99] R. R. Holmes, T.-Y. Tam, *Distance to the convex hull of an orbit under the action of a compact Lie group*. J. Austral. Math. Soc. Ser. A 66 (1999), no. 3, 331–357.

[Kna02] A. W. Knapp. *Lie groups beyond an introduction*. Second edition. Progress in Mathematics, 140. Birkhäuser Boston, Inc., Boston, MA, 2002.

[LaMcD95a] F. Lalonde, D. McDuff. *Hofer's L^∞-geometry: energy and stability of Hamiltonian flows. I, II.* Invent. Math. 122 (1995), no. 1, 1–33, 35–69.

[LaMcD95b] F. Lalonde, D. McDuff. *The geometry of symplectic energy*. Ann. of Math. (2) 141 (1995), no. 2, 349–371.

[Lar19] G. Larotonda. *Metric geometry of infinite dimensional Lie groups and their homogeneous spaces*, Forum Math. 31 (2019), no. 6, 1567–1605.

[Lew98] A. S. Lewis. *Eigenvalue-constrained faces*. Linear Algebra Appl. 269 (1998), 159–181.

[Lew00] A. S. Lewis. *Convex analysis on Cartan subspaces*. Nonlinear Anal. 42 (2000), no. 5, Ser. A: Theory Methods, 813–820.

[LiTsi89] C.-K. Li, N.-M. Tsing. *Norms that are invariant under unitary similarities and the C-numerical radii*. Linear and Multilinear Algebra 24 (1989), no. 3, 209–222.

[McD02] D. McDuff. *Geometric variants of the Hofer norm*. J. Symplectic Geom. 1 (2002), no. 2, 197–252.

[MOZZ03] N. McCarthy, D. Ogilvie, N. Zobin, V. Zobin. *Convex geometry of Coxeter-invariant polyhedra.* Trends in Banach spaces and operator theory (Memphis, TN, 2001), 153–179, Contemp. Math., 321, Amer. Math. Soc., Providence, RI, 2003.

[McDSa17] D. McDuff, D. Salamon. *Introduction to symplectic topology*. Third edition. Oxford Graduate Texts in Mathematics. Oxford University Press, Oxford, 2017.

[Mil76] J. Milnor. *Curvatures of left invariant metrics on Lie groups*. Advances in Math. 21 (1976), no. 3, 293–329.

[MoSha18] J. Montaldi, A. Shaddad *Non-Abelian momentum polytopes for products of CP^2*, J. Geom. Mech. 11 (2019), no. 4, 575–599.

[Pol01] L. Polterovich. *The geometry of the group of symplectic diffeomorphisms.* Lectures in Mathematics ETH Zurich. Birkhauser Verlag, Basel, 2001.

[PolShe15] L. Polterovich, E. Shelukhin. *Autonomous Hamiltonian flows, Hofer's geometry and persistence modules*. Selecta Math. (N.S.) 22 (2016), no. 1, 227–296.

[Rock70] R. T. Rockafellar, *Convex analysis*. Princeton Mathematical Series, No. 28 Princeton University Press, Princeton, N.J. 1970

[Serre01] J.-P. Serre, *Complex semisimple Lie algebras*. Translated from the French by G. A. Jones. Reprint of the 1987 edition. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2001.