Invariant Metrics on Lie Groups

Bi-invariance and One-Parametre Subgroups

V.A. Purho

Master's Thesis

Department of Mathematics and Statistics
University of Jyväskylä
Finland
2022/9/15

Invariant Metrics on Lie Groups

Bi-invariance and One-Parametre Subgroups

V.A. Purho

Abstract

We show that an admissible left-invariant geodetic metric on a connected Lie group is bi-invariant if and only if every one-parametre subgroup $t \mapsto \exp(tX)$ is a geodesic.

Mathematics Subject Classification – MSC2020: `22E, 53C60, 51F`.

Contents

0 Johdanto (in Finnish)

Tässä pro gradu -tutkielmassa todistan Enrico Le Donne -professorin ja Sebastiano Nicolussi Golo -tutkijatohtori -ohjaajieni kanssa, että Lien ryhmällä oleva vasemmastasainen, geodeettinen ja moniston topologian virittävä etäisyys on kahdestasainen, jos ja vain jos Lien ryhmän yksimuuttujaiset aliryhmät ovat geodeesejä. Milnorin tutkielmasta [7] seuraa, että tälläinen ryhmä on $\mathbb{R}^n$ Euklidelan ja K tiheän Lien ryhmän Cartesilainen tulo. Tulos yleistää yksimuuttujaisten aliryhmien geodeettisyyden ja kahdestasaisuuden yhtäpätevyyden Latifin ja Toomanian todistamasta [5] yleisille vasemmastasaisille, geodeettisille ja moniston topologian virittäville etäisyyskuvauksille.

Tutkielma alkaa eräitä alkeellisia kahdestasaisten etäkkeiden eli metriikoiden ominaisuuksia kerraten. Sitten Finslerin tapaiset etäkkeet esitellään. Kolmosluvussa johdetaan kaksi kahdestasaisuuden ominaistusta. Nelosluvussa tavataan Berestovskiin lause, joka kertoo vasemmastasaisten, geodeettisten ja moniston topologian virittävien etäisyyksien olevan aliFinslerin etäkkeitä. Viitosluku on lyhyt Minkowskilaisiin etäkkeisiin vievä johdanto, joille Latifin ja Toomanian oman vastaavan lauseensa todistivat. Kuutos- ja seitsosluvuissa yhtäpätevyyden päälause todistetaan.

Kiitän ohjaajiani heidän antamastaan tuesta ja kärsivällisyydestä.

1 Introduction (englanniksi)

In this thesis we will derive some basic results on invariant distances on Lie groups. We will show that a Lie group that admits a bi-invariant distance function also admits a bi-invariant Riemannian metric and so, by a theorem in Milnor's paper [7], the group must be the Cartesian group product of a Euclidean space $\mathbb{R}^n$ and a compact group. Afterwards we will consider left-invariant geodetic metrics that generate the manifold topology on Lie groups. Berestovskii showed in the 1980s that on each Lie group these are in one-to-one correspondance with bracket-generating subspaces W of the Lie algebra $\mathfrak{g}$ of G, such that W is equipped with a norm N.

The main result of the thesis is Theorem 7.2 which shows that an admissible left-invariant geodetic metric is bi-invariant if and only if all the one-parametre subgroups $t \mapsto \exp(tX)$ are geodesics. By geodesic we mean a locally length minimising path. This is a generalisation of the same characterisation for Minkowskian Finsler metrics given by Latifi and Toomanian in [5], and of the classical Riemannian result found in Milnor [7].

Theorem 7.2. *Let d be an admissible left-invariant geodetic metric on a connected Lie group G. Then the following are equivalent*

1. *d is bi-invariant;*

2. *All the one-parametre subgroups $t \mapsto \exp(tX)$ for $X \in \mathfrak{g}$ are geodesics.*

The reader is expected to know basic theory of differentiable manifolds, and that of Lie groups.

The thesis was supervised by Professor Enrico Le Donne, and Postdoctoral Researcher Sebastiano Nicolussi Golo. I am grateful for their time, guidance, and advice.

2 Bi-invariant Distances and Metrics

Definition 2.1. We call a distance d on a group G *left-invariant* if it is preserved by left-translations, that is, if for all $g, x, y \in G$

$$d(gx, gy) = d(x, y).$$

We call d *right-invariant* if it is preserved by right-translations. We call the distance *bi-invariant* if it is both left- and right-invariant.

Proposition 2.2. If d is a left-invariant distance on a group then the following are equivalent:

1. d is bi-invariant;
2. d is inversion invariant, that is, $d(x, y) = d(x^{-1}, y^{-1})$;
3. conjugations are isometries.

Proof. **1 implies 2:** If the distance is bi-invariant, then

$$d(x^{-1}, y^{-1}) = d(e, xy^{-1}) = d(y, x) = d(x, y).$$

2 implies 3: Denote left-multiplication with g by l_g. Conjugation by g can be written as

$$c_g(x) = gxg^{-1} = l_g \circ (y \mapsto y^{-1}) \circ l_g \circ (y \mapsto y^{-1})(x)$$

and is therefore an isometry if inversion is an isometry.

3 implies 1: Finally, if conjugations are isometries, we calculate

$$\begin{aligned} d(hg, kg) &= d(ghgg^{-1}, gkgg^{-1}) \\ &= d(gh, gk) = d(h, k) \end{aligned}$$

which shows that the distance is bi-invariant. □

Heuristically, on a Lie group it makes sense to ask for lengths of paths to be integrals of a potential of their derivative. For this end, we define the notions of Finsler and subFinsler metrics. Later on, in Section 4, we discover that subFinsler metrics are, in fact, the most general admissible left-invariant geodetic distances on Lie groups.

Definition 2.3 (Finsler Metric)**.** On a Lie group G, for each $x \in G$ let M_x be a norm on T_xG. If the mapping $x \mapsto M_x$ is continuous, that is, if

$$x \mapsto M_x(X_x)$$

is continuous for all continuous vector fields X, then $x \mapsto M_x$ is called a *Finsler metric.*

Definition 2.4. A *polarisation* on a Lie group G is a constant rank subbundle of the tangent bundle TG. Polarisations are also called *distributions.*

Definition 2.5 (subFinsler Metric)**.** Let $\Delta \subset TG$ be a polarisation on a Lie group G. For each $x \in G$, let N_x be a norm on $\Delta_x \subset T_xG$. If the mapping $x \mapsto N_x$ is continuous with respect to *horizontal* vector fields, that is, if

$$x \mapsto N_x(X_x)$$

is continuous for all continuous vector fields X such that $X_x \in \Delta_x$ for all $x \in G$, then $x \mapsto N_x$ is called a *subFinsler metric.*

Obviously every Finsler metric is also subFinsler. A Riemannian metric is a Finsler metric when one considers the norm generated by the inner product on every tangent space. In the context of left-invariant subFinsler metrics on Lie groups we will usually identify the subFinsler metric with the norm of the metric on the Lie algebra $\mathfrak{g}$ of G.

Next we define the length of paths with respect to subFinsler metrics:

Definition 2.6 (Length of Paths)**.** Let N be a subFinsler metric defined on a distribution $\Delta \subset TG$, and let $\gamma\colon [a,b] \to G$ be a horizontal piecewise-C^1 path, that is, a path such that for almost every $t \in [a,b]$ the derivative $\dot\gamma(t)$ exists and $\dot\gamma(t) \in \Delta_{\gamma(t)}$. If γ is everywhere C^1 the length ℓ of γ is defined to be

$$\ell(\gamma) \coloneqq \int N_{\gamma(t)}\left(\dot\gamma(t)\right) \mathrm{d}t.$$

Otherwise γ can be written as the concatenation of C^1 paths $\gamma_{0\cdots n}$ and we define the length of γ to be

$$\ell(\gamma) \coloneqq \sum_i \ell(\gamma_i).$$

If γ isn't horizontal almost everywhere we define its length to be infinite.

Definition 2.7 (Carnot-Carathéodory distances)**.** The *Carnot-Carathéodory distance* induced by a subFinsler metric N is defined as

$$d(x,y) \coloneqq \inf\left\{\ell(\gamma) \mid x, y \in \mathrm{Im}(\gamma)\right\}$$

where $\ell(\gamma)$ is defined using the subFinsler metric N, and γ ranges over paths on G that are piecewise-C^1. We implicitly identify a subFinsler metric and the Carnot-Carathéodory distance induced by the metric.

It is a well known result that any two points on a connected component of G can be joined by a finite-length path if the distribution of the subFinsler metric is bracket-generating, that is, if the Lie algebra generated by horizontal vector fields is the whole tangent bundle. These kinds of distributions are sometimes also referred to as *non-holonomic*, and in the context of partial differential equations the condition is called the *Hörmander condition*. By abuse of notation, we will refer to subFinsler metrics whose distribution is bracket-generating as *bracket-generating*.

We state the following useful lemma:

Lemma 2.8. *Let ℓ be a lower semi continuous length structure, d_ℓ the distance function induced by ℓ and $\tilde{\ell}$ the length structure induced by d_ℓ, then*

$$\tilde{\ell} = \ell.$$

The lengths of paths in the sense of metric geometry induced by a Carnot-Carathéodory distance agree with the subFinsler lengths ℓ used to define the Carnot-Carathéodory distance. In other words, the following formula holds for all piecewise-C^1 paths:

$$\sup_{t_0<\cdots<t_n} \sum d(\gamma(t_i), \gamma(t_{i+1})) = \ell(\gamma).$$

A length structure induced by a metric is lower semi continuous.

Proof. See for example Gromov 1.6. Proposition [3], and Le Donne Proposition 2.3.6. [6]. □

Lemma 2.9. *Let W be a subspace of the Lie algebra $\mathfrak{g}$, and let N be a norm defined on W. Extending W and N to a left-invariant subFinsler metric on G, we obtain a subFinsler distance. If W is bracket-generating, then this distance generates the manifold topology on G.*

Proof. We refer the reader to Berestovskii Theorem 1. [2]. □

The previous two lemmas tell us that the Carnot-Carathéodory distance induced by a *bracket-generating* subFinsler metric N is a length metric in the sense of metric geometry, and that this distance generates the manifold topology on G.[1] Because Lie groups are locally compact complete spaces, any two points on the Lie group can be joined by a geodesic of the induced distance by the Hopf-Rinow theorem.

A (sub)Finsler metric is a structure on the tangent bundle TG so it can be pulled back using left and right translations. Therefore, it makes sense to ask for metrics to be left-, right- or bi-invariant, too:

[1] We note that this nomenclature is unfortunate as a metric is what we call a distance; however, check Theorem 2.13.

Definition 2.10. We call a Finsler or a subFinsler *metric N left-invariant* if it is preserved by pushforwards of left-translations, that is, if for all $g, h \in G$ and $X \in T_h G$

$$N_{gh}(l_{g*}(X)) = N_h(X)$$

where l_{g*} is the pushforward of the left-translation by g. We call the metric right-invariant if it preserved by pushforwards of right-translations. We call the metric bi-invariant if it is left- and right-invariant.

Lemma 2.11. *A left-invariant metric N is bi-invariant if and only if N_e is invariant under* Ad_G*.*

Proof. The adjoint map Ad_g is the composition of the pushforwards of l_g and $r_{g^{-1}}$. If the norm is bi-invariant, then it is invariant under Ad_g.

On the other hand, assume that N_e is invariant under Ad_G. Let $g, h \in G$, and let $X = l_{h*}(Y) \in T_h G$. Then, because left- and right-translations commute,

$$\begin{aligned} N_{hg}(r_{g*}(X)) &= N_e(l_{(hg)^{-1}}{}_* \circ r_{g*}(X)) \\ &= N_e(\mathrm{Ad}_g \circ l_{g^{-1}}{}_* \circ l_{h^{-1}*} \circ r_{g*}(X)) \\ &= N(X). \end{aligned}$$

□

Lemma 2.12. *If $\Delta \subset TG$ is a bi-invariant bracket-generating polarisation then it is the whole tangent bundle $\Delta = TG$. Especially, every bi-invariant subFinsler metric is Finsler.*

Proof. Let Δ_e be the polarisation at the identity. Because Δ is bi-invariant

$$\Delta_e = (l_{g*} \circ r_{g^{-1}}{}_*)\big|_e (\Delta_e) = c_{g*}\big|_e (\Delta_e) = \mathrm{Ad}_g \Delta_e.$$

Let us calculate the bracket of two vectors $X \in \mathfrak{g}$, and $Y \in \Delta_e$:

$$[X, Y] = \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{Ad}_{\exp(tX)} Y.$$

Since the map $t \mapsto \mathrm{Ad}_{\exp(tX)} Y$ takes values in Δ_e, which is an embedded submanifold of the Lie algebra $\mathfrak{g}$, the bracket $[X, Y]$ is in Δ_e too. The distribution Δ is bracket-generating so $[\mathfrak{g}, \Delta_e] \subset \Delta_e$ implies that $\Delta_e = \mathfrak{g}$. □

We have now defined two different notions of invariance: one for distance functions, and one for metrics on the tangent bundle. Because every subFinsler metric induces a distance on the Lie group it is natural to ask if these two concepts agree for Finsler metrics. This is, in fact, the case for bi-invariant Finsler metrics and their induced distances:

Proposition 2.13. The Carnot-Carathéodory distance induced by a Finsler metric N on a connected Lie group G is a bi-invariant distance if and only if the Finsler metric N is bi-invariant.

Proof. Because pullbacks of the tangents of one-parametre subgroups to the identity element are constant:

$$l^*_{\exp(tX)} \frac{\mathrm{d}}{\mathrm{d}t} \exp(tX) = X$$

the length of the path $\gamma_X(t) = \exp(tX) \colon [0,1] \to G$ equals the norm of X at the identity T_eG

$$\ell(\gamma_X) = \int_0^1 N(X)\mathrm{d}t = N(X).$$

Using Lemma 2.8, it can readily be seen from the length of a path as a sum of distances of points on the curve that isometries preserve lengths of paths. We conclude by applying this to the isometry $c_g(x) := gxg^{-1}$:

$$\begin{aligned} N(X) = \int_0^1 N(X)\mathrm{d}t &= \ell(\gamma_X) \\ &= \ell(c_g \circ \gamma_X) \end{aligned}$$

using the formula $c_g(\exp(X)) = \exp(\mathrm{Ad}_g(X))$:

$$= \int_0^1 N(\mathrm{Ad}_g X)\mathrm{d}t = N(\mathrm{Ad}_g X).$$

In the other direction: Let $B^d_e(\varepsilon)$ be an open ball for the Carnot-Carathéodory distance centered at e such that the exponential map is a diffeomorphism from an open subset $\mathfrak{b} \subset \mathfrak{g}$ to $B^d_e(\varepsilon)$. Because $\mathfrak{b}$ is open, there exists an open ball $B^N_0(\varepsilon') \subset \mathfrak{b}$ for the Finsler metric N. Define $B := \exp(B^N_0(\varepsilon'))$. Let $\gamma \colon [0,1] \to G$ be a geodesic joining e and $x \in B$. We can write γ using the exponential map as $\gamma(t) = \exp(X(t))$. We will use Duhamel's formula, which is the coordinate expression for the derivative of the exponential map pulled back to the origin. We refer the reader to, for example, Varadarajan Formula 2.14.5. [8]. By abuse of notation, we leave the pullback implicit and simply write $(\mathrm{D}\exp)_X$.

(Duhamel's Formula) $$(\mathrm{D}\exp)_X = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} (\mathrm{ad}_X)^n.$$

Because the Lie algebra automorphisms Ad_g are isomorphisms for N_e, the path $t \mapsto \exp(\mathrm{Ad}_g X(t)) \colon [0,1] \to G$ is contained in the set B for all

$g \in G$. We calculate $\mathrm{D}(\exp \circ \mathrm{Ad}_g X(t))$:

$$\begin{aligned}
\mathrm{D}(\exp \circ \mathrm{Ad}_g X(t)) &= (\mathrm{D} \exp)_{\mathrm{Ad}_g X(t)} \circ \mathrm{D}(\mathrm{Ad}_g X(t)) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} (\mathrm{ad}_{\mathrm{Ad}_g X(t)})^{\circ n} (\mathrm{D}(\mathrm{Ad}_g X(t))) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} (\mathrm{ad}_{\mathrm{Ad}_g X(t)})^{\circ n} (\mathrm{Ad}_g \circ \mathrm{D}(X(t))) \\
&= \mathrm{Ad}_g \left(\sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} (\mathrm{ad}_{X(t)})^{\circ n} (\mathrm{D}(X(t))) \right).
\end{aligned}$$

Likewise calculating $\mathrm{D}(\exp \circ X(t))$ we see that

$$\begin{aligned}
\mathrm{D}(\exp \circ X(t)) &= \mathrm{D} \exp_{X(t)} \circ \mathrm{D}(X(t)) \\
&= \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!} (\mathrm{ad}_{X(t)})^{\circ n} (\mathrm{D}(X(t))),
\end{aligned}$$

and we obtain the equality $\mathrm{D}(\exp \circ \mathrm{Ad}_g X(t)) = \mathrm{Ad}_g(\mathrm{D} \exp \circ X(t))$.

Let γ' be the path $t \mapsto \exp(\mathrm{Ad}_g X(t))\colon [0,1] \to G$ still denoting the path $t \mapsto \exp(X(t))\colon [0,1] \to G$ by γ. Because Ad_g is an isometry for the norm, the paths γ and γ' have the same length. The distance between e and $c_g(x)$ is the infimum of lengths of paths joining the two points so

$$\begin{aligned}
d(e, c_g(x)) &\leq \ell(c_{g*}\gamma) \\
&= \ell(\gamma') \\
&= \ell(\gamma) \\
&= d(e, x).
\end{aligned}$$

Because $x = c_{g^{-1}} \circ c_g(x)$, a similar argument shows that $d(e,x) \leq d(e, gxg^{-1})$ and therefore $d(e, gxg^{-1}) = d(e,x)$ for all $x \in B$.

Let y be an arbitrary point of G. By geodecity of the metric, for any $n \in \mathbb{N}_+$ we can find n points $x_i \in G$ for $1 \leq i \leq n$ such that $d(x_i, x_j) = \frac{|i-j|}{n+1} \cdot d(e,y)$ for all $0 \leq i,j \leq n+1$, where we let $x_0 \coloneqq e$ and $x_{n+1} \coloneqq y$. Let n be so large that for any i, $x_i^{-1} x_{i+1} \in B$. Then calculating for arbitrary

$g \in G$

$$\begin{aligned} d(e,y) &= \sum_{i=0}^{n} d(x_i, x_{i+1}) \\ &= \sum_{i=0}^{n} d(e, x_i^{-1} x_{i+1}) \\ &= \sum_{i=0}^{n} d(e, c_g(x_i^{-1} x_{i+1})) \\ &= \sum_{i=0}^{n} d(e, c_g(x_i^{-1}) c_g(x_{i+1})) \\ &= \sum_{i=0}^{n} d(c_g(x_i), c_g(x_{i+1})) \\ &\geq d(e, c_g(y)) \end{aligned}$$

and because $y = c_{g^{-1}} \circ c_g(y)$ we obtain the opposite inequality. We conclude with Proposition 2.2 which showed equivalence between bi-invariance and conjugations being isometries.

□

From now on, we will make no distinction between a Finsler metric being bi-invariant, and the Carnot-Carathéodory distance induced by the metric being bi-invariant. By abuse of notation, we will refer to a Finsler metric and the Carnot-Carathéodory distance induced by the metric as *metrics* whenever this is convenient to us.

3 Two Characterisations

Proposition 3.1. Continuing the classification of Proposition 2.2: Let d be the Carnot-Carathéodory distance induced by a bracket-generating left-invariant subFinsler metric N. The following statements are equivalent on *connected* Lie groups:

1.' d is bi-invariant;

4. Ad_g is an isometry of the Finsler norm for all elements $g \in G$;

5. The norm is constant on Ad-orbits;

6. For almost every $X \in \mathfrak{g}$ and for all $A \in \mathfrak{g}$,

$$DN|_X([A, X]) = 0.$$

Proof. Using Lemma 2.11 and Proposition 2.13 we already know that the conditions 4, and 1' are equivalent. The other equivalences are trivial:

4 is equivalent to 5: This follows because all orbits are of the form $\mathrm{Ad}_G(X)$ for some $X \in \mathfrak{g}$.

5 implies 6: Due to Rademacher's theorem the norm, being a Lipschitz function, is differentiable almost everywhere, and because N is constant on Ad-orbits the derivative of N along an orbit must equal 0:

Using the chain rule and evaluating the derivative at $t = 0$ we calculate:

$$\begin{aligned} 0 &= \frac{d}{dt}\bigg|_{t=0} N(\mathrm{Ad}_{\exp(tA)} X) \\ &= \frac{d}{dt}\bigg|_{t=0} N(e^{\mathrm{ad}\, tA} X) \\ &= \frac{d}{dt}\bigg|_{t=0} N(X + t[A, X] + O(t^2)) \\ &= DN\big|_X([A, X]). \end{aligned}$$

6 implies 5: By analysis, a Lipschitz-function that is almost everywhere differentiable, and has 0 derivative is constant. Using the same calculation as in the previous case, and the fact that G is connected, we see that N is constant on Ad-orbits.

□

We give two applications of Proposition 3.1:

Corollary 3.2. *If G is a non-Abelian connected Lie group then it admits a non-bi-invariant left-invariant Riemannian metric.*

Proof. A metric N is non-bi-invariant if and only if there is an element X of the Lie algebra whose adjoint orbit isn't stable for the norm: that is there is an element $h \in G$ such that

$$N(\mathrm{Ad}_h X) \neq N(X).$$

We begin by observing that there exists an element $h \in G$ such that $\mathrm{Ad}_h \notin \{\mathrm{Id}, -\mathrm{Id}\}$. In fact, because the group isn't Abelian the map Ad_g can't identically be the identity operator. Because $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$ is continuous there must exist an element $h \in G$ such that $\mathrm{Ad}_h \notin \{\mathrm{Id}, -\mathrm{Id}\}$.

Fix an element $X \in \mathfrak{g}$ such that $\mathrm{Ad}_h(X) \notin \{X, -X\}$. Now we can find an inner product $\langle -, - \rangle$ such that the unit sphere of $\langle -, - \rangle$ contains X but not $\mathrm{Ad}_h(X)$. □

Corollary 3.3. *On the sphere $\mathbf{S}^3$ when realised as the Lie group of units of the quaternions, up to scalar multiplication, there is only one unique bi-invariant Finsler structure, and it is the usual Riemannian sphere structure.*

Proof. The orbit of any non-zero $X \in T_p\mathbf{S}^3$ under the differentials of the isotropy group of p spans the whole tangent space under multiplication by positive scalars so this must be a λ-sphere for our norm for some $\lambda > 0$. We know that the sphere $\mathbf{S}^3$ admits a bi-invariant Riemannian structure so by uniqueness of the adjoint orbits the statement follows. □

Next, we prove another classification theorem, but this time for the *existence* of a bi-invariant distance on a connected Lie group. We start with a couple of lemmas:

Lemma 3.4 (Milnor. Lemma 7.5.)**.** *A connected Lie group G 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, that is $\mathbb{R}^n$.*

Proof. We refer the reader to Milnor Lemma 7.5. [7]. □

Lemma 3.5. *Let G be a connected Lie group. If the adjoint group Ad_G is precompact in the space $\mathrm{GL}(\mathfrak{g})$, then G admits a bi-invariant Riemannian metric.*

If G admits a bi-invariant Riemannian metric, then its adjoint group Ad_G is compact.

Proof. Take any inner product $\langle\cdot,\cdot\rangle$ on $\mathfrak{g}$. Because Ad_G is a precompact Lie subgroup of the general linear group of $\mathfrak{g}$ we can define a new function on $\mathfrak{g}$ by integrating over the compact Lie group $K := \overline{\mathrm{Ad}_G} \subset \mathrm{GL}(\mathfrak{g})$:

$$\langle X, Y\rangle' := \int_{\varphi\in K} \langle \varphi(X), \varphi(Y)\rangle \,\mathrm{d}\varphi.$$

Extend this function to the whole tangent space TG using left-translations $l_{g^{-1}}$. It is trivial to check that the new function $\langle\cdot,\cdot\rangle'$ is invariant under Ad_G, is symmetric, bilinear, and positive definite. Hence it is a bi-invariant Riemannian metric on G.

For the second part, if $\langle\cdot,\cdot\rangle$ is a bi-invariant Riemannian metric on G, then Ad_g is an isometry for the inner product $\langle\cdot,\cdot\rangle$; therefore, Ad_G belongs to the orthogonal group $O(\langle\cdot,\cdot\rangle)$ so Ad_G is precompact. To show that it is, in fact, compact we will use Lemma 3.4: If G admits a bi-invariant Riemannian metric, then G factors as $G = K \times \mathbb{R}^n$, where K is a compact Lie group. Conjugation in G can be written as

$$(a, b)(x, y)(a^{-1}, -b) = (axa^{-1}, y)$$

so the adjoint group of G is isomorphic to the adjoint group of K. The adjoint map $\mathrm{Ad}\colon K \to \mathrm{Aut}(\mathfrak{k})$ is a smooth map, and a continuous image of a compact set is compact; therefore, Ad_G is compact. □

Definition 3.6. We call a metric on a Lie group G *admissible* if it generates the manifold topology on G.

Theorem 3.7. *Let G be a connected Lie group. Then the following are equivalent:*

1. *There exists an admissible bi-invariant distance on G;*
2. *There exists a bi-invariant Riemannian metric on G;*
3. *G is the direct product of a compact group and a vector group, that is, $\mathbb{R}^n$ for some $n \in \mathbb{N}$;*
4. *Ad_G is compact in the space of linear transformations of the Lie algebra $\mathfrak{g}$.*

Note that Iwasawa's theorem tells us that any Lie group is topologically the product of a compact group and a vector group, but if (and only if) the group admits an admissible bi-invariant metric, then this product can be made to respect the group law.

Proof. The implication $2 \implies 1$ follows from Proposition 2.13. The equivalence between 2 and 3 is Lemma 3.4, and the equivalence between 2 and 4 is Lemma 3.5. Therefore; the only implication we have to show is the implication $1 \implies 2$.

Let $\mathfrak{b}$ be an open bounded subset of $\mathfrak{g}$ such that the exponential map is a diffeomorphism from $\mathfrak{b}$ onto an open neighbourhood of $e \in G$. Because the distance d is admissible there exists an open ball $B_e^d(\lambda) \subset \exp(\mathfrak{b})$. Because d is bi-invariant conjugations are isometries for d. Take an element $x \in B_e^d(\lambda)$. The points $c_g \circ \exp(X) = \exp(\mathrm{Ad}_g X)$, and $\exp(X)$ are equidistant from the identity e. Therefore, $\mathrm{Ad}_g X$ belongs to the bounded set $\mathfrak{b}$ for all $g \in G$. In order to establish that precompactness of Ad_G in $\mathrm{GL}(n)$ agrees with the precompactness of the inclusion of Ad_G to the manifold of all matrices $\mathrm{M}_{n\times n}$, we need to show that Ad_G stays away from the variety of zero-determinental matrices;

Hadamard's inequality tells us that the determinant of a matrix is bounded by the norms of its columns. Because all the image vectors of the linear transformation Ad_g are bounded, the determinant of Ad_g is bounded by some positive real number S. Because Ad_G forms a group this means that the determinants of the inverses $\mathrm{Ad}_{g^{-1}}$ are bounded by S^{-1}. Therefore $\mathrm{Ad}_G \subset \mathrm{GL}(n) \subset \mathrm{M}_{n\times n}$ forms a bounded subgroup and hence it is precompact. The adjoint representation of G is precompact and Lemma 3.5 tells us that the group admits a bi-invariant Riemannian metric. □

As an application of Theorem 3.7 we mention the following:

Corollary 3.8. *The Heisenberg group doesn't admit an admissible bi-invariant distance.*

Proof. In fact, the Heisenberg group is homeomorphic to $\mathbb{R}^3$, but isn't Abelian. □

We note that we can't drop the requirement that the distance be admissible, as the degenerate separating distance $d(x, y) := 1$ whenever $x \neq y$ is bi-invariant on all groups.

4 Berestovskii's Theorem

Theorem 4.1 (Berestovskii's Theorem)**.** *Let d be an admissible left-invariant geodetic metric on a connected Lie group. Then d is subFinsler, and bracket-generating.*

Every left-invariant bracket-generating subFinsler distance on a connected Lie group is admissible, geodetic, and (trivially) left-invariant.

Proof. We refer the reader to Berestovskii 1984 [2]. □

Proposition 4.2. Let G be a Lie group equipped with an admissible left-invariant geodetic distance d. If d is bi-invariant, then d is Finsler.

Proof. By Theorem 4.1, d is a left-invariant subFinsler metric. Because d is subFinsler, it is defined on a polarisation $\Delta \subset TG$. By Proposition 2.12, the polarisation Δ must equal the whole tangent bundle. □

Proposition 4.3. Let G be a Lie group equipped with a left-invariant geodetic distance d. If all one-parametre subgroups of G are geodesics, then d is Finsler.

Proof. Assume to the contrary that $X \in T_eG$ has infinite norm. Let γ_X be the one-parametre subgroup $t \mapsto \exp(tX) \colon [0, 1] \to G$. Then

$$\begin{aligned} l(\gamma_X) &= \int_0^1 N_e(X)\,\mathrm{d}t \\ &= N_e(X) \\ &= \infty \end{aligned}$$

which contradicts the assumption that γ_X is a geodesic. □

5 Minkowskian Finsler Metrics

Definition 5.1. Let G be a differentiable manifold. A continuous real function N on the tangent bundle TG is called a *Minkowskian Finsler* metric if

1. $N(\lambda X) = \lambda N(X)$ for $X \in T_g G$, and $\lambda \geq 0$;[2]

2. N is smooth on TM outside of the zero section;

3. The fundamental tensor

$$g_v(X,Y) \coloneqq \frac{1}{2}\frac{\partial^2}{\partial s \partial t}\left((N(v+sX+tY))^2\right)\Big|_{s=t=0}$$

is positive-definite on any tangent space $T_g G$ for every $0 \neq v \in T_g G$.

A norm N on vector space V is called *Minkowskian* if it satisfies the conditions 1, and 3, and is smooth outside of $0 \in V$. A left-invariant Finsler metric is a Minkowskian Finsler metric if and only if the norm on the Lie algebra is a Minkowskian norm. The positive-definiteness of the fundamental tensor implies subadditivity of the norm, and that its unit ball is strongly convex in the sense explained below in Proposition 5.3. *This further implies that all geodesics are uniquely determined by their derivative at a point* (Check for example the book by Bao, Chern, and Shen [1].)

The following proof is from a paper by Latifi and Toomanian [5].

Theorem 5.2 (Latifi-Toomanian)**.** *If (G,d) is a Lie group with an admissible bi-invariant geodetic metric such that the induced norm on $\mathfrak{g}$ is Minkowskian, then all the one-parametre subgroups are geodesics.*

Proof. Let γ be a geodesic through the identity element e, with $\gamma(0) = e$. For all $g \in G$ define $\sigma_g(x) \coloneqq g x^{-1} g$. These are isometries by Proposition 2.2. The conjugation σ_g is involutive because $\sigma_g^2 = \mathrm{id}$, and has g as an isolated fixed point because the differential at g is $-\mathrm{id}$:

$$\begin{aligned} \mathrm{d}\sigma_g \circ \mathrm{d}l_g(X) &= \mathrm{d}l_g \circ \mathrm{d}(y \mapsto y^{-1}) \circ \mathrm{d}l_{g^{-1}} \circ \mathrm{d}l_g\, X \\ &= \mathrm{d}l_g\,(-X) \\ &= -\mathrm{d}l_g\, X. \end{aligned}$$

Because all the geodesics are uniquely determined by their tangent vector at a point, and because the involutive symmetry has differential $-\mathrm{id}$ at the point e, the geodesic γ must be reversible, that is, γ traversed backwards

[2] While in Finsler geometry it is commonplace to assume that the involved norms need not be symmetric, note that because we are dealing with distance functions that are symmetric, that is $d(x,y) = d(y,x)$, λ can be taken to be any real number, assuming we replace the right-hand-side by $|\lambda| N(X)$.

is still a geodesic. Applying the involutive symmetry σ_e to the point $\gamma(t)$ we see that $\gamma(t)$ must be reflected a distance $d(e, \gamma(t))$ about the identity element. We obtain $\sigma_e(\gamma(t)) = \gamma(-t)$ for t small enough for the geodesic to be distance minimising. Applying the symmetry $\sigma_{\gamma(c)}$ to the point $\gamma(-t)$ it must be reflected a distance $d(\gamma(c), \gamma(-t))$ about $\gamma(c)$ this time obtaining

$$\sigma_{\gamma(c)} \circ \sigma_e(\gamma(t)) = \gamma(c - (-t - c)) = \gamma(2c + t)$$

once again for c, and t small enough for the geodesic to be distance minimising on the required interval. On the other hand, by definition of the maps $\sigma_{(\cdot)}$, this point must equal

$$\begin{aligned} \sigma_{\gamma(c)} \circ \sigma_e(\gamma(t)) &= \sigma_{\gamma(c)}(e.\gamma(t)^{-1}.e) \\ &= \gamma(c)\gamma(t)\gamma(c) \\ &= \gamma(c)^2\gamma(t). \end{aligned}$$

Substituting in $t = 0$ and using induction, we arrive at the formula

$$\gamma(nc) = \gamma(c)^n$$

and if $c_1/c_2 \in \mathbb{Q}$ such that $c_1 = \alpha n_1$, $c_2 = \alpha n_2$ with n_{12} integers then

$$\gamma(c_1 + c_2) = \gamma((n_1 + n_2)\alpha) = \gamma(\alpha)^n_1\gamma(\alpha)^{n_2} = \gamma(c_1)\gamma(c_2).$$

By continuity this is true for all c_{12} small enough. Hence $\gamma(t)$ coincides with a one-parametre subgroup in a neighbourhood of $e \in G$. Because the metric is left-invariant the geodesic is a one-parametre subgroup globally.

Now let us prove that every one-parametre subgroup is a geodesic: Under the Minkowskian norm hypothesis the geodesics leaving a point are characterised one-to-one by their derivatives. Let $\exp(tX)$ be a one-parametre subgroup and let γ be a geodesic leaving the identity with derivative X. We know that γ must be a one-parametre subgroup having derivative X at the identity, hence it must equal $\exp(tX)$, because any one-parametre subgroup is characterised by its derivative in the Lie algebra. This shows that all the one parametre subgroups are geodesics. □

Proposition 5.3. Let N be a norm on a vector space V, and let B be its closed unit ball. The norm N is Minkowskian if and only if ∂B is a smooth hypersurface embedded in V and B is strongly convex, that is, the second fundamental form σ^ξ of ∂B is positive definite with respect to any transverse vector ξ pointing out to $\mathring{B}$.

Proof. We refer the reader to Javaloyes and Sánchez Proposition 2.3. [4]. □

6 Minkowskian Approximations of Norms

Lemma 6.1. *Let N be a norm on a vector space V. Let $G \subset \mathrm{GL}(V)$ be a group. If N is G-invariant, then there exists a Euclidean norm $|\cdot|$ that is also G-invariant, and any Lebesque measure of V is G-invariant.*

Proof. Since N is G-invariant the elements of G are isometries of (V, N) that fix 0. Using the Ascoli-Arzelà theorem G is a precompact subgroup of $\mathrm{GL}(V)$. Maximal compact subgroups of $\mathrm{GL}(V)$ are conjugate copies of $O(n)$; therefore, there exists a scalar product $\langle -, - \rangle$ such that G is a group of isometries of $(V, \langle -, - \rangle)$. Finally, observe that $|\det(g)| = 1$ for $g \in O(n)$. $\square$

Proposition 6.2. If N is a G-invariant norm on a vector space V, then there exists a sequence (N_i) of norms on V such that:

1. N_i is smooth on $V \setminus \{0\}$;
2. $N_i \to N$ uniformly on compact sets;
3. N_i is Minkowskian;
4. N_i is G-invariant.

In other words, the norm N can be approximated by a sequence of G-invariant Minkowskian norms.

Proof. Using Lemma 6.1 there exists a G-invariant Euclidean norm $|\cdot|$ on V. Let ρ_ε a family of mollifiers such that $\mathrm{spt}(\rho_\varepsilon) \subset B_0^{|\cdot|}(\varepsilon)$, and $\rho_\varepsilon(x) = \rho_\varepsilon(y)$ for all x, and y such that $|x| = |y|$. Define

$$\tilde{N}_\varepsilon(v) := \int_V N(v - w)\rho_\varepsilon(w)\mathrm{d}w.$$

Notice that $\tilde{N}_\varepsilon$ is smooth but it is not a norm, because norms aren't smooth at the origin. For this reason, we will use the level set of $\tilde{N}_\varepsilon$ to define a norm N_ε.

Let $L > 0$ be a real number such that $|v|/L \leq N(v) \leq L|v|$. We claim that for all $v \in V$

$$N(v) - L\varepsilon \leq \tilde{N}_\varepsilon(v) \leq N(v) + L\varepsilon. \tag{1}$$

Indeed, since $\int \rho_\varepsilon(w)\mathrm{d}w = 1$, and $\mathrm{spt}(\rho_\varepsilon) \subset B_0^{|\cdot|}(\varepsilon)$, we immediately obtain

$$\inf\{N(x) : |v - x| \leq \varepsilon\} \leq \tilde{N}_\varepsilon(v) \leq \sup\{N(x) : |v - x| \leq \varepsilon\}. \tag{2}$$

By the reverse triangle inequality, we also have for all v, x with $|v - x| \leq \varepsilon$

$$N(v) - L\varepsilon \leq N(x) \leq N(v) + L\varepsilon.$$

Thus, we obtain (1).

Fix $\varepsilon < 1/L$ and set

$$B_\varepsilon \coloneqq \{v : \tilde{N}_\varepsilon(v) \le 1\}.$$

We claim that B_ε is the unit ball of a norm: This is equivalent to showing that 0 belongs to the interior of B_ε, and that B_ε is closed, symmetric, convex, and bounded. Indeed, B_ε is closed because $\tilde{N}_\varepsilon$ is continuous, and by (1), $\tilde{N}_\varepsilon(0) \le L\varepsilon < 1$ since $\varepsilon < 1/L$, thus 0 belongs to the interior of B_ε. The set B_ε is symmetric because

$$\begin{aligned} \tilde{N}_\varepsilon(-v) &= \int N(-v-w)\rho_\varepsilon(w)\mathrm{d}w \\ &= \int N(v+w)\rho_\varepsilon(w)\mathrm{d}w \end{aligned}$$

with a change of variables $w \mapsto -w$:

$$= \int N(v-w)\rho_\varepsilon(-w)\mathrm{d}w$$

using the fact that ρ_ε is symmetric:

$$\begin{aligned} &= \int N(v-w)\rho_\varepsilon(w)\mathrm{d}w \\ &= \tilde{N}_\varepsilon(v). \end{aligned}$$

Moreover, B_ε is convex because $\tilde{N}_\varepsilon$ is convex; in other words, for all $a, b \ge 0$ with $a + b = 1$ and for all $u, v \in B_\varepsilon$,

$$\begin{aligned} \tilde{N}_\varepsilon(au+bv) &= \int N(au+bv-w)\rho_\varepsilon(w)\mathrm{d}w \\ &= \int N(a(u-w)+b(v-w))\rho_\varepsilon(w)\mathrm{d}w \\ &\le a\int N(u-w)\rho_\varepsilon(w)\mathrm{d}w + b\int N(v-w)\rho_\varepsilon(w)\mathrm{d}w \\ &= a\tilde{N}_\varepsilon(u) + b\tilde{N}_\varepsilon(v) \le 1. \end{aligned}$$

Finally, B_ε is bounded because, by (1), if $\tilde{N}_\varepsilon(v) \le 1$, then $N(v) \le 1 + L\varepsilon$, and thus $B_\varepsilon \subset B_0^N(1 + L\varepsilon)$.

This concludes the proof that B_ε is the unit ball of a norm, which we shall denote by N_ε.

Next, we claim that N_ε is smooth on $V \setminus \{0\}$: This is equivalent to the statement that ∂B_ε is a smooth graph in the radial direction by

Lemma 6.3. To prove our claim, we will show that $\tilde{N}_\varepsilon$ has strictly positive radial derivative on ∂B_ε, that is, for all v with $\tilde{N}_\varepsilon(v) = 1$ we have

$$\left.\frac{\mathrm{d}}{\mathrm{d}h}\right|_{h=0} \tilde{N}_\varepsilon(v + hv) > 0. \tag{3}$$

Since N is Lipschitz,

$$\left.\frac{\mathrm{d}}{\mathrm{d}h}\right|_{h=0} \tilde{N}_\varepsilon(v + hv) = \int \left.\frac{\mathrm{d}}{\mathrm{d}h}\right|_{h=0} N(v + hv - w)\rho_\varepsilon(w)\mathrm{d}w.$$

Notice that, whenever N is differentiable at $v-w$, with $w \in \mathrm{spt}(\rho_\varepsilon) \subset B_0^{|\cdot|}(\varepsilon)$, we have

$$\begin{aligned}
\left.\frac{\mathrm{d}}{\mathrm{d}h}\right|_{h=0} N(v + hv - w) &= \lim_{h\to 0} \frac{N(v - w + hv) - N(v - w)}{h} \\
&= \lim_{h\to 0} \Bigg(\frac{N(v - w + h(v - w)) - N(v - w)}{h} \\
&\qquad\qquad + \frac{N(v - w + hv) - N(v - w + h(v - w))}{h} \Bigg) \\
&\geq N(v - w) - \limsup_{h\to 0} \frac{N(v - w + hv) - N(v - w + h(v - w))}{h}
\end{aligned}$$

grouping terms and using the reverse triangle inequality

$$\begin{aligned}
&\geq N(v - w) - N(w) \\
&\geq |v - w|/L - L|w| \\
&\geq |v|/L - (L + 1/L)\varepsilon.
\end{aligned}$$

Using the inequality $|v| \geq N(v)/L$ one can easily check that, if $\varepsilon < \frac{1}{L\cdot(1+L^2)}$, then $|v|/L - (L + 1/L)\varepsilon > 0$ and thus the strict inequality (3) holds.

Next, we claim that the spheres ∂B_N and $\partial B_{N_\varepsilon}$ are $(L^2\varepsilon)$-close in the Hausdorff sense with respect to the chosen norm $|\cdot|$:

On the one hand, let $x_0 \in \partial B_{N_\varepsilon}$ and consider Inequality 2 again:

$$\inf\left\{N(x_0 - w) : w \in B_0^{|\cdot|}(\varepsilon)\right\} \leq \tilde{N}(x_0) \leq \sup\left\{N(x_0 - w) : w \in B_0^{|\cdot|}(\varepsilon)\right\}.$$

Using continuity of N and the Intermediate Value Theorem, if $\tilde{N}_\varepsilon(x_0) = 1$, then there must be a point $y_0 \in B_{x_0}^{|\cdot|}(\varepsilon)$ so that $N(y_0) = 1$. Hence, for any point x_0 on $\partial B_{N_\varepsilon}$, there is a point y_0 on ∂B_N, that is ε-close to x_0. On the other hand, let $y_0 \in \partial B_0^N(1)$. By (1), we have that for $h > 0$,

$$h - L\varepsilon \leq \tilde{N}_\varepsilon(hy_0) \leq h + L\varepsilon.$$

It follows that $\tilde{N}_\varepsilon(hy_0) \geq 1$ if $h = 1+L\varepsilon$, where $|hy_0 - y_0| = L\varepsilon|y_0|$. Similarly, $\tilde{N}_\varepsilon(hy_0) \leq 1$ if $h = 1 - L\varepsilon$, where $|hy_0 - y_0| = L\varepsilon|y_0|$.
Now, since $\sup\{|y| : y \in \partial B_0^N(1)\} \leq L$, we have obtained that there are two points y_+ and y_- in $B_{y_0}^{|\cdot|}(L^2\varepsilon)$ such that $\tilde{N}_\varepsilon(y_+) \geq 1 \geq \tilde{N}_\varepsilon(y_-)$. By the Intermediate Value Theorem, there exists a point $x_0 \in B_{y_0}^{|\cdot|}(L^2\varepsilon)$ with $\tilde{N}_\varepsilon(x_0) = 1$, i.e., $x_0 \in \partial B_0^{N_\varepsilon}(1)$.

Next, we claim that the norm N_ε is G-invariant: Let g be an element of G, and let $v \in V$ such that $N_\varepsilon(gv) = 1$. Then

$$\begin{aligned}
N_\varepsilon(gv) &= \tilde{N}_\varepsilon(gv)\\
&= \int_V N(g(v) - w)\rho_\varepsilon(w)\,\mathrm{d}w\\
&= \int_V N(g(v) - g\cdot g^{-1}(w))\rho_\varepsilon(g\cdot g^{-1}(w))\,\mathrm{d}w\\
&= \int_V N(g(v) - g(w))\rho_\varepsilon(g(w))\,\mathrm{d}w\\
&= \int_V N(v - w)\rho_\varepsilon(w)\,\mathrm{d}w\\
&= \tilde{N}_\varepsilon(v) = N_\varepsilon(v)
\end{aligned}$$

which shows that the norms N_ε are G-invariant.

Finally: The fact that $N_\varepsilon \to N$ uniformly on compact sets as $\varepsilon \to 0$ follows immediately from Lemma 6.4.

For any $i \in \mathbb{N}_+$ the norm

$$N_i' := \sqrt{(1 - 1/i)N^2_{\varepsilon=1/i} + 1/i|\cdot|^2}$$

is Minkowskian; In fact, we check that the norm satisfies condition 3 of Definition 5.1:

$$\begin{aligned}
g_v(X,Y) &:= \frac{1}{2}\frac{\partial^2}{\partial s\partial t}\left((N_i'(v + sX + tY))^2\right)\Big|_{s=t=0}\\
&= \frac{1}{2}\frac{\partial^2}{\partial s\partial t}\left((1-\varepsilon)N_\varepsilon^2 + \varepsilon|\cdot|^2\right)\\
&= \frac{1}{2}(1-\varepsilon)\frac{\partial^2}{\partial s\partial t}N_\varepsilon^2 + \frac{1}{2}\varepsilon\frac{\partial^2}{\partial s\partial t}|\cdot|^2\\
&= \frac{1}{2}(1-\varepsilon)\frac{\partial^2}{\partial s\partial t}N_\varepsilon^2 + \varepsilon\sum_i X_iY_i.
\end{aligned}$$

The latter summand is obviously positive-definite. It suffices to show that the first summand is not strictly negative: in fact, N_ε^2 is the composition of the non-decreasing convex real function $x \mapsto x^2$ and the function N_ε is convex because it is a norm; therefore the Hessian is positive semi-definite at all non-zero points.

We conclude that the sequence of norms

$$N_i' := \sqrt{(1-1/i)N_{\varepsilon=1/i}^2 + 1/i|\cdot|^2}$$

converges to N uniformly on compact sets, that the norms N_i' are strongly convex, are smooth on $V \setminus \{0\}$, and are G-invariant. This concludes the proof of Proposition 6.2. $\square$

Lemma 6.3. *Let N be a norm on a vector space V and let $\mathbb{S} \subset V$ be the unit sphere of a Euclidean norm $|\cdot|$. Then the unit sphere ∂B_N of N is a radial graph, that is, there is a map $r\colon \mathbb{S} \to \,]0,+\infty[$ such that*

$$\partial B_N = \{r(p)p : p \in \mathbb{S}\}.$$

In fact, we have $r(p) = 1/N(p)$ and thus, for all $v \in V \setminus \{0\}$,

$$N(v) = \frac{|v|}{r(v/|v|)}.$$

In particular, N is smooth on $V \setminus \{0\}$ if and only if the function r is smooth.

Proof. By definition of the function r, if $v \neq 0$ then

$$1 = N\left(\frac{r(v/|v|)}{|v|}v\right) = \frac{r(v/|v|)}{|v|}N(v).$$

$\square$

Lemma 6.4. *Let $|\cdot|$ be a Euclidean norm on a vector space V. Suppose that N_1 and N_2 are norms on V whose unit spheres have Hausdorff distance less than η with respect to $|\cdot|$. Let L_1 and L_2 be constants such that*

$$|x|/L_j \leq N_j(x) \leq L_j|x|$$

for all $x \in V$. Then

$$|N_1(x) - N_2(x)| \leq L_1L_2\eta|x|. \tag{4}$$

Proof. Firstly, if $N_2(\hat{x}) = 1$, then there exists $\hat{y}$ such that $N_1(\hat{y}) = 1$ and $|\hat{x} - \hat{y}| \leq \eta$. Therefore,

$$\begin{aligned} N_1(\hat{x}) &\leq N_1(\hat{y}) + N_1(\hat{x} - \hat{y}) \\ &\leq 1 + L_1|\hat{x} - \hat{y}| \\ &\leq 1 + L_1\eta. \end{aligned}$$

Similarly, $N_1(\hat{x}) \geq 1 - L_1\eta$.

Secondly, if $x \in V \setminus \{0\}$, then

$$N_1(x) = N_2(x) N_1\left(\frac{x}{N_2(x)}\right) \leq N_2(x)(1 + L_1\eta),$$

and $N_1(x) \geq N_2(x)(1 - L_1\eta)$.

Finally, for all $x \in V \setminus \{0\}$,

$$|N_1(x) - N_2(x)| = N_2(x) \cdot \left|\frac{N_1(x)}{N_2(x)} - 1\right| \leq L_2|x| \cdot L_1\eta,$$

which proves (4). □

7 Main Theorem

Lemma 7.1. *Let $|\cdot|$ be a Euclidean norm on $\mathfrak{g}$. For a fixed path, the length functional $\gamma \mapsto \ell_N(\gamma) = \int N(\dot\gamma)\,\mathrm{d}t$ is continuous in the norm variable N with respect to the Hausdorff distance induced by the norm $|\cdot|$ on $\mathfrak{g}$.*

Proof. Let γ be a C^1-path and fix a norm N on the Lie algebra $\mathfrak{g}$. Let M be another norm on $\mathfrak{g}$ such that the unit balls of N and M are $\varepsilon > 0$ apart. We calculate

$$\begin{aligned} |\ell_N(\gamma) - \ell_M(\gamma)| &= \left|\int N(\dot\gamma(t))\mathrm{d}t - \int M(\dot\gamma(t))\mathrm{d}t\right| \\ &\leq \int |N(\dot\gamma(t)) - M(\dot\gamma(t))|\,\mathrm{d}t \end{aligned}$$

By Lemma 6.4:

$$\begin{aligned} &\leq \int L_1 L_2 \varepsilon |\dot\gamma(t)| \mathrm{d}t \\ &\leq L_1 L_2 \varepsilon \ell_{|\cdot|}(\gamma), \end{aligned}$$

where L_1 and L_2 are the constant in Lemma 6.4. This shows that the length functional is continuous in norm. □

Theorem 7.2. *Let d be an admissible left-invariant geodetic metric on a connected Lie group G. Then the following are equivalent*

1. *d is bi-invariant;*
2. *All the one-parametre subgroups $t \mapsto \exp(tX)$ for $X \in \mathfrak{g}$ are geodesics.*

Proof. **Proof that bi-invariance implies that the one-parametre subgroups are geodesics:** By Lemma 4.2, d is a Finsler metric. If d is Minkowskian, then the theorem by Latifi and Toomanian (Theorem 5.2) tells us that all the one-parametre subgroups are geodesics.

Otherwise use Proposition 6.2 to find a sequence (N_i) of bi-invariant Minkowskian Finsler norms that approach the norm N of d in the Hausdorff sense on the Lie algebra. The norms N_i extend to bi-invariant distances d_i on the Lie group. Let $\mathfrak{u} \subset \mathfrak{g}$, and $U \subset G$ be two open neighbourhoods of $0 \in \mathfrak{g}$, and $e \in G$ such that the exponential map is a diffeomorphism from $\mathfrak{u}$ onto U. The distance $d(\{e\}, U^{\complement})$ is strictly positive so Lemma 7.1 tells us that the sequence of real numbers $d_i(\{e\}, U^{\complement})$ converges to $d(\{e\}, U^{\complement})$; therefore, the number $R \coloneqq \inf_i \left\{ d_i(\{e\}, U^{\complement}) \right\}$ is strictly positive.

Let L be a positive real number such that $1/LN < |\cdot| < LN$. We know that the ball $B_0^{(\cdot)}(R)$ of any norm N_i or N is included in the set $\mathfrak{u}$. For every norm N_i there exists a maximal $r_i \in \mathbb{R}_+^*$ such that $B_0^N(r_i) \subset B_0^{N_i}(R)$. Because the norms N_i converge to N in the Hausdorff sense the radii must converge to R; In fact, the intersection of $B_0^N(r_i)$ and $B_0^{N_i}(R)$ is non-empty so the Hausdorff distance of the unit balls of N_i and N must be greater than

$$\frac{1}{LR} \cdot |R - r_i|.$$

This means that the sequence (r_i) is bounded from below and there exists a strictly positive real number r' such that $B_0^N(r')$ is included in all the balls $B_0^{N_i}(R)$.

Let $t \mapsto \exp(tX)$ be an arbitrary one-parametre subgroup, and let t_* be a strictly positive real number such that t_*X is in the set $B_0^N(r')$. For all metrics d_i the length minimising path joining e and the point $p_* \coloneqq \exp(t_*X)$ is a piece of a one-parametre subgroup $t \mapsto \exp(tY)$. Because exp is a local diffeomorphism from $\mathfrak{u}$ onto U, if $Y \neq X$ the one-parametre subgroup must leave the set U to connect to the point p_*. The relation $B_0^{N_i}(R) \subset \mathfrak{u}$ implies that the length of such a path must be greater than R. The ball $B_0^N(r')$ is included in the ball $B_0^{N_i}(R)$ so the length of $t \mapsto \exp(tX)\colon [0, t_*] \to G$ is bounded above by R. This shows that the length minimising path joining e and p_* is the path $t \mapsto \exp(tX)\colon [0, t_*] \to G$ for all metrics d_i.

Suppose that γ is any other piecewise-C^1 path connecting the points e and p_*. Then we have the following strict inequality for any metric d_i

$$\ell_{N_i}(\gamma) > \ell_{N_i}(t \mapsto \exp(tX)\colon [0, t_*] \to G).$$

Using the continuity of the length functional we pass to the limit as i tends to infinity, and obtain

$$\ell_N(\gamma) \geq \ell_N(t \mapsto \exp(tX)\colon [0, t_*] \to G).$$

This shows that the one-parametre subgroup $t \mapsto \exp(tX)$ is a geodesic for the limit distance. We remark that because the strict inequality turns into a non-strict one we lose local uniqueness of geodesics. As an example, one can take the L^∞ norm on the Abelian Lie group $(R^2, +)$ and write it as a limit of L^p norms. For each $1 < p < \infty$ the geodesics are unique and are the usual straight lines. However for the limit norm L^∞ any graph of a 1-Lipschitz function is a geodesic.

Proof that if all the one-parametre subgroups are geodesics, then the metric d is bi-invariant By Proposition 3.1 it suffices to show that

$$\mathrm{d}N|_v\left([X,v]\right) = 0 \qquad \forall X \in \mathfrak{g},\ \text{for a.e. } v \in \mathfrak{g}.$$

Because the derivative operator is linear and the Lie bracket bilinear, by replacing v with λv when necessary we can assume that the one-parametre subgroup $t \mapsto \exp(tv)$ minimises the distance between $\exp(0v) = e$ and $\exp(1v)$.

Let $\gamma(t) \coloneqq \exp(tv)\colon [0,1] \to G$. We define a variation η of γ:

$$\eta(t) = \exp(tv + \varepsilon\varphi(t)X)$$

where $\varepsilon > 0$, and φ is a smooth function such that $\varphi(0) = 0 = \varphi(1)$. We have the relations $\eta(0) = \gamma(0)$, $\eta(1) = \gamma(1)$, and $l(\gamma) \leq l(\eta)$. We will calculate the derivative of the variation η:

$$\eta'(t) = \mathrm{d}\exp|_{tv+\varepsilon\varphi(t)X}\left[v + \varepsilon\varphi'(t)X\right]$$

Using Duhamel's Formula this equals

$$\eta'(t) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}\,\mathrm{ad}^n_{tv+\varepsilon\varphi X}\left[v + \varepsilon\varphi' X\right].$$

Because η is a variation of γ and the length of γ is the infimum of lengths of paths joining $\gamma(0)$ and $\gamma(1)$ we must have that:

$$\begin{aligned}
0 &= \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \int_0^1 N\left(\eta'(t)\right)\mathrm{d}t \\
&= \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \int_0^1 N\left(\sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}\,\mathrm{ad}^n_{tv+\varepsilon\varphi X}\left[v + \varepsilon\varphi' X\right]\right)\mathrm{d}t
\end{aligned}$$

because $N(\eta'(t))$ is a Lipschitz-function of ε we can exchange the derivative and the integral:

$$\begin{aligned}
&= \int_0^1 \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} N\left(\sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)!}\,\mathrm{ad}^n_{tv+\varepsilon\varphi X}\left[v + \varepsilon\varphi' X\right]\right)\mathrm{d}t \\
&= \int_0^1 \mathrm{d}N|_{\sum_{n=0}^{\infty}\frac{(-1)^n}{(n+1)!}\mathrm{ad}^n_{tv}[v]}\Bigg(\varphi' X + \sum_{n=1}^{\infty}\frac{(-1)^n}{(n+1)!}\Bigg(\\
&\qquad \sum_{j=1}^{n} \mathrm{ad}^{j-1}_{tv} \circ \mathrm{ad}_{\varphi X} \circ \mathrm{ad}^{n-j}_{tv}[v] + \mathrm{ad}^n_{tv}[\varphi' X]\Bigg)\Bigg)\mathrm{d}t,
\end{aligned}$$

where all the terms with more than one ε-variable were killed inside the brackets

$$= \int_0^1 \mathrm{d}N|_v \left(\varphi' X + \sum_{n=1}^{\infty} \frac{(-1)^n}{(n+1)!} (t^n \varphi' - t^{n-1}\varphi) \operatorname{ad}_v^n(X) \right) \mathrm{d}t$$

using integration by parts, and grouping terms that depend on t and v:

$$\begin{aligned}
&= \int_0^1 \mathrm{d}N|_v \left(\varphi' X\right) \mathrm{d}t + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!} \int_0^1 t^{n-1}\varphi(t)\, \mathrm{d}t \cdot \mathrm{d}N|_v \left(\operatorname{ad}_v^n(X)\right) = 0 \\
&= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!} \int_0^1 t^{n-1}\varphi(t)\, \mathrm{d}t \cdot \mathrm{d}N|_v \left(\operatorname{ad}_v^n(X)\right) = 0.
\end{aligned}$$

Fix a non-zero positive smooth function φ such that $\varphi(0) = 0$, $\varphi(t) = 0$ for all $t \geq 1$, and such that $\varphi \geq 0$. For any $0 < \lambda < 1$, define $\varphi_\lambda(t) := \varphi(t/\lambda)$; the new functions also satisfy the conditions $\varphi_\lambda(0) = \varphi(0/\lambda) = 0$, and $\varphi_\lambda(1) = \varphi(1/\lambda) = 0$, and they are smooth as compositions of two smooth functions. Therefore: (this is, in fact, a Mellin transform of φ)

$$\int_0^1 t^{n-1}\varphi_\lambda(t)\, \mathrm{d}t = \lambda^n \int_0^1 t^{n-1}\varphi\, \mathrm{d}t.$$

This gives us a geometric series in λ:

$$\begin{aligned}
&\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n!} \left(\int_0^1 t^{n-1}\varphi_\lambda(t)\, \mathrm{d}t \right) \cdot \mathrm{d}N|_v \left(\operatorname{ad}_v^n(X)\right) \\
&= \sum_{n=1}^{\infty} \lambda^n \frac{(-1)^{n+1}}{n!} \left(\int_0^1 t^{n-1}\varphi(t)\, \mathrm{d}t \right) \cdot \mathrm{d}N|_v \left(\operatorname{ad}_v^n(X)\right)
\end{aligned}$$

and it is a fact that if a geometric series vanishes everywhere on an open set then its coefficients are 0. Therefore, for each $n \in \mathbb{N}_+$ we have

$$\mathrm{d}N|_v \left(\operatorname{ad}_v^n(X)\right) = 0$$

especially for $n = 1$ we have

$$\mathrm{d}N|_v \left(\operatorname{ad}_v(X)\right) = 0$$

which is what we wanted to show.

$\square$

References

[1] Bao D. Chern S.-S. Shen Z. *An Introduction to Riemann-Finsler Geometry.* Springer. 2000.

[2] Berestovskii V. *Homogeneous Manifolds with Intrinsic Metric. I.* Sibirskii Matematicheskii Zhurnal, Vol. 29. 1988.

[3] Gromov M. *Metric Structures for Riemannian and Non-Riemannian Spaces.* Birkhäuser. 2001.

[4] Javaloyes M. Sánchez M. *On the Definition and Examples of Finsler Metrics.* Arxiv, 2011.

[5] Latifi D. Toomanian M. *On the existence of bi-invariant Finsler metrics on Lie groups.* Mathematical Sciences. 2013.

[6] Le Donne E. *Lecture notes on sub-Riemannian geometry.* Version of February 2021.

[7] Milnor J. *Curvatures of Left Invariant Metrics on Lie Groups.* Advances in Mathematics 21. 1976.

[8] Varadarajan V.S. *Lie Groups, Lie Algebras, and Their Representations.* Springer-Verlag. 1984.