JYU DISSERTATIONS 103

Eero Hakavuori

Sub-Riemannian Geodesics

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Mattilanniemen salissa MaA211
elokuun 10. päivänä 2019 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
in Mattilanniemi, auditorium MaA211, on August 10, 2019 at 12 o'clock noon.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2019

Editors
Enrico Le Donne
Department of Mathematics and Statistics, University of Jyväskylä
Ville Korkiakangas
Open Science Centre, University of Jyväskylä

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-7810-5

ISBN 978-951-39-7810-5 (PDF)
URN:ISBN:978-951-39-7810-5
ISSN 2489-9003

Acknowledgements

I want to thank my advisor Enrico Le Donne for all of his help and support. I have learned much – mathematically, meta-mathematically, and otherwise – from all of our discussions over the past several years.

I also wish to thank all of my friends and coworkers at the department currently and previously, particularly the bubble of sub-Riemannian people: Francesca, the two Sebastianos, Terhi, Thibaut and Ville.

I am grateful to the Vilho, Yrjö and Kalle Väisälä Foundation of the Finnish Academy of Science and Letters for financial support during the bulk of my doctorate studies, allowing me to concentrate my efforts on research.

Finally, a huge thanks to my wife Johanna, without whom I could not have stayed sane during the more arduous times of my studies.

Jyväskylä, June 25, 2019
Department of Mathematics and Statistics
Eero Hakavuori

List of included articles

This dissertation consists of an introductory part and the following three articles:

[A] Eero Hakavuori and Enrico Le Donne, *Non-minimality of corners in subriemannian geometry*, Invent. Math. **206** (2016), no. 3, 693–704. MR 3573971

[B] Eero Hakavuori and Enrico Le Donne, *Blowups and blowdowns of geodesics in Carnot groups*, arXiv e-prints (2018), arXiv:1806.09375.

[C] Eero Hakavuori, *Infinite geodesics and isometric embeddings in Carnot groups of step 2*, arXiv e-prints (2019), arXiv:1905.03214.

The author of this dissertation has actively taken part in the research of the joint articles [A] and [B].

Contents

1. Sub-Riemannian geodesics

1.1. **Sub-Riemannian manifolds.** The notion of a sub-Riemannian manifold is a generalization of the notion of a Riemannian manifold, where the main difference is that some directions of travel are explicitly forbidden. From a practical perspective, one may consider the examples of riding a bicycle, where the motion is always in the direction of the front wheel, or a robotic arm whose freedom of movement is constrained by the configuration of its joints. Such mechanical systems are one of the motivations for the study of sub-Riemannian geometry, and in particular, for the study of optimal control on sub-Riemannian manifolds.

A Riemannian manifold consists of a pair $(M, \langle\cdot,\cdot\rangle)$, where M is a smooth manifold and $\langle\cdot,\cdot\rangle$ a Riemannian metric. A sub-Riemannian manifold is a triple $(M, \Delta, \langle\cdot,\cdot\rangle)$, where the extra structure Δ is a smooth *bracket-generating* distribution $\Delta \subset TM$ determining the allowed directions of travel. For convenience all manifolds will be assumed to be connected. Bracket-generating means that the induced flag of (possibly singular) distributions eventually reaches the entire tangent bundle:

$$\Delta = \Delta_1 \subset \Delta_2 \subset \cdots \subset \Delta_s = TM.$$

Here the (singular) distributions Δ_k are determined by the at most k-th order iterated Lie brackets of sections of Δ, defined recursively by

$$\Delta_{k+1}(p) := \Delta_k(p) \oplus [\Delta, \Delta_k](p) \subset T_pM. \tag{1}$$

Distributions are often described by their *growth vector* at a point $p \in M$, i.e., the increasing sequence of integers

$$(\dim \Delta_1(p), \dim \Delta_2(p), \dots, \dim \Delta_s(p)).$$

Two other simple measures of complexity of the distribution are its *rank* $\dim \Delta(p)$ and *step*, which is the minimal integer $s \in \mathbb{N}$ such that $\Delta_s(p) = T_pM$ for all $p \in M$. A sub-Riemannian manifold is said to be *equiregular* when its growth vector is independent from the basepoint $p \in M$, i.e. when all Δ_k are non-singular distributions. Comprehensive treatises for more in depth information on sub-Riemannian geometry may be found in any of the books [31, 37, 3].

An absolutely continuous curve $\gamma\colon [0,1] \to (M, \Delta, \langle\cdot,\cdot\rangle)$ is called *horizontal*, if it is almost everywhere tangent to Δ, i.e., if $\dot\gamma(t) \in \Delta(\gamma(t))$ for almost every $t \in [0,1]$. The length $\ell(\gamma)$ of a horizontal curve $\gamma\colon [a,b] \to M$ is defined as in the Riemannian case by the integral of its speed

$$\ell(\gamma) := \int_a^b \|\dot\gamma(t)\|\, dt,$$

where the norm is $\|v\| = \sqrt{g(v,v)}$. The sub-Riemannian distance d_{SR} is defined similarly to the Riemannian distance, but with the extra horizontality restriction

$$d_{\mathrm{SR}}(x,y) = \inf\{\ell(\gamma) : \gamma \text{ horizontal curve from } x \text{ to } y\}. \tag{2}$$

It follows from the definition that the only finite length curves (in the usual metric space sense) are the horizontal curves. Indeed, it is sometimes helpful to think that $\|X\|^2 = g(X,X) = \infty$ for any non-horizontal vector $X \in TM \setminus \Delta$. In the definition of a sub-Riemannian manifold, the inner product g is usually only required to be defined on the distribution Δ. The restriction is mainly for convenience since there is no mathematical distinction – an inner product on the distribution Δ always has a Riemannian extension to an inner product on TM.

By the Chow-Rashevsky Theorem [35, 15], the bracket-generating assumption guarantees that the distance d_{SR} is finite. When the sub-Riemannian manifold is complete, the infimum in (2) is realized, so the distance d_{SR} is a geodesic distance. A curve $\gamma\colon [a,b] \to M$ realizing the infimum, i.e., a curve for which

$$d_{\mathrm{SR}}(\gamma(a), \gamma(b)) = \ell(\gamma),$$

is called a *length-minimizer*, or *geodesic*. A fundamental problem in sub-Riemannian geometry and the focal point of this dissertation is the following:

Geodesic Regularity Problem.
What is the regularity of geodesics in sub-Riemannian manifolds?

A priori, a sub-Riemannian geodesic is nothing more than a Lipschitz curve, but all known examples are much more regular – all known sub-Riemannian geodesics are smooth.

1.2. **Geodesics and control theory.** The language of control theory gives a way to locally parametrize horizontal curves. In any small enough open set $U \subset M$, the distribution Δ has an orthonormal frame of vector fields $X_1, \ldots, X_r$. Writing the derivative of a horizontal curve $\gamma\colon [a,b] \to U$ in this frame as

$$\dot{\gamma}(t) = \sum_{i=1}^{r} u_i(t) X_i(\gamma(t))$$

determines the *control* $u = (u_1, \ldots, u_r) \in \mathrm{L}^1([a,b];\mathbb{R}^r)$ of the horizontal curve γ. Conversely, fixing a basepoint $p \in M$, any control $u \in \mathrm{L}^1([0,1];\mathbb{R}^r)$ determines a unique absolutely continuous curve $\gamma_u\colon [0,1] \to M$ with $\gamma_u(0) = p$ and $\dot{\gamma}_u(t) = \sum_{i=1}^{r} u_i(t) X_i(\gamma_u(t))$. This defines the *endpoint map* based at p:

$$\mathrm{End}_p\colon \mathrm{L}^1([0,1];\mathbb{R}^r) \to M, \quad \mathrm{End}_p(u) = \gamma_u(1).$$

Orthonormality of the vector fields $X_1, \ldots, X_r$ implies that the sub-Riemannian norm can be computed by $\|\dot{\gamma}_u(t)\|^2 = \sum_{i=1}^{r} |u_i(t)|^2 = \|u(t)\|^2$, so the length of the curve γ_u is also given by the $\mathrm{L}^1(\mathrm{L}^2)$-norm of its control:

$$\ell(\gamma_u) = \int_0^1 \|u(t)\|\, dt = \|u\|_{\mathrm{L}^1(\mathrm{L}^2)}.$$

This rephrases the problem of finding a geodesic from p to q as an optimal control problem

$$\|u\|_{\mathrm{L}^1(\mathrm{L}^2)} \to \min, \quad \mathrm{End}_p(u) = q.$$

By reparametrizing the horizontal curve to have constant speed, minimizing the length $\|u\|_{\mathrm{L}^1(\mathrm{L}^2)}$ is equivalent to minimizing the L^2-energy $\|u\|_{\mathrm{L}^2}^2 = \int_0^1 \|u(t)\|^2\, dt$. This allows the problem to be studied in the Hilbert space $\mathrm{L}^2([0,1];\mathbb{R}^r)$ where the endpoint map is a smooth map $\mathrm{End}\colon \mathrm{L}^2([0,1];\mathbb{R}^r) \to M$, see e.g. [31, Appendix E] or [3, Section 8].

1.3. **Normal and abnormal extremals.** The optimal control problem

$$\|u\|_2 \to \min, \quad u \in \mathrm{L}^2([0,1];\mathbb{R}^r), \quad \mathrm{End}_p(u) = q$$

fits into the framework of problems covered by the Pontryagin Maximum Principle, see [5, Section 12] for the general statement, or [3, Section 3] for the sub-Riemannian application. The PMP gives a first order necessary criterion for optimality of a control. It turns out that there are two classes of curves satisfying this criterion with quite different behavior – the *normal* and the *abnormal* trajectories.

The classification into normal and abnormal trajectories is based on the notion of the *cotangent lift*, see [3, Section 4] for the details of the following description. For every control $u\colon [0,1] \to \mathbb{R}^r$, the integral curve $\gamma_u\colon [0,1] \to M$ can be described as the flowline from the initial point $\gamma_u(0) = p$ of the non-autonomous vector field $\sum_{i=1}^r u_i X_i$. Letting the initial point p vary, the non-autonomous vector field determines a flow $\Phi_t\colon M \to M$, which defines the cotangent lift: for a fixed initial covector $\lambda_0 \in T_p^*M$, the curve

$$\lambda\colon [0,1] \to T^*M, \quad \lambda(t) = (\Phi_t^{-1})^*\lambda_0$$

is called a cotangent lift of the curve γ_u. By construction $\lambda(t) \in T^*_{\Phi_t(p)}M = T^*_{\gamma_u(t)}M$, which is why the term lift is used. If u is an optimal control, then the PMP implies that the initial covector λ_0 can be chosen such that one of two conditions hold:

(N) $u_i(t) = \langle \lambda(t), X_i(\gamma(t)) \rangle$ for all $i = 1, \ldots, r$, or
(A) λ is a section of $\Delta^\perp$.

Here $\langle \cdot, \cdot \rangle : T_p^*M \times T_pM \to \mathbb{R}$ is the dual pairing and $\Delta^\perp \subset T^*M$ is the annihilator subbundle of Δ defined pointwise by

$$\Delta^\perp(p) = \{\lambda \in T_p^*M : \Delta(p) \subset \ker \lambda\}.$$

A cotangent lift $\lambda\colon [0,1] \to T^*M$ satisfying (N) is called a normal extremal, and a cotangent lift satisfying (A) is called an abnormal extremal. The trajectory $\gamma = \pi_*\lambda\colon [0,1] \to M$ is correspondingly called normal or abnormal. An important remark is that a trajectory may be both normal and abnormal, since a trajectory may admit several different extremal cotangent lifts for different choices of the initial covector λ_0.

The normal extremals behave similarly to Riemannian geodesics. Indeed, (N) is referred to as a *sub-Riemannian geodesic equation.* It follows that the normal extremals are locally length-minimizing and smooth, see [3, Theorem 4.65] for a proof based on calibrations, or [20] for a more direct geometric argument. It was originally believed due to a mistaken application of the PMP that all sub-Riemannian geodesics are normal trajectories, see e.g. [39]. Later this claim was corrected [40], showing that the earlier proof that all geodesics are normal holds under the *strongly bracket-generating* assumption that $TM = \Delta \oplus [X, \Delta]$ for any non-zero horizontal section $X \in \Gamma(\Delta)$.

Since normal extremals are well behaved, the Geodesic Regularity Problem is all about understanding the abnormal extremals (A) that are minimizers. The abnormal trajectories have an alternate characterization as the critical points of the endpoint map, i.e., points $u \in \mathrm{L}^2([0,1];\mathbb{R}^r)$ where the differential $d\,\mathrm{End}_u \colon \mathrm{L}^2([0,1];\mathbb{R}^r) \to T_{\mathrm{End}_p(u)}M$ is not surjective. The correspondence between the characterizations is that the initial covectors $\lambda_0 \in T_p^*M$ of abnormal extremals given by the PMP are such that $\lambda(1) = (\Phi_1^{-1})^*\lambda_0$ annihilates the image $\mathrm{im}\, d\,\mathrm{End}_u \subset T_{\mathrm{End}_p(u)}M$.

A peculiarity of the abnormals (A) is that they only depend on the distribution Δ, so they are completely independent from the metric d_{SR}. Nonetheless there exist *strictly abnormal* minimizers: abnormal minimizers which do not admit any normal lift. The first example of a strictly abnormal minimizer was found by Montgomery [30] within a Martinet distribution. Later Liu and Sussmann showed that strictly abnormal minimizers are ubiquitous for rank two distributions [27] by showing that a generic distribution of rank two admits so called *regular abnormal extremals*. A relatively explicit example of such a curve can be found in [17].

The regular abnormal extremals of a rank two distribution Δ are defined as sections of $\Delta_2^\perp \setminus \Delta_3^\perp$, and have several interesting properties. Among the most interesting is one of the only general positive optimality criteria in sub-Riemannian geometry: every regular abnormal extremal is locally length minimizing [27]. However, from the point of view of the regularity problem they give no extra information, as the curves are smooth. The smoothness follows from the observation that the constraint $\lambda(t) \in \Delta_2^\perp \setminus \Delta_3^\perp$ combined with the horizontality condition $\dot\gamma(t) \in \Delta$ defines a unique smooth vector field Y on T^*M such that every regular abnormal extremal is up to reparametrization an integral curve of Y.

An earlier study by Bryant and Hsu [12] showed that the regular abnormal extremals are *rigid* curves: they are isolated in the $\mathrm{W}^{1,\infty}$-topology (a.k.a. the C^1-topology) of horizontal curves. This rigidity phenomenon of regular abnormals has recently been further clarified for corank one regular abnormals by Agrachev and Boarotto in [4], showing in particular that the isolation in the $\mathrm{W}^{1,\infty}([0,1];\mathbb{R}^2)$-topology happens also in a less restrictive product topology $\mathrm{W}^{1,\infty}([0,1];\mathbb{R}) \times \mathrm{W}^{1,2}([0,1];\mathbb{R})$.

For more general abnormal minimizers $\gamma\colon [a,b] \to M$, Sussmann proved in [41] that if the sub-Riemannian manifold M is analytic, there exists an open dense subset $U \subset [a,b]$ such that $\gamma|_U : U \to M$ is analytic. The argument consists of constructing a suitable stratification of the abnormal set $\Delta^\perp \subset T^*M$ into analytic submanifolds of varying dimensions each with their own sub-Riemannian structure that is of lower rank than the original sub-Riemannian structure. Sussmann shows that on an open dense subset $U_1 \subset [a,b]$, the geodesic γ is also a geodesic within some of these sub-Riemannian structures of lower rank. By induction, there exists an open dense subset $U \subset \cdots \subset U_2 \subset U_1$ where the geodesic γ is also the geodesic of some analytic sub-Riemannian manifolds of rank 1. That is, the geodesic γ is on each subinterval $I \subset U$ the integral curve of an analytic vector field, and hence is analytic on U.

1.4. **Ignoring abnormals.** As the example of Montgomery [30] and the family of examples of Liu and Sussmann [27] show, strictly abnormal sub-Riemannian minimizers exist, so the study of abnormals cannot be avoided in general. Nonetheless there exist results showing that sometimes there is no harm in ignoring abnormals, either because abnormals do not exist, or because they are in a reasonable sense insignificant.

There are two important cases of distributions where strictly abnormal minimizers are completely non-existent. Both are based on second order optimality conditions showing that any strictly abnormal sub-Riemannian geodesic has a cotangent lift $\lambda\colon [0,1] \to T^*M$ which satisfies the Goh condition, see [2] or [5, Section 20]. The Goh condition is that the abnormal extremal $\lambda\colon [0,1] \to T^*M$ annihilates brackets of horizontal vectors as well as the horizontal vectors of Δ, i.e., λ is a section of $\Delta_2^\perp$.

The first case when abnormals can be ignored is when the distribution has step 2, i.e., $\Delta_2 = TM$, so the annihilator bundle $\Delta_2^\perp$ is only the zero-bundle. That is, there are no non-zero sections of $\Delta_2^\perp$, so no Goh extremals, and therefore no strictly abnormal minimizers.

The second case is generic distributions of rank ≥ 3. To make the statement precise, let M be a fixed manifold, and let $\mathcal{D}_r$ be the topological space of all rank r distributions Δ on M in the Whitney C^∞ topology. Chitour, Jean, and Trélat showed that there exists an open dense subset $O_r \subset \mathcal{D}_r$ such that no distribution $\Delta \in O_r$ admits any Goh extremals [14, Corollary 2.5]. Consequently every sub-Riemannian manifold (M, Δ, g) with $\Delta \in O_r$ does not have any strictly abnormal minimizers.

In general sub-Riemannian manifolds there are many abnormals, but it is still conjectured that from any fixed point $p \in M$ abnormal trajectories can only reach a small portion of the manifold. For a fixed basepoint $p \in M$, denote by Σ the *abnormal set*, i.e., the collection of all endpoints $\gamma(1) \in M$ of abnormal trajectories $\gamma\colon [0,1] \to M$ starting from the point p. Denote by $\Sigma_{\text{opt}} \subset \Sigma$ the subset of endpoints of all minimizing abnormal trajectories.

Sub-Riemannian Sard Conjecture.
The set Σ has zero volume with respect to any volume measure on M.

The Sard Conjecture is an open problem in sub-Riemannian geometry, see [31, Section 10.2] for a brief overview, or [25, 11, 10] for some recent progress.

Although the answer to the Sard Conjecture is in most cases unknown, it is known that the minimizing abnormal set Σ_{opt} is topologically small. By results of Rifford and Trélat [36] and Agrachev [1], on any complete real-analytic sub-Riemannian manifold $(M, \Delta, \langle \cdot, \cdot \rangle)$, there exists for any basepoint $p \in M$ an open dense subset $U \subset M$ such that every point $q \in U$ is connected to p by a unique geodesic, and the unique geodesic is normal. That is, the minimizing abnormal set Σ_{opt} is at least contained in the closed nowhere dense subset $M \setminus U$.

1.5. **The sub-Finsler generalization.** Sub-Finsler geometry is a generalization of sub-Riemannian geometry analogously to Finsler geometry as a generalization of Riemannian geometry. Namely the only difference in the definition is that in the triple $(M, \Delta, \langle \cdot, \cdot \rangle)$, the Riemannian metric $\langle \cdot, \cdot \rangle$ on the distribution Δ is replaced with a smoothly varying norm $\|\cdot\|$ on Δ.

In general sub-Finsler manifolds geodesics have no regularity beyond the Lipschitz-class, since examples of non-differentiable geodesics are available already in normed spaces such as $(\mathbb{R}^n, \|\cdot\|_\infty)$. However with the added requirement of a strict convexity of the sub-Finsler norm, also the regularity of sub-Finsler geodesics is an unanswered open problem.

2. Carnot groups

A *Carnot group* is the homogeneous model space of sub-Riemannian geometry. That is, a Carnot group is a sub-Riemannian manifold $(G, \Delta, \langle \cdot, \cdot \rangle)$, where the manifold is a stratified Lie group G and the distribution Δ and inner product $\langle \cdot, \cdot \rangle$ are left-invariant. More precisely, a Carnot group consists of

(i) a connected and simply connected Lie group G,
(ii) a *horizontal subspace* $V_1 \subset \mathfrak{g}$ of the Lie algebra $\mathfrak{g}$, and

(iii) an inner product $\langle\cdot,\cdot\rangle$ on the horizontal space V_1,

such that $\mathfrak{g}$ is stratified with the horizontal space as its first layer:

$$V_{k+1} := [V_1, V_k] \implies \begin{cases} \mathfrak{g} = V_1 \oplus V_2 \oplus \cdots \oplus V_s \\ V_{s+1} = \{0\} \end{cases} .$$

The left-invariant distribution Δ is given by left translating the horizontal space, $\Delta(g) = (L_g)_* V_1$, where $L_g\colon G \to G$ is the left translation map $L_g(h) = gh$. Similarly the inner product on each subspace $\Delta(g)$ is the pull-back inner product $\langle\cdot,\cdot\rangle_g = (L_{g^{-1}})^* \langle\cdot,\cdot\rangle$. The bracket-generating assumption is inbuilt into the stratification assumption, and the higher order distributions (1) are given at the identity element $e \in G$ by

$$\Delta_k(e) = V_1 \oplus V_2 \oplus \cdots \oplus V_k.$$

For more in depth information on Carnot groups, see [22] and the plethora of references within.

One of the more important features of a Carnot group G is its homogeneous structure, i.e., its *dilations*. The dilations are Lie group automorphisms $\delta_\alpha\colon G \to G$, $\alpha \in \mathbb{R}_+$, defined on the Lie algebra $\mathfrak{g} = V_1 \oplus \cdots \oplus V_s$ as the graded automorphisms such that

$$\delta_\alpha(X) = \alpha^k X \quad \forall X \in V_k.$$

The dilations form a one-parameter group of automorphisms: $\delta_{\alpha_1} \circ \delta_{\alpha_2} = \delta_{\alpha_1\alpha_2}$. The name dilation comes from its interaction with the sub-Riemannian metric:

$$d_{\mathrm{SR}}(\delta_\alpha(g), \delta_\alpha(h)) = \alpha d_{\mathrm{SR}}(g,h) \quad \forall g,h \in G.$$

2.1. **Carnot groups as tangent spaces.** Carnot groups play a similar role for sub-Riemannian manifolds as Euclidean spaces do for Riemannian manifolds. That is, not only are Carnot groups simple homogeneous model spaces, but they also capture the *infinitesimal geometry* of sub-Riemannian manifolds.

Rothschild and Stein developed a method to approximate a bracket-generating system of vector fields $X_1,\dots,X_r$ with left-invariant vector fields $Y_1,\dots,Y_r$ generating a free nilpotent Lie algebra [38]. A more intrinsic viewpoint was later adopted by Mitchell, who showed in [29] that the *metric tangent* in the Gromov-Hausdorff sense of an equiregular sub-Riemannian manifold is a Carnot group, see also [28] for an overview.

The notion of a metric tangent, see [18] or [13, Section 8.2], is a purely metric concept capturing the infinitesimal structure of a metric space (X,d) at some point $p \in X$ up to isometry. The idea is to take larger and larger dilation factors $\alpha \to \infty$ and consider the

Gromov-Hausdorff limit of the pointed metric spaces $(X, \alpha d, p)$. If the Gromov-Hausdorff limit $(X_\infty, d_\infty, p_\infty) = \lim_{\alpha\to\infty}(X, \alpha d, p)$ exists, it is called the metric tangent of the metric space (X, d) at the point $p \in X$. The content of Mitchell's Theorem [29] is that for an equiregular sub-Riemannian manifold $(M, \Delta, \langle\cdot,\cdot\rangle)$, the metric tangent at every point $p \in M$ exists, and is a sub-Riemannian Carnot group $(G, V_1, \langle\cdot,\cdot\rangle)$ with $\dim V_1 = \dim \Delta(p)$ and step equal to the step of Δ.

Mitchell's Theorem was later generalized to arbitrary sub-Riemannian manifolds by Bellaïche [9], see also [19, Section 2]. For a non-equiregular sub-Riemannian manifold $(M, \Delta, \langle\cdot,\cdot\rangle)$, the statement is that at every point $p \in M$ the metric tangent exists, but instead of a Carnot group, the tangent space is a homogeneous space G/H: the quotient of a Carnot group by some closed subgroup $H < G$, equipped with a natural sub-Riemannian structure.

For a sub-Riemannian Carnot group $(G, V_1, \langle\cdot,\cdot\rangle)$ the construction of a metric tangent becomes trivial. This follows from the observation that the dilation $\delta_\alpha\colon G \to G$ defines an isometry from the pointed metric space $(G, \alpha d_{\mathrm{SR}}, e)$ to (G, d_{SR}, e). Since left translations are isometries, it follows that the metric tangent of a Carnot group G at any point $g \in G$ is the group G itself.

2.2. **Normal extremals.** In a sub-Riemannian Carnot group, the nilpotency of the group G implies that the geodesic equation (N) becomes a polynomial ODE on the cotangent bundle T^*G. Two particularly convenient expressions for the ODE are obtained by trivializing the cotangent bundle as $T^*G \simeq G \times \mathfrak{g}^*$.

2.2.1. *Left-trivialized geodesic equation.* If the cotangent bundle is left-trivialized, i.e., $\lambda \in T^*_gG$ is identified with $(L_g)^*\lambda \in \mathfrak{g}^*$, then in coordinates on $G \times \mathfrak{g}^*$ the geodesic equation on the fibers $\mathfrak{g}^*$ becomes a quadratic ODE which is completely independent from the component in the group G.

For each $X \in \mathfrak{g}$, denote by $h_X\colon \mathfrak{g}^* \to \mathbb{R}$ the element of the double dual defined by the dual pairing $h_X(\lambda) = \langle \lambda, X\rangle$. Let $X_1, \dots, X_n$ be a basis of the Lie algebra $\mathfrak{g}$ such that $X_1, \dots, X_r$ is a basis of the horizontal space V_1. The functions $h_i := h_{X_i}$, $i = 1, \dots, n$, define coordinates on $\mathfrak{g}^*$.

The Hamiltonian description of the geodesic equation, see [3, Section 4], implies that the dual functions h_X evolve along the normal extremal $\lambda\colon [0,1] \to T^*G$ as

$$\dot{h}_X = \sum_{j=1}^{r} h_{[X,X_j]} h_{X_j}.$$

Writing the Lie brackets $[X_i, X_j]$ in terms of the structure coefficients $c_{ij}^k \in \mathbb{R}$ as $[X_i, X_j] = \sum_{k=1}^n c_{ij}^k X_k$, the fiber component of the geodesic equation takes the coordinate form

$$\dot{h}_i = \sum_{j=1}^r \sum_{k=1}^n c_{ij}^k h_k h_j, \quad i = 1, \ldots, n.$$

If the solution to the fiber component $(h_1, \ldots, h_n)$ is known, the group component is obtained by integrating the control $u\colon [0,1] \to \mathbb{R}^r$ given by $u = (h_1, \ldots, h_r)$.

2.2.2. *Right-trivialized geodesic equation.* If the cotangent bundle is right-trivialized, then the geodesic equation on the fiber $\mathfrak{g}^*$ becomes completely trivial: $\dot{\lambda} = 0$. The price to pay is that the equation on the group is a more complicated polynomial ODE. Namely, the equation is generically a polynomial ODE of degree $s-1$, where s is the step of the group G. The explicit geodesic equation is

$$\dot{\gamma}(t) = \sum_{i=1}^r \lambda(\operatorname{Ad}_{\gamma(t)} X_i) X_i(\gamma(t)). \tag{3}$$

2.3. **Abnormal extremals.** The abnormal extremals in a Carnot group are determined by the vanishing of a polynomial expression. This viewpoint was first presented by Le Donne, Leonardi, Monti, and Vittone in [23] for the case of a free nilpotent Lie group and in the followup [24] for more general stratified groups.

The *abnormal varieties*, as they were referred to in [25], are determined by the same polynomial expressions as in the right-trivialized geodesic equation (3). Le Donne, Montgomery, Ottazzi, Pansu, and Vittone [25, Corollary 2.14] showed that a curve $\gamma\colon [0,1] \to G$ is abnormal if and only if there exists a covector $\lambda \in \mathfrak{g}^*$ such that

$$\lambda(\operatorname{Ad}_{\gamma(t)} X_i) = 0 \quad \forall i = 1, \ldots, r,\ \forall t \in [0,1],$$

or more succinctly,

$$\lambda(\operatorname{Ad}_{\gamma(t)} V_1) = 0 \quad \forall t \in [0,1].$$

The polynomials $g \mapsto \lambda(\operatorname{Ad}_g X_i)$ are generically of degree $s-1$, where s is the step of the Carnot group. The simple expression of the abnormality criterion has been used to prove regularity of geodesics in Carnot groups of step 2 and 3.

Golé and Karidi [17] proved that in Carnot groups of step 2, any abnormal curve is contained in a Carnot subgroup of lower rank. This follows from the observation that $\lambda(\operatorname{Ad}_g V_1) = 0$ is in coordinates a linear system in $g \in G$, and defines a lower dimensional subspace $W_1 \subset V_1$ containing the projection of the abnormal curve. Hence the abnormal curve is contained in the Carnot subgroup generated by the smaller horizontal space W_1.

A similar argument for strictly abnormal trajectories in Carnot groups of step 3 was used in [25]. In step 3, the condition $\lambda(\mathrm{Ad}_g V_1) = 0$ is not linear but quadratic, but the Goh condition

$$\lambda(\mathrm{Ad}_{\gamma(t)} V_2) = 0 \quad \forall t \in [0,1],$$

is linear. Since any non-smooth geodesic would have to be strictly abnormal and hence satisfy the Goh condition, the same induction argument as in [17] can be applied to prove smoothness of geodesics in step 3 Carnot groups.

3. The tangent cone

3.1. **The tangent cone of a curve.** The *tangent cone* $\mathrm{Tang}(\gamma, t_0)$ of a curve $\gamma\colon [a,b] \to M$ at a point $t_0 \in [a,b]$ is the collection of all *(metric) tangents* to γ at the point t_0. The notion is similar to that of a metric tangent of a metric space, with the only difference that one needs to keep track of the position of the curve with respect to the ambient space M. The tangent of a curve in a sub-Riemannian manifold M is a curve in the metric tangent of M.

For a Carnot group G, the metric tangent is the group itself, and the tangents can be described as limits of dilated curves within the group G. Let $\gamma\colon [-1,1] \to G$ be a horizontal curve and consider the dilated and reparametrized curve

$$\gamma_\alpha\colon [-\alpha, \alpha] \to G, \quad \gamma_\alpha(t) := \delta_\alpha\Big(\gamma(0)^{-1}\gamma(t/\alpha)\Big). \tag{4}$$

The reparametrization guarantees that if γ is a Lipschitz curve, then γ_α is a curve with the same Lipschitz constant. The translation $\gamma(0)^{-1}$ fixes the basepoint $\gamma_\alpha(0) = e$, so Arzelà-Ascoli implies that up to taking a subsequence, there exists a limit

$$\gamma_\infty\colon \mathbb{R} \to G, \quad \gamma_\infty(t) = \lim_{\alpha_k \to \infty} \delta_{\alpha_k}\Big(\gamma(0)^{-1}\gamma(t/\alpha_k)\Big), \tag{5}$$

with uniform convergence on compact sets. Any such curve γ_∞ is called a (metric) tangent of the curve γ at the time 0. The collection of all such tangents forms the tangent cone $\mathrm{Tang}(\gamma, 0)$ of γ at 0. The tangent cones $\mathrm{Tang}(\gamma, t)$ at other points t are defined by considering limits of the curves

$$\tau \mapsto \delta_\alpha\Big(\gamma(t)^{-1}\gamma((\tau - t)/\alpha)\Big).$$

The limit (5) can be viewed as a non-abelian version of the usual difference quotient in $\mathbb{R}^n$. When the limit (5) exists for all sequences $\alpha_k \to \infty$, the unique limit is called the *Pansu derivative* [34] of the curve γ at $t = 0$. For horizontal curves, the notion of Pansu differentiability is equivalent to the usual manifold notion of differentiability. For more general curves Pansu differentiability is more restrictive, e.g. for $X, Y \in V_1$ with $[X, Y] \neq 0$, the vertical line $t \mapsto \exp(t[X, Y])$ is differentiable, but not Pansu differentiable.

In a sub-Riemannian manifold M, there is no intrinsic notion of dilation δ_α. Instead one considers the reparametrized curves $t \mapsto \gamma(t/\alpha)$ as curves in the dilated space $(M, \alpha d_{\mathrm{SR}})$, and the limits as curves in the limiting metric tangent space. This can be made precise through the construction of the metric tangent of a sub-Riemannian manifold using privileged coordinates [33].

An important, albeit simple, result from metric geometry is that tangents of tangents are tangents. Namely, a diagonal argument shows that if $\sigma \in \mathrm{Tang}(\gamma, 0)$ and $\eta \in \mathrm{Tang}(\sigma, 0)$, then also $\eta \in \mathrm{Tang}(\gamma, 0)$. Consequently, defining the iterated tangent cones

$$\mathrm{Tang}^{k+1}(\gamma) := \mathrm{Tang}(\mathrm{Tang}^k(\gamma), 0) = \bigcup_{\sigma \in \mathrm{Tang}^k(\gamma)} \mathrm{Tang}(\sigma, 0),$$

gives a decreasing sequence of curve families

$$\mathrm{Tang}(\gamma) \supset \mathrm{Tang}^2(\gamma) \supset \mathrm{Tang}^3(\gamma) \supset \dots$$

The above observation is particularly relevant for studying geodesics in sub-Riemannian manifolds, since it allows linking properties of Carnot group geodesics to more general sub-Riemannian geodesics. Namely, any tangent of a sub-Riemannian geodesic is a geodesic in a (quotient of a) Carnot group. It follows that the tangent cone of a sub-Riemannian geodesic contains the tangent cone of a Carnot geodesic.

Another useful simple observation is that the tangent cone of a geodesic contains only geodesics. This follows from the fact that any limit of geodesics is a geodesic.

3.2. **Corner-type singularities.** The first two articles [A] and [B] included in the thesis are part of a sequence of articles [26]-[A]-[32]-[B] each giving strictly stronger restrictions on the tangent cone of geodesics in sub-Riemannian manifolds.

The beginning of the chain of results is the study of the possibility of geodesics with *corner-type singularities.* A curve $\gamma \colon [-1, 1] \to M$ has a corner-type singularity at 0 if the one-sided derivatives at 0 exist, but are not equal, i.e., if in coordinates

$$\dot{\gamma}_-(0) := \lim_{t \nearrow 0} \frac{\gamma(t) - \gamma(0)}{t} \neq \lim_{t \searrow 0} \frac{\gamma(t) - \gamma(0)}{t} =: \dot{\gamma}_+(0).$$

If the geodesic γ has one-sided derivatives at 0, then $\mathrm{Tang}(\gamma, 0)$ consists of a unique curve $\tilde{\gamma} \colon \mathbb{R} \to G/H$ in the metric tangent space

G/H that has a control which is constant both on $\mathbb{R}_+$ and $\mathbb{R}_-$ [33, Theorem 3.5]. The values of the control are determined by the one-sided derivatives. All geodesics in the quotient G/H lift to geodesics in the Carnot group G, so one finds a geodesic $\sigma\colon \mathbb{R} \to G$ in a Carnot group G which is the concatenation of two half-lines:

$$\sigma(t) = \begin{cases} \exp(tX), & t \leq 0 \\ \exp(tY), & t > 0 \end{cases}$$

If the original curve γ had a corner-type singularity, then the directions $X, Y \in V_1$ are distinct, i.e., the curve σ is a *corner*. Finally, since a corner is necessarily contained in a rank 2 Carnot subgroup, the conclusion is that if there exists a sub-Riemannian geodesic with a corner-type singularity, then there also exists such an example in the simpler setting of a rank 2 Carnot group.

Leonardi and Monti [26] studied corner-type singularities in equiregular manifolds under an additional technical restriction on the flag of distributions Δ_k that

$$[\Delta_i, \Delta_j] \subset \Delta_{i+j-1}, \quad i, j \geq 2, i + j \geq 5. \tag{6}$$

The technical restriction was required so that after the reduction to the rank 2 Carnot group, the layers of the stratification satisfy

$$[V_i, V_j] = 0, \quad \text{for} \quad i, j \geq 2, i + j \geq 5.$$

The relevant consequence of this restriction is that in exponential coordinates of the second kind, the horizontal vector fields X_1, X_2 only depend on the horizontal coordinates x_1, x_2. This enabled an iterative *cut-and-correct* argument to prove the non-minimality of corners under the technical restriction (6).

The idea of the cut-and-correct construction is to make changes in the horizontal projection $\pi \circ \gamma\colon \mathbb{R} \to V_1$ of the curve, and to study what happens when the horizontal projection is lifted back to a curve in the group G. The key points are that the length of a curve in a Carnot group is given by the length of its horizontal projection, and the non-horizontal coordinates of the curve are given by iterated integrals of the horizontal coordinates. The construction is the following:

(1) Replace a short piece of the corner with a line segment in the horizontal projection, giving a shorter curve, but causing an error in the non-horizontal coordinates.
(2) Perturb the curve by replacing a line-segment away from the corner by *correction-curves*, adding some extra length, fixing some of the error, but causing further errors in other coordinates.
(3) Repeat (2), carefully quantifying the added length, the errors fixed, and the errors caused, so that if the original corner-cut was small

enough, all errors can eventually be eliminated and the length addition is still smaller than the initial length gain.

In [26], the correction-curves were all rectangles with width depending on a parameter. By carefully choosing these parameters, a countable number of correcting rectangles was used to find another curve with exactly the same endpoints as the corner, but with shorter length, proving non-minimality of corners under the restriction (6).

In [A], the technical restriction (6) was proved to be unnecessary by a refinement of the choice of correcting-curves. The key difference is the introduction of iteration on the step of the Carnot group, changing the construction from a countable procedure to a finite one. By choosing the correcting curves among geodesics in a Carnot group of one step lower, it is possible to fix the errors caused by the cut first in the layer V_2, then in V_3, and so on.

Making use of the different behavior of the dilations on different layers V_k, the quantification of added length vs. errors created becomes independent from the actual curve used for the correction. This is the key point that enabled the use of more complicated correcting curves.

The non-minimality of corners can be rephrased as the statement that if a sub-Riemannian geodesic is piecewise C^1, then it is globally C^1. This type of patching-up of piecewise regularity has been used as a concluding tool in sub-Riemannian regularity results [8, 10].

In [8], Barilari, Chitour, Jean, Prandi, and Sigalotti proved that in sub-Riemannian manifolds of rank 2, all abnormal minimizers that are sections of $\Delta_2^\perp \setminus \Delta_4^\perp$ are C^1. The proof involves a careful study of the planar dynamics induced by the abnormal equations to conclude that at every point along the curve, the one-sided derivatives exist.

In [10], Belotto da Silva, Figalli, Parusiński, and Rifford proved that in arbitrary sub-Riemannian manifolds of rank 2 and dimension 3, all injective abnormal curves are semianalytic curves. They further remarked [10, Remark B.4] that the existence of a Whitney regular stratification implies that a semianalytic curve has a piecewise analytic parametrization. Consequently any sub-Riemannian minimizer in dimension 3 must be C^1.

3.3. **Existence of tangent lines.** The inductive cut-and-correct method was refined by Monti, Pigati, and Vittone to show that for any sub-Riemannian geodesic γ, at every time t, the tangent cone $\mathrm{Tang}(\gamma, t)$ contains at least one line $t \mapsto \exp(tX)$ [32].

The refinement was based on the observation that in the study of corners it was not actually important that the horizontal projection of the curve consists exactly of two linearly independent halflines, but instead the important feature is the existence of sufficiently linearly independent pieces along the curve. The linear independence was quantified through the notion of *excess*, measuring the L^2-average deviation from any fixed horizontal direction. More precisely, for a horizontal curve $\gamma\colon [a,b] \to G$, the excess on the interval $[a,b]$ is

$$\operatorname{Exc}(\gamma,[a,b]) := \inf_{\substack{X\in V_1\\ \|X\|=1}} \left(\fint_a^b \langle \dot\gamma(t), X\rangle^2 \, dt \right)^{1/2}.$$

The result in [32] is that the cut-and-correct method can be used to show that for every geodesic $\gamma\colon [a,b] \to G$ and every point $t \in [a,b]$, there exists a shrinking sequence of neighborhoods $[t-\epsilon_k, t+\epsilon_k] \subset [a,b]$ such that $\operatorname{Exc}(\gamma,[t-\epsilon_k, t+\epsilon_k]) \to 0$. Consequently a tangent along the corresponding sequence of scales has zero excess, and thus must be a line.

A further refinement of the cut-and-correct technique is in [B], where the excess was replaced with a discrete measure of linear independence. For a horizontal curve $\gamma\colon [a,b] \to G$ in a rank r Carnot group, fix $r+1$ times $a \le t_0 < \cdots < t_r \le b$ and consider the connecting line segments

$$v_j := \pi\circ\gamma(t_j) - \pi\circ\gamma(t_{j-1}) \in V_1, \quad j = 1,\dots,r$$

in the horizontal projection. Define the *size* of the configuration of points as

$$\operatorname{Size}(\gamma, t_0,\dots,t_r) := \min_{j=1,\dots,r} d(v_j, \operatorname{span}\{v_1,\dots,\hat v_j,\dots,v_r\}),$$

i.e., as the smallest height of the simplex with vertices $v_1,\dots,v_r$.

In the study of tangents, the notions of excess and size are equivalent for the purposes of the cut-and-correct method. The extra improvement in [B] is to quantify the relation between geodesics in step s Carnot groups and their projections to the quotient groups of step $s-1$. Namely, it is shown that for a unit-speed geodesic $\gamma\colon I \to G$, the projection $\pi_{s-1}\circ\gamma\colon I \to G/\exp(V_s)$ is locally a *sublinear bilipschitz embedding* in the terminology of [16]. That is, for any $t \in I$ there exists a small δ-neighborhood and a constant $C > 0$ such that for any $a, b \in (t-\delta, t+\delta)$ the projection satisfies the distance estimate

$$|a-b| - C|a-b|^{\frac{s}{s-1}} \le d(\pi_{s-1}\circ\gamma(a), \pi_{s-1}\circ\gamma(b)) \le |a-b|. \tag{7}$$

The constant C is inversely proportional to the size of simplices outside the interval $(t-\delta, t+\delta)$:

$$C \simeq \frac{1}{\max\limits_{\substack{t_0<\cdots<t_r\\ t_i\notin(t-\delta,t+\delta)}} \operatorname{Size}(\gamma, t_0,\dots,t_r)^{1/(s-1)}}. \tag{8}$$

This quantified projection behavior gives a slightly stronger form of the existence of a tangent line. The bound (7) implies that every tangent of a step s geodesic is also a geodesic when projected to the step $s-1$ group $G/\exp(V_s)$. By iteration it follows that all curves in the iterated tangent cone $\mathrm{Tang}^s(\gamma)$ are geodesics in the inner product space $G/[G,G] \simeq V_1$, and hence must be lines.

An auxiliary benefit of the discrete linear independence notion is that it is well suited to large scale study. In [B] it is also proved that in Carnot groups $(1,C)$-*quasi-geodesics* $\gamma\colon \mathbb{R} \to G$, i.e., not-necessarily-continuous curves for which

$$|a-b| - C \le d(\gamma(a),\gamma(b)) \le |a-b| + C \quad \forall a,b \in \mathbb{R},$$

have their horizontal projections coarsely contained in a proper subspace of the horizontal space V_1. More precisely, for any $(1,C)$-quasi-geodesic there exists a Carnot subgroup $H < G$ and a radius $R > 0$ such that

$$d(\pi\circ\gamma(t),\pi(H)) = \inf_{h\in H} d(\pi\circ\gamma(t),\pi(h)) < R \quad \forall t \in \mathbb{R}.$$

In the chain of articles [26]-[A]-[32]-[B], the sub-Riemannian metric is not critical. For the most part, the inner product is used only in the initial step of the induction, to argue that the only geodesics are lines in the normed space case of step 1. The results are valid also more generally for sub-Finsler manifolds with strictly convex norms.

4. Isometric embeddings

The final included article [C] deals with the study of *infinite geodesics* and more general *isometric embeddings* in Carnot groups of step 2. An infinite geodesic is a length-minimizer $\gamma\colon I \to G$, where the interval $I \subset \mathbb{R}$ is unbounded. An isometric embedding is a map $\varphi\colon (H,d_H) \to (G,d_G)$ between metric spaces such that

$$d_G(\varphi(h_1),\varphi(h_2)) = d_H(h_1,h_2) \quad \forall h_1,h_2 \in H.$$

In [C] it is proved that if G is a sub-Finsler Carnot group with a strictly convex norm, then all isometric embeddings $H \hookrightarrow G$ from other Carnot groups H are affine, i.e., are compositions of left translations and group homomorphisms. The same is true also in the slightly more general setting of stratified groups of step 2 equipped with arbitrary *homogeneous distances*, i.e. distances d which satisfy the homogeneity condition

$$d(\delta_\alpha(g),\delta_\alpha(h)) = \alpha d(g,h) \quad \forall g,h \in G.$$

For homogeneous distances, the strict convexity assumption is a condition on the *projection norm*

$$\|\cdot\|\colon V_1 \to \mathbb{R}, \quad \|X\| = d(e, \exp(X)).$$

The affinity result of [C] is the Carnot-group analogue of the classical result of affinity of isometric embeddings in real normed spaces with strictly convex norms. The case of sub-Riemannian Carnot groups of step 2 was already proved by Kishimoto in [21]. For more general homogeneous distances, the result was known in the Heisenberg groups by the results of Balogh, Fässler and Sobrino [7], and Balogh and Calogero [6]. The proof strategy in the general case is similar to the strategy used in the Heisenberg setting.

In [7], Balogh, Fässler and Sobrino showed that that if $\mathbb{H}^n$ is a Heisenberg group with a homogeneous distance such that all infinite geodesics are lines, then all isometric embeddings $\mathbb{R}^m \hookrightarrow \mathbb{H}^n$ and $\mathbb{H}^m \hookrightarrow \mathbb{H}^n$ are affine. Their proof works also for more general stratified groups with homogeneous distances, reducing the problem of affinity of isometric embeddings to the study of infinite geodesics.

The first simplification of the study of infinite geodesics is that instead of more general homogeneous distances, it suffices to consider sub-Finsler metrics. This follows from the observation that all geodesics in a homogeneous distance are also geodesics of the sub-Finsler distance induced by the projection norm, see [7, Proposition 2.19] for the Heisenberg case, or [C, Lemma 5.1] for the general homogeneous group case. A second simplification is that by the second order optimality criterion of the Goh condition, all geodesics are normal in step 2. Although in the sub-Finsler case this does not lead to an explicit geodesic equation as in the sub-Riemannian case, there is still enough information to study the asymptotics of normal extremals.

The asymptotic study of infinite geodesics in Heisenberg groups was handled by Balogh and Calogero in [6] by making use of tools from convex analysis. A careful study of the information given by the normal case of the Pontryagin Maximum Principle was used to show that every normal trajectory that is not a line trajectory has a bounded horizontal component. Having a bounded horizontal component implies that *asymptotic cone* – the collection of the limits as $\alpha \to 0$ of the shrinked down curves $\gamma_\alpha = \delta_\alpha \circ \gamma \circ \delta_{1/\alpha}$ (4) – consists of only the trivial curve $t \mapsto e$. If the original trajectory was a geodesic, then the limit should also be a geodesic, so such a degeneration could not happen. Consequently the only infinite geodesics in the Heisenberg groups are lines.

The asymptotic study of infinite geodesics in the general step 2 case in [C] follows a roughly similar idea. Although in the general case the horizontal component does not have to be bounded, the normal trajectory may be thought of as consisting of a bounded component and a line component. This is made precise in [C, Section 4] by an integral average formulation of the control of a curve in the asymptotic cone of the geodesic. The proof is concluded by the observation that the existence of a non-trivial non-line component would cause some degeneration in the limit, which would be impossible for a geodesic.

References

[1] A. A. Agrachëv. Any sub-Riemannian metric has points of smoothness. *Dokl. Akad. Nauk*, 424(3):295–298, 2009.

[2] A. A. Agrachev and A. V. Sarychev. Abnormal sub-Riemannian geodesics: Morse index and rigidity. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 13(6):635–690, 1996.

[3] Andrei Agrachev, Davide Barilari, and Ugo Boscain. A comprehensive introduction to sub-riemannian geometry, 2019.

[4] Andrei A. Agrachev and Francesco Boarotto. Structure of the endpoint map near nice singular curves. *arXiv e-prints*, page arXiv:1810.12662, Oct 2018.

[5] Andrei A. Agrachev and Yuri L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. Control Theory and Optimization, II.

[6] Z. M. Balogh and A. Calogero. Infinite geodesics of sub-Finsler distances in the Heisenberg groups. *ArXiv e-prints*, page arXiv:1807.10369, July 2018.

[7] Zoltán M. Balogh, Katrin Fässler, and Hernando Sobrino. Isometric embeddings into Heisenberg groups. *Geom. Dedicata*, 195:163–192, 2018.

[8] D. Barilari, Y. Chitour, F. Jean, D. Prandi, and M. Sigalotti. On the regularity of abnormal minimizers for rank 2 sub-Riemannian structures. *ArXiv e-prints*, April 2018.

[9] André Bellaïche. The tangent space in sub-Riemannian geometry. In *Sub-Riemannian geometry*, volume 144 of *Progr. Math.*, pages 1–78. Birkhäuser, Basel, 1996.

[10] André Belotto da Silva, Alessio Figalli, Adam Parusiński, and Ludovic Rifford. Strong Sard Conjecture and regularity of singular minimizing geodesics for analytic sub-Riemannian structures in dimension 3. *ArXiv e-prints*, October 2018.

[11] André Belotto da Silva and Ludovic Rifford. The Sard conjecture on Martinet surfaces. *Duke Math. J.*, 167(8):1433–1471, 2018.

[12] Robert L. Bryant and Lucas Hsu. Rigidity of integral curves of rank 2 distributions. *Invent. Math.*, 114(2):435–461, 1993.

[13] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[14] Y. Chitour, F. Jean, and E. Trélat. Genericity results for singular curves. *J. Differential Geom.*, 73(1):45–73, 2006.

[15] Wei-Liang Chow. Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. *Math. Ann.*, 117:98–105, 1939.

[16] Yves Cornulier. On sublinear bilipschitz equivalence of groups. *arXiv e-prints*, page arXiv:1702.06618, Feb 2017.

[17] Chr. Golé and R. Karidi. A note on Carnot geodesics in nilpotent Lie groups. *J. Dynam. Control Systems*, 1(4):535–549, 1995.

[18] Mikhael Gromov. *Structures métriques pour les variétés riemanniennes*, volume 1 of *Textes Mathématiques [Mathematical Texts]*. CEDIC, Paris, 1981. Edited by J. Lafontaine and P. Pansu.

[19] Frédéric Jean. *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*. SpringerBriefs in Mathematics. Springer, Cham, 2014.

[20] MichałJóźwikowski and Witold Respondek. Why are normal sub-Riemannian extremals locally minimizing? *Differential Geom. Appl.*, 60:174–189, 2018.

[21] Iwao Kishimoto. Geodesics and isometries of Carnot groups. *J. Math. Kyoto Univ.*, 43(3):509–522, 2003.

[22] Enrico Le Donne. A primer on Carnot groups: homogenous groups, Carnot-Carathéodory spaces, and regularity of their isometries. *Anal. Geom. Metr. Spaces*, 5(1):116–137, 2017.

[23] Enrico Le Donne, Gian Paolo Leonardi, Roberto Monti, and Davide Vittone. Extremal curves in nilpotent Lie groups. *Geom. Funct. Anal.*, 23(4):1371–1401, 2013.

[24] Enrico Le Donne, Gian Paolo Leonardi, Roberto Monti, and Davide Vittone. Extremal polynomials in stratified groups. *Comm. Anal. Geom.*, 26(4):723–757, 2018.

[25] Enrico Le Donne, Richard Montgomery, Alessandro Ottazzi, Pierre Pansu, and Davide Vittone. Sard property for the endpoint map on some Carnot groups. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 33(6):1639–1666, 2016.

[26] Gian Paolo Leonardi and Roberto Monti. End-point equations and regularity of sub-Riemannian geodesics. *Geom. Funct. Anal.*, 18(2):552–582, 2008.

[27] Wensheng Liu and Héctor J. Sussman. Shortest paths for sub-Riemannian metrics on rank-two distributions. *Mem. Amer. Math. Soc.*, 118(564):x+104, 1995.

[28] G. A. Margulis and G. D. Mostow. Some remarks on the definition of tangent cones in a Carnot-Carathéodory space. *J. Anal. Math.*, 80:299–317, 2000.

[29] John Mitchell. On Carnot-Carathéodory metrics. *J. Differential Geom.*, 21(1):35–45, 1985.

[30] Richard Montgomery. Abnormal minimizers. *SIAM J. Control Optim.*, 32(6):1605–1620, 1994.

[31] Richard Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

[32] Roberto Monti, Alessandro Pigati, and Davide Vittone. Existence of tangent lines to Carnot-Carathéodory geodesics. *Calc. Var. Partial Differential Equations*, 57(3):Art. 75, 18, 2018.

[33] Roberto Monti, Alessandro Pigati, and Davide Vittone. On tangent cones to length minimizers in Carnot-Carathéodory spaces. *SIAM J. Control Optim.*, 56(5):3351–3369, 2018.

[34] Pierre Pansu. Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un. *Ann. of Math. (2)*, 129(1):1–60, 1989.

[35] P. K. Rashevsky. Any two points of a totally nonholonomic space may be connected by an admissible line. *Uch. Zap. Ped. Inst. im. Liebknechta, Ser. Phys. Math*, 2:83–94, 1938.

[36] L. Rifford and E. Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.

[37] Ludovic Rifford. *Sub-Riemannian geometry and optimal transport*. SpringerBriefs in Mathematics. Springer, Cham, 2014.

[38] Linda Preiss Rothschild and E. M. Stein. Hypoelliptic differential operators and nilpotent groups. *Acta Math.*, 137(3-4):247–320, 1976.

[39] Robert S. Strichartz. Sub-Riemannian geometry. *J. Differential Geom.*, 24(2):221–263, 1986.

[40] Robert S. Strichartz. Corrections to: "Sub-Riemannian geometry" [J. Differential Geom. **24** (1986), no. 2, 221–263; MR0862049 (88b:53055)]. *J. Differential Geom.*, 30(2):595–596, 1989.

[41] H. J. Sussmann. A regularity theorem for minimizers of real-analytic subriemannian metrics. In *53rd IEEE Conference on Decision and Control*, pages 4801–4806, Dec 2014.

Included articles

[A]

Non-minimality of corners in subriemannian geometry

E. Hakavuori and E. Le Donne

First published in *Inventiones mathematicae* 206 (2016), no. 3

https://doi.org/10.1007/s00222-016-0661-9

NON-MINIMALITY OF CORNERS IN SUBRIEMANNIAN GEOMETRY

EERO HAKAVUORI AND ENRICO LE DONNE

Abstract. We give a short solution to one of the main open problems in subriemannian geometry. Namely, we prove that length minimizers do not have corner-type singularities. With this result we solve Problem II of Agrachev's list, and provide the first general result toward the 30-year-old open problem of regularity of subriemannian geodesics.

Contents

1. Introduction

One of the major open problems in subriemannian geometry is the regularity of length-minimizing curves (see [Mon02, Section 10.1] and [Mon14b, Section 4]). This problem has been open since the work of Strichartz [Str86, Str89] and Hamenstädt [Ham90].

Contrary to Riemannian geometry, where it is well known that all length minimizers are C^∞-smooth, the problem in the subriemannian case is significantly more difficult. The primary reason for this difficulty is the existence of abnormal curves (see [AS04,

Date: March 8, 2016.
2010 *Mathematics Subject Classification.* 53C17, 49K21, 28A75.
Key words and phrases. Corner-type singularities, geodesics, sub-Riemannian geometry, Carnot groups, regularity of length minimizers.

ABB15]), which we know may be length minimizers since the work of Montgomery [Mon94]. Nowadays, many more abnormal length minimizers are known [BH93, LS94, LS95, GK95, Sus96].

Abnormal curves, when parametrized by arc-length, need only have Lipschitz-regularity (see [LDLMV14, Section 5]), which is why, a priori, no further regularity can be assumed from an arbitrary length minimizer in a subriemannian space. However, a recent result of Sussmann states that in the analytic setting, every length minimizer is analytic on an open dense subset of its domain, see [Sus14]. Nonetheless, even including all known abnormal minimizers, no example of a non-smooth length minimizer has yet been shown.

A considerable effort has been made to find examples of non-smooth minimizers (or to prove the non-existence thereof) in the simple case of curves where the lack of continuity of the derivative is at a single point. Partial results for the non-minimality of corners can be found, e.g., in [Mon14a, Mon14c, LDLMV13].

In this paper, we prove the non-minimality of curves with a corner-type singularity in complete generality. Thus we solve Problem II of Agrachev's list of open problems in subriemannian geometry [Agr14], by proving the following result (definitions are recalled in Section 1.2):

Theorem 1.1. *Length-minimizing curves in subriemannian manifolds do not have corner-type singularities.*

In fact, our proof also shows that the same result holds even if instead of subriemannian manifolds, we consider the slightly more general setting of Carnot-Carathéodory spaces with strictly convex norms.

1.1. **The idea of the argument.** The argument builds on ideas of the two papers [LM08, LDLMV15]. Up to a desingularization, blow-up, and reduction argument, it is sufficient to consider the case of a corner in a Carnot group of rank 2. For Carnot groups, we prove the result (Theorem 3.1) by induction on the step s of the group, starting with $s = 2$, i.e., the Heisenberg group, see Lemma 2.1.

For an arbitrary step $s \geq 3$ we project the corner into a Carnot group of step $s-1$. The inductive argument then gives us the existence of a shorter curve in the group of step $s-1$. Lifting this curve back to the original group, we get a curve shorter than the initial corner, but with an error in the endpoint by an element of degree s, see Lemma 2.2.

We correct the error by a system of curves placed along the corner. In fact, we prove that this is possible with a system of three curves with endpoints in the subspace of degree $s-1$. This last fact is the core of the argument (see Lemma 2.3) and is a crucial consequence of the fact that the space is a nilpotent and stratified group.

Finally, we consider the situation at smaller scales by modifying the initial corner using an ϵ-dilation of the lifted curve and suitable dilations of the three correcting curves. By Lemma 2.3 the suitable factor to correct the error of the dilated corner-cut is $\epsilon^{s/(s-1)}$, essentially due to the fact that the error scales with order s and the correction scales with order $s-1$. Hence, the length of the new curve is the length of the corner plus a term of the form

$$-a\epsilon + b\epsilon^{s/(s-1)},$$

for some positive constants a, b. We conclude that for ϵ small enough the new curve is shorter than the corner.

1.2. **Definitions.** Let M be a smooth Riemannian manifold and Δ a smooth subbundle of the tangent bundle. We consider the length functional L_Δ on curves in M that for a curve γ is defined as the Riemannian length of γ if $\dot\gamma \in \Delta$ almost everywhere, and ∞ otherwise. Analogously to the Riemannian setting, let d_Δ be the distance associated to L_Δ. We assume that Δ is bracket generating, in which case d_Δ is finite and its length functional is L_Δ. In this paper, we call (M, d_Δ) a *subriemannian manifold*. For more on the subject see [Gro96, Gro99, Mon02, Jea14, Rif14, ABB15]. If instead of a Riemannian structure, we use a continuously varying norm on the tangent bundle, we call the resulting metric space a *Carnot-Carathéodory space* (*C-C space*, for short).

Let $\gamma : [-1, 1] \to M$ be an absolutely continuous curve on a manifold M. We say that γ has a *corner-type singularity* at time 0, if the left and right derivatives at 0 exist and are linearly independent.

Let G be a Lie group. We say that a curve $\gamma : [-1, 1] \to G$ is a *corner* if there exist linearly independent vectors X_1, X_2 in the Lie algebra of G such that

$$\gamma(t) = \begin{cases} \exp(-tX_1) & \text{if } t \in [-1, 0] \\ \exp(tX_2) & \text{if } t \in (0, 1]\,. \end{cases}$$

In such a case, we will say that γ is the corner from $\exp(X_1)$ to $\exp(X_2)$. Notice that at 0 the left derivative of γ is $-X_1$, while the right derivative is X_2. Hence, a corner has a corner-type singularity at 0.

Let G be a simply connected Lie group with a Lie algebra $\mathfrak{g}$ admitting a stratification, i.e., $\mathfrak{g} = V_1 \oplus \cdots \oplus V_s$, where $V_j \subset \mathfrak{g}$ are disjoint vector subspaces of the algebra, such that $V_{j+1} = [V_1, V_j]$ for all $j = 1, \ldots, s$ with $V_{s+1} = \{0\}$. The subspaces V_j are called the *layers* of the stratification. Let $|\cdot|$ be a norm on the first layer V_1 of the Lie algebra. The Lie group G together with a stratification $\mathfrak{g} = V_1 \oplus \cdots \oplus V_s$ of its algebra and a norm $|\cdot|$ on the first layer V_1 is called a *Carnot group*. See [Mon02, LD15] for more discussion on Carnot groups.

A Carnot group has a natural structure of C-C space where the subbundle Δ is the left-translation of the first layer V_1 and the norm is extended left-invariantly. Then by construction the C-C distance $d = d_\Delta$ on a Carnot group is left-invariant. In addition, a Carnot group also has a family of Lie group automorphisms $\{\delta_\epsilon\}_{\epsilon>0}$

adapted to the stratification. Namely, each δ_ϵ is determined by $(\delta_\epsilon)_*(X) = \epsilon^j X$, for $X \in V_j$. Moreover, each map δ_ϵ behaves as an ϵ-dilation for the C-C distance, i.e., $d(\delta_\epsilon(g), \delta_\epsilon(h)) = \epsilon d(g,h)$, for all points $g, h \in G$.

In a Carnot group, the curves $t \mapsto \exp(tX)$, with $X \in \mathfrak{g}$, have locally finite length if and only if $X \in V_1$. Actually, such curves are length minimizing and $d(e, \exp(X)) = |X|$, where e denotes the identity element of G.

A norm $|\cdot|$ is *strictly convex* if in its unit sphere there are no non-trivial segments. Equivalently, if $|x| = |y| = 1$ and $|x+y| = 2$, strict convexity implies $x = y$.

2. Preliminary lemmas

The following lemma is the base of our inductive argument. In particular, it proves Theorem 1.1 for the Heisenberg group equipped with a strictly convex norm.

Lemma 2.1. *Let G be a step-2 Carnot group with a distance d associated to a strictly convex norm. Then in (G,d) no corner is length minimizing.*

Proof. Let X_1 and X_2 be linearly independent vectors of the first layer V_1 of G. For $\epsilon > 0$, consider the group elements

$$g_1 = \exp((\epsilon - 1)X_1), \qquad g_2 = \exp(\epsilon(X_2 - X_1)), \qquad g_3 = \exp((\tfrac{1}{2} - \epsilon)X_2),$$
$$g_4 = \exp(-\epsilon^2 X_1), \qquad g_5 = \exp(\tfrac{1}{2}X_2), \qquad g_6 = \exp(\epsilon^2 X_1).$$

Using the Baker-Campbell-Hausdorff Formula, which in step 2 is $\exp(X)\exp(Y) = \exp(X + Y + \frac{1}{2}[X,Y])$, one can verify that $\exp(X_2) = \exp(X_1)\, g_1 \cdots g_6$. We may assume that $|X_1| = |X_2| = 1$. Since X_1 and X_2 are linearly independent and the norm is strictly convex, the distance

$$D = d(e, \exp(X_2 - X_1)) = |X_2 - X_1|$$

is strictly smaller than 2. By left-invariance of the distance and the triangle inequality, we get the upper bound

$$d(\exp(X_1), \exp(X_2)) = d(e, g_1 \cdots g_6) \leq \sum_{j=1}^{6} d(e, g_j),$$

which we can explicitly calculate as

$$\begin{aligned}\sum_{j=1}^{6} d(e, g_j) &= (1-\epsilon) + \epsilon D + (\tfrac{1}{2} - \epsilon) + \epsilon^2 + \tfrac{1}{2} + \epsilon^2 \\ &= 2 - (2-D)\epsilon + 2\epsilon^2.\end{aligned}$$

Since $-(2-D) < 0$, taking small enough $\epsilon > 0$ we deduce $d(\exp(X_1), \exp(X_2)) < 2$. Hence the corner from $\exp(X_1)$ to $\exp(X_2)$ is not length minimizing in G. □

The geometric interpretation of the next lemma is the following. Curves from a quotient group can be isometrically lifted. Thus in our inductive argument we can use a geodesic from the previous step to get a curve that is shorter than the corner and has an error only in the last layer.

Lemma 2.2. *Let G be a Carnot group of step s. Assume that there are no minimizing corners in any Carnot group of step $s-1$ with first layer isometric to the first layer of G. For all linearly independent $X_1, X_2 \in V_1$ there exists $h \in \exp(V_s)$ such that*

$$d(h\exp(X_1), \exp(X_2)) < |X_1| + |X_2|.$$

Proof. Consider the closed central subgroup $H = \exp(V_s)$. The quotient G/H is a Carnot group of step $s-1$ with first layer $\pi_*(V_1)$. Note that the norm on $\pi_*(V_1)$ is exactly the one that makes the projection $\pi_* : V_1 \to \pi_*(V_1)$ an isometry. Therefore the first layer $\pi_*(V_1)$ of G/H is isometric to V_1, so by assumption there are no minimizing corners in G/H.

If X_1 and X_2 are linearly independent, then so are $\pi_*(X_1)$ and $\pi_*(X_2)$. Thus, by assumption, the corner in G/H from $\exp(\pi_*(X_1))$ to $\exp(\pi_*(X_2))$ is not length minimizing. Observe that since π is a Lie group homomorphism, we have $\exp(\pi_*(X)) = \pi(\exp(X))$. Hence,

$$d(\pi(\exp(X_1)), \pi(\exp(X_2))) < |X_1| + |X_2|.$$

Using left-invariance of the distance on G we see that

$$\begin{aligned} d(\pi(\exp(X_1)), \pi(\exp(X_2))) &= d(H\exp(X_1), H\exp(X_2)) \\ &= \inf_{h\in H} d(h\exp(X_1), \exp(X_2)). \end{aligned}$$

Combining the above equality with the previous inequality, we conclude that there exists a point $h \in H$ for which the statement of the lemma holds. □

The next lemma is the technical core of our argument. It shows that any error coming from Lemma 2.2 can be corrected using vectors in the layer $s-1$. It also quantifies how the corrections change when scaling the error. In what follows, we consider the conjugation map $\mathrm{C}_p(q) = pqp^{-1}$.

Lemma 2.3. *Let G be a Carnot group of step $s \geq 3$ and let X_1 and X_2 be vectors spanning V_1. Then for any $h \in \exp(V_s)$ there exist vectors $Y_1, Y_2, Y_3 \in V_{s-1}$ such that*

$$\mathrm{C}_{\exp(X_1)}(\exp(\epsilon^s Y_1)) \cdot \mathrm{C}_{\exp\left(\frac{1}{2}X_2\right)}(\exp(\epsilon^s Y_2)) \cdot \mathrm{C}_{\exp(X_2)}(\exp(\epsilon^s Y_3)) = \delta_\epsilon(h),$$

for all $\epsilon > 0$.

Proof. Consider first for some $Z \in V_s$ the equation

$$\mathrm{C}_{\exp(X_1)}(\exp(Y_1)) \cdot \mathrm{C}_{\exp\left(\frac{1}{2}X_2\right)}(\exp(Y_2)) \cdot \mathrm{C}_{\exp(X_2)}(\exp(Y_3)) = \exp(Z) \tag{2.4}$$

in the variables $Y_1, Y_2, Y_3 \in V_{s-1}$. Since the step of the group G is s, each conjugation can be expanded by the Baker-Campbell-Hausdorff Formula[1] as

$$\mathrm{C}_{\exp(X)}(\exp(Y)) = \exp(X)\exp(Y)\exp(-X) = \exp(Y + [X, Y]).$$

We remark that the subgroup $\exp(V_{s-1} \oplus V_s)$, containing the above conjugations, is commutative because of the assumption $s \geq 3$. Hence exp is a homomorphism on $V_{s-1} \oplus V_s$. Consequently, since exp is also injective, we see that (2.4) is equivalent to the linear equation

$$Y_1 + Y_2 + Y_3 + [X_1, Y_1] + [X_2, \tfrac{1}{2}Y_2 + Y_3] = Z. \tag{2.5}$$

Since the vectors X_1 and X_2 span the first layer V_1, and $V_s = [V_1, V_{s-1}]$, for any $Z \in V_s$ there exist $W_1, W_2 \in V_{s-1}$ such that

$$Z = [X_1, W_1] + [X_2, W_2].$$

Therefore, to solve the linear equation (2.5), it is sufficient to solve the linear system

$$\begin{aligned} Y_1 + Y_2 + Y_3 &= 0 \\ Y_1 &= W_1 \\ \tfrac{1}{2}Y_2 + Y_3 &= W_2, \end{aligned}$$

which has the solution $Y_1 = W_1, Y_2 = -2W_1 - 2W_2, Y_3 = W_1 + 2W_2$. Hence for any data $Z \in V_s$, equation (2.4) has a solution $Y_1, Y_2, Y_3 \in V_{s-1}$.

Consider a fixed $h \in \exp(V_s)$ and let $Z \in V_s$ be such that $\exp(Z) = h$. Note that then $\delta_\epsilon(h) = \exp(\epsilon^s Z)$ for any $\epsilon > 0$. Recalling that the solution Y_1, Y_2, Y_3 for the data Z is given by a linear equation, we have that for any $\epsilon > 0$ the vectors $\epsilon^s Y_1, \epsilon^s Y_2, \epsilon^s Y_3$ give a solution for the data $\epsilon^s Z$, resulting in the statement of the lemma. □

3. The main result

3.1. **Reduction to Carnot groups.** The proof of Theorem 1.1 can be reduced to the corresponding result for Carnot groups. Due to the possibility of the manifold not being equiregular (see [Jea14] for the definition), we first consider a desingularization of the manifold near the corner-type singularity. Then we perform a blow-up, giving a corner in the metric tangent, which is a Carnot group by Mitchell's Theorem.

Let M be a subriemannian manifold with subbundle Δ, and let γ be a curve in M. Fix a local orthonormal frame $X_1, \ldots, X_r$ for Δ near $\gamma(0)$. By [Jea14, Lemma 2.5, page 49] there exists an equiregular subriemannian manifold N with an orthonormal frame $\xi_1, \ldots, \xi_r$ and a map $\pi : N \to M$ onto a neighborhood of $\gamma(0)$ such that

[1] Alternatively, one can use the formula [War83, page 114]

$$\mathrm{C}_{\exp(X)}(\exp(Y)) = \exp(\mathrm{Ad}_{\exp(X)} Y) = \exp(e^{\mathrm{ad}_X} Y).$$

$\pi_* \xi_i = X_i$. We observe that π is 1-Lipschitz with respect to the subriemannian distances.

Assume that γ is length minimizing, has a corner-type singularity at 0, and is contained in $\pi(N)$. Let u_j be integrable functions such that $\dot\gamma = \sum_j u_j X_j$ almost everywhere.

Let σ be a curve in N such that $\dot\sigma = \sum_j u_j \xi_j$ almost everywhere. Hence $\pi \circ \sigma = \gamma$ and the two curves σ and γ have the same length, see the proof of [Jea14, Lemma 2.5, page 49]. Since π does not stretch distances, we conclude that σ is length minimizing.

Since the vector fields X_j form a frame, the coefficients u_j are uniquely determined from $\dot\gamma$, and the existence of the left and right derivatives at 0 is equivalent to 0 being a left and right Lebesgue point for u_j. Therefore σ also admits[2] left and right derivatives at 0. Noting that $\pi_*\dot\sigma = \dot\gamma$ and that γ has a corner-type singularity at 0, we conclude that σ also has a corner-type singularity at 0.

The curve σ is now a length-minimizing curve with a corner-type singularity on an equiregular subriemannian manifold N. The metric tangent of N is a Carnot group G, see a detailed proof in [Jea14, Proposition 2.4, page 39]. The blow-up of σ on the Carnot group G is length minimizing and is given by the concatenation of two half-lines, see [LM08, Proposition 2.4].

3.2. **The inductive non-minimality argument.** By the previous argument, to show that a length-minimizing curve in a subriemannian manifold cannot have a corner-type singularity, it suffices to prove the corresponding result for Carnot groups. In fact, we prove the slightly stronger statement:

Theorem 3.1. *Corners are not length minimizing in any Carnot group equipped with a Carnot-Carathéodory distance coming from a strictly convex norm.*

In the above, the distance is only coming from a strictly convex norm, as opposed to an inner product as in the subriemannian case. The argument at the beginning of this section is however not dependent on the chosen distance. Thus it shows that Theorem 1.1 also holds for C-C spaces with strictly convex norms.

Proof of Theorem 3.1. We remark that it suffices to consider the case of rank-2 Carnot groups. Indeed, any corner is contained in some rank-2 subgroup, and if a curve is length minimizing, it must also be length minimizing in any subgroup containing it. The theorem will then be proven for rank-2 Carnot groups by induction on the step s of the group. The base of induction is the case $s = 2$, where the result is verified by Lemma 2.1.

[2]We remark that for rank-varying distributions, desingularizations of curves with corner-type singularities need not have one-sided derivatives.

Let G be a rank-2 Carnot group of step s with a Carnot-Carathéodory distance coming from a strictly convex norm. Consider the corner from $\exp(X_1)$ to $\exp(X_2)$, for some linearly independent $X_1, X_2 \in V_1$ with $|X_1| = |X_2| = 1$.

Taking the quotient of G by the central subgroup $\exp(V_s)$, we get a Carnot group of step $s-1$ whose first layer is isometric to the first layer of G. Note that the projection of a corner is still a corner in the quotient, where by induction we assume that corners are not length minimizing. Hence, by Lemma 2.2, there exists $h \in \exp(V_s)$ such that

$$d(h\exp(X_1), \exp(X_2)) < 2. \tag{3.2}$$

By Lemma 2.3, for this fixed $h \in \exp(V_s)$, there exist vectors $Y_1, Y_2, Y_3 \in V_{s-1}$ satisfying the equation

$$\delta_\epsilon(h)^{-1}\,\mathrm{C}_{\exp(X_1)}\left(\exp(\epsilon^s Y_1)\right)\cdot \mathrm{C}_{\exp\left(\frac{1}{2}X_2\right)}\left(\exp(\epsilon^s Y_2)\right)\cdot \mathrm{C}_{\exp(X_2)}\left(\exp(\epsilon^s Y_3)\right) = e. \tag{3.3}$$

For a given $\epsilon > 0$, consider the following points

$$\begin{aligned}
g_1 &= \exp(\epsilon^s Y_1) = \delta_{\epsilon^{s/(s-1)}}(\exp(Y_1)),\\
g_2 &= \exp(-(1-\epsilon)X_1) = \delta_{1-\epsilon}(\exp(-X_1)),\\
g_3 &= \exp(-\epsilon X_1)\delta_\epsilon(h)^{-1}\exp(\epsilon X_2) = \delta_\epsilon\left(\exp(-X_1)h^{-1}\exp(X_2)\right),\\
g_4 &= \exp((\tfrac{1}{2}-\epsilon)X_2) = \delta_{\frac{1}{2}-\epsilon}(\exp(X_2)),\\
g_5 &= \exp(\epsilon^s Y_2) = \delta_{\epsilon^{s/(s-1)}}(\exp(Y_2)),\\
g_6 &= \exp(\tfrac{1}{2}X_2) = \delta_{\frac{1}{2}}(\exp(X_2)), \quad \text{and}\\
g_7 &= \exp(\epsilon^s Y_3) = \delta_{\epsilon^{s/(s-1)}}(\exp(Y_3)).
\end{aligned}$$

We claim that

$$\exp(X_2) = \exp(X_1)\, g_1 \cdots g_7, \tag{3.4}$$

and that for small enough $\epsilon > 0$

$$\sum_{j=1}^{7} d(e, g_j) < 2, \tag{3.5}$$

from which the result of the theorem will follow. Regarding (3.4), writing explicitly the definitions of the points g_j, we have

$$\begin{aligned}
\exp(X_1)\, g_1 \cdots g_7 = \exp(X_1)\exp(\epsilon^s Y_1)\exp(-(1-\epsilon)X_1)\exp(-\epsilon X_1)\delta_\epsilon(h)^{-1}\\
\cdot \exp(\epsilon X_2)\exp((\tfrac{1}{2}-\epsilon)X_2)\exp(\epsilon^s Y_2)\exp(\tfrac{1}{2}X_2)\exp(\epsilon^s Y_3).
\end{aligned}$$

Then, using the fact that h is in $Z(G)$, we rewrite the right-hand side in terms of conjugations as

$$\delta_\epsilon(h)^{-1}\,\mathrm{C}_{\exp(X_1)}\left(\exp(\epsilon^s Y_1)\right)\cdot \mathrm{C}_{\exp\left(\frac{1}{2}X_2\right)}\left(\exp(\epsilon^s Y_2)\right)\cdot \mathrm{C}_{\exp(X_2)}\left(\exp(\epsilon^s Y_3)\right)\exp(X_2).$$

Since Y_1, Y_2, Y_3 were chosen to satisfy (3.3), the above term reduces to $\exp(X_2)$, thus showing (3.4). To show (3.5), we note that as the points g_j are all dilations of some fixed points, the individual distances are given by

$$\begin{aligned}
d(e, g_1) &= \epsilon^{s/(s-1)} d(e, \exp(Y_1)),\\
d(e, g_2) &= 1 - \epsilon\\
d(e, g_3) &= \epsilon d(e, \exp(-X_1) h^{-1} \exp(X_2)) = \epsilon d(h \exp(X_1), \exp(X_2)),\\
d(e, g_4) &= \frac{1}{2} - \epsilon,\\
d(e, g_5) &= \epsilon^{s/(s-1)} d(e, \exp(Y_2)),\\
d(e, g_6) &= \frac{1}{2} \quad \text{and}\\
d(e, g_7) &= \epsilon^{s/(s-1)} d(e, \exp(Y_3)).
\end{aligned}$$

Summing all the above distances, we get

$$\sum_{j=1}^{7} d(e, g_j) = 2 - (2 - D)\epsilon + o(\epsilon), \quad \text{as } \epsilon \to 0,$$

where

$$D = d(h \exp(X_1), \exp(X_2)).$$

By the choice of h from (3.2), we have $-(2 - D) < 0$. Therefore, for small enough $\epsilon > 0$, we deduce (3.5).

We finally estimate using left-invariance, equations (3.4) and (3.5), and the triangle inequality, that

$$d(\exp(X_1), \exp(X_2)) = d(e, g_1 \cdots g_7) \leq \sum_{i=1}^{7} d(e, g_i) < 2,$$

for small enough $\epsilon > 0$. Since the considered corner from $\exp(X_1)$ to $\exp(X_2)$ has length equal to 2, where X_1 and X_2 were arbitrary linearly independent unit-norm vectors of the first layer V_1, we conclude that corners in the group G of step s are not length minimizing. □

Acknowledgement. The authors thank A. Ottazzi, D. Vittone, and the anonymous referees for their helpful remarks. E.L.D. acknowledges the support of the Academy of Finland project no. 288501.

A Séminaire Bourbaki presentation including the content of this paper and related work has been subsequently given by L. Rifford [Rif16].

References

[ABB15] Andrei Agrachev, Davide Barilari, and Ugo Boscain, *Introduction to Riemannian and Sub-Riemannian geometry*, Manuscript (2015).

[Agr14] Andrei A. Agrachev, *Some open problems*, Geometric control theory and sub-Riemannian geometry, Springer INdAM Ser., vol. 5, Springer, Cham, 2014, pp. 1–13.

[AS04] Andrei A. Agrachev and Yuri L. Sachkov, *Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences, vol. 87, Springer-Verlag, Berlin, 2004, Control Theory and Optimization, II.

[BH93] Robert L. Bryant and Lucas Hsu, *Rigidity of integral curves of rank 2 distributions*, Invent. Math. **114** (1993), no. 2, 435–461.

[GK95] Chr. Golé and R. Karidi, *A note on Carnot geodesics in nilpotent Lie groups*, J. Dynam. Control Systems **1** (1995), no. 4, 535–549.

[Gro96] Mikhail Gromov, *Carnot-Carathéodory spaces seen from within*, Sub-Riemannian geometry, Progr. Math., vol. 144, Birkhäuser, Basel, 1996, pp. 79–323.

[Gro99] ———, *Metric structures for Riemannian and non-Riemannian spaces*, Progress in Mathematics, vol. 152, Birkhäuser Boston Inc., Boston, MA, 1999, Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French version by Sean Michael Bates.

[Ham90] Ursula Hamenstädt, *Some regularity theorems for Carnot-Carathéodory metrics*, J. Differential Geom. **32** (1990), no. 3, 819–850.

[Jea14] Frédéric Jean, *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*, Springer Briefs in Mathematics, Springer, Cham, 2014.

[LD15] Enrico Le Donne, *A primer of Carnot groups*, Manuscript (2015).

[LDLMV13] Enrico Le Donne, Gian Paolo Leonardi, Roberto Monti, and Davide Vittone, *Extremal curves in nilpotent Lie groups*, Geom. Funct. Anal. **23** (2013), no. 4, 1371–1401.

[LDLMV14] Enrico Le Donne, Gian Paolo Leonardi, Roberto Monti, and Davide Vittone, *Extremal polynomials in stratified groups*, Preprint, submitted (2014).

[LDLMV15] Enrico Le Donne, Gian Paolo Leonardi, Roberto Monti, and Davide Vittone, *Corners in non-equiregular sub-Riemannian manifolds*, ESAIM Control Optim. Calc. Var. **21** (2015), no. 3, 625–634.

[LM08] Gian Paolo Leonardi and Roberto Monti, *End-point equations and regularity of sub-Riemannian geodesics*, Geom. Funct. Anal. **18** (2008), no. 2, 552–582.

[LS94] Wensheng Liu and Héctor J. Sussman, *Abnormal sub-Riemannian minimizers*, Differential equations, dynamical systems, and control science **152** (1994), xl+946, A Festschrift in honor of Lawrence Markus.

[LS95] ———, *Shortest paths for sub-Riemannian metrics on rank-two distributions*, Mem. Amer. Math. Soc. **118** (1995), no. 564, x+104.

[Mon94] Richard Montgomery, *Abnormal minimizers*, SIAM J. Control Optim. **32** (1994), no. 6, 1605–1620.

[Mon02] ———, *A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, vol. 91, American Mathematical Society, Providence, RI, 2002.

[Mon14a] Roberto Monti, *A family of nonminimizing abnormal curves*, Ann. Mat. Pura Appl. (4) **193** (2014), no. 6, 1577–1593.

[Mon14b] ———, *The regularity problem for sub-Riemannian geodesics*, Geometric control theory and sub-Riemannian geometry, Springer INdAM Ser., vol. 5, Springer, Cham, 2014, pp. 313–332.

[Mon14c] ———, *Regularity results for sub-Riemannian geodesics*, Calc. Var. Partial Differential Equations **49** (2014), no. 1-2, 549–582.

[Rif14] Ludovic Rifford, *Sub-Riemannian geometry and optimal transport*, Springer Briefs in Mathematics, Springer, Cham, 2014.

[Rif16] ———, *Singulières minimisantes en géométrie sous-riemannienne [d'après Hakavuori, Le Donne, Leonardi, Monti...]*, Astérisque (2016), Exp. No. 1113.

[Str86] Robert S. Strichartz, *Sub-Riemannian geometry*, J. Differential Geom. **24** (1986), no. 2, 221–263.

[Str89] ———, *Corrections to: "Sub-Riemannian geometry" [J. Differential Geom.* **24** *(1986), no. 2, 221–263; (88b:53055)]*, J. Differential Geom. **30** (1989), no. 2, 595–596.

[Sus96] Héctor J. Sussmann, *A cornucopia of four-dimensional abnormal sub-Riemannian minimizers*, Sub-Riemannian geometry, Progr. Math., vol. 144, Birkhäuser, Basel, 1996, pp. 341–364.

[Sus14] H. J. Sussmann, *A regularity theorem for minimizers of real-analytic subriemannian metrics*, Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on, Dec 2014, pp. 4801–4806.

[War83] Frank W. Warner, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, vol. 94, Springer-Verlag, New York, 1983, Corrected reprint of the 1971 edition.

E-mail address: `eero.j.hakavuori@jyu.fi`

E-mail address: `enrico.ledonne@jyu.fi`

(Hakavuori and Le Donne) Department of Mathematics and Statistics, University of Jyväskylä, 40014 Jyväskylä, Finland

[B]

Blowups and blowdowns of geodesics in Carnot groups

E. Hakavuori and E. Le Donne

Preprint

BLOWUPS AND BLOWDOWNS OF GEODESICS IN CARNOT GROUPS

EERO HAKAVUORI AND ENRICO LE DONNE

Abstract. This paper provides some partial regularity results for geodesics (i.e., isometric images of intervals) in arbitrary sub-Riemannian and sub-Finsler manifolds. Our strategy is to study infinitesimal and asymptotic properties of geodesics in Carnot groups equipped with arbitrary sub-Finsler metrics. We show that tangents of Carnot geodesics are geodesics in some groups of lower nilpotency step. Namely, every blowup curve of every geodesic in every Carnot group is still a geodesic in the group modulo its last layer. Then as a consequence we get that in every sub-Riemannian manifold any s times iterated tangent of any geodesic is a line, where s is the step of the sub-Riemannian manifold in question. With a similar approach, we also show that blowdown curves of geodesics in sub-Riemannian Carnot groups are contained in subgroups of lower rank. This latter result is also extended to rough geodesics.

Contents

Date: October 5, 2018.

2010 *Mathematics Subject Classification.* 53C17, 49K21, 28A75.

Key words and phrases. Geodesics, tangent cones, asymptotic cones, sub-Riemannian geometry, sub-Finsler geometry, Carnot groups, regularity of length minimizers.

E.H. was supported by the Vilho, Yrjö and Kalle Väisälä Foundation. E.L.D. was partially supported by the Academy of Finland (grant 288501 '*Geometry of subRiemannian groups*') and by the European Research Council (ERC Starting Grant 713998 GeoMeG '*Geometry of Metric Groups*').

1. Introduction

In sub-Riemannian geometry, one of the major open problems is the regularity of geodesics, i.e., of isometric embeddings of intervals. Because of the presence of abnormal curves, a priori sub-Riemannian geodesics only have Lipschitz regularity, yet all known examples are C^∞. For a modern introduction to the topic we refer to [Vit14].

We approach the differentiability problem by considering the infinitesimal geometry, which is given by sub-Riemannian Carnot groups, and within them studying limits of dilated curves, called tangents or blowups. The main aim of this paper is to show that iterating the process of taking tangents one necessarily obtains only lines:

Theorem 1.1. *If γ is a geodesic in a sub-Riemannian manifold, then every s times iterated tangent of γ is a line, where s is the step of the sub-Riemannian manifold.*

This is a generalization of other partial results that have already been attained using a similar approach: In [HL16] we showed that tangents of geodesics are not corners, and in [MPV18a] it is shown that among all tangents at a point, one of the tangents is a line. Here "line" means a left translation of a one-parameter subgroup and "corner" means two half-lines joined together not forming a line.

A basic fact from metric geometry is that tangents of geodesics are themselves infinite geodesics. Therefore knowledge about infinite geodesics can help understand the regularity problem. For this reason, in this present work in addition to tangents we consider asymptotic cones, also called blowdowns, of infinite geodesics in Carnot groups.

Before stating our more specific results, we shall specify the notion of tangents. The notion is the same as previously used in [HL16, MPV18a]. We shall mainly restrict our considerations to Carnot groups, while allowing arbitrary length distances. For the notion of tangents within manifolds, we refer to [MPV18b].

Let G be a sub-Finsler Carnot group, cf. the standard definition in [LD17]. In G we have a Carnot-Carathéodory distance d defined by a norm on the horizontal space V_1 of G, and we have a one-parameter family of dilations, denoted by $(\delta_h)_{h>0}$. Let I be an open interval in $\mathbb{R}$, possibly $I = \mathbb{R}$. Let $\gamma : I \to G$ be a 1-Lipschitz curve and fix $\bar{t} \in I$. Denote by $\gamma_h : I_h \to G$ the curve defined on $I_h := \frac{1}{h}(I - \bar{t})$ by

$$\gamma_h(t) := \delta_{\frac{1}{h}} \left(\gamma(\bar{t})^{-1}\gamma(\bar{t} + ht)\right).$$

Notice that the last definition is just the non-abelian version of the difference quotient used in the definition of derivatives. It is trivial to check that γ_h is 1-Lipschitz and $\gamma_h(0) = 1_G$ for all $h \in (0, \infty)$. Consequently, by Ascoli-Arzelá, for every sequence $h_j \to 0$ there is a subsequence h_{j_k} and a curve $\sigma : \mathbb{R} \to G$ such that $\gamma_{h_{j_k}} \to \sigma$ uniformly on compact sets of $\mathbb{R}$. Hence, we define the collection of *tangents* as the nonempty set

$$\operatorname{Tang}(\gamma,\bar{t}) := \left\{\sigma \mid \exists h_j \to 0 : \gamma_{h_j} \to \sigma\right\}.$$

In the case where $\gamma : I \to M$ is a Lipschitz curve on a sub-Riemannian or sub-Finsler manifold M, we will also denote by $\operatorname{Tang}(\gamma,\bar{t})$ the collection of metric tangents of γ at $\bar{t}$. In this case, the elements of $\operatorname{Tang}(\gamma,\bar{t})$ are no longer curves in M, but instead curves in the metric tangent space $\tilde{M}$, also called the nilpotent approximation of M. We refer to [Jea14, Section 2.3.1] and [MPV18b] for details on this more general construction.

When $I = \mathbb{R}$, we will also consider limits of the curves γ_h for sequences $h_j \to \infty$. Similarly to the case $h_j \to 0$, for every sequence $h_j \to \infty$ there is a subsequence h_{j_k} and a curve $\sigma : \mathbb{R} \to G$ such that $\gamma_{h_{j_k}} \to \sigma$ uniformly on compact sets of $\mathbb{R}$, so we define the collection of *asymptotic cones* as the nonempty set

$$\operatorname{Asymp}(\gamma) := \left\{\sigma \mid \exists h_j \to \infty : \gamma_{h_j} \to \sigma\right\}.$$

The definition of $\operatorname{Asymp}(\gamma)$ is independent on the choice of $\bar{t}$ and technically the assumption that $I = \mathbb{R}$ is not necessary if we use the domains I_h as in the definition of tangents. However if I is bounded, the domains I_h degenerate to a point, and in the case where I is a half-line, all arguments are only superficially different from the line case.

Finally, we define the iterated tangent cones as the set of all tangents of (iterated) tangents at 0, i.e., for each $k \geq 1$ we define

$$\operatorname{Tang}^{k+1}(\gamma,\bar{t}) := \bigcup_{\sigma \in \operatorname{Tang}^k(\gamma,\bar{t})} \operatorname{Tang}(\sigma, 0).$$

The elements $\sigma \in \operatorname{Tang}^k(\gamma,\bar{t})$ for any $\bar{t}$ are called *k times iterated tangents of γ*. We remark that a simple diagonal argument[1] shows that iterated tangents are also tangents, i.e., that

$$\cdots \subset \operatorname{Tang}^{k+1}(\gamma,\bar{t}) \subset \operatorname{Tang}^{k}(\gamma,\bar{t}) \subset \cdots \subset \operatorname{Tang}(\gamma,\bar{t}).$$

Assume $\gamma : I \to G$ is a geodesic, i.e., $d(\gamma(a),\gamma(b)) = |a-b|$, for all $a, b \in I$. Our main results in the Carnot group setting are that every element in $\operatorname{Tang}(\gamma,\bar{t})$ is a geodesic also when projected into some quotient group of lower step, and that every element in $\operatorname{Asymp}(\gamma)$ is a geodesic inside some subgroup of lower rank (see Theorem 1.2 and Corollary 1.5, respectively).

[1] If $\gamma_{h_j} \to \sigma$ and $\sigma_{k_j} \to \eta$ for some $h_j, k_j \to 0$, then for all ℓ we have $\gamma_{k_\ell h_j} \to \sigma_{k_\ell}$ and so, by a diagonal argument, there is a sequence ℓ_j such that $\gamma_{k_{\ell_j} h_j} \to \eta$.

1.1. **Statement of the results.** Unless otherwise stated, in what follows G will be a sub-Finsler Carnot group of nilpotency step s and $V_1 \oplus \cdots \oplus V_s = \mathfrak{g}$ will be the stratification of the Lie algebra $\mathfrak{g}$ of G. We denote by $\pi : G \to G/[G,G]$ the projection on the abelianization and by $\pi_{s-1} : G \to G/\exp(V_s)$ the projection modulo the last layer V_s of $\mathfrak{g}$.

Both groups $G/[G,G]$ and $G/\exp(V_s)$ are canonically equipped with structures of sub-Finsler Carnot groups (see Proposition 2.1). The normed vector space $G/[G,G]$ is also further canonically identified with the first layer V_1 and its dimension is the rank of G. The group $G/\exp(V_s)$ has nilpotency step $s-1$, one lower than the original group G.

Theorem 1.2 (Blowup of geodesics)**.** *If $\gamma : I \to G$ is a geodesic and $t \in I$, then for every $\sigma \in \mathrm{Tang}(\gamma, t)$, the curve $\pi_{s-1} \circ \sigma : \mathbb{R} \to G/\exp(V_s)$ is a geodesic.*

This result implies the previously known ones from [HL16] that corners are not minimizing and from [MPV18a] that in the sub-Riemannian case one of the tangents is a line. In fact, iterating the above result, we get the following corollary.

Corollary 1.3. *If $\gamma : I \to G$ is a geodesic and $t \in I$, then for every $\sigma \in \mathrm{Tang}^{s-1}(\gamma, t)$, the horizontal projection $\pi \circ \sigma$ is a geodesic. In particular, if G is sub-Riemannian then every $\sigma \in \mathrm{Tang}^{s-1}(\gamma, t)$ is a line.*

In the sub-Riemannian setting, since all infinite geodesics in step 2 are lines (see Theorem 5.6), Theorem 1.1 and Corollary 1.3 can be improved slightly, decreasing the number of iterations needed from s and $s-1$ to $s-1$ and $s-2$, respectively.

As applications of the existence of a line tangent, we show that in every non-Abelian Carnot group where in the abelianization the infinite geodesics are lines, there is always a geodesic that loses optimality whenever it is extended (see Proposition 6.1), and show that the non-minimality of corners holds also in the non-constant rank case (see Proposition 6.2).

As mentioned in the introduction, every element in $\mathrm{Tang}(\gamma, t)$ is an infinite geodesic. We provide next other results that are valid for any infinite geodesics regardless of whether or not they are tangents.

Theorem 1.4. *If $\gamma : \mathbb{R} \to G$ is a geodesic such that $\pi \circ \gamma : \mathbb{R} \to G/[G,G]$ is not a geodesic, then there exist $R > 0$ and a hyperplane $W \subset V_1$ such that $\mathrm{Im}(\pi \circ \gamma) \subset B_{V_1}(W, R)$.*

In the above theorem, we denote by $B_{V_1}(W, R)$ the R-neighborhood of W within V_1. To prove Theorem 1.4 we shall adopt a wider viewpoint. In fact, we will consider rough geodesics and still have the same rigidity result (see Theorem 4.2).

It is possible that the claim of Theorem 1.4 could be strengthened to say that the projection of the geodesic is asymptotic to the hyperplane. In Corollary 7.20, we show that this is true for the only known family of examples of non-line infinite geodesics, arising from the explicit study of geodesics in the Engel group, see [AS15]. We also show that each of these geodesics is in a finite neighborhood of a line in the Engel group itself. However, by lifting the same geodesics to a step 4 Carnot group, we show that there exist infinite geodesics that are not in a finite neighborhood of any line (see Corollary 7.28).

Since in Euclidean spaces the only infinite geodesics are the straight lines, an immediate consequence of Theorem 1.4 is the following.

Corollary 1.5 (Blowdown of geodesics)**.** *If γ is a geodesic in a sub-Riemannian Carnot group $G \neq \mathbb{R}$, then there exists a proper Carnot subgroup $H < G$ containing every element of* $\operatorname{Asymp}(\gamma)$*.*

As with Theorem 1.4, Corollary 1.5 admits a generalization for rough geodesics (see Corollary 4.10). As a stepping stone to this generalization, we also prove that rough geodesics in Euclidean spaces have unique blowdowns (see Proposition 4.7).

Similarly as with Theorem 1.2, we can iterate Corollary 1.5 and deduce that some blowdown of an infinite geodesic in a sub-Riemannian Carnot group must be a line. Furthermore, we show that in sub-Riemannian Carnot groups, every blowdown of an infinite geodesic is a line or an abnormal geodesic (see Proposition 5.5).

1.2. **Organization of the paper.** In Section 2 we discuss technical lemmas based on linear algebra and our error correction procedure. We introduce the concepts of minimal height and size. Proposition 2.24 is the crucial estimate and is a variant of a triangle inequality with an error term depending on the notion of size. This proposition is the key ingredient for both the proof of Theorem 1.2 and the proof of Theorem 1.4.

In Sections 3 and 4 we prove our main results. Section 3 covers our results about tangents of geodesics: Theorems 1.1 and 1.2, and Corollary 1.3. We also give a quantified version in Theorem 3.1, which expresses the extent to which the projection of a geodesic may fail to be minimizing. Section 4 covers our results about infinite geodesics and blowdowns: Theorem 1.4 and Corollary 1.5, and their rough counterparts: Theorem 4.2 and Corollary 4.10.

In Sections 5 and 6 we discuss some applications of our main results. In Section 5 we consider the Hamiltonian point of view of geodesics as normal or abnormal extremals, and prove the statements about abnormality of blowdowns (Proposition 5.5) and infinite geodesics in step 2 Carnot groups (Proposition 5.6). In Section 6 we consider applications of the existence of a line tangent. We prove the existence of non-extendable geodesics in non-abelian Carnot groups (Proposition 6.1) and the non-minimality of corners in non-constant rank sub-Riemannian manifolds (Proposition 6.2).

In Section 7 we discuss to which extent one can expect an improvement of the blowdown result Theorem 1.4, restricting our attention to rank-2 Carnot groups. In

Section 7.1 we cover preliminaries on lines in Carnot groups and study when two lines are at bounded distance. In Section 7.2 we consider the example of an infinite non-line geodesic in the Engel group. We use this curve to find a counter-example to one possible strengthening of Theorem 1.4.

2. Preliminaries: minimal height, size, and error correction

2.1. Carnot structures on quotients.

Proposition 2.1. *On $G/[G,G]$ and on $G/\exp(V_s)$ there are canonical structures of sub-Finsler Carnot groups such that the projections $\pi : G \to G/[G,G]$ and $\pi_{s-1} : G \to G/\exp(V_s)$ are submetries. In particular, for any $g_1, g_2 \in G$ there exists $h \in \exp(V_s)$ such that*

$$d(\pi_{s-1}(g_1), \pi_{s-1}(g_2)) = d(g_1, hg_2).$$

Proof. This proof is well known. It probably goes back to Berestovskii [Ber89, Theorem 1]. The key point here is that both $\exp(V_s)$ and $[G,G]$ are normal subgroups. Thus one can define the distance of two points in the quotient as the distance between their preimages. The reader can find the details in [LR16, Corollary 2.11]. □

2.2. Minimal height of a parallelotope and its properties.

Definition 2.2 (Minimal height of a parallelotope) Let V be a normed vector space with distance d_V. The *minimal height* of an m-tuple of points $(a_1, \ldots, a_m) \in V^m$ is the smallest height of the parallelotope generated by the points, i.e.,

$$\mathrm{MinHeight}(a_1, \ldots, a_m) = \min_{j \in \{1,\ldots,m\}} d_V(a_j, \mathrm{span}\{a_1, \ldots, \hat{a}_j, \ldots, a_m\}).$$

Remarks 2.3.1 Points $a_1, \ldots, a_m$ in a normed vector space are linearly independent if and only if $\mathrm{MinHeight}(a_1, \ldots, a_m) \neq 0$.
2.3.2 Assume V is a Euclidean space $\mathbb{R}^r$ and denote by vol_m the usual m-dimensional volume. Let $\mathcal{P}(a_1, \ldots, a_m)$ denote the parallelotope generated by the vectors $a_1, \ldots, a_m$. Notice that the volume of $\mathcal{P}(a_1, \ldots, a_m)$ equals the volume of any base $\mathcal{P}(a_1, \ldots, \hat{a}_j, \ldots, a_m)$ times the corresponding height, which is $d(a_j, \mathrm{span}\{a_1, \ldots, \hat{a}_j, \ldots, a_m)$. Hence, we have

$$\begin{aligned}\mathrm{MinHeight}(a_1, \ldots, a_m) &= \min_{j \in \{1,\ldots,m\}} \frac{\mathrm{vol}_m \mathcal{P}(a_1, \ldots, a_m)}{\mathrm{vol}_{m-1} \mathcal{P}(a_1, \ldots, \hat{a}_j, \ldots, a_m)} \\ &= \frac{\mathrm{vol}_m \mathcal{P}(a_1, \ldots, a_m)}{\max_{j \in \{1,\ldots,m\}} \mathrm{vol}_{m-1} \mathcal{P}(a_1, \ldots, \hat{a}_j, \ldots, a_m)}.\end{aligned}$$

Hence, if $\mathcal{P}^* := \mathcal{P}(a_1, \ldots, \hat{a}_j, \ldots, a_m)$ is a face of the parallelotope with maximal $(m-1)$-dimensional volume, then

$$\mathrm{MinHeight}(a_1, \ldots, a_m) = \frac{\mathrm{vol}_m \mathcal{P}(a_1, \ldots, a_m)}{\mathrm{vol}_{m-1} \mathcal{P}^*} = d(a_j, \mathrm{span}\, \mathcal{P}^*).$$

We next prove a basic lemma that uses the notion of minimal height to bound the entries of the inverse of a matrix. This bound will then be used in Lemma 2.11.

Lemma 2.4. *Let A be a matrix with columns $A_1, \dots, A_r \in \mathbb{R}^r$. If* $\mathrm{MinHeight}(A_1, \dots, A_r) > 0$*, then A is invertible and its inverse B has entries B_{kj} bounded by*

$$|B_{kj}| \le \frac{1}{\mathrm{MinHeight}(A_1, \dots, A_r)}, \qquad \forall k, j = 1, \dots, r.$$

Proof. The fact that A is invertible follows from Remark 2.3.1. For the estimate on the entries of the inverse, we will use a well-known formula from linear algebra (see [Lan71, page 219]): If $A_{(k,j)}$ denotes the matrix A with row k and column j removed, then the entries of B can be calculated by

$$B_{kj} = (-1)^{k+j} \frac{\det A_{(k,j)}}{\det A}. \tag{2.5}$$

Fix $j, k \in \{1, \dots, r\}$. Let $P_k : \mathbb{R}^r \to \mathbb{R}^{r-1}$ be the projection that forgets the k-th coordinate:

$$P_k(y_1, \dots, y_r) := (y_1, \dots, \hat{y}_k, \dots, y_r).$$

Consider the following parallelotopes: Let $\mathcal{P}$ be the r-parallelotope in $\mathbb{R}^r$ determined by the points $A_1, \dots, A_r$, let $\mathcal{P}_j$ be the $(r-1)$-parallelotope in $\mathbb{R}^r$ determined by the same points excluding the vertex A_j, and let $\mathcal{P}_j^k = P_k(\mathcal{P}_j)$, which is an $(r-1)$-parallelotope in $\mathbb{R}^{r-1}$.

The geometric interpretation of the determinant states that

$$|\det A| = \mathrm{vol}_r(\mathcal{P}) \quad \text{and} \quad \left|\det A_{(k,j)}\right| = \mathrm{vol}_{r-1}(\mathcal{P}_j^k).$$

Moreover, since $P_k(\mathcal{P}_j) = \mathcal{P}_j^k$ and the projection P_k is 1-Lipschitz, we have

$$\mathrm{vol}_{r-1}(\mathcal{P}_j^k) \le \mathrm{vol}_{r-1}(\mathcal{P}_j).$$

By these last two observations, we have that

$$\frac{\left|\det A_{(k,j)}\right|}{|\det A|} = \frac{\mathrm{vol}_{r-1}(\mathcal{P}_j^k)}{\mathrm{vol}_r(\mathcal{P})} \le \frac{\mathrm{vol}_{r-1}(\mathcal{P}_j)}{\mathrm{vol}_r(\mathcal{P})}. \tag{2.6}$$

Let $L_j := \mathrm{span}\{a_1, \dots, \hat{a}_j, \dots, a_m\}$. Since L_j is the span of $\mathcal{P}_j$ and $\mathcal{P}_j$ is a face of $\mathcal{P}$, we calculate the volume of $\mathcal{P}$ as in Remark 2.3.2 as

$$\mathrm{vol}_r(\mathcal{P}) = d(a_j, L_j)\, \mathrm{vol}_{r-1}(\mathcal{P}_j). \tag{2.7}$$

By the definition of MinHeight as the minimum of the distances $d(a_j, L_j)$, we conclude that

$$|B_{kj}| \overset{(2.5)}{=} \frac{\left|\det A_{(k,j)}\right|}{|\det A|} \overset{(2.6)}{\le} \frac{\mathrm{vol}_{r-1}(\mathcal{P}_j)}{\mathrm{vol}_r(\mathcal{P})} \overset{(2.7)}{=} \frac{1}{d(a_j, L_j)} \le \frac{1}{\mathrm{MinHeight}(A_1, \dots, A_r)}. \qquad \square$$

2.3. Size of a configuration and error correction.

Definition 2.8 (Size of a configuration) Let G be a Carnot group. The *size* of an $(m+1)$-tuple of points $(g_0, \dots, g_m) \in G^{m+1}$ is

(2.9)
$$\mathrm{Size}(g_0, \dots, g_m) = \mathrm{MinHeight}(\pi(g_1) - \pi(g_0), \pi(g_2) - \pi(g_1), \dots, \pi(g_m) - \pi(g_{m-1})).$$

Remark 2.10 Remark 2.3.1 states that non-zero MinHeight characterizes linear independence of points. Analogously, $\mathrm{Size}(g_0, \dots, g_m) \neq 0$ if and only if the horizontal projections $\pi(g_0), \dots, \pi(g_m) \in G/[G,G]$ are in general position.

The reason to consider this notion of size stems from Lemma 2.20 below, which describes our error correction procedure. Within this lemma, we need to bound the norms of solutions to a certain linear system. A convenient bound is given in Lemma 2.11 in terms of the size of a configuration of points. This dependence of the bound of solutions on the size of a configuration is the reason we are able to give restrictions on the behavior of tangents and asymptotic cones of geodesics.

Lemma 2.11 (Linear system of corrections)**.** *For every Carnot group G of rank r and step $s \geq 2$, there exists a constant $K > 0$ with the following property:*

Let $x_0, \dots, x_r \in G$ and $X_j := \log(x_{j-1}^{-1} x_j)$, for $j = 1, \dots, r$. If $\mathrm{Size}(x_0, \dots, x_r) > 0$, then for every $Z \in V_s$ there exist $Y_1, \dots, Y_r \in V_{s-1}$ such that

$$[Y_1, X_1] + \dots + [Y_r, X_r] = Z \tag{2.12}$$

and

$$d(1_G, \exp(Y_j))^{s-1} \leq K \frac{d(1_G, \exp(Z))^s}{\mathrm{Size}(x_0, \dots, x_r)}, \qquad \forall j \in 1, \dots, r. \tag{2.13}$$

Proof. Fix arbitrary norms on the vector spaces V_{s-1} and V_s, and denote them generically as $\|\cdot\|$. Observe that the functions

$$W \mapsto d(1_G, \exp(W))^{s-1} \quad \text{and} \quad Z \mapsto d(1_G, \exp(Z))^s$$

are 1-homogeneous with respect to scalar multiplication. Therefore there exists a constant $C_1 > 1$ such that

$$\|W\| \simeq_{C_1} d(1_G, \exp(W))^{s-1} \quad \text{and} \quad \|Z\| \simeq_{C_1} d(1_G, \exp(Z))^s, \tag{2.14}$$

where $a \simeq_c b$ stands for $b/c \leq a \leq cb$.

Fix a basis $\bar{X}_1, \dots, \bar{X}_r$ of V_1. Observe that the map $(W_1, \dots, W_r) \mapsto [W_1, \bar{X}_1] + \dots + [W_r, \bar{X}_r]$ is a linear surjection between the normed vector spaces $(V_{s-1})^r$ and V_s, where on $(V_{s-1})^r$ we use the norm $\max_{i=1,\dots,r}\{\|W_i\|\}$. Thus the map can be restricted to some subspace so that it becomes a biLipschitz linear isomorphism. In other words, there exists a constant $C_2 > 1$ such that for all $Z \in V_s$ there exist vectors $W_1, \dots, W_r \in V_{s-1}$ such that

$$Z = [W_1, \bar{X}_1] + \dots + [W_r, \bar{X}_r] \tag{2.15}$$

and

$$\max_{i=1,\dots,r}\{\|W_i\|\} \simeq_{C_2} \|Z\| . \tag{2.16}$$

The choice of the basis $\bar{X}_1, \dots, \bar{X}_r$ lets us identify $G/[G,G]$ with $\mathbb{R}^r$ via the linear isomorphism $\phi : \mathbb{R}^r \to G/[G,G]$ defined by

$$\phi(a_1, \dots, a_r) := \exp(a_1\bar{X}_1 + \cdots + a_r\bar{X}_r + \mathfrak{g}^2),$$

where $\mathfrak{g}^2 = V_2 \oplus \cdots \oplus V_s$ and so $\exp(\mathfrak{g}^2) = [G,G]$. As a linear isomorphism, for some $C_3 > 1$, the map ϕ is a C_3-biLipschitz equivalence between $\mathbb{R}^r$ with the standard metric and $G/[G,G]$ with the quotient metric. Consequently, we have

$$\operatorname{MinHeight}(a_1, \dots, a_r) \simeq_{C_3} \operatorname{MinHeight}(\phi(a_1), \dots, \phi(a_r)) \qquad \forall a_1, \dots, a_r \in \mathbb{R}^r. \tag{2.17}$$

We now show that the constant $K := rC_1^2C_2C_3$ satisfies the conclusion of the lemma. Take an arbitrary $Z \in V_s$ and write it as in (2.15) for some $W_1, \dots, W_r \in V_{s-1}$ satisfying the bound (2.16).

Given points $x_0, \dots, x_r \in G$ with $\operatorname{Size}(x_0, \dots, x_r) > 0$, let $v_0, \dots, v_r \in \mathbb{R}^r$ be such that $\phi(v_j) = \pi(x_j)$ and write $v_j = (v_{j,1}, \dots, v_{j,r})$. In other words,

$$x_j \in \exp\Big(\sum_{k=1}^r v_{j,k}\bar{X}_k + \mathfrak{g}^2\Big).$$

Let A be the $r \times r$ matrix whose j-th column is $A_j := v_j - v_{j-1}$, so that

$$x_{j-1}^{-1}x_j \in \exp\Big(\sum_{k=1}^r (v_{j,k} - v_{j-1,k})\bar{X}_k + \mathfrak{g}^2\Big) = \exp\Big(\sum_{k=1}^r A_{kj}\bar{X}_k + \mathfrak{g}^2\Big). \tag{2.18}$$

The bound (2.17) combined with linearity of ϕ implies that $\operatorname{MinHeight}(A_1, \dots, A_r)$ is comparable to $\operatorname{Size}(x_0, \dots, x_r)$:

$$\begin{aligned}
\operatorname{MinHeight}(A_1, \dots, A_r) &= \operatorname{MinHeight}(v_1 - v_0, \dots, v_r - v_{r-1}) \\
&\simeq_{C_3} \operatorname{MinHeight}(\phi(v_1 - v_0), \dots, \phi(v_r - v_{r-1})) \\
&= \operatorname{MinHeight}(\phi(v_1) - \phi(v_0), \dots, \phi(v_r) - \phi(v_{r-1})) \\
&= \operatorname{MinHeight}(\pi(x_1) - \pi(x_0), \dots, \pi(x_r) - \pi(x_{r-1})) \\
&= \operatorname{Size}(x_0, \dots, x_r).
\end{aligned}$$

In particular, $\operatorname{MinHeight}(A_1, \dots, A_r) > 0$ so we further deduce by Lemma 2.4 that A is invertible and its inverse B satisfies

$$|B_{jl}| \le \frac{1}{\operatorname{MinHeight}(A_1, \dots, A_r)} \le \frac{C_3}{\operatorname{Size}(x_0, \dots, x_r)}. \tag{2.19}$$

Set $Y_j := \sum_{l=1}^r B_{jl}W_l$. We shall verify that this choice of Y_j's satisfies the conclusion of the lemma, i.e., the properties (2.12) and (2.13).

The first property is deduced from bilinearity of the Lie bracket and the fact that AB is the identity matrix. By (2.18), we can write the vectors X_j as sums

$$X_j = \log(x_{j-1}^{-1} x_j) = \sum_{k=1}^{r} A_{kj} \bar{X}_k + \mathfrak{g}^2.$$

Since $[V_{s-1}, \mathfrak{g}^2] = [V_{s-1}, V_2 \oplus \cdots \oplus V_s] = 0$, it follows by bilinearity of the bracket that

$$\sum_{j=1}^{r} [Y_j, X_j] = \sum_{j=1}^{r} \left[\sum_{l=1}^{r} B_{jl} W_l, \sum_{k=1}^{r} A_{kj} \bar{X}_k \right] = \sum_{k=1}^{r} \sum_{l=1}^{r} \sum_{j=1}^{r} A_{kj} B_{jl} \left[W_l, \bar{X}_k \right].$$

Using the fact that AB is the identity matrix, we have $\sum_{j=1}^{r} A_{kj} B_{jl} = \delta_{kl}$, so the sum simplifies to

$$\sum_{j=1}^{r} [Y_j, X_j] = \sum_{k=1}^{r} \left[W_k, \bar{X}_k \right] \overset{(2.15)}{=} Z,$$

showing property (2.12).

Regarding, property (2.13), we first observe that estimating each $\|W_l\|$ by (2.16) and each $|B_{jl}|$ by (2.19), we can bound $\|Y_j\|$ by

$$\|Y_j\| = \|B_{jl} W_l\| \leq \sum_{l=1}^{r} |B_{jl}| \, \|W_l\| \leq \sum_{l=1}^{r} \frac{C_2 C_3}{\mathrm{Size}(x_0, \dots, x_r)} \|Z\| = \frac{r C_2 C_3}{\mathrm{Size}(x_0, \dots, x_r)} \|Z\|.$$

Then, using (2.14) to give bounds for $\|Y_j\|$ and $\|Z\|$, we conclude that

$$C_1^{-1} d(1_G, \exp(Y_j))^{s-1} \leq \|Y_j\| \leq \frac{r C_2 C_3}{\mathrm{Size}(x_0, \dots, x_r)} \|Z\| \leq \frac{r C_1 C_2 C_3}{\mathrm{Size}(x_0, \dots, x_r)} d(1_G, \exp(Z))^s.$$

Hence the lemma holds with the proposed constant $K = r C_1^2 C_2 C_3$. □

As mentioned before, the following lemma describes our error correction procedure. The strategy is the same as used before in [LM08, HL16, MPV18a]. The geometric idea is that given a horizontal curve we perturb it adding an amount of length that depends on two factors:

(i) the desired change ($k \in G$) in the endpoint of the curve, and
(ii) the size of configuration of points ($x_0, \dots, x_r \in G$) that the curve passes through.

However, instead of writing the argument using the language of curves, we write it as a form of a triangle inequality. The horizontal curve should be thought of as replaced by the points $x_0, \dots, x_r$ along the curve. The benefits of this approach are twofold. First, we avoid having to worry about some technicalities, such as the parametrization of the curve or the concept of inserting one curve within another. Second, a triangle-inequality form is well suited to large-scale geometry, where the local behavior of horizontal curves is irrelevant. This allows us to immediately apply our argument in the asymptotic case not only to geodesics, but to rough geodesics as well.

Lemma 2.20. *For every Carnot group G of rank r and step $s \geq 2$, there exists a constant $C > 0$ with the following property:*

Let $x_0, \dots, x_r \in G$ and $k \in \exp(V_s)$. If $\mathrm{Size}(x_0, \dots, x_r) > 0$, then

$$d(x_0, kx_r) \leq C \left(\frac{d(1_G, k)^s}{\mathrm{Size}(x_0, \dots, x_r)} \right)^{\frac{1}{s-1}} + \sum_{j=1}^{r} d(x_{j-1}, x_j).$$

Proof. Let K be the constant from Lemma 2.11 for the group G. We claim that the constant $C := 2(r+1)K^{\frac{1}{s-1}}$ will satisfy the statement of the current lemma. Given $x_0, \dots, x_r \in G$ and $k \in \exp(V_s)$, we apply Lemma 2.11 with $Z := \log(k)$ and $X_j := \log(x_{j-1}^{-1} x_j)$, for $j = 1, \dots, r$. We get the existence of $Y_1, \dots, Y_r \in V_{s-1}$ satisfying (2.12) and the bound (2.13).

Define the following points in G:

$$\begin{aligned} &&& y_j := \exp(Y_j), \quad \text{for } j = 1, \dots, r; \\ &\alpha_0 := x_0, && \alpha_j := x_{j-1}^{-1} x_j, \quad \text{for } j = 1, \dots, r; \\ &\beta_0 := y_1, && \beta_j := y_{j-1}^{-1} y_j, \quad \text{for } j = 1, \dots, r-1, \qquad \beta_r := y_r^{-1}. \end{aligned}$$

Since $Y_j \in V_{s-1}$, by the BCH formula we have

$$\mathrm{C}_{y_j}(\alpha_j) = y_j \alpha_j y_j^{-1} = y_j \alpha_j y_j^{-1} \alpha_j^{-1} \alpha_j = \exp([Y_j, X_j]) \alpha_j,$$

where C_y denotes the conjugation by y. Consequently, since $\exp([Y_j, X_j]) \in \exp(V_s)$ commutes with everything, we have

$$\begin{aligned} \prod_{j=0}^{r} (\alpha_j \beta_j) &= \alpha_0 \beta_0 \alpha_1 \beta_1 \alpha_2 \beta_2 \cdots \alpha_r \beta_r \\ &= \alpha_0 y_1 \alpha_1 y_1^{-1} y_2 \alpha_2 y_2^{-1} \cdots y_r \alpha_r y_r^{-1} \\ &= \alpha_0 \mathrm{C}_{y_1}(\alpha_1) \mathrm{C}_{y_2}(\alpha_2) \cdots \mathrm{C}_{y_r}(\alpha_r) \\ &= \alpha_0 \exp([Y_1, X_1]) \alpha_1 \exp([Y_2, X_2]) \alpha_2 \cdots \exp([Y_r, X_r]) \alpha_r \\ &= \exp([Y_1, X_1]) \exp([Y_2, X_2]) \cdots \exp([Y_r, X_r]) \alpha_0 \alpha_1 \alpha_2 \cdots \alpha_r. \end{aligned}$$

Observe that a product of exponentials is the exponential of a sum for elements in V_s and that the points α_j form the telescopic product $x_r = \alpha_0 \alpha_1 \alpha_2 \cdots \alpha_r$. Thus the above identity simplifies to

$$\prod_{j=0}^{r} (\alpha_j \beta_j) = \exp([Y_1, X_1] + [Y_2, X_2] + \ldots + [Y_r, X_r]) x_r \overset{(2.12)}{=} \exp(Z) x_r = k x_r. \tag{2.21}$$

By the definition of the points α_j for $j = 1, \dots, r$, we have

$$d(1_G, \alpha_j) = d(x_{j-1}, x_j), \tag{2.22}$$

and for the points β_j for $j = 0, \dots, r$, we have from (2.13) the distance estimate

$$d(1_G, \beta_j) \leq 2K^{\frac{1}{s-1}} \left(\frac{d(1_G, k)^s}{\text{Size}(x_0, \dots, x_r)} \right)^{\frac{1}{s-1}}. \tag{2.23}$$

Combining (2.21), (2.22) and (2.23) we have that

$$\begin{aligned} d(x_0, kx_r) &\overset{(2.21)}{=} d(x_0, \Pi_{j=0}^r(\alpha_j\beta_j)) \\ &= d(1_G, \beta_0\Pi_{j=1}^r(\alpha_j\beta_j)) \\ &\leq d(1_G, \beta_0) + \sum_{j=1}^r d(1_G, \beta_j) + \sum_{j=1}^r d(1_G, \alpha_j) \\ &\overset{(2.22)\&(2.23)}{\leq} 2(r+1)K^{\frac{1}{s-1}} \left(\frac{d(1_G, k)^s}{\text{Size}(x_0, \dots, x_r)} \right)^{\frac{1}{s-1}} + \sum_{j=1}^r d(x_{j-1}, x_j). \end{aligned}$$

Hence the lemma holds with the proposed constant $C = 2(r+1)K^{\frac{1}{s-1}}$. □

The following proposition contains the particular form of triangle inequality that allows us to deduce our results for both tangents and asymptotic cones of geodesics. For any set of points $x_0, \dots, x_m \in G$ the standard triangle inequality states that

$$d(x_0, x_m) \leq \sum_{k=1}^m d(x_{k-1}, x_k).$$

The following proposition states that we can replace one of the terms of the sum with the distance $d(\pi_{s-1}(x_{\ell-1}), \pi_{s-1}(x_\ell))$ in the quotient group $G/\exp(V_s)$, if we pay a correction term coming from Lemma 2.20.

Theorem 1.2 for tangents will follow from the numerator of the correction term being related to the removed distance with a power $1+\epsilon$, which implies that in the tangential limit, the correction term is irrelevant. Theorem 1.4 on the other hand will follow from the correction term being inversely related to the size of the configuration of the other points. This will allow us to apply Lemma 2.29 to constrain the behavior of geodesics on the large scale.

Proposition 2.24. *For every Carnot group G of rank r and step $s \geq 2$, there exists a constant $K > 0$ such that for any $E = (y_0, \dots, y_{r+2}) \in G^{r+3}$, $\ell \in \{1, \dots, r+2\}$ and $E_\ell := (y_0, \dots, \hat{y}_{\ell-1}, \hat{y}_\ell, \dots y_{r+2}) \in G^{r+1}$ the following modified triangle inequality holds:*

$$d(y_0, y_{r+2}) \leq d(\pi_{s-1}(y_{\ell-1}), \pi_{s-1}(y_\ell)) + K \left(\frac{d(y_{\ell-1}, y_\ell)^s}{\text{Size}(E_\ell)} \right)^{\frac{1}{s-1}} + \sum_{j \neq \ell} d(y_{j-1}, y_j).$$

Proof. Since the claim of the proposition is degenerate when $\mathrm{Size}(E_\ell) = 0$, we can assume that $\mathrm{Size}(E_\ell) > 0$. Let C be the constant of Lemma 2.20 for the group G. We claim that the constant $K := 2^{\frac{s}{s-1}} C$ will satisfy the statement of the proposition.

By Proposition 2.1 there exists $h \in \exp(V_s)$ such that

$$d(y_{\ell-1}, hy_\ell) = d(\pi_{s-1}(y_{\ell-1}), \pi_{s-1}(y_\ell)). \tag{2.25}$$

We consider the points $x_j := y_j$ for $j < \ell - 1$ and $x_j := hy_{j+2}$ for $j \geq \ell - 1$. Since translation by h does not change the horizontal projection,

$$\mathrm{Size}(x_0, \dots, x_r) = \mathrm{Size}(E_\ell) > 0.$$

Applying Lemma 2.20 with $k := h^{-1}$ and the points $x_0, \dots, x_r$, we obtain the estimate

$$d(x_0, h^{-1}x_r) \leq C \left(\frac{d(1_G, h^{-1})^s}{\mathrm{Size}(x_0, \dots, x_r)} \right)^{\frac{1}{s-1}} + \sum_{j=1}^{r} d(x_{j-1}, x_j). \tag{2.26}$$

By the definition of the points x_j, for $j \neq \ell - 1$, we have

$$d(x_{j-1}, x_j) = \begin{cases} d(y_{j-1}, y_j), & \text{if } j < \ell - 1 \\ d(hy_{j+1}, hy_{j+2}), & \text{if } j > \ell - 1 \end{cases}$$

so

$$\sum_{j<\ell-1} d(x_{j-1}, x_j) = \sum_{j<\ell-1} d(y_{j-1}, y_j) \quad \text{and} \quad \sum_{j>\ell-1} d(x_{j-1}, x_j) = \sum_{j>\ell+1} d(y_{j-1}, y_j).$$

For $j = \ell - 1$ on the other hand, applying the identity (2.25) through a triangle inequality, we have

$$\begin{aligned} d(x_{\ell-2}, x_{\ell-1}) = d(y_{\ell-2}, hy_{\ell+1}) &\leq d(y_{\ell-2}, y_{\ell-1}) + d(y_{\ell-1}, hy_\ell) + d(hy_\ell, hy_{\ell+1}) \\ &= d(y_{\ell-2}, y_{\ell-1}) + d(\pi_{s-1}(y_{\ell-1}), \pi_{s-1}(y_\ell)) + d(y_\ell, y_{\ell+1}), \end{aligned}$$

filling in the missing terms $d(y_{j-1}, y_j)$ for $j = \ell - 1$ and $j = \ell + 1$. Combining the cases, we get the estimate

$$\sum_{j=1}^{r} d(x_{j-1}, x_j) \leq d(\pi_{s-1}(y_\ell), \pi_{s-1}(y_{\ell+1})) + \sum_{j\neq\ell} d(y_{j-1}, y_j). \tag{2.27}$$

We combine the identity (2.25) with the fact that the projection π_{s-1} is 1-Lipschitz, and we get that $d(y_{\ell-1}, hy_\ell) \leq d(y_{\ell-1}, y_\ell)$. Thus since h is in the center of G, the distance $d(1_G, h^{-1})$ can be estimated by

$$d(1_G, h^{-1}) = d(hy_{\ell-1}, y_{\ell-1}) \leq d(hy_{\ell-1}, hy_\ell) + d(hy_\ell, y_{\ell-1}) \leq 2d(y_{\ell-1}, y_\ell). \tag{2.28}$$

Combining (2.26) with (2.27) and (2.28) results in the desired inequality

$$d(y_0, y_{r+2}) \leq 2^{\frac{s}{s-1}} C \left(\frac{d(y_{\ell-1}, y_\ell)^s}{\mathrm{Size}(x_0, \dots, x_r)} \right)^{\frac{1}{s-1}} + d(\pi_{s-1}(y_{\ell-1}), \pi_{s-1}(y_\ell)) + \sum_{j\neq\ell} d(y_{j-1}, y_j).$$

□

2.4. **Geometric lemmas about minimal height and size.** None of the estimates of the rest of this section will be used for Theorem 1.2, so the reader interested in just the results about tangents can skip the following two lemmas. For the proof of Theorem 1.4 (and its generalization Theorem 4.2) we need to describe how the boundedness of the previously defined notions of Size and MinHeight relate to uniform neighborhoods of hyperplanes in the abelianization $G/[G,G]$. Lemma 2.29 describes how MinHeight and hyperplane neighborhoods are related and Lemma 2.30 gives a lower bound for Size in terms of MinHeight of a translation of the vertices.

We will only need the implications and estimates in one direction, however all of these lemmas can be generalized to include also the opposite inequalities (with possibly worse constants) and the reverse implications.

Lemma 2.29. *Let Γ be a subset of $\mathbb{R}^r$. If there exists $K > 0$ such that* $\mathrm{MinHeight}(P) \leq K$ *for all $P \in \Gamma^m$, then there exists an $(m-1)$-plane $W \subset \mathbb{R}^r$ such that $\Gamma \subset \bar{B}_{\mathbb{R}^r}(W, K)$.*

Proof. Consider first the case when Γ is a finite set. We take $P^* \in \Gamma^{m-1}$ so that the parallelotope $\mathcal{P}(P^*)$ generated by P^* maximizes $\mathrm{vol}_{m-1}\,\mathcal{P}(P')$ among all $P' \in \Gamma^{m-1}$. We claim that $\Gamma \subset \bar{B}_{\mathbb{R}^r}(\mathrm{span}(P^*), K)$. Indeed, for every $a \in \Gamma$, since P^* has maximal volume, we have by Remark 2.3.2 that

$$d(a, \mathrm{span}(P^*)) = \frac{\mathrm{vol}_m\,\mathcal{P}(P^*, a)}{\mathrm{vol}_{m-1}\,\mathcal{P}(P^*)} \overset{2.3.2}{=} \mathrm{MinHeight}(P^*, a) \leq K.$$

Consider then the case of an infinite set Γ, and let $(p_n)_{n\in\mathbb{N}}$ be a countable dense set in Γ. Applying the lemma for the finite sets $\{p_1, \ldots, p_n\}$, we have the existence of $(m-1)$-planes W_n such that $\{p_1, \ldots, p_n\} \subset \bar{B}_{\mathbb{R}^r}(W_n, K)$. By compactness there exist an $(m-1)$-plane W and a diverging sequence n_j such that $W_{n_j} \to W$, as $j \to \infty$.

We want to prove that $\Gamma \subset \bar{B}_{\mathbb{R}^r}(W, K)$. It is enough to show that $\{p_1, \ldots, p_n\} \subset \bar{B}_{\mathbb{R}^r}(W, K+\epsilon)$, for all $n \in \mathbb{N}$ and $\epsilon > 0$. Fix such n and ϵ and fix R_n so that $\{p_1, \ldots, p_n\} \subset \bar{B}_{\mathbb{R}^r}(0, R_n)$. Then we take j large enough that $n_j > n$ and

$$\bar{B}_{\mathbb{R}^r}(W_{n_j}, K) \cap \bar{B}_{\mathbb{R}^r}(0, R_n) \subset \bar{B}_{\mathbb{R}^r}(W, K+\epsilon),$$

which is possible since $W_{n_j} \to W$, and so $\bar{B}_{\mathbb{R}^r}(W_{n_j}, K) \to \bar{B}_{\mathbb{R}^r}(W, K)$ on compact sets in the Hausdorff sense. Thus we conclude the proof of the claim:

$$\{p_1, \ldots, p_n\} \subset \{p_1, \ldots, p_{n_j}\} \cap \bar{B}_{\mathbb{R}^r}(0, R_n) \subset \bar{B}_{\mathbb{R}^r}(W_{n_j}, K) \cap \bar{B}_{\mathbb{R}^r}(0, R_n) \subset \bar{B}_{\mathbb{R}^r}(W, K+\epsilon).$$

□

For convenience of applying Lemma 2.29 within the proof of Theorem 1.4, we give a lower bound for Size in terms of MinHeight. We will not need this bound for the proof of Theorem 1.2.

Lemma 2.30. *In any Carnot group G, there exists a constant $c > 0$ such that the following holds:*

For any $E = (g_0, \dots, g_r) \in G^{r+1}$ *and* $\ell \in \{0, \dots, r\}$, *let* $\Gamma_\ell \in (G/[G,G])^r$ *be the tuple of the points* $\pi(g_j) - \pi(g_\ell)$, $j \neq \ell$. *Then*

$$\mathrm{Size}(E) \geq c \cdot \mathrm{MinHeight}(\Gamma_\ell).$$

Proof. In $\mathbb{R}^n$, consider for each $\ell \in \{0, \dots, r\}$ the map $A^\ell : (\mathbb{R}^n)^r \to (\mathbb{R}^n)^r$, whose component functions $A^\ell_k : (\mathbb{R}^n)^r \to \mathbb{R}^n$ are defined by

$$A^\ell_k(x_1, \dots, x_r) = \sum_{j=k}^{\ell} x_j \quad \text{for } k = 1, \dots, \ell$$

and by

$$A^\ell_k(x_1, \dots, x_r) = \sum_{j=\ell+1}^{k} x_j \quad \text{for } k = \ell+1, \dots, r.$$

In block-matrix form, the linear map A^ℓ has the form $A^\ell = \begin{bmatrix} U & 0 \\ 0 & L \end{bmatrix}$, where

$$U = \begin{bmatrix} I & I & \dots & I \\ 0 & I & \dots & I \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I \end{bmatrix} \quad \text{and} \quad L = \begin{bmatrix} I & 0 & \dots & 0 \\ I & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ I & I & \dots & I \end{bmatrix}$$

are themselves $\ell \times \ell$ and $(r-\ell) \times (r-\ell)$ upper and lower triangular block-matrices, whose $n \times n$-blocks are all either the $n \times n$ identity matrix I or zero.

From the above description, it is clear that A^ℓ is a linear bijection, so there exists a constant $C_\ell > 0$ such that A^ℓ is a C_ℓ-biLipschitz map. Thus for any set $\mathcal{P} \subset \mathbb{R}^r$, we have

$$C_\ell^{-m} \operatorname{vol}_m(\mathcal{P}) \leq \operatorname{vol}_m(A^\ell(\mathcal{P})) \leq C_\ell^m \operatorname{vol}_m(\mathcal{P}).$$

By the characterization of MinHeight as volume quotients in Remark 2.3.2, it follows that

$$\mathrm{MinHeight}(A^\ell(x_1, \dots, x_r)) \leq C_\ell^{2r-1} \cdot \mathrm{MinHeight}(x_1, \dots, x_r) \tag{2.31}$$

The abelianization $G/[G,G]$ is a normed space, so there exists for some $C > 0$ and $n \in \mathbb{N}$ a C-biLipschitz isomorphism $\phi : G/[G,G] \to \mathbb{R}^n$. We claim that the constant

$$c := \min_{\ell \in \{0, \dots, r\}} C^{-2} C_\ell^{1-2r} \tag{2.32}$$

satisfies the claim of the lemma.

Let $y_j := \pi(g_j) - \pi(g_{j-1})$, $j = 1, \dots, r$ so that the definition (2.9) of Size is written as

$$\mathrm{Size}(E) = \mathrm{Size}(g_0, \dots, g_r) = \mathrm{MinHeight}(y_1, \dots, y_r). \tag{2.33}$$

Apply the map $(\phi^{-1})^r \circ A^\ell \circ (\phi)^r : (G/[G,G])^r \to (G/[G,G])^r$ to the tuple $(y_1, \dots, y_r) \in (G/[G,G])^r$. For $k \le \ell$, we have

$$\begin{aligned}(\phi^{-1})^r \circ A_k^\ell(\phi(y_1), \dots, \phi(y_r)) &= (\phi^{-1})^r \left(\sum_{j=k}^{\ell} (\phi \circ \pi(g_j) - \phi \circ \pi(g_{j-1})) \right) \\ &= (\phi^{-1})^r \left(\phi \circ \pi(g_\ell) - \phi \circ \pi(g_{k-1})\right) \\ &= \pi(g_\ell) - \pi(g_{k-1}).\end{aligned}$$

Similarly for $k \ge \ell + 1$, we have

$$(\phi^{-1})^r \circ A_k^\ell(\phi(y_1), \dots, \phi(y_r)) = \pi(g_k) - \pi(g_\ell).$$

That is, up to the sign of the elements $k \le \ell$ components, the components of $(\phi^{-1})^r \circ A^\ell(\phi(y_1), \dots, \phi(y_r))$ form exactly the tuple Γ_ℓ.

For any C-Lipschitz map f, we have

$$\text{MinHeight}(f(y_1), \dots, f(y_r)) \le C \cdot \text{MinHeight}(y_1, \dots, y_r).$$

Since both ϕ and ϕ^{-1} are C-Lipschitz, by (2.31) we get

$$\begin{aligned}\text{MinHeight}(\Gamma_\ell) &= \text{MinHeight}((\phi^{-1})^r \circ A^\ell(\phi(y_1), \dots, \phi(y_r))) \\ &\le C \cdot \text{MinHeight}(A^\ell(\phi(y_1), \dots, \phi(y_r))) \\ &\overset{(2.31)}{\le} C C_\ell^{2r-1} \cdot \text{MinHeight}(\phi(y_1), \dots, \phi(y_r)) \\ &\le C^2 C_\ell^{2r-1} \cdot \text{MinHeight}(y_1, \dots, y_r).\end{aligned}$$

By (2.33) and (2.32) we end up with the desired estimate

$$\text{Size}(E) \overset{(2.33)}{=} \text{MinHeight}(y_1, \dots, y_r) \ge \frac{1}{C^2 C_\ell^{2r-1}} \text{MinHeight}(\Gamma_\ell) \overset{(2.32)}{\ge} c \cdot \text{MinHeight}(\Gamma_\ell).$$

□

3. Blowups of geodesics

We next prove the results on blowups of geodesics (Theorems 1.1 and 1.2). In fact, instead of the qualitative claim of Theorem 1.2, we will prove a slightly stronger quantified statement. We show that $\pi_{s-1} \circ \gamma$ satisfies a sublinear distance estimate on some small enough interval, implying that any tangent of $\pi_{s-1} \circ \gamma$ is a geodesic. The estimate shall follow by applying the triangle inequality of Proposition 2.24 with tuples $E = (y_0, \dots, y_{r+2})$ where only two of the points $y_{\ell-1}$ and y_ℓ will vary.

Theorem 3.1. *Let G be a Carnot group of step s and let $\gamma : I \to G$ be a geodesic. Then for any $\bar{t} \in I$, there exist constants $C > 0$ and $\delta > 0$ such that for all $a, b \in (\bar{t} - \delta, \bar{t} + \delta)$,*

$$|a - b| - C\,|a - b|^{\frac{s}{s-1}} \le d(\pi_{s-1} \circ \gamma(a), \pi_{s-1} \circ \gamma(b)) \le |a - b|\,.$$

Proof. The upper bound follows directly from the projection $\pi_{s-1} : G \to G/\exp(V_s)$ being 1-Lipschitz. The non-trivial statement is the lower bound, which will follow from Proposition 2.24.

Translating the parametrization if necessary, we may assume that $\bar{t} = 0$. Since any geodesic is still a geodesic within every Carnot subgroup containing it, we may also assume that G is the smallest Carnot subgroup containing $\gamma(I)$. Hence, if r is the rank of G, there exist $t_0, \ldots, t_r \neq 0$ such that the points $\pi \circ \gamma(t_0), \ldots, \pi \circ \gamma(t_r)$ are in general position. By Remark 2.10, we have that

$$\Delta := \operatorname{Size}(\gamma(t_0), \ldots, \gamma(t_r)) > 0. \tag{3.2}$$

Let K be the constant given by Proposition 2.24 for the Carnot group G. We claim that the constants $C := K\Delta^{-\frac{1}{s-1}}$ and $\delta := \min(|t_0|, \ldots, |t_r|)$ will satisfy the claim of the theorem.

Fix $a, b \in (-\delta, \delta)$. Consider the set of points

$$E := \{y_0, \ldots, y_{r+2}\} = \{\gamma(t_j) : \ j = 0, \ldots, r\} \cup \{\gamma(a), \gamma(b)\},$$

where the points y_j are indexed by the order in which they appear along γ. By the choice of δ, the points $\gamma(a)$ and $\gamma(b)$ are consecutive in this ordering, so there is some $\ell \in \{1, \ldots, r+2\}$ such that $y_{\ell-1} = \gamma(a)$ and $y_\ell = \gamma(b)$.

We apply Proposition 2.24 with the above E and ℓ. By (3.2), we get the estimate

$$d(y_0, y_{r+2}) \leq d(\pi_{s-1}\circ\gamma(a), \pi_{s-1}\circ\gamma(b)) + K\left(\frac{d(\gamma(a), \gamma(b))^s}{\Delta}\right)^{\frac{1}{s-1}} + \sum_{j\neq\ell} d(y_{j-1}, y_j). \tag{3.3}$$

By the choice of the points y_j as sequential points along the geodesic γ, we have

$$\sum_{j\neq\ell} d(y_{j-1}, y_j) = d(y_0, y_{r+2}) - d(y_{\ell-1}, y_\ell) = d(y_0, y_{r+2}) - d(\gamma(a), \gamma(b)). \tag{3.4}$$

We then apply the identity (3.4) to (3.3), we use the fact that $\gamma|_{[a,b]}$ is a geodesic, and we reorganize the terms. This gives the lower bound

$$d(\pi_{s-1} \circ \gamma(a), \pi_{s-1} \circ \gamma(b)) \geq |a-b| - K\Delta^{-\frac{1}{s-1}} |a-b|^{\frac{s}{s-1}},$$

proving the claim of the theorem. □

Theorem 1.2 shall follow immediately from Theorem 3.1 by taking any limit of dilations $h_k \to 0$.

Proof of Theorem 1.2. Reparametrizing and left-translating if necessary, we may assume that $t = 0$ and $\gamma(0) = 1_G$. Then $\sigma \in \operatorname{Tang}(\gamma, 0)$ is given by some sequence $h_k \to 0$ as $\sigma = \lim_{k\to\infty} \gamma_{h_k}$.

For any $h > 0$ and $a, b \in I_h$, expanding the definition of the dilated curve $\gamma_h = \delta_{1/h} \circ \gamma \circ \delta_h$, we get

$$d(\gamma_h(a), \gamma_h(b)) = \frac{1}{h} d(\gamma(ha), \gamma(hb)). \tag{3.5}$$

Let $C > 0$ and $T > 0$ be the constants of Theorem 3.1. Rephrasing the statement of Theorem 3.1 for γ_h using (3.5), we get for all $a, b \in (-T/h, T/h)$ that

$$|a-b| - Ch^{\frac{1}{s-1}} |a-b|^{\frac{s}{s-1}} \leq d(\pi_{s-1} \circ \gamma_h(a), \pi_{s-1} \circ \gamma_h(b)) \leq |a-b|. \tag{3.6}$$

For any $a, b \in \mathbb{R}$, the condition $a, b \in (-T/h_k, T/h_k)$ is satisfied for any large enough indices $k \in \mathbb{N}$. Thus taking the limit of (3.6) as $h = h_k \to 0$, we get for the limit curve $\pi_{s-1} \circ \sigma = \lim_{k \to \infty} \pi_{s-1} \circ \gamma_{h_k}$ the estimate

$$|a-b| \leq d(\pi_{s-1} \circ \sigma(a), \pi_{s-1} \circ \sigma(b)) \leq |a-b|,$$

showing that $\pi_{s-1} \circ \sigma$ is a geodesic. □

Corollary 1.3 follows directly from Theorem 1.2 by induction on the step of the Carnot group. We prove next that Theorem 1.1 follows from the fact that the metric tangent of a sub-Riemannian manifold is a quotient of a sub-Riemannian Carnot group, which is a well known theorem attributed to Bellaïche [Bel96].

Proof of Theorem 1.1. Let M be a sub-Riemannian manifold. Let s be the step of the sub-Riemannian manifold, i.e., Lie brackets of length s of the horizontal vector fields in M span the tangent spaces T_pM at each point $p \in M$. Let $\gamma : I \to M$ be a geodesic.

Following [MPV18b, Theorem 3.6], we see that any metric tangent σ of γ is a geodesic in the nilpotent approximation $\tilde{M}$ of M. By [Jea14, Theorem 2.7], the nilpotent approximation of a sub-Riemannian manifold is a homogeneous space G/H, where G is a Carnot group of step s and $H < G$ is a closed dilation invariant Lie subgroup. In particular, any iterated tangent of σ gives another geodesic in $\tilde{M} = G/H$.

On the other hand, since the projection $\pi : G \to G/H$ is a submetry, the geodesic σ can be lifted to a geodesic $\tilde{\sigma}$ in G. Applying Corollary 1.3 we see that any $s-1$ times iterated tangent of $\tilde{\sigma}$ is a line. Projecting back to G/H, we see that also necessarily any $s-1$ times iterated of σ must be a line. Since σ was an arbitrary tangent of γ, it follows that any s times iterated tangent of γ is a line. □

4. Blowdowns of rough geodesics

In this section we prove Theorem 1.4 and Corollary 1.5. Due to our formulation of the core of the argument (Proposition 2.24) as a triangle inequality, we are able to prove the stronger claims of Theorem 4.2 and Corollary 4.10 for rough geodesics.

To make the terminology precise, by *rough geodesic*, we mean a not-necessarily-continuous curve that is a $(1, C)$-quasi-geodesic for some $C \geq 0$. By a $(1, C)$-*quasi-geodesic* we mean a $(1, C)$-*quasi-isometric embedding*, i.e., some $\gamma : I \to G$ such that

$$|t_1 - t_2| - C \leq d(\gamma(t_1), \gamma(t_2)) \leq |t_1 - t_2| + C, \qquad \forall t_1, t_2 \in I. \tag{4.1}$$

Thus a $(1, 0)$-quasi-geodesic is exactly a geodesic.

Theorem 4.2. *If $\gamma : \mathbb{R} \to G$ is a $(1, C)$-quasi-geodesic, then one of the following holds:*

(4.2.i) There exist a hyperplane $W \subset V_1$ and some $R > 0$ such that $\mathrm{Im}(\pi \circ \gamma) \subset B_{V_1}(W, R)$.
(4.2.ii) There exists $C' \geq 0$ such that $\pi \circ \gamma : \mathbb{R} \to G/[G, G]$ is a $(1, C')$-quasi-geodesic.

Moreover, one can take $C' = (r + 2)^{s-1} C$.

Proof. Assume (4.2.i) does not hold. We claim that it is enough to show that $\pi_{s-1} \circ \gamma$ is a $(1, C_1)$-quasi-geodesic with $C_1 := (r + 2)C$. Indeed, then we can iterate: the curve $\pi_{s-1} \circ \gamma$ has the same projection as γ on $G/[G, G]$. Thus, (4.2.i) does not hold for $\pi_{s-1} \circ \gamma$ either, and we have that $\pi_{s-2} \circ \pi_{s-1} \circ \gamma$ is a $(1, C_2)$-quasi-geodesic in $G/\exp(V_{s-1} \oplus V_s)$ with $C_2 = (r + 2)C_1 = (r + 2)^2 C$. We repeat until after $(s - 1)$ steps we get that $\pi \circ \gamma = \pi_1 \circ \cdots \circ \pi_{s-1} \circ \gamma$ is a $(1, (r + 2)^{s-1} C)$-quasigeodesic.

As with Theorem 3.1, the upper bound follows immediately from the projection π_{s-1} being 1-Lipschitz. Thus it is enough to show the lower bound $|b - a| - C_1 \leq d(\pi_{s-1} \circ \gamma(a), \pi_{s-1} \circ \gamma(b))$, for all $a, b \in \mathbb{R}$.

Set $\Gamma := \gamma(\mathbb{R} \setminus [a, b])$ and fix an arbitrary basepoint $\bar{t} \in \mathbb{R} \setminus [a, b]$. Since (4.2.i) does not hold for γ, the same is true for any translation of γ. Therefore we can assume without loss of generality that $\gamma(\bar{t}) = 1_G$.

Fix an arbitrary $\epsilon > 0$. Let $K > 0$ be the constant of Proposition 2.24 and let $c > 0$ be the constant of Lemma 2.30. Since $\gamma([a, b])$ is a bounded set, the failure of (4.2.i) for γ implies that Γ is also not contained in any neighborhood of any hyperplane. Since $G/[G, G]$ and $\mathbb{R}^r$ are biLipschitz equivalent, Lemma 2.29 implies that $\mathrm{MinHeight}(\pi(P))$ is not bounded as P varies in Γ^r. In particular, we may fix some $P \in \Gamma^r$ such that

$$\mathrm{MinHeight}(\pi(P)) > \frac{K^{s-1} d(\gamma(a), \gamma(b))^s}{c\epsilon^{s-1}}. \tag{4.3}$$

Consider the tuple $E := (\gamma(t_0), \ldots, \gamma(t_{r+2}))$, where

$$\{\gamma(t_0), \ldots, \gamma(t_{r+2})\} = P \cup \{\gamma(\bar{t}), \gamma(a), \gamma(b)\},$$

with the times t_j ordered so that $t_0 < \cdots < t_{r+2}$.

By the definition of Γ and $\bar{t}$, the points $\gamma(a)$ and $\gamma(b)$ are necessarily consecutive in this ordering, so there is some $\ell \in \{1, \ldots, r+2\}$ such that $t_{\ell-1} = a$ and $t_\ell = b$. Denote by $E_P \in \Gamma^{r+1}$ the tuple E without $\gamma(a)$ and $\gamma(b)$, i.e.,

$$E_P := (\gamma(t_0), \ldots, \gamma(t_{\ell-2}), \gamma(t_{\ell+1}), \ldots, \gamma(t_{r+2})).$$

Applying Proposition 2.24 with the above E and ℓ, we get the bound

$$d(\gamma(t_0), \gamma(t_{r+2})) \leq d(\pi_{s-1} \circ \gamma(a), \pi_{s-1} \circ \gamma(b)) + \sum_{j \neq \ell} d(\gamma(t_{j-1}), \gamma(t_j)) + K\left(\frac{d(\gamma(a), \gamma(b))^s}{\mathrm{Size}(E_P)}\right)^{\frac{1}{s-1}}. \tag{4.4}$$

Estimating the distances along γ by (4.1) gives

$$\sum_{j \neq \ell} d(y_{j-1}, y_j) \leq \sum_{j \neq \ell} |t_{j-1} - t_j| + (r+1)C = |t_0 - t_{r+2}| - |a - b| + (r+1)C$$

and

$$d(\gamma(t_0), \gamma(t_{r+2})) \geq |t_0 - t_{r+2}| - C.$$

Applying the above distance estimates to (4.4) and reorganizing terms, we get the lower bound

$$d(\pi_{s-1} \circ \gamma(a), \pi_{s-1} \circ \gamma(b)) \geq |a - b| - (r+2)C - K\left(\frac{d(\gamma(a), \gamma(b))^s}{\mathrm{Size}(E_P)}\right)^{\frac{1}{s-1}}, \tag{4.5}$$

which is exactly the desired lower bound except for the final term.

However, since $\gamma(\bar{t}) = 1_G$, applying Lemma 2.30 with ℓ such that $t_\ell = \bar{t}$ gives

$$\mathrm{Size}(E_P) \geq c \cdot \mathrm{MinHeight}(\pi(P)). \tag{4.6}$$

Bounding $\mathrm{Size}(E_P)$ by (4.6) and $\mathrm{MinHeight}(\pi(P))$ by (4.3), the lower bound (4.5) is simplified to

$$d(\pi_{s-1} \circ \gamma(a), \pi_{s-1} \circ \gamma(b)) \geq |a - b| - (r+2)C - \epsilon.$$

Since $\epsilon > 0$ was arbitrary, we have the desired quasi-geodesic lower bound. ☐

The second possible conclusion (4.2.ii) in Theorem 4.2 is that $\pi \circ \gamma$ is a quasi-geodesic in the normed space $G/[G,G]$. We next show that in the case of an inner product space, quasi-geodesics are well behaved on the large scale. Namely, every rough geodesic in $\mathbb{R}^n$ has a unique asymptotic cone and this asymptotic cone is a line.

Proposition 4.7. *Every $(1, C)$-quasi-geodesic in Euclidean n-space has a unique blowdown and the blowdown is a line.*

Proof. Let $\gamma : \mathbb{R} \to \mathbb{R}^n$ be a $(1, C)$-quasi-geodesic. Translating and reparametrizing if necessary, we may assume that $\gamma(0) = 0$. We denote by $\angle(t, s)$ the angle formed by $\gamma(t)$ and $\gamma(s)$ at 0. Its magnitude is given by the standard inner product on $\mathbb{R}^n$ via

$$\cos\angle(t,s) = \frac{\gamma(t)\cdot\gamma(s)}{|\gamma(t)|\,|\gamma(s)|}. \tag{4.8}$$

We first show that as $t, s \to \infty$, the angle vanishes, i.e., $1 - \cos\angle(t, s) \to 0$. By symmetry we can assume that $t \geq s \geq 0$.

In an inner product space we have for all x, y the identity

$$2\,|x|\,|y| - 2x\cdot y = |x-y|^2 - (|x|-|y|)^2.$$

Combining (4.8) and the above identity for $x = \gamma(t)$ and $y = \gamma(s)$, we get

$$\begin{aligned} 1-\cos\angle(t,s) &= \frac{2\,|\gamma(t)|\,|\gamma(s)| - 2\gamma(t)\cdot\gamma(s)}{2\,|\gamma(t)|\,|\gamma(s)|} \\ &= \frac{|\gamma(t)-\gamma(s)|^2 - (|\gamma(t)|-|\gamma(s)|)^2}{2\,|\gamma(t)|\,|\gamma(s)|}. \end{aligned} \tag{4.9}$$

The quasi-geodesic bound (4.1) and the assumption $t \geq s \geq 0$ imply that

$$\begin{aligned} |\gamma(t)-\gamma(s)|^2 - (|\gamma(t)|-|\gamma(s)|)^2 &\leq (t-s+C)^2 - (t-s-2C)^2 \\ &= 6C(t-s) - 3C^2 \leq 6Ct. \end{aligned}$$

Moreover, the bound (4.1) implies also that when $t, s \geq 2C$ we have

$$|\gamma(t)|\,|\gamma(s)| \geq (t-C)(s-C) \geq \frac{1}{4}ts.$$

Estimating (4.9) using the above two inequalities, we get for all $t, s \geq 2C$ the upper bound

$$1-\cos\angle(t,s) \leq \frac{6t}{\frac{1}{4}ts} = \frac{24}{s}$$

and hence $\angle(t, s) \to 0$ as $t \geq s \to \infty$. Repeating a similar argument for $t \leq s \leq 0$, we see also that $\angle(t, s) \to 0$ as $t \leq s \to -\infty$.

From this estimate of angles we conclude that the limit directions $v_+ = \lim\limits_{t\to\infty} \gamma(t)/|\gamma(t)|$ and $v_- = \lim\limits_{t\to-\infty} \gamma(t)/|\gamma(t)|$ always exist. We claim that this implies that the asymptotic cone $\lim\limits_{h\to\infty} \gamma_h$ exists without taking any subsequences, thus proving uniqueness.

First, observe that the existence of the limit direction and γ being a quasi-geodesic implies that also $\lim\limits_{t\to\infty} \gamma(t)/t = v_+$. Indeed, for any $t > C$, by (4.1), we have

$$\left|\frac{\gamma(t)}{t} - \frac{\gamma(t)}{|\gamma(t)|}\right| = \frac{|\gamma(t)|\,||\gamma(t)|-t|}{t\,|\gamma(t)|} \leq \frac{(t+C)C}{t(t-C)} \to 0$$

as $t \to \infty$. This implies that $\lim_{h\to\infty} \gamma_h(1) = v_+$. For arbitrary $t > 0$,

$$\lim_{h\to\infty} \gamma_h(t) = \lim_{h\to\infty} \frac{\gamma(ht)}{h} = t \lim_{h\to\infty} \frac{\gamma(ht)}{ht} = tv_+.$$

Similarly $\lim_{h\to\infty} \gamma_h(t) = -tv_-$ for all $t < 0$, proving existence and uniqueness of the blowdown.

To see that the unique limit is a line, i.e., that $v_- = -v_+$, it suffices to observe that any blowdown of a $(1, C)$-quasi-geodesic in $\mathbb{R}^n$ is a geodesic in $\mathbb{R}^n$, and geodesics in $\mathbb{R}^n$ are lines. □

Combining Theorem 4.2 with Proposition 4.7 allows us to conclude the lower rank subgroup containment for blowdowns of rough geodesics in sub-Riemannian Carnot groups:

Corollary 4.10. *If γ is a $(1, C)$-quasi-geodesic in a sub-Riemannian Carnot group $G \neq \mathbb{R}$, then there exists a proper Carnot subgroup $H < G$ containing every element of* $\mathrm{Asymp}(\gamma)$.

Proof. Consider the two cases of Theorem 4.2. In the first case (4.2.i), the horizontal projection is in a finite neighborhood of a hyperplane, $\mathrm{Im}(\pi \circ \gamma) \subset B_{V_1}(W, R)$. Thus any blowdown $\sigma \in \mathrm{Asymp}(\gamma)$ has its horizontal projection completely contained in W. Since $\sigma(0) = 1_G$, it follows that σ is contained in the Carnot subgroup H generated by W. The rank of H is by construction the dimension of W, which is smaller than the rank of G.

In the second case (4.2.ii), the horizontal projection $\pi \circ \gamma$ is a $(1, C')$-quasi-geodesic. Thus by Proposition 4.7 it has a unique blowdown σ, which is a line. But then $H := \sigma(\mathbb{R})$ is itself a one-parameter subgroup containing all blowdowns, proving the claim. □

5. Dilations of geodesics from the Hamiltonian viewpoint

Let G be a sub-Riemannian Carnot group, so that on the first layer V_1 of the Lie algebra $\mathfrak{g}$ we have an inner product $\langle \cdot, \cdot \rangle$. Every geodesic $\gamma : I \to G$ on a finite interval $I \subset \mathbb{R}$ is then a solution of the Pontryagin maximum principle. In sub-Riemannian Carnot groups the principle takes the form

$$\text{(PMP)} \qquad \lambda \left(\int_I \mathrm{Ad}_{\gamma(t)} v(t)\, dt \right) = \xi \langle u_\gamma, v \rangle \quad \forall v \in L^2(I; V_1),$$

for some $\lambda \in \mathfrak{g}^*$ and $\xi \in \mathbb{R}$ such that $(\lambda, \xi) \neq (0, 0)$, see [LMO+16] for the calculation of the differential of the endpoint map. Here, $u_\gamma \in L^2(I; V_1)$ denotes the control of γ.

A curve is abnormal exactly when it satisfies PMP with $\xi = 0$ for some $\lambda \in \mathfrak{g}^* \setminus \{0\}$. In the case of a geodesic $\gamma : J \to \mathbb{R}$ on an unbounded interval $J \subset \mathbb{R}$, there exists a pair $(\lambda, \xi) \neq (0, 0)$ for which PMP is satisfied for every bounded subinterval $I \subset J$.

In this section we will consider properties of asymptotic cones of geodesics from the point of view of the Pontryagin maximum principle. The next lemma describes what happens to PMP for dilations of geodesics.

Lemma 5.1. *Let $\gamma : I \to G$ be a horizontal curve in G that satisfies PMP for a pair (λ, ξ). Then for any $h > 0$, the dilated curve $\gamma_h : I_h \to G$ satisfies PMP for the pair $(\delta_h^* \lambda, h\xi)$.*

Proof. We may suppose without loss of generality that the interval I is bounded and the dilation

$$\gamma_h(t) = \delta_{1/h} \circ \gamma(\bar{t} + ht)$$

is happening at $\bar{t} = 0$. The dilations are homomorphisms, so by the definition of Ad_g as the differential of the conjugation $x \mapsto gxg^{-1}$, the map $\mathrm{Ad}_{\gamma(t)}$ can be written in terms of $\mathrm{Ad}_{\gamma_h(t)}$ as

$$\mathrm{Ad}_{\gamma(t)} = \mathrm{Ad}_{\delta_h \circ \gamma_h(t/h)} = (\delta_h)_* \circ \mathrm{Ad}_{\gamma_h(t/h)} \circ (\delta_{1/h})_*.$$

Therefore, PMP for γ gives the identity

$$\xi \langle u_\gamma, v \rangle = \lambda \left(\int_I \mathrm{Ad}_{\gamma(t)} v(t)\, dt \right) = (\delta_h^* \lambda) \left(\int_I \mathrm{Ad}_{\gamma_h(t/h)} \frac{1}{h} v(t)\, dt \right). \tag{5.2}$$

Denote for each $v \in L^2(I; V_1)$ by $\tilde{v} \in L^2(I_h; V_1)$ the reparametrized function $\tilde{v}(t) = v(ht)$. Then after a change of variables, the right hand side of (5.2) is

$$\int_I \mathrm{Ad}_{\gamma_h(t/h)} \frac{1}{h} v(t)\, dt = \int_{I_h} \mathrm{Ad}_{\gamma_h(t)} \tilde{v}(t)\, dt. \tag{5.3}$$

Since the control u_h of the dilated curve γ_h is

$$u_h(t) = (\delta_{1/h})_* u_\gamma(ht) \cdot h = u_\gamma(ht),$$

a similar change of variables as in (5.3) shows that

$$\langle u_\gamma, v \rangle = \int_I u_\gamma(t) v(t)\, dt = \int_I u_h(t/h) \tilde{v}(t/h)\, dt = h \int_{I_h} u_h(t) \tilde{v}(t)\, dt = h \langle u_h, \tilde{v} \rangle. \tag{5.4}$$

Applying both changes of variables (5.3) and (5.4) to (5.2) gives the identity

$$h\xi \langle u_h, \tilde{v} \rangle = (\delta_h^* \lambda) \left(\int_{I_h} \mathrm{Ad}_{\gamma_h(t)} \tilde{v}(t)\, dt \right).$$

Since every element of $L^2(I_h; V_1)$ can be written as $\tilde{v}$ for some $v \in L^2(I; V_1)$, the above shows that γ_h satifies PMP for the pair $(\delta_h^* \lambda, h\xi)$. ☐

5.1. **Abnormality of blowdowns of geodesics.** In every sub-Finsler Carnot group horizontal lines through the identity are infinite geodesics that are dilation invariant. Hence, the unique blowdown of any horizontal line is the line itself translated to the identity, which may or may not be abnormal. For all other curves however, every blowdown is necessarily an abnormal curve:

Proposition 5.5. *In sub-Riemannian Carnot groups asymptotic cones of non-line infinite geodesics are abnormal curves.*

Proof. The argument is partially inspired by [Agr98]. Let γ be a geodesic in G and let $(\lambda, \xi) \in \mathfrak{g}^* \times \mathbb{R}$ be a pair for which γ satisfies PMP. We decompose λ as $\lambda = \lambda^{(1)} + \cdots + \lambda^{(s)} \in V_1^* \oplus \cdots \oplus V_s^* \simeq \mathfrak{g}^*$ and let $j \in \{1, \ldots, s\}$ be the largest index for which $\lambda^{(j)} \neq 0$.

If $\lambda^{(2)} = \cdots = \lambda^{(s)} = 0$, then PMP reduces to

$$\lambda\Big(\int_I v\Big) = \xi \, \langle u_\gamma, v\rangle \quad \forall v \in L^2(I; V_1)$$

on every finite interval $I \subset \mathbb{R}$. Thus if $\lambda^{(2)} = \cdots = \lambda^{(s)} = 0$, then u_γ is constant and γ is a line. Assume from now on that γ is not a line, so $j \geq 2$.

By Lemma 5.1 the dilated curve γ_h satisfies PMP for the pair $(\delta_h^*\lambda, h\xi)$. In terms of the decomposition into layers, we have

$$\delta_h^*\lambda = \delta_h^*(\lambda^{(1)} + \cdots + \lambda^{(j)}) = h\lambda^{(1)} + \cdots + h^j\lambda^{(j)}.$$

Note that PMP is scale invariant with respect to the covector pair. Therefore scaling by $\frac{1}{h^j}$, we see that γ_h satisfies PMP also for the pair $\frac{1}{h^j}(\delta_h^*\lambda, h\xi) = (\frac{1}{h^j}\delta_h^*\lambda, \frac{1}{h^{j-1}}\xi)$. These pairs form a convergent sequence as $h \to \infty$:

$$\lim_{h\to\infty}(\frac{1}{h^j}\delta_h^*\lambda, \frac{1}{h^{j-1}}\xi) = (\lambda_\infty, 0),$$

where

$$\lambda_\infty := \lim_{h\to\infty}(h^{1-j}\lambda^{(1)} + h^{2-j}\lambda^{(2)} + \cdots + \lambda^{(j)}) = \lambda^{(j)} \neq 0.$$

Let $\sigma \in \mathrm{Asymp}(\gamma)$, so there exists some sequence $h_j \to \infty$ for which $\sigma = \lim_{j\to\infty} \gamma_{h_j}$. By continuity, it follows that σ satisfies PMP for the pair $(\lambda_\infty, 0)$, so σ is an abnormal curve. □

5.2. **Infinite geodesics in step 2 sub-Riemannian Carnot groups.**

Proposition 5.6. *The only infinite geodesics in sub-Riemannian Carnot groups of step 2 are the horizontal lines.*

Proof. Let $\gamma : \mathbb{R} \to G$ be an infinite geodesic in a rank r step 2 Carnot group G. By lifting γ, we may assume that G is the free Carnot group of rank r and step 2.

In step 2 Carnot groups, every geodesic is normal, so γ satisfies PMP for some pair $(\lambda, 1)$. For normal geodesics, PMP can be rewritten as an ODE for γ by renormalizing so that $\xi = 1$. In step 2 Carnot groups, the ODE is affine, and in the specific case of a free Carnot group of step 2 we get the following form:

Decompose $\lambda = \lambda_H + \lambda_V \in V_1^* + V_2^*$ and fix an orthonormal basis of V_1. Then the horizontal projection $\pi \circ \gamma$ of the curve satisfies the ODE

$$\dot{x} = A_{\lambda_V} x + \lambda_H^*,$$

where $A_{\lambda_V} \in \mathfrak{so}(r)$ is a skew-symmetric matrix whose elements are (up to sign) the components of the vertical part λ_V, and $\lambda_H^* \in V_1$ is the dual of $\lambda_H \in V_1^*$ with respect to the sub-Riemannian inner product. By linearity we can translate the curve γ by some element $g \in G$ such that the projection $\pi(g \cdot \gamma) = \pi(g) + \pi \circ \gamma$ satisfies the ODE

$$\dot{x} = A_{\lambda_V} x + b_{\lambda_H}, \tag{5.7}$$

where $b_{\lambda_H} \in V_1$ is the projection of λ_H^* to the orthogonal complement of $\mathrm{Im}(A_{\lambda_V}) \subset V_1$. Furthermore, renormalizing the ξ component given by Lemma 5.1, we see that the horizontal projection of a dilation $\gamma_h := \delta_{1/h} \circ (g \cdot \gamma) \circ \delta_h$ satisfies a similar ODE, where λ is replaced by $\frac{1}{h}\delta_h^* \lambda = \lambda_H + h\lambda_V$. Explicitly, since the matrix A_{λ_V} depends linearly on λ_V, we have

$$\dot{x} = A_{h\lambda_V} x + b_{\lambda_H} = hA_{\lambda_V} x + b_{\lambda_H}.$$

The solution of the above with the initial condition $x(0) = \pi \circ \gamma_h(0) = \frac{1}{h}\pi(g\gamma(0))$ is

$$x(t) = \frac{1}{h} e^{hA_{\lambda_V} t} \pi(g\gamma(0)) + b_{\lambda_H} t. \tag{5.8}$$

Consider any blowdown of the curve $g \cdot \gamma$, i.e., a limit $\sigma = \lim_{j\to\infty} \gamma_{h_j}$ along some sequence $h_j \to 0$. By independence from the basepoint of a blowdown, σ is also a blowdown of γ for the same sequence h_j. Taking the limit of (5.8) as $h_j \to \infty$, we see that the limit curve is the line $\sigma(t) = b_{\lambda_H} t$.

Since γ is a geodesic, the ODE (5.7) implies that $\|A_{\lambda_V} x + b_{\lambda_H}\|^2 = 1$. On the other hand, the vector b_{λ_H} is by construction orthogonal to $\mathrm{Im}(A_{\lambda_V})$, so for any point $x = \pi(g) + \pi(\gamma(t))$, $t \in \mathbb{R}$, we have

$$1 = \|A_{\lambda_V} x + b_{\lambda_H}\|^2 = \|A_{\lambda_V} x\|^2 + \|b_{\lambda_H}\|^2.$$

That is, either $A_{\lambda_V} x = 0$ for all $x = \pi(g) + \pi(\gamma(t))$, in which case the ODE (5.7) implies that γ is a line, or $\|b_{\lambda_H}\| < 1$. But in the latter case we would have

$$\|\dot{\sigma}\| = \|b_{\lambda_H}\| < 1,$$

so the blowdown σ would not be parametrized with unit speed. This would contradict the assumption that γ is an infinite geodesic, so we see that γ must be a line. $\square$

Remark 5.9 Proposition 5.6 can be used to prove that in fact any isometric embedding of any Carnot group into any sub-Riemannian Carnot group of step 2 is affine. This follows by replicating the proof of [BFS18, Theorem 1.1] that all infinite geodesics being lines is sufficient to conclude that arbitrary isometric embeddings from other Carnot groups are affine. Although the result of Balogh, Fässler, and Sobrino is stated in the setting of Heisenberg groups, their proof (with only superficial modifications) works also in the general setting of arbitrary step 2 Carnot groups.

6. Applications of the existence of a line tangent

We next provide some consequences of the existence of a line tangent.

6.1. **Loss of optimality.**

We prove that there are geodesics that lose optimality whenever they are extended.

Proposition 6.1. *In every non-Abelian sub-Finsler Carnot group defined by a strictly convex norm (e.g., in every sub-Riemannian Carnot group) there exist finite-length geodesics that cannot be extended as geodesics.*

Proof. For every such group G, we know that the only infinite geodesics in $G/[G,G]$ are lines. Therefore, by Corollary 1.3 every geodesic has an iterated tangent that is a line. Since iterated tangents are tangents, we have that every geodesic in G has a line tangent.

Fix a nonzero element $v \in V_2$, which exists since G is not Abelian. Let $\gamma : [0,T] \to G$ be a geodesic with $\gamma(0) = 1_G$ and $\gamma(T) = \exp(v)$. We claim that any such geodesic cannot be extended to a geodesic $\tilde{\gamma} : [-\epsilon, T] \to G$ such that $\tilde{\gamma}|_{[0,T]} = \gamma$ for any $\epsilon > 0$.

Let $\delta_{-1} : G \to G$ be the group homomorphism such that $(\delta_{-1})_*(v) = (-1)^j v$ for all $v \in V_j$. The map δ_{-1} is an isometry, since $(\delta_{-1})_*|_{V_1}$ is an isometry. Notice that $\delta_{-1} \circ \gamma$ is another[2] geodesic from 1_G to $\exp(v)$.

Suppose that an extension $\tilde{\gamma} : [\epsilon, T] \to G$ of γ existed. By the existence of a line tangent outlined in the first paragraph, we have that there exists a sequence $h_j \to 0$ such that

$$\tilde{\gamma}_{h_j} = \delta_{\frac{1}{h_j}} \circ \tilde{\gamma} \circ \delta_{h_j} \to \sigma,$$

with $\sigma(t) = \exp(tX)$ for some $X \in V_1$. Replace γ by $\delta_{-1} \circ \gamma$ in the extension $\tilde{\gamma}$, i.e., consider the concatenated curve

$$\eta := \tilde{\gamma}|_{[-\epsilon,0]} * (\delta_{-1} \circ \gamma).$$

[2]We learned this trick for proving non-uniqueness of geodesics in Carnot groups from [Ber16, Proposition 3.2]

Since γ and $\delta_{-1} \circ \gamma$ are both geodesics with the same endpoints, and $\tilde{\gamma}$ was a geodesic extension of γ, the curve η is also a geodesic. However, η has a blowup at 0 that is not injective: for $t < 0$,

$$\eta_{\epsilon_j}(t) = (\delta_{\frac{1}{\epsilon_j}} \circ \eta \circ \delta_{\epsilon_j})(t) = \tilde{\gamma}_{\epsilon_j}(t) \to \exp(tX)$$

whereas for $t > 0$,

$$\eta_{\epsilon_j}(t) = (\delta_{\frac{1}{\epsilon_j}} \circ \delta_{-1} \circ \gamma \circ \delta_{\epsilon_j})(t) = \delta_{-1} \circ \tilde{\gamma}_{\epsilon_j}(t) \to \delta_{-1}\exp(tX) = \exp(-tX).$$

Any blowup of the geodesic η would have to be a geodesic, but this blowup is not even injective, so we get a contradiction. $\square$

6.2. **Non-minimality of corners for distributions of non-constant rank.** In the previous work [HL16] we proved that corners cannot be length-minimizing in any sub-Riemannian manifold, in which by standard definition the distribution has constant rank. We will next show that the existence of a line tangent implies that the assumption that the distribution has constant rank can be omitted.

Proposition 6.2. *Let M be a generalized sub-Riemannian manifold with a distribution not necessarily of constant rank. If $\gamma : I \to M$ is a geodesic, then γ cannot have a corner-type singularity.*

Proof. Reparametrizing and translating, it suffices to consider the case when $I = (-\epsilon, \epsilon)$ and show that γ cannot have a corner-type singularity at 0.

Let $\tilde{M}$ be a desingularization of M as in [Jea14, Lemma 2.5], that is an equiregular sub-Riemannian manifold equipped with a canonical projection $\pi : \tilde{M} \to M$. Since the projection is a submetry, the geodesic γ can be lifted to a geodesic $\tilde{\gamma} : I \to \tilde{M}$. Let $u : (-\epsilon, \epsilon) \to \mathbb{R}^r$ be the control of $\tilde{\gamma}$ with respect to a fixed frame $\tilde{X}_1, \dots, \tilde{X}_r$ of horizontal vector fields on $\tilde{M}$. Then u is also the control of γ with respect to the projected horizontal frame $\pi_*\tilde{X}_1, \dots, \pi_*\tilde{X}_r$ on M.

By Theorem 1.1, the curve $\tilde{\gamma}$ has an iterated tangent that is a line, and thus also a tangent that is a line. By [MPV18b, Remark 3.12] it follows that there exist a constant $v \in \mathbb{R}^r$ and a sequence of scales $h_j \to 0$ such that for the rescaled controls $u^{(j)} : (-\epsilon/h_j, \epsilon/h_j) \to \mathbb{R}^r$, $u^{(j)}(t) = u(h_j t)$, we have $u^{(j)} \to v$ in $L^2_{\mathrm{loc}}(\mathbb{R};\mathbb{R}^r)$.

On the other hand, in coordinates near $\gamma(0)$ on M, for any small enough h_j we have

$$\frac{\gamma(h_j)}{h_j} = \int_0^{h_j} \sum_{k=1}^r u_k(t) X_k(\gamma(t)) \frac{dt}{h_j} = \int_0^1 \sum_{k=1}^r u_k^{(j)}(t) X_k(\gamma(h_j t))\, dt \quad \text{and}$$

$$\frac{\gamma(-h_j)}{-h_j} = \int_{-h_j}^0 \sum_{k=1}^r u_k(t) X_k(\gamma(t)) \frac{dt}{h_j} = \int_{-1}^0 \sum_{k=1}^r u_k^{(j)}(t) X_k(\gamma(h_j t))\, dt.$$

By continuity of the vector fields $X_1, \dots, X_r$ and the convergence $u^{(j)} \to v$ in L^2_{loc}, we see that

$$\lim_{j\to\infty} \frac{\gamma(h_j)}{h_j} = \sum_{k=1}^{r} v_k X_k(\gamma(0)) = \lim_{j\to\infty} \frac{\gamma(-h_j)}{-h_j}.$$

In particular, if γ has one-sided derivatives at 0, then they must be equal, so γ cannot have a corner-type singularity. □

7. On sharpness of Theorem 1.4

In this section we want to consider whether Theorem 1.4 can be improved. In particular, we will show that in the statement of the theorem, taking the horizontal projection is essential. That is, there exist geodesics that are not in a finite neighborhood of any proper Carnot subgroup (see Corollary 7.28).

A possible improvement of Theorem 1.4 would be to strengthen the claim in the horizontal projection. Namely, the following might be true.

Conjecture 7.1. *If $\gamma : \mathbb{R} \to G$ is a geodesic such that $\pi \circ \gamma : \mathbb{R} \to G/[G,G]$ is not a geodesic, then there exists a hyperplane $W \subset V_1$ such that $\lim_{t\to\pm\infty} d(\pi \circ \gamma(t), W) = 0$.*

Toward this conjecture, we shall consider the case of rank 2 Carnot groups, where proper Carnot subgroups are simply lines. For this reason, in the next subsection we first prove some general statements about lines that are a finite distance apart.

7.1. Lines in Carnot groups.

A *line* in a Lie group is a left-translation of a one-parameter subgroup, i.e., a curve $L : \mathbb{R} \to G$ such that $L(t) = g\exp(tX)$ for some $g \in G$ and $X \in \mathfrak{g}$. We stress that in case G is a Carnot group the vector X is not assumed to be horizontal.

The distance between lines will be measured by the Hausdorff distance: The Hausdorff distance of two subsets $A, B \subset G$ is

$$d_H(A,B) := \max\Big(\sup_{a\in A} d(a,B), \sup_{b\in B} d(b,A)\Big).$$

In Lemma 7.5 we will give two equivalent algebraic conditions for two lines to be at a bounded distance from each other. In the proof we will want to use also the notion of distance of lines given by the sup-norm, which is parametrization dependent. For this reason we first prove a sufficient condition (Lemma 7.2) for the equivalence of boundedness of Hausdorff distance and boundedness of sup-norm. This result is naturally stated in much more generality than just lines in Carnot groups.

Lemma 7.2. *Let X and Y be metric spaces, and let $\alpha : X \to Y$ and $\beta : X \to Y$ be maps such that the following conditions hold.*

(a) The map β is bornologous: For every $R < \infty$, there exists $R' < \infty$ such that $\beta(B_X(x,R)) \subset B_Y(\beta(x), R')$ for any $x \in X$.

(b) The map $\alpha \times \beta : X^2 \to Y^2$ maps distant points to distant points: For every $M < \infty$, there exists $R < \infty$ such that $d(\alpha(x_1), \beta(x_2)) > M$ for any x_1, x_2 with $d_X(x_1, x_2) > R$.

Then $d_H(\alpha(X), \beta(X)) < \infty$ if and only if $\sup_{x \in X} d(\alpha(x), \beta(x)) < \infty$.

Proof. Clearly $d_H(\alpha(X), \beta(X)) \leq \sup_{x \in X} d(\alpha(x), \beta(x))$ so it suffices to prove the "only if" implication. That is, we assume that $M := d_H(\alpha(X), \beta(X)) < \infty$ and we will show that also $\sup_{x \in X} d(\alpha(x), \beta(x)) < \infty$.

By the definition of the Hausdorff distance, we have $d(\alpha(x), \beta(X)) \leq M$ for every $x \in X$. Therefore there exists a (possibly discontinuous) map $f : X \to X$ choosing roughly closest points from $\beta(X)$, i.e., a map such that

$$d(\alpha(x), \beta \circ f(x)) \leq M + 1 \quad \forall x \in X. \tag{7.3}$$

Let $R < \infty$ be the constant given by the assumption (b) such that $d(\alpha(x_1), \beta(x_2)) > M + 1$ for any $x_1, x_2 \in X$ with $d(x_1, x_2) > R$. Then the bound (7.3) implies that

$$d(x, f(x)) \leq R \quad \forall x \in X.$$

Assumption (a) then implies that there exists $R' < \infty$ such that

$$d(\beta(x), \beta \circ f(x)) \leq R' \quad \forall x \in X. \tag{7.4}$$

Combining the bounds (7.3) and (7.4), we get for any $x \in X$ the uniform bound

$$d(\alpha(x), \beta(x)) \leq d(\alpha(x), \beta \circ f(x)) + d(\beta \circ f(x), \beta(x)) \leq M + 1 + R' < \infty,$$

proving the claim. □

Lemma 7.5. *Assume G is a Carnot group and let $L_1(t) = g\exp(tX)$ and $L_2(t) = h\exp(tY)$ be two lines in the group. The following are equivalent:*

(i) There exists a constant $c > 0$ such that $X = c\operatorname{Ad}_{g^{-1}h} Y$.
(ii) There exist a constant $c > 0$ and an element $k \in G$ such that $L_1(t) = L_2(ct)k$.
(iii) $d_H(L_1(\mathbb{R}_+), L_2(\mathbb{R}_+)) < \infty$.

Proof. The equivalence of (i) and (ii) is an algebraic computation: For any $k \in G$ and $Z \in \mathfrak{g}$, we have the identity

$$k\exp(Z) = k\exp(Z)k^{-1}k = C_k(\exp(Z)) \cdot k = \exp(\operatorname{Ad}_k Z)k.$$

For any $c > 0$, we apply the above with $k = g^{-1}h$ and $Z = ctY$. This gives the identity

$$L_1(t)^{-1}L_2(ct) = (g\exp(tX))^{-1} \cdot (h\exp(ctY)) = \exp(-tX)\exp(ct\operatorname{Ad}_{g^{-1}h} Y)g^{-1}h. \tag{7.6}$$

If (i) holds, then (7.6) implies that $L_1(t)^{-1}L_2(ct)$ is constant, proving (ii). Vice versa, if (ii) holds, then $L_1(t)^{-1}L_2(ct)$ is constant, so (7.6) is constant. But this is only possible if (i) holds.

That (ii) implies (iii) is immediate from the left-invariance of the distance on G. It remains to prove that (iii) implies (ii). The claim is equivalent to saying that the product $L_1(t)^{-1}L_2(ct)$ is constant for some $c > 0$. Since the product is in exponential coordinates a polynomial expression, it suffices to show that

$$\sup_{t\in\mathbb{R}_+} d(L_1(t), L_2(ct)) < \infty. \tag{7.7}$$

We will prove this by induction on the step of the group G. In a normed space, two half-lines are a finite distance apart if and only if they are parallel, so the claim holds in step 1.

Suppose that the claim is true for all Carnot groups of step at most $s-1$ and suppose that G is of step s. We will prove (7.7) by applying Lemma 7.2 to the curves $\alpha(t) = L_2(ct)$ and $\beta(t) = L_1(t)$ for some $c > 0$ to be fixed later.

From the identity

$$d(L_1(t_2), L_1(t_1)) = d(L_1(0), L_1(t_1 - t_2))$$

we see that $R' = \sup_{|t|\leq R} d(L_1(0), L_1(t)) < \infty$ satisfies assumption (a) of Lemma 7.2.

For assumption (b) of Lemma 7.2, we need a lower bound for $d(L_1(t_1), L_2(ct_2))$. We consider first the case when the lines degenerate under the projection $\pi_{s-1} : G \to G/\exp(V_s)$ to step $s-1$, i.e., when $X, Y \in V_s$. Since elements in $\exp(V_s)$ commute with everything, for any $t_1, t_2 \in \mathbb{R}_+$ we have that

$$d(L_1(t_1), L_2(t_2)) = d(1_G, g^{-1}h\exp(t_2Y - t_1X)).$$

If $Y = cX$ for some $c > 0$, then condition (i) is satisfied, which implies (7.7) by the first part of the proof. Otherwise, $t_2Y - t_1X$ escapes any compact subset of V_s as $|t_2 - t_1| \to \infty$. Recall that in Carnot groups the exponential map is a global diffeomorphism and the distance function is proper. Hence, the lower bound

$$d(L_1(t_1), L_2(t_2)) \geq d(1_G, \exp(t_2Y_2 - t_1Y_1)) - d(1_G, g^{-1}h)$$

implies that assumption (b) of Lemma 7.2 is satisfied for any $c > 0$. By Lemma 7.2 we conclude that in this case (7.7) is satisfied for any $c > 0$.

Next we consider the case when at least one of the lines does not degenerate under the projection $\pi_{s-1} : G \to G/\exp(V_s)$. Since the projection π_{s-1} is 1-Lipschitz, we have

$$d_H(\pi_{s-1} \circ L_1(\mathbb{R}_+), \pi_{s-1} \circ L_2(\mathbb{R}_+)) \leq d_H(L_1(\mathbb{R}_+), L_2(\mathbb{R}_+)) < \infty.$$

Note that the above implies that also the other line cannot degenerate to a constant.

By the inductive assumption in the step $s-1$ Carnot group $G/\exp(V_s)$, we can fix $c > 0$ such that

$$M := \sup_{t\in\mathbb{R}_+} d(\pi_{s-1} \circ L_1(t), \pi_{s-1} \circ L_2(ct)) < \infty.$$

It follows that for any $t_1, t_2 \in \mathbb{R}_+$ we get the lower bound

$$\begin{aligned} d(L_1(t_1), L_2(ct_2)) &\geq d(\pi_{s-1} \circ L_1(t_1), \pi_{s-1} \circ L_2(ct_2)) \\ &\geq d(\pi_{s-1} \circ L_1(t_1), \pi_{s-1} \circ L_1(t_2)) - d(\pi_{s-1} \circ L_1(t_2), \pi_{s-1} \circ L_2(ct_2)) \\ &\geq d(\pi_{s-1} \circ L_1(t_1), \pi_{s-1} \circ L_1(t_2)) - M. \end{aligned} \tag{7.8}$$

Decompose the direction vector of L_1 into homogeneous components as $X = X_{(1)} + \cdots + X_{(s)} \in V_1 \oplus \cdots \oplus V_s$ and let k be the smallest index for which $X_{(k)} \neq 0$. Since $\pi_{s-1} \circ L_1$ is non-constant, we have $k \leq s-1$. By homogeneity of the distance in the projection to step k we get the lower bound

$$\begin{aligned} d(\pi_{s-1} \circ L_1(t_1), \pi_{s-1} \circ L_1(t_2)) &\geq d(\pi_k \circ L_1(0), \pi_k \circ L_1(t_2 - t_1)) \\ &= d(\pi_k(1_G), \pi_k \circ \exp((t_2 - t_1)X_{(k)})) \\ &= |t_2 - t_1|^{1/k} \, d(\pi_k(1_G), \pi_k \circ \exp(X_{(k)})). \end{aligned} \tag{7.9}$$

Combining (7.8) and (7.9) and denoting $C := d(\pi_k(1_G), \pi_k \circ \exp(X_{(k)})) > 0$, we have that

$$d(L_1(t_1), L_2(ct_2)) \geq C\, |t_2 - t_1|^{1/k} - M.$$

This shows that assumption (b) of Lemma 7.2 holds for $\alpha(t) = L_2(ct)$ and $\beta(t) = L_1(t)$, so Lemma 7.2 implies that we have (7.7). ☐

7.2. **An explicit infinite non-line geodesic in the Engel group.** The sub-Riemannian Engel group E is a sub-Riemannian Carnot group of rank 2 and step 3 of dimension 4. Its Lie algebra $\mathfrak{g}$ has a basis $X_1, X_2, X_{12}, X_{112}$ whose only non-zero commutators are

$$[X_1, X_2] = X_{12} \quad \text{and} \quad [X_1, X_{12}] = X_{112}.$$

In [AS15], Ardentov and Sachkov studied the cut time for normal extremals in the Engel group, and found a family of infinite geodesics that are not lines. These geodesics have a property stronger than that implied by Theorem 1.4. Namely, instead of merely having their horizontal projections contained in a finite neighborhood of a hyperplane, their horizontal projections are in fact asymptotic to a line.

To study these infinite geodesics explicitly, we will consider exponential coordinates

$$\mathbb{R}^4 \to E, \quad x = (x_1, x_2, x_{12}, x_{112}) \mapsto \exp(x_1X_1 + x_2X_2 + x_{12}X_{12} + x_{112}X_{112}).$$

By the BCH formula, the group law is given by $x \cdot y = z$, where

$$\begin{aligned} z_1 &= x_1 + y_1 \\ z_2 &= x_2 + y_2 \\ z_{12} &= x_{12} + y_{12} + \frac{1}{2}(x_1y_2 - x_2y_1) \\ z_{112} &= x_{112} + y_{112} + \frac{1}{2}(x_1y_{12} - x_{12}y_1) + \frac{1}{12}(x_1^2y_2 - x_1x_2y_1 - x_1y_1y_2 + x_2y_1^2). \end{aligned} \tag{7.10}$$

The left-invariant extensions of the horizontal basis vectors X_1, X_2 are

$$X_1(x) = \partial_1 - \frac{1}{2}x_2\partial_{12} - (\frac{1}{12}x_1x_2 + \frac{1}{2}x_{12})\partial_{112} \quad \text{and} \tag{7.11}$$
$$X_2(x) = \partial_2 + \frac{1}{2}x_1\partial_{12} + \frac{1}{12}x_1^2\partial_{112}.$$

Note that the coordinates used in [AS15] are not exponential coordinates, but the two coordinate systems agree in the horizontal (x_1 and x_2) components.

Given a covector written in the dual basis as $\lambda = (\lambda_1, \lambda_2, \lambda_{12}, \lambda_{112}) \in \mathfrak{g}^*$, the normal equation given by PMP takes the form

$$u_\gamma(t) = \lambda\left(\operatorname{Ad}_{\gamma(t)} X_1\right) X_1 + \lambda\left(\operatorname{Ad}_{\gamma(t)} X_2\right) X_2. \tag{7.12}$$

In [AS15], the space of covectors $\mathfrak{g}^*$ is stratified into 7 different classes $C_1, \dots, C_7$ based on the different types of trajectories of the corresponding normal extremals. For our purposes the relevant class is C_3, which consists of the non-line infinite geodesics. In [AS15], the class was parametrized by

$$C_3 = \{(\cos(\theta + \pi/2), \sin(\theta + \pi/2), c, \alpha) : \ \alpha \neq 0, \frac{c^2}{2} - \alpha\cos\theta = |\alpha|, c \neq 0\}.$$

An example of a covector $\lambda \in \mathfrak{g}^*$ in the class C_3 is $\lambda = (0, 1, 2, 1)$. However, instead of integrating the normal equation (7.12) with this covector, we will consider a translation of the curve to simplify the asymptotic study of the resulting curve. Instead of considering the geodesic starting from $(0, 0, 0, 0)$, we will consider the translated geodesic starting from $(2, 0, 0, 0)$.

If $\gamma : \mathbb{R} \to E$ satisfies (7.12) with the covector λ, then a left-translation $\beta = g\gamma : \mathbb{R} \to E$ by $g \in E$ satisfies

$$\begin{aligned} u_\beta(t) &= \lambda\left(\operatorname{Ad}_{\gamma(t)} X_1\right) X_1 + \lambda\left(\operatorname{Ad}_{\gamma(t)} X_2\right) X_2 \\ &= \lambda\left(\operatorname{Ad}_{g^{-1}\beta(t)} X_1\right) X_1 + \lambda\left(\operatorname{Ad}_{g^{-1}\beta(t)} X_2\right) X_2. \end{aligned} \tag{7.13}$$

Using the formula $\operatorname{Ad}_{\exp(Y)} X = e^{\operatorname{ad}(Y)}X$, we compute for $x = (x_1, x_2, x_{12}, x_{112}) \in E$ that

$$\operatorname{Ad}_x X_1 = X_1 - x_2X_{12} - (x_{12} + \frac{1}{2}x_1x_2)X_{112} \quad \text{and}$$
$$\operatorname{Ad}_x X_2 = X_2 + x_1X_{12} + \frac{1}{2}x_1^2X_{112}.$$

Evaluated for the covector $\lambda = (0, 1, 2, 1)$, we get

$$\lambda(\operatorname{Ad}_x X_1) = -2x_2 - x_{12} - \frac{1}{2}x_1x_2, \tag{7.14}$$
$$\lambda(\operatorname{Ad}_x X_2) = 1 + 2x_1 + \frac{1}{2}x_1^2.$$

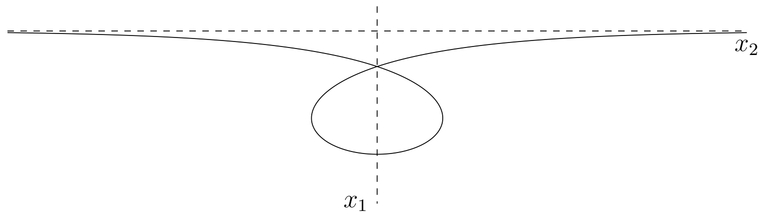

FIGURE 1. Horizontal projection of the non-line Engel infinite geodesic β (rotated 90° clockwise)

By the group law (7.10), the translation of the curve in which we are interested is

$$(2,0,0,0)^{-1}\cdot\beta = \left(\beta_1 - 2,\ \beta_2,\ \beta_{12}-\beta_2,\ \beta_{112}-\beta_{12}+\frac{1}{3}\beta_2+\frac{1}{6}\beta_1\beta_2\right).$$

Substituting the points $x = (2,0,0,0)^{-1}\cdot\beta(t)$ into (7.13) using (7.14), we get the ODE

$$\dot\beta_1 = -\frac{1}{2}\beta_1\beta_2 - \beta_{12} \qquad \dot\beta_2 = \frac{1}{2}\beta_1^2 - 1. \tag{7.15}$$

Lemma 7.16. *The horizontal curve $\beta : \mathbb{R} \to E$ satisfying the ODE* (7.15) *with the initial condition $\beta(0) = (2,0,0,0)$ has the explicit form (see Figure 1)*

$$\beta_1(t) = \frac{2}{\cosh(t)},\quad \beta_2(t) = 2\tanh(t) - t,\quad \beta_{12}(t) = \frac{t}{\cosh(t)},\quad \beta_{112}(t) = \frac{2}{3}\tanh(t) - \frac{t}{3\cosh(t)^2}.$$

Proof. The proof of the lemma is a direct computation. First we shall verify horizontality of β, i.e., that $\dot\beta(t) = \dot\beta_1(t)X_1(\beta(t)) + \dot\beta_2(t)X_2(\beta(t))$. By the coordinate form (7.11) of the left-invariant frame, we need to check that

$$\dot\beta_{12} = \frac{1}{2}(\beta_1\dot\beta_2 - \beta_2\dot\beta_1) \quad \text{and} \tag{7.17}$$

$$\dot\beta_{112} = \frac{1}{12}\beta_1^2\dot\beta_2 - \left(\frac{1}{12}\beta_1\beta_2 + \frac{1}{2}\beta_{12}\right)\dot\beta_1. \tag{7.18}$$

From the given explicit form of β, we compute the derivatives

$$\begin{aligned}\dot\beta_1 &= -\frac{2\,\sinh(t)}{\cosh(t)^2},\\ \dot\beta_2 &= 2(1-\tanh(t)^2)-1 = 1-2\,\tanh(t)^2,\\ \dot\beta_{12} &= \frac{\cosh(t)-t\sinh(t)}{\cosh(t)^2},\\ \dot\beta_{112} &= \frac{2}{3\cosh(t)^2}-\frac{\cosh(t)-2t\sinh(t)}{3\cosh(t)^3} = \frac{\cosh(t)+2t\sinh(t)}{3\cosh(t)^3}.\end{aligned}$$

Expanding the right hand side $\frac{1}{2}(\beta_1\dot\beta_2-\beta_2\dot\beta_1)$ of (7.17), we get

$$\begin{aligned}&\frac{1}{2}\left(\frac{2}{\cosh(t)}\Big(1-\frac{2\sinh(t)^2}{\cosh(t)^2}\Big)-\Big(\frac{2\sinh(t)}{\cosh(t)}-t\Big)\Big(-\frac{2\sinh(t)}{\cosh(t)^2}\Big)\right)\\ &=\frac{1}{\cosh(t)}-\frac{2\sinh(t)^2}{\cosh(t)^3}+\frac{2\sinh(t)^2}{\cosh(t)^3}-\frac{t\sinh(t)}{\cosh(t)^2}\\ &=\frac{\cosh(t)-2t\sinh(t)}{\cosh(t)^2},\end{aligned}$$

which is exactly $\dot\beta_{12}$. Similarly, expanding the right hand side $\frac{1}{12}\beta_1^2\dot\beta_2-\Big(\frac{1}{12}\beta_1\beta_2+\frac{1}{2}\beta_{12}\Big)\dot\beta_1$ of (7.18), we get

$$\begin{aligned}&\frac{1}{12}\frac{4}{\cosh(t)^2}\Big(1-\frac{2\sinh(t)^2}{\cosh(t)^2}\Big)-\Big(\frac{1}{12}\frac{2}{\cosh(t)}\Big(\frac{2\sinh(t)}{\cosh(t)}-t\Big)+\frac{1}{2}\frac{t}{\cosh(t)}\Big)\Big)\Big(-\frac{2\sinh(t)}{\cosh(t)^2}\Big)\\ &=\frac{1}{3\cosh(t)^2}-\frac{2\sinh(t)^2}{3\cosh(t)^4}+\frac{2\sinh(t)^2}{3\cosh(t)^4}+\frac{2t\sinh(t)}{3\cosh(t)^3}\\ &=\frac{\cosh(t)+2t\sinh(t)}{3\cosh(t)^3},\end{aligned}$$

which is exactly $\dot\beta_{112}$, proving horizontality of the curve β.

Finally, we verify that β satisfies the ODE (7.15). Once again, expanding the right hand sides we get

$$\begin{aligned}-\frac{1}{2}\beta_1\beta_2-\beta_{12} &= -\frac{1}{2}\frac{2}{\cosh(t)}\Big(2\frac{\sinh(t)}{\cosh(t)}-t\Big)-\frac{t}{\cosh(t)} = -\frac{2\sinh(t)}{\cosh(t)^2} = \dot\beta_1 \quad \text{and}\\ \frac{1}{2}\beta_1^2-1 &= \frac{1}{2}\frac{4}{\cosh(t)^2}-1 = \frac{2}{\cosh(t)^2}-1 = 1-2\tanh(t)^2 = \dot\beta_2.\end{aligned}$$ □

From the explicit form of the infinite geodesic β we can deduce two properties stronger than that of Theorem 1.4: its horizontal projection is asymptotic to a line and the curve itself is in a finite neighborhood of a line.

Proposition 7.19. *Let $L : \mathbb{R} \to E$, $L(t) = \exp(-tX_2)$, which is the abnormal line in the Engel group, and let $\beta : \mathbb{R} \to E$ be the infinite geodesic of Lemma 7.16. Then*

$$\lim_{t\to\infty} d(\beta(t), \exp(\tfrac{2}{3}X_{112})L(t-2)) = 0 \quad \textit{and} \quad \lim_{t\to-\infty} d(\beta(t), \exp(-\tfrac{2}{3}X_{112})L(t+2)) = 0.$$

Proof. To prove the claim, we will consider the distances $d(\exp(bX_{112})\exp(-(t+a)X_2), \beta(t))$, where $a, b \in \mathbb{R}$ are some constants. This distance is zero exactly when the product

$$z(t) = (0, t+a, 0, -b) \cdot \beta(t)$$

vanishes.

By the group law (7.10) and the explicit form of the components given in Lemma 7.16, we see that the components of the product $z(t)$ are

$$\begin{aligned} z_1(t) &= \beta_1(t) = \frac{2}{\cosh(t)}, \\ z_2(t) &= \beta_2(t) + t + a = 2\tanh(t) + a, \\ z_{12}(t) &= \beta_{12}(t) - \frac{1}{2}(t+a)\beta_1(t) = -\frac{a}{\cosh(t)} \quad \text{and} \\ z_{112}(t) &= \beta_{112}(t) + \frac{1}{12}(t+a)\beta_1(t)^2 - b = \frac{2}{3}\tanh(t) + \frac{a}{3\cosh(t)^2} - b. \end{aligned}$$

From the above we deduce that

$$\lim_{t\to\infty} z(t) = (0, 2+a, 0, 2/3-b) \quad \text{and} \quad \lim_{t\to-\infty} z(t) = (0, -2+a, 0, -2/3-b).$$

and the claim of the proposition follows. □

Corollary 7.20. *Let $L : \mathbb{R} \to E$, $L(t) = \exp(-tX_2)$, which is the abnormal line in the Engel group, and let $\beta : \mathbb{R} \to E$ be the infinite geodesic of Lemma 7.16. Then*

$$\lim_{t\to\pm\infty} d(\pi\circ\beta(t), \pi\circ L(t)) = 0 \quad \textit{and} \quad \sup_{t\in\mathbb{R}} d(\beta(t), L) < \infty.$$

Proof. Since the horizontal projections of the elements $\exp(\pm\frac{2}{3}X_{112})$ are zero, the lines in Proposition 7.19 have the same horizontal projection as the abnormal line L, and the claim $\lim_{t\to\pm\infty} d(\pi\circ\beta(t), \pi\circ L(t)) = 0$ follows.

On the other hand, the elements $\exp(\pm\frac{2}{3}X_{112})$ are also in the center of the Engel group, so for all $t \in \mathbb{R}$ we have

$$\begin{aligned} d(L(t), \exp(\tfrac{2}{3}X_{112})L(t-2)) &\le d(L(t), L(t-2)) + d(L(t-2), \exp(\tfrac{2}{3}X_{112})L(t-2)) \\ &= 2 + d(1_E, \exp(\tfrac{2}{3}X_{112})). \end{aligned}$$

Thus Proposition 7.19 implies that

$$\sup_{t\in\mathbb{R}_+} d(\beta(t), L(t)) \le \sup_{t\in\mathbb{R}_+} d(\beta(t), \exp(\tfrac{2}{3}X_{112})L(t-2)) + d(\exp(\tfrac{2}{3}X_{112})L(t-2), L(t)) < \infty.$$

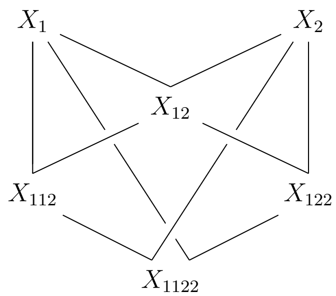

FIGURE 2. Diagram of relations in the step 4 Lie algebra $\mathfrak{g}$ with the Engel Lie algebra as a quotient.

Similarly using the triangle inequality through $\exp(-\frac{2}{3}X_{112})L(t+2)$ instead of $\exp(\frac{2}{3}X_{112})L(t-2)$, we see that $\sup_{t\in\mathbb{R}_-} d(\beta(t), L(t)) < \infty$, proving the claim. □

7.3. **Lift of the infinite non-line geodesic to step 4.** We shall next show that Theorem 1.4 cannot be improved to say that every sub-Riemannian geodesic is at a finite distance from a lower rank subgroup. Although by Corollary 7.20, this stronger claim is true for the Engel group, the claim is no longer true for the lift of the geodesic β from Lemma 7.16 to a specific Carnot group of rank 2 and step 4.

We will prove the claim by showing that the mismatched limits

$$\lim_{t\to\infty}\beta_{112}(t) = \frac{2}{3} \neq -\frac{2}{3} = \lim_{t\to-\infty}\beta_{112}(t)$$

will cause the lift of β to have different lines as asymptotes as $t\to\infty$ and as $t\to-\infty$ (Proposition 7.25). The claim will then follow from Lemma 7.5, where we proved that the only lines a finite distance apart are right translations of one another.

The specific Carnot group G where we will consider a lift of the Engel geodesic β is the one whose Lie algebra $\mathfrak{g}$ has the basis $X_1, X_2, X_{12}, X_{112}, X_{122}, X_{1122}$, whose only non-zero commutators are (see Figure 2 for a visual description)

$$[X_1, X_2] = X_{12}, \quad [X_1, X_{12}] = X_{112}, \quad [X_{12}, X_2] = X_{122}, \quad [X_1, X_{122}] = [X_{112}, X_2] = X_{1122}.$$

The Lie algebra of the Engel group is a quotient of $\mathfrak{g}$ by the ideal generated by X_{122}, so the Engel group is the quotient of G by the subgroup $H = \exp(\mathrm{span}\{X_{122}, X_{1122}\})$. The metric on G is the sub-Riemannian metric such that the projection $\pi_E : G \to E = G/H$ to the Engel group is a submetry.

Let $\beta : \mathbb{R} \to E$ be the geodesic in the Engel group E given in Lemma 7.16. In exponential coordinates on the Engel group, $\beta(0) = (2, 0, 0, 0)$, so for any initial point $x_0 = (2, 0, 0, 0, x_{122}, x_{1122}) \in G$ there exists a horizontal lift of β to G starting from x_0. Let $\alpha : \mathbb{R} \to G$ be the horizontal lift with the initial point $\alpha(0) = (2, 0, 0, 0, 2/3, 0)$. As with $\beta_1(0) = 2$, the initial coordinate $\alpha_{122}(0) = 2/3$ will simplify the asymptotic behavior. Since the projection $\pi_E : G \to E$ is a submetry and $\pi_E \circ \alpha = \beta$ is an infinite geodesic, the curve α is an infinite geodesic in G.

To study the lift α, we will work in exponential coordinates. The group law is once again given by the BCH formula, which in a nilpotent Lie algebra of step 4 takes the form (for computation of the coefficients, see e.g. [Var84, 2.15])

$$\log(\exp(X)\exp(Y)) = X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}([X,[X,Y]]+[[X,Y],Y])+\frac{1}{24}[X,[[X,Y],Y]].$$

In the first four coordinates, the group law $z = x\cdot y$ is the same as in the Engel group, so the components $z_1, z_2, z_{12}, z_{112}$ are given by (7.10). In the last two coordinates, we have

(7.21)
$$\begin{aligned} z_{122} &= x_{122} + y_{122} + \frac{1}{2}(x_{12}y_2 - x_2y_{12}) + \frac{1}{12}(x_1y_2^2 - x_1x_2y_2 - x_2y_1y_2 + x_2^2y_1), \\ z_{1122} &= x_{1122} + y_{1122} + \frac{1}{2}(x_1y_{122} - x_{122}y_1 - x_2y_{112} + x_{112}y_2) - \frac{1}{6}(x_1x_2y_{12} + x_{12}y_1y_2) \\ &\quad + \frac{1}{12}(x_1x_{12}y_2 + x_1y_2y_{12} + x_2x_{12}y_1 + x_2y_1y_{12}) + \frac{1}{24}(x_1^2y_2^2 - x_2^2y_1^2). \end{aligned}$$

The left-invariant extensions of the horizontal vectors X_1 and X_2 are

(7.22)
$$\begin{aligned} X_1(x) &= \partial_1 - \frac{1}{2}x_2\partial_{12} - (\frac{1}{12}x_1x_2 + \frac{1}{2}x_{12})\partial_{112} + \frac{1}{12}x_2^2\partial_{122} + (\frac{1}{12}x_{12}x_2 - \frac{1}{2}x_{122})\partial_{1122}, \\ X_2(x) &= \partial_2 + \frac{1}{2}x_1\partial_{12} + \frac{1}{12}x_1^2\partial_{112} - (\frac{1}{12}x_1x_2 - \frac{1}{2}x_{12})\partial_{122} + (\frac{1}{12}x_1x_{12} + \frac{1}{2}x_{112})\partial_{1122}. \end{aligned}$$

Lemma 7.23. *In exponential coordinates, the second coordinate of degree 3 of $\alpha : \mathbb{R} \to G$ is*

$$\alpha_{122}(t) = \frac{t^2+4}{6\cosh(t)} + \frac{t\sinh(t)}{3\cosh(t)^2}.$$

Proof. By the explicit form of the left-invariant frame given in (7.22), we need to show that the given expression for α_{122} satisfies both the horizontality condition

$$\dot\alpha_{122} = \dot\alpha_1X_1(\alpha) + \dot\alpha_2X_2(\alpha) = \frac{1}{12}\alpha_2^2\dot\alpha_1 - (\frac{1}{12}\alpha_1\alpha_2 - \frac{1}{2}\alpha_{12})\dot\alpha_2 \tag{7.24}$$

and the initial condition $\alpha_{122}(0) = \frac{2}{3}$. The initial condition is immediately verified, since $\alpha_{122}(0) = \frac{4}{6\cosh(0)} = \frac{2}{3}$.

Since α and β agree in the first four coordinates, we get by Lemma 7.16 that

$$\begin{aligned}\frac{1}{12}\alpha_2^2\dot{\alpha}_1 &= \frac{1}{12}\Big(2\frac{\sinh(t)}{\cosh(t)} - t\Big)^2\Big(-\frac{2\sinh(t)}{\cosh(t)^2}\Big)\\ &= -\frac{2\sinh(t)^3}{3\cosh(t)^4} + \frac{2t\sinh(t)^2}{3\cosh(t)^3} - \frac{t^2\sinh(t)}{6\cosh(t)^2}\\ &= \frac{-4\sinh(t)^3 + 4t\sinh(t)^2\cosh(t) - t^2\sinh(t)\cosh(t)^2}{6\cosh(t)^4},\end{aligned}$$

$$\begin{aligned}-\frac{1}{12}\alpha_1\alpha_2\dot{\alpha}_2 &= -\frac{1}{12}\frac{2}{\cosh(t)}\Big(\frac{2\sinh(t)}{\cosh(t)} - t\Big)\Big(1 - \frac{2\sinh(t)^2}{\cosh(t)^2}\Big)\\ &= -\frac{\sinh(t)}{3\cosh(t)^2} + \frac{2\sinh(t)^3}{3\cosh(t)^4} + \frac{t}{6\cosh(t)} - \frac{t\sinh(t)^2}{3\cosh(t)^3}\\ &= \frac{-2\sinh(t)\cosh(t)^2 + 4\sinh(t)^3 + t\cosh(t)^3 - 2t\sinh(t)^2\cosh(t)}{6\cosh(t)^4},\end{aligned}$$

and

$$\begin{aligned}\frac{1}{2}\alpha_{12}\dot{\alpha}_2 &= \frac{1}{2}\frac{t}{\cosh(t)}\Big(1 - \frac{2\sinh(t)^2}{\cosh(t)^2}\Big)\\ &= \frac{t}{2\cosh(t)} - \frac{t\sinh(t)^2}{\cosh(t)^3}\\ &= \frac{3t\cosh(t)^3 - 6t\sinh(t)^2\cosh(t)}{6\cosh(t)^4}.\end{aligned}$$

Summing up the above, we get

$$\begin{aligned}\frac{1}{12}\alpha_2^2\dot{\alpha}_1 - (\frac{1}{12}\alpha_1\alpha_2 - \frac{1}{2}\alpha_{12})\dot{\alpha}_2 &= \frac{-4t\sinh(t)^2 - t^2\sinh(t)\cosh(t) - 2\sinh(t)\cosh(t) + 4t\cosh(t)^2}{6\cosh(t)^3}\\ &= \frac{2t}{3\cosh(t)^3} - \frac{(t^2+2)\sinh(t)}{6\cosh(t)^2}.\end{aligned}$$

On the other hand, by differentiating the given expression for α_{122}, we also get

$$\begin{aligned}\frac{d}{dt}\Big(\frac{t^2+4}{6\cosh(t)} + \frac{t\sinh(t)}{3\cosh(t)^2}\Big) &= \frac{12t\cosh(t) - 6(t^2+4)\sinh(t)}{36\cosh(t)^2}\\ &\quad + \frac{3(\sinh(t) + t\cosh(t))\cosh(t)^2 - 6t\sinh(t)^2\cosh(t)}{9\cosh(t)^4}\\ &= \frac{2t}{3\cosh(t)^3} - \frac{(t^2+2)\sinh(t)}{6\cosh(t)^2},\end{aligned}$$

so the horizontality condition (7.24) is satisfied. □

Proposition 7.25. *Let* $L_\pm(t) = \exp(-(tX_2 \pm \frac{2}{3}tX_{1122}))$*. Then*

$$\sup_{t\in\mathbb{R}_+} d(\alpha(t), L_+(t)) < \infty \quad \textit{and} \quad \sup_{t\in\mathbb{R}_-} d(\alpha(t), L_-(t)) < \infty.$$

Proof. As in Proposition 7.19, we compute the distances $d(\alpha(t), L_\pm(t))$ directly by the considering the products $L_\pm(t)^{-1}\alpha(t)$. Since the lines L_+ and L_- only differ by the sign of $\frac{2}{3}tX_{1122}$, we will combine the computations. That is, we will consider the product

$$z(t) := \exp(tX_2 \pm \tfrac{2}{3}tX_{1122})\alpha(t).$$

The group law in the first four coordinates is exactly the group law of the Engel group (7.10), so the first four components $z_1, z_2, z_{12}, z_{112}$ are bounded by Corollary 7.20. It remains to consider the components z_{122} and z_{1122}.

By the group law (7.21), we have

$$z_{122}(t) = \alpha_{122}(t) - \frac{1}{2}t\alpha_{12}(t) - \frac{1}{12}t\alpha_1(t)\alpha_2(t) + \frac{1}{12}t^2\alpha_1(t) \quad \text{and}$$
$$z_{1122}(t) = \alpha_{1122}(t) \pm \frac{2}{3}t - \frac{1}{2}t\alpha_{112}(t) + \frac{1}{12}t\alpha_1(t)\alpha_{12}(t) - \frac{1}{24}t^2\alpha_1(t)^2.$$

By the explicit expressions given in Lemma 7.16, we see that the components $\alpha_1 = \beta_1$ and $\alpha_{12} = \beta_{12}$ are both exponentially asymptotically vanishing, i.e., for any polynomial $P : \mathbb{R} \to \mathbb{R}$, we have

$$\lim_{t\to\pm\infty} P(t)\alpha_1(t) = \lim_{t\to\pm\infty} P(t)\alpha_{12}(t) = 0.$$

Therefore there exists a constant $C > 0$ such that

$$|z_{122}(t)| \le |\alpha_{122}(t)| + C \quad \text{and} \quad |z_{1122}(t)| \le \left|\alpha_{1122}(t) \pm \frac{2}{3}t - \frac{1}{2}t\alpha_{112}(t)\right| + C. \tag{7.26}$$

By the explicit form in Lemma 7.23, we see that α_{122} is bounded, so the same is true for z_{122}. For z_{1122}, we will consider the term

$$w(t) := \alpha_{1122}(t) \pm \frac{2}{3}t - \frac{1}{2}t\alpha_{112}(t)$$

separately for $t > 0$ and $t < 0$.

Instead of explicitly computing α_{1122}, we will consider the derivative $\dot{w}$. Since α is a horizontal curve, from the explicit form (7.22) of the left-invariant frame, we get the identity

$$\dot{\alpha}_{1122} = (\frac{1}{12}\alpha_{12}\alpha_2 - \frac{1}{2}\alpha_{122})\dot{\alpha}_1 + (\frac{1}{12}\alpha_1\alpha_{12} + \frac{1}{2}\alpha_{112})\dot{\alpha}_2.$$

By the explicit expressions given in Lemmas 7.16 and 7.23, we see that as $t \to \pm\infty$ the terms $\alpha_1, \alpha_{12}, \alpha_{122}, \dot{\alpha}_{112}$ and $\dot{\alpha}_2 + 1$, are all exponentially vanishing. It follows that

$$\dot{w}(t) = -\alpha_{112}(t) \pm \frac{2}{3} + \epsilon(t), \tag{7.27}$$

where $\epsilon : \mathbb{R} \to \mathbb{R}_+$ is some smooth function such that $\epsilon(t) = O(e^{-|t|})$ as $t \to \pm\infty$.

Finally, we observe that as $t \to \infty$, $\alpha_{112}(t) - \frac{2}{3} = O(e^{-t})$, and as $t \to -\infty$, $\alpha_{112}(t) + \frac{2}{3} = O(e^{t})$. Therefore from (7.27) we conclude that as $t \to \infty$ we have

$$\left|\alpha_{1122}(t) + \frac{2}{3}t - \frac{1}{2}t\alpha_{112}(t)\right| \leq \int_0^t \left|-\alpha_{112}(s) + \frac{2}{3} + \epsilon(s)\right| \, ds = O(e^{-t}).$$

It follows from (7.26) that also the final coordinate of $L_+(t)^{-1}\alpha(t)$ is bounded on $\mathbb{R}_+$. Thus the product $L_+(t)^{-1}\alpha(t)$ is bounded on $\mathbb{R}_+$.

Similarly for $t \to -\infty$ we conclude that

$$\left|\alpha_{1122}(t) - \frac{2}{3}t - \frac{1}{2}t\alpha_{112}(t)\right| = O(e^{t}),$$

from which it follows that the product $L_-(t)^{-1}\alpha(t)$ is bounded on $\mathbb{R}_-$, proving the claim. ☐

Corollary 7.28. *Let $L : \mathbb{R} \to G$ be any line. Then $d_H(\alpha(\mathbb{R}), L(\mathbb{R})) = \infty$.*

Proof. The corollary follows from combining Lemma 7.5 and Proposition 7.25. Suppose there existed a line $L \subset G$ such that $d_H(\alpha(\mathbb{R}), L(\mathbb{R})) < \infty$. Then also $d_H(\pi \circ \alpha(\mathbb{R}), \pi \circ L(\mathbb{R})) \leq M$, so from the explicit form of the horizontal components of α given in Lemma 7.16, we see that $\pi \circ L$ must be parallel to the x_2-axis.

Up to reparametrizing L we can then assume that $\pi \circ L(t) = (C, -t)$ for some $C \in \mathbb{R}$. In particular, we have

$$d(\alpha(t), L(s)) \geq d(\pi \circ \alpha(t), \pi \circ L(s)) \geq |t - s| - 2.$$

Then by Lemma 7.2, since $d_H(\alpha(\mathbb{R}), L(\mathbb{R})) < \infty$, we have that also $\sup_t d(\alpha(t), L(t)) < \infty$. In particular $d_H(\alpha(\mathbb{R}_+), L(\mathbb{R}_+)) < \infty$ and $d_H(\alpha(\mathbb{R}_-), L(\mathbb{R}_-)) < \infty$.

Let $L_\pm(t) = \exp(tY_\pm)$ be the lines of Proposition 7.25. Proposition 7.25 and the triangle inequality for the Hausdorff distance imply that

$$d_H(L(\mathbb{R}_+), L_+(\mathbb{R}_+)) \leq d_H(L(\mathbb{R}_+), \alpha(\mathbb{R}_+)) + d_H(\alpha(\mathbb{R}_+), L_-(\mathbb{R}_+)) < \infty$$

and similarly that $d_H(L(\mathbb{R}_-), L_2(\mathbb{R}_-)) < \infty$. By applying Lemma 7.5 to both halves of the line L, we get the existence of constants $c_\pm > 0$ such that

$$X = c_- \operatorname{Ad}_{g^{-1}} Y_- = c_+ \operatorname{Ad}_{g^{-1}} Y_+,$$

where X and g are such that $L(t) = g \exp(tX)$. This implies that Y_+ and Y_- are linearly dependent, which is a contradiction. ☐

Corollary 7.28 shows that $\alpha : \mathbb{R} \to G$ is a geodesic that is not in a finite neighborhood of any line, showing that the claim of Theorem 1.4 cannot hold without considering the projection $\pi : G \to G/[G, G]$. Still Conjecture 7.1 may be true.

References

[Agr98] A. Agrachev, *Compactness for sub-Riemannian length-minimizers and subanalyticity*, Rend. Sem. Mat. Univ. Politec. Torino **56** (1998), no. 4, 1–12 (2001), Control theory and its applications (Grado, 1998).

[AS15] A. A. Ardentov and Yu. L. Sachkov, *Cut time in sub-Riemannian problem on Engel group*, ESAIM Control Optim. Calc. Var. **21** (2015), no. 4, 958–988. MR 3395751

[Bel96] André Bellaïche, *The tangent space in sub-Riemannian geometry*, Sub-Riemannian geometry, Progr. Math., vol. 144, Birkhäuser, Basel, 1996, pp. 1–78.

[Ber89] Valeriĭ N. Berestovskiĭ, *Homogeneous manifolds with an intrinsic metric. II*, Sibirsk. Mat. Zh. **30** (1989), no. 2, 14–28, 225.

[Ber16] ———, *Busemann's results, ideas, questions and locally compact homogeneous geodesic spaces*, Preprint, written for Busemann's Selected Works (2016).

[BFS18] Zoltán M. Balogh, Katrin Fässler, and Hernando Sobrino, *Isometric embeddings into Heisenberg groups*, Geom. Dedicata **195** (2018), 163–192.

[HL16] Eero Hakavuori and Enrico Le Donne, *Non-minimality of corners in subriemannian geometry*, Invent. Math. **206** (2016), no. 3, 693–704.

[Jea14] Frédéric Jean, *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*, Springer Briefs in Mathematics, Springer, Cham, 2014.

[Lan71] Serge Lang, *Linear algebra*, Second edition, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1971.

[LD17] Enrico Le Donne, *A Primer on Carnot Groups: Homogenous Groups, Carnot-Carathéodory Spaces, and Regularity of Their Isometries*, Anal. Geom. Metr. Spaces **5** (2017), 116–137.

[LM08] Gian Paolo Leonardi and Roberto Monti, *End-point equations and regularity of sub-Riemannian geodesics*, Geom. Funct. Anal. **18** (2008), no. 2, 552–582.

[LMO+16] Enrico Le Donne, Richard Montgomery, Alessandro Ottazzi, Pierre Pansu, and Davide Vittone, *Sard property for the endpoint map on some Carnot groups*, Ann. Inst. H. Poincaré Anal. Non Linéaire **33** (2016), no. 6, 1639–1666.

[LR16] Enrico Le Donne and Séverine Rigot, *Remarks about the Besicovitch Covering Property in Carnot groups of step 3 and higher*, Proc. Amer. Math. Soc. **144** (2016), no. 5, 2003–2013.

[MPV18a] Roberto Monti, Alessandro Pigati, and Davide Vittone, *Existence of tangent lines to Carnot-Carathéodory geodesics*, Calc. Var. Partial Differential Equations **57** (2018), no. 3, Art. 75, 18.

[MPV18b] ———, *On Tangent Cones to Length Minimizers in Carnot–Carathéodory Spaces*, SIAM J. Control Optim. **56** (2018), no. 5, 3351–3369.

[Var84] V. S. Varadarajan, *Lie groups, Lie algebras, and their representations*, Graduate Texts in Mathematics, vol. 102, Springer-Verlag, New York, 1984, Reprint of the 1974 edition. MR 746308

[Vit14] Davide Vittone, *The regularity problem for sub-Riemannian geodesics*, Geometric measure theory and real analysis, CRM Series, vol. 17, Ed. Norm., Pisa, 2014, pp. 193–226.

Email address: eero.j.hakavuori@jyu.fi

Email address: enrico.ledonne@jyu.fi

(Hakavuori and Le Donne) Department of Mathematics and Statistics, University of Jyväskylä, 40014 Jyväskylä, Finland

[C]

Infinite geodesics and isometric embeddings in Carnot groups of step 2

E. Hakavuori

Preprint

INFINITE GEODESICS AND ISOMETRIC EMBEDDINGS IN CARNOT GROUPS OF STEP 2

EERO HAKAVUORI

Abstract. In the setting of step 2 sub-Finsler Carnot groups with strictly convex norms, we prove that all infinite geodesics are lines. It follows that for any other homogeneous distance, all geodesics are lines exactly when the induced norm on the horizontal space is strictly convex. As a further consequence, we show that all isometric embeddings between such homogeneous groups are affine.

Contents

Date: May 8, 2019.

2010 *Mathematics Subject Classification.* 30L05, 53C17, 49K21, 22E25.

Key words and phrases. Carnot groups, isometries, isometric embeddings, geodesics, sub-Riemannian geometry, sub-Finsler geometry.

The author was has been partially supported by the Vilho, Yrjö and Kalle Väisälä Foundation, by the Academy of Finland (grant 288501 '*Geometry of sub-Riemannian groups*'), and by the European Research Council (ERC Starting Grant 713998 GeoMeG '*Geometry of Metric Groups*').

1. Introduction

Carnot groups have rich algebraic and metric structures, and share many properties with normed spaces. Recently several articles have generalized classical regularity results of isometric embeddings in normed spaces into the setting of Carnot groups. In real normed spaces, there are two simple criteria for an isometric embedding to be affine: surjectivity or strict convexity of the norm on the target. Both regularity criteria have analogues for isometric embeddings of Carnot groups.

Surjective isometric embeddings behave in the Carnot group case similarly as they do in the normed-space case. Namely, isometries between arbitrary (open subsets of) Carnot groups are affine [LDO16], i.e., compositions of left translations and a group homomorphisms. For globally defined isometries, there is an even more general result that isometries between connected nilpotent metric Lie groups are affine [KLD17].

For non-surjective isometric embeddings, it was proved in [Kis03] that if G is a sub-Riemannian Carnot group of step 2, then all isometric embeddings $\mathbb{R} \hookrightarrow G$, i.e., all *infinite geodesics*, are affine. This property was coined the *geodesic linearity property* in [BFS18], and was used as an alternative to the strict convexity criterion as the two conditions are equivalent in normed spaces. More precisely, it was shown in [BFS18] that if $\mathbb{H}^n$ is a Heisenberg group with a homogeneous distance satisfying the geodesic linearity property, then all isometric embeddings $\mathbb{R}^m \hookrightarrow \mathbb{H}^n$ and $\mathbb{H}^m \hookrightarrow \mathbb{H}^n$ are affine.

It was conjectured in [BFS18] and subsequently proved in [BC18] that for Heisenberg groups the geodesic linearity property is equivalent to strict convexity of the projection norm (see Definition 2.4). The main result of this paper is to generalize the same characterization to arbitrary Carnot groups of step 2:

Theorem 1.1. *In every sub-Finsler Carnot group of step 2 with a strictly convex norm, every infinite geodesic is affine.*

Corollary 1.2. *Let G be a stratified group of step 2 equipped with a homogeneous distance d such that the projection norm of d is strictly convex. Then every infinite geodesic in (G, d) is affine.*

The necessity of the strict convexity assumption is a direct consequence of the necessity of strict convexity for linearity of geodesics in the normed-space case, see Proposition 5.2. The restriction to step 2 is motivated by the known counterexample in the simplest Carnot group of step 3, the sub-Riemannian Engel group. The complete study of geodesics in the sub-Riemannian Engel group in [AS15] gives the first (and to date essentially only) known example of a non-affine infinite geodesic in a sub-Riemannian Carnot group.

The proof for Heisenberg groups in [BFS18] that the geodesic linearity property of the target implies that all isometric embeddings are affine works also more generally for stratified groups. Consequently, Corollary 1.2 leads to the corresponding regularity for arbitrary isometric embeddings:

Theorem 1.3. *Let (H, d_H) and (G, d_G) be stratified groups with homogeneous distances such that G has step 2 and the projection norm of d_G is strictly convex. Then every isometric embedding $(H, d_H) \hookrightarrow (G, d_G)$ is affine.*

It is worth remarking that although there are no explicit restrictions on the domain (H, d_H) in Theorem 1.3, the mere existence of an isometric embedding $(H, d_H) \hookrightarrow (G, d_G)$ implies some restrictions. In particular, Pansu's Rademacher theorem [Pan89] implies that there must exist an injective homogeneous homomorphism $H \to G$. It follows that H has step at most 2 and rank at most the rank of G.

1.1. **Structure of the paper.** Section 2 presents the relevant definitions that will be used throughout the rest of the paper and some basic lemmas. The main points of interest are properties of blowdowns of geodesics, i.e., geodesics "viewed from afar", and the collection of observations about subdifferentials of convex functions.

Sections 3–5 are devoted to the proofs of Theorem 1.1 and Corollary 1.2 about infinite geodesics. Section 3 rephrases the classical first order optimality condition of the Pontryagin Maximum Principle in the setting of a step 2 sub-Finsler Carnot group. In the sub-Riemannian case the PMP reduces to a linear ODE for the controls. This is no longer true in the sub-Finsler case, making explicit solution of the system unfeasible. Nonetheless, the PMP has a form (Proposition 3.1) that is well suited to the study of asymptotic behavior of optimal controls. The key object is the bilinear form $B\colon V_1 \times V_1 \to \mathbb{R}$.

Section 4 covers the aforementioned asymptotic study. The goal of the section is to study blowdowns of infinite geodesics through the behavior of their controls. Using integral averages of controls, it is shown that any blowdown control must in fact be contained in the kernel of the bilinear form B.

Section 5 wraps up the proof of Theorem 1.1 using the conclusions of the previous sections. This section is the only place where strict convexity appears. The importance of the assumption is that any linear map has a unique maximum on the ball. By observing that any element of $\ker B$ defines an invariant along the corresponding optimal control, the uniqueness is exploited to prove that infinite geodesics must be invariant under blowdowns. Corollary 1.2 follows from the sub-Finsler case by the observation that the length metric associated to a homogeneous norm is always a sub-Finsler metric.

Section 6 covers the proof of Theorem 1.3 about isometric embeddings as a consequence of Corollary 1.2. The link between geodesics and general isometric embeddings arises from considering a foliation by horizontal lines in the domain and studying the induced foliation by infinite geodesics in the image. The affinity of isometric embeddings follows from the observation that two lines are at a sublinear distance from each other if and only if they are parallel.

2. Preliminaries

2.1. Stratified groups and homogeneous distances.

Definition 2.1. A *stratified group* is a Lie group G whose Lie algebra has a decomposition $\mathfrak{g} = V_1 \oplus V_2 \oplus \cdots \oplus V_s$ such that $V_s \neq \{0\}$ and $[V_1, V_k] = V_{k+1}$ for all $k = 1, \ldots, s$, with the convention that $V_{s+1} = \{0\}$. The *rank* and *step* of the stratified group G are the integers $r = \dim V_1$ and s respectively.

Definition 2.2. A *dilation* by a factor $h \in \mathbb{R}$ on a stratified group G is the Lie group automorphism $\delta_h \colon G \to G$ defined for any $X = X_1 + \cdots + X_s \in V_1 \oplus \cdots \oplus V_s$ by

$$\delta_h \exp(X_1 + X_2 + \cdots + X_s) = \exp(hX_1 + h^2 X_2 + \cdots + h^s X_s).$$

Definition 2.3. A *homogeneous distance* on a stratified group G is a left-invariant distance d, which is one-homogeneous with respect to the dilations, i.e., which satisfies

$$d(\delta_h(g), \delta_h(h)) = hd(g, h) \quad \forall h > 0, \forall g, h \in G.$$

2.2. The projection norm.

Definition 2.4. Let G be a stratified group and let d be a homogeneous distance on G. The *projection norm* associated to the homogeneous distance d is the function

$$\|\cdot\|_d : V_1 \to \mathbb{R}, \quad \|X\|_d = d(e, \exp(X)),$$

where e is the identity element of the group G.

It is not immediate that $\|\cdot\|_d$ defines a norm. In the setting of the Heisenberg groups, this is proved in [BFS18, Proposition 2.8]. Their proof works with minor modification for any homogeneous distances in arbitrary stratified groups and is captured in the following lemmas.

The triangle inequality of $\|\cdot\|_d$ is the only non-trivial part. In order to make use of the triangle inequality of the distance d, the following distance estimate is required. The estimate relies on the existence of a dilation, and may fail for non-homogeneous left-invariant distances.

Lemma 2.5. *Let $\pi_{V_1}\colon \mathfrak{g} = V_1 \oplus \cdots \oplus V_s \to V_1$ be the projection with respect to the direct sum decomposition. Then*

$$\|X\|_d \leq d(e, \exp(X+Y)) \quad \forall X \in V_1,\ \forall Y \in [\mathfrak{g},\mathfrak{g}],$$

so the horizontal projection $\pi = \pi_{V_1} \circ \log\colon (G,d) \to (V_1, \|\cdot\|_d)$ is a submetry.

Proof. Observe first that for any $X \in V_1$ and $Y = Y_2 + \cdots + Y_s \in V_2 \oplus \cdots \oplus V_s = [\mathfrak{g},\mathfrak{g}]$, and any $n \in \mathbb{N}$, homogeneity and the triangle inequality imply that

$$\begin{aligned} nd(e, \exp(X + \frac{1}{n}Y_2 + \cdots + \frac{1}{n^{s-1}}Y_s)) &= d(e, \exp(nX + nY)) \\ &\leq nd(e, \exp(X+Y)). \end{aligned}$$

Continuity of the distance then gives the bound

$$\begin{aligned} d(e, \exp(X)) &= \lim_{n\to\infty} d(e, \exp(X + \frac{1}{n}Y_2 + \cdots + \frac{1}{n^{s-1}}Y_s)) \\ &\leq d(e, \exp(X+Y)) \end{aligned}$$

for any $X \in V_1$ and $Y \in [\mathfrak{g},\mathfrak{g}]$ as claimed.

The previous estimate implies the containment $\pi(B(e,r)) \subset B_{\|\cdot\|_d}(0,r)$ for the projection of any ball $B(e,r) \subset G$. On the other hand, Definition 2.4 of the projection norm directly implies the opposite containment

$$B_{\|\cdot\|_d}(0,r) = V_1 \cap \log B(e,r) \subset \pi(B(e,r)).$$

By left-invariance of the distance d it follows that the map π is a submetry. □

Lemma 2.6. *The projection norm is a norm.*

Proof. Positivity and homogeneity of the projection norm $\|\cdot\|_d$ follow immediately from positivity and homogeneity of the homogeneous distance d. For the triangle inequality, let $X, X' \in V_1$ and let $Y \in [\mathfrak{g},\mathfrak{g}]$ be the element given by the Baker-Campbell-Hausdorff formula such that

$$\exp(X)\exp(X') = \exp(X + X' + Y).$$

Lemma 2.5 gives the bound $\|X+X'\|_d \leq d(e, \exp(X+X'+Y))$. By the choice of Y, the left-invariance and triangle inequality of d conclude the claim:

$$d(e, \exp(X+X'+Y)) = d(e, \exp(X)\exp(X')) \leq \|X\|_d + \|X'\|_d. \quad \square$$

2.3. **Length structures and sub-Finsler Carnot groups.**

Definition 2.7. Let (X,d) be a metric space. Let Ω be the space of rectifiable curves of X and let $\ell_d\colon \Omega \to \mathbb{R}$ be the length functional. For points $x,y \in X$, denote by $\Omega(x,y) \subset \Omega$ the space of all rectifiable curves connecting the points x and y. The *length metric associated to the metric* d is the map $d_\ell\colon X \times X \to \mathbb{R} \cup \{\infty\}$ defined by

$$d_\ell(x,y) := \inf\{\ell_d(\gamma) : \gamma \in \Omega(x,y)\}.$$

If $d = d_\ell$, then the metric d is called a *length metric.*

See [BBI01, Section 2.3] for further information about length structures induced by metrics. For the purposes of this paper, only the special case of the length metric associated to a homogeneous distance will be relevant. Such a length metric always determines a sub-Finsler Carnot group, see Definition 2.9 and Lemma 5.1.

Definition 2.8. Let G be a stratified group. Denote by $L_g\colon G \to G$ the left-translation $L_g(h) = gh$. An absolutely continuous curve $\gamma\colon [0,T] \to G$ is a *horizontal curve* if $(L_{\gamma(t)^{-1}})_*\dot{\gamma}(t) \in V_1$ for almost every $t \in [0,T]$. The *control* of a horizontal curve γ is its left-trivialized derivative, i.e., the map

$$u\colon [0,T] \to V_1, \quad u(t) = (L_{\gamma(t)^{-1}})_*\dot{\gamma}(t).$$

Definition 2.9. A *sub-Finsler Carnot group* is a stratified group G equipped with a norm $\|\cdot\| : V_1 \to \mathbb{R}$. The norm induces a homogeneous distance d_{SF} via the length structure induced by $\|\cdot\|$ over horizontal curves.

More explicitly, for a horizontal curve $\gamma\colon [0,T] \to G$ with control $u\colon [0,T] \to V_1$, define the length

$$\ell_{\|\cdot\|}(\gamma) = \int_0^T \|u(t)\| \, dt.$$

For $g,h \in G$, let $\Omega(g,h)$ be the family of all horizontal curves connecting g and h. The sub-Finsler distance d_{SF} is defined as

$$d_{SF}(g,h) := \inf\{\ell_{\|\cdot\|}(\gamma) : \gamma \in \Omega(g,h)\}.$$

2.4. **Geodesics and blowdowns.**

Definition 2.10. Let G be a stratified group equipped with a homogeneous distance d. A *geodesic* is an isometric embedding $\gamma\colon [0,T] \to (G,d)$. That is, a geodesic satisfies

$$d(\gamma(t),\gamma(s)) = |t-s| \quad \forall t,s \in [0,T].$$

In the proof of Theorem 1.3 it will be convenient to consider also curves which preserve distances up to a constant factor. A curve $\gamma\colon [0,T] \to (G,d)$ for which there exists some constant $C>0$ such that

$$d(\gamma(t),\gamma(s)) = C\,|t-s| \quad \forall t,s \in [0,T]$$

will be called a *geodesic with speed C*.

Lemma 2.11. *Let $\gamma\colon [0,\infty) \to G$ be a horizontal curve with control $u\colon [0,\infty) \to V_1$ and let $h>0$ be a dilating factor. Then the dilated and reparametrized curve*

$$\gamma_h\colon [0,\infty) \to G, \quad \gamma_h(t) := \delta_{1/h}\gamma(ht),$$

has the control

$$u_h\colon [0,\infty) \to V_1, \quad u_h(t) := u(ht).$$

Proof. Since the dilations are group homomorphisms, the claim follows directly by the chain rule and Definition 2.8 of a control:

$$\frac{d}{dt}\gamma_h(t) = (\delta_{1/h})_*\frac{d}{dt}\gamma(ht) = (\delta_{1/h})_*(L_{\gamma(ht)})_*u(ht)h = (L_{\gamma_h(t)})_*u_h(t). \quad \square$$

Definition 2.12. Let $\gamma\colon [0,\infty) \to G$ be a horizontal curve. Suppose for some sequence of scales $h_k \to \infty$ the pointwise limit

$$\tilde\gamma\colon [0,\infty) \to G, \quad \tilde\gamma(t) = \lim_{k\to\infty}\gamma_{h_k}(t) = \lim_{k\to\infty}\delta_{1/h_k}\gamma(h_k t)$$

exists. Such a curve $\tilde\gamma$ is called a *blowdown* of the curve γ along the sequence of scales h_k.

Remark 2.13. If the curve γ is L-Lipschitz, then the curves γ_h are also all L-Lipschitz. Hence by Arzelà-Ascoli, up to taking a subsequence a blowdown along a sequence of scales will always exist.

Lemma 2.14. *Let $\gamma\colon [0,\infty) \to G$ be an infinite geodesic and let $\tilde\gamma = \lim_{k\to\infty}\gamma_{h_k}$ be any blowdown of the curve γ. Let u and $\tilde u$ be the controls of the curves γ and $\tilde\gamma$ respectively. Then*

(i) The curve $\tilde\gamma$ is an infinite geodesic.

(ii) Up to taking a subsequence, the dilated controls u_{h_k} converge to the control $\tilde u$ in $\mathrm{L}^2_{loc}([0,\infty);V_1)$.

Proof. **(i).** The curve $\tilde{\gamma}$ is a geodesic as the pointwise limit of geodesics.

(ii). The claim follows from [MPV18, Remark 3.13]. The point is that by weak compactness of closed balls in $\mathrm{L}^2_{\mathrm{loc}}([0,\infty); V_1)$ there exists a weakly convergent subsequence $u_h \rightharpoonup v$ to some $v \in \mathrm{L}^2_{\mathrm{loc}}([0,\infty); V_1)$. The definitions of control and weak compactness imply that v is a control for $\tilde{\gamma}$, so in particular $\tilde{u}(t) = v(t)$ for almost every t. Finally, the geodesic assumption implies that $\|u(t)\| \equiv 1 \equiv \|\tilde{u}(t)\|$, so the weak convergence is upgraded to strong convergence $u_h \to \tilde{u}$ in $\mathrm{L}^2_{\mathrm{loc}}([0,\infty); V_1)$.

□

Lemma 2.15. *Let G be a sub-Finsler Carnot group with a strictly convex norm and let $\gamma\colon [0,\infty) \to G$ be an infinite geodesic. Then there exists a sequence $h_k \to \infty$ such that the blowdown $\tilde{\gamma} = \lim_{k\to\infty} \gamma_{h_k}$ is affine.*

Proof. If the geodesic γ is itself affine, then the claim is immediate. Otherwise, consider the horizontal projection $\pi\circ\gamma\colon [0,\infty) \to G/[G,G]$. Since $G/[G,G]$ is a normed space with a strictly convex norm, and the geodesic γ is not affine, the projection curve $\pi \circ \gamma$ is not affine and hence not a geodesic. Then [HLD18, Theorem 1.4] states that there exists a Carnot subgroup $H < G$ of lower rank such that all blowdowns of the geodesic γ are contained in H.

Let the curve $\beta\colon [0,\infty) \to H < G$ be any blowdown. By Lemma 2.14(i), β is a geodesic. If β is also not affine, then iterating the above there exists a Carnot subgroup $K < H < G$ of even lower rank such that all blowdowns of β are in K. Blowdowns of the geodesic β are also blowdowns of the geodesic γ by a diagonal argument, so the claim follows by induction, since a Carnot subgroup of rank 1 is just a one parameter subgroup. □

2.5. **Subdifferentials.** In this section, let V be some fixed finite dimensional vector space and let $E\colon V \to \mathbb{R}$ be a convex continuous function. In the application in Section 5, the space V will be the horizontal layer $V_1 \subset \mathfrak{g}$, and the convex function of interest will be a squared norm $\frac{1}{2}\|\cdot\|^2$.

Definition 2.16. A linear function $a\colon V \to \mathbb{R}$ is a *subdifferential* of the function E at a point $Y \in V$ if

$$a(X - Y) \leq E(X) - E(Y) \quad \forall X \in V.$$

The collection of all subdifferentials a at a point $Y \in V$ is denoted $\partial E(Y) \subset V^*$.

The following lemmas are all simple properties of convex functions and their subdifferentials. They will be utilized in the proof of Theorem 1.1 in Section 5. The first lemma is the continuity of subdifferentials as a set valued map $V \to \mathcal{P}(V^*)$, $Y \mapsto \partial E(Y)$.

Lemma 2.17. *Let $Y_k \to Y \in V$ be a converging sequence and let $a_k \in \partial E(Y_k)$. Then there exists a converging subsequence $a_k \to a \in \partial E(Y)$.*

Proof. [Roc70, Theorem 24.7] shows (among other things) that since the set of points $\mathcal{S} = \{Y_i : i \in \mathbb{N}\} \cup \{Y\}$ is closed and bounded, the family of subdifferentials

$$\partial E(\mathcal{S}) := \bigcup_{X \in \mathcal{S}} \{a \in \partial E(X)\}$$

is also closed and bounded, and the subdifferentials $a \in \partial E(\mathcal{S})$ are equicontinuous. Hence the existence of a converging subsequence $a_k \to a$ to some linear map $a\colon V \to \mathbb{R}$ is a consequence of Arzelà-Ascoli.

The claim is concluded by [Roc70, Theorem 24.4], which shows that the convergences $Y_k \to Y$ and $a_k \to a$ with $a_k \in \partial E(Y_k)$ imply that $a \in \partial E(Y)$. □

The next two lemmas contain the maximality argument that will eventually be used to allow a blowdown argument to conclude that all infinite geodesics are lines.

Lemma 2.18. *Suppose the map E is strictly convex and let $a \in \partial E(Y)$. Then the point Y is the unique maximizer of the linear function a in the sublevel set $\{X \in V : E(X) \leq E(Y)\}$.*

Proof. For $E(X) \leq E(Y)$, the subdifferential condition $a \in \partial E(Y)$ gives the bound

$$a(X) - a(Y) = a(X - Y) \leq E(X) - E(Y) \leq 0, \tag{1}$$

proving maximality of Y.

Let $X \in V$ be another maximum. That is, suppose $a(X) = a(Y)$ and $E(X) \leq E(Y)$, so the bound (1) implies that necessarily $E(X) = E(Y)$. By linearity also $a((X+Y)/2) = a(Y)$, so the bound (1) further implies that also $E(X) = E((X+Y)/2) = E(Y)$. By strict convexity, this is only possible if $X = Y$, proving uniqueness of the maximizer Y. □

Lemma 2.19. *Let $\|\cdot\|$ be a norm on V and let $a\colon V \to \mathbb{R}$ be a subdifferential of the map $E = \frac{1}{2}\|\cdot\|^2$ at a point $Y \in V$. Then $|a(X)| \leq \|X\|\,\|Y\|$ for all $X \in V$, and $a(Y) = \|Y\|^2$.*

Proof. For any points $X, Y \in V$ and any $\epsilon > 0$, the subdifferential condition $a \in \partial E(Y)$ implies that

$$\begin{aligned}\epsilon a(X) &= a(Y + \epsilon X - Y) \leq E(Y + \epsilon X) - E(Y) \\ &\leq \frac{1}{2}\Big((\|Y\| + \epsilon \|X\|)^2 - \|Y\|^2\Big) = \epsilon \|X\| \|Y\| + \frac{1}{2}\epsilon^2 \|X\|^2 .\end{aligned}$$

Letting $\epsilon \to 0$ proves the bound $a(X) \leq \|X\| \|Y\|$. Repeating the same consideration for $-X$, gives the opposite bound $-a(X) \leq \|X\| \|Y\|$.

For the equality $a(Y) = \|Y\|^2$, let $\epsilon > 0$, and observe that a similar computation as before shows that

$$\begin{aligned}-\epsilon a(Y) &= a((1-\epsilon)Y - Y) \leq E((1-\epsilon)Y) - E(Y) \\ &= \frac{1}{2}((1-\epsilon)^2 - 1) \|Y\|^2 = (-\epsilon + \frac{1}{2}\epsilon^2) \|Y\|^2 .\end{aligned}$$

That is, $a(Y) \geq (1 - \frac{1}{2}\epsilon) \|Y\|^2$. The limit as $\epsilon \to 0$ and the previous upper bound prove the claim. □

3. Step 2 sub-Finsler Pontryagin Maximum Principle

In this section, the Pontryagin Maximum Principle will be rephrased in a convenient form for the purposes of Theorem 1.1. The precise statement to be proved is the following:

Proposition 3.1 (Step 2 sub-Finsler PMP.)**.** *Let G be a step 2 sub-Finsler Carnot group with an arbitrary norm $\|\cdot\| : V_1 \to \mathbb{R}$ and let $0 \leq T \leq \infty$. If $u\colon [0,T] \to V_1$ is the control of a geodesic, then there exists an absolutely continuous curve $a\colon [0,T] \to V_1^*$ and a skew-symmetric bilinear form $B\colon V_1 \times V_1 \to \mathbb{R}$ such that*

(i) At almost every $t \in [0,T]$, the curve a has the derivative

$$\frac{d}{dt}a(t)Y = B(u(t), Y) \quad \forall Y \in V_1.$$

(ii) At almost every $t \in [0,T]$, the linear map $a(t)\colon V_1 \to \mathbb{R}$ is a subdifferential of the squared norm $\frac{1}{2}\|\cdot\|^2$ at the point $u(t) \in V_1$.

Remark 3.2. In the sub-Riemannian case, the the squared norm $\frac{1}{2}\|\cdot\|^2$ is differentiable at every point, and the unique subdifferential is the inner product $a(t) = \langle u(t), \cdot\rangle$. The derivative condition (i) then gives the usual linear ODE of controls in implicit form

$$\langle \dot{u}(t), Y\rangle = \frac{d}{dt}\langle u(t), Y\rangle = B(u(t), Y) \quad \forall Y \in V_1.$$

3.1. **General statement of the PMP.** For the rest of Section 3, let G be a fixed sub-Finsler Carnot groupof step 2 with a generic norm $\|\cdot\| : V_1 \to \mathbb{R}$, and let $u\colon [0,T] \to V_1$ be the control of a geodesic $\gamma\colon [0,T] \to G$.

Consider first the finite time $T < \infty$ case. By Definition 2.9 of the sub-Finsler distance, the control u minimizes the length functional $\int_0^T \|u(t)\|\, dt$ among all controls defining curves with the same endpoints as γ. Since a geodesic has by definition constant speed, it follows that u is also a minimizer of the energy functional $\frac{1}{2}\int_0^T \|u(t)\|^2\, dt$.

Define the left-trivialized Hamiltonian

$$h\colon V_1 \times \mathbb{R} \times \mathfrak{g}^* \to \mathbb{R}, \quad h(u,\xi,\lambda) = \lambda(u) + \frac{1}{2}\xi \|u\|^2 . \tag{2}$$

By the Pontryagin Maximum Principle as presented in [AS04, Theorem 12.10], the control $u\colon [0,T] \to V_1$ can minimize the energy $\frac{1}{2}\int_0^T \|u(t)\|^2\, dt$ only if there is an almost everywhere non-zero absolutely continuous dual curve $t \mapsto (\xi, \lambda(t)) \in \mathbb{R} \times T^*_{\gamma(t)}G$ such that

$$\xi \leq 0 \tag{3}$$

$$\dot{\lambda} = \vec{h}_{u(t),\xi}(\lambda) \quad \text{a.e. } t \in [0,T], \tag{4}$$

$$h_{u(t),\xi}(\lambda(t)) \geq h_{v,\xi}(\lambda(t)) \quad \forall v \in V_1 \quad \text{a.e. } t \in [0,T]. \tag{5}$$

Here $h_{v,\xi}$ and $\vec{h}_{v,\xi}$, for $v \in V_1$, are the left-invariant Hamiltonian and the associated Hamiltonian vector field respectively.

More explicitly, $h_{v,\xi}\colon T^*G \to \mathbb{R}$ is the function defined from the left-trivialized Hamiltonian (2) in the natural way by

$$h_{v,\xi}(\lambda) = h(v,\xi,L_g^*\lambda), \quad \forall \lambda \in T_g^*G, \tag{6}$$

and $\vec{h}_{v,\xi}$ is the Hamiltonian vector field associated to the left-invariant Hamiltonian $h_{v,\xi}$ and the canonical symplectic form ω on the cotangent bundle T^*G by the duality

$$\omega(w, \vec{h}_{v,\xi}(\lambda)) = dh_{v,\xi}(w) \quad \forall w \in T_\lambda(T^*G). \tag{7}$$

Observe that if $(\xi, \lambda(t))$ is a dual curve satisfying the conditions (3)–(5) of the PMP, then also any scalar multiple $(C\xi, C\lambda(t))$ for any $C > 0$ satisfies the conditions (3)–(5) of the PMP. This observation allows the infinite time case $T = \infty$ to be handled as a limit of the finite time case. Namely, if $u\colon [0,\infty) \to V_1$ is the control of a geodesic, then all its finite restrictions $u|_{[0,k]} : [0,k] \to V_1$ for $k \in \mathbb{N}$ are also controls of geodesics, so by the above they have corresponding dual curves $t \mapsto (\xi_k, \lambda_k(t))$. By taking suitable rescalings of the (ξ_k, λ_k), there exists a non-zero limit $(\xi_\infty, \lambda_\infty)$, which then satisfies the conditions (3)–(5) of the PMP on the entire interval $[0,\infty)$.

3.2. **The normality/abnormality condition.** Condition (3) is a binary condition $\xi = 0$ or $\xi \neq 0$. The case $\xi = 0$ is the case of an abnormal control u, and may be ignored in the step 2 setting.

Indeed, suppose to the contrary that there does not exist a dual curve (ξ, λ) with $\xi \neq 0$, but only some dual curve with $\xi = 0$. In this case, the maximality condition (5) states that

$$\lambda(t)((L_{\gamma(t)})_* u(t)) \geq \lambda(t)((L_{\gamma(t)})_* v) \quad \forall v \in V_1 \quad \text{a.e. } t \in [0,T],$$

which is only possible when $(L_{\gamma(t)})_* V_1 \subset \ker \lambda(t)$. Moreover, second order optimality conditions from [AS04, Section 20] further imply that (possibly changing the dual curve λ) the Goh condition

$$\lambda(t)((L_{\gamma(t)})_*[v,w]) = 0, \quad \forall v, w \in V_1 \quad \text{a.e. } t \in [0,T]$$

is also satisfied. Since the group G has step 2, its Lie algebra has the decomposition $\mathfrak{g} = V_1 \oplus [V_1, V_1]$. The above would then imply that $\lambda(t) = 0$ almost everywhere, which would contradict the assumption that (ξ, λ) is almost everywhere non-zero.

Therefore without loss of generality it suffices to consider the normal case $\xi < 0$. By rescaling (ξ, λ) it further suffices to consider the case $\xi = -1$.

3.3. **The Hamiltonian ODE in left-trivialized coordinates.** The normal Hamiltonian vector field $\vec{h}_{u(t),-1}(\lambda)$ appearing in the ODE (4) is straight-forward to compute in left-trivialized coordinates on T^*G. The explicit expression will be given in Lemma 3.3.

Let $X_1, \ldots, X_r$ be a basis of V_1. Fix a basis $X_{r+1}, \ldots, X_n$ for $V_2 = [V_1, V_1]$ by choosing a maximal linearly independent subset of the Lie brackets $\{[X_i, X_j] : 1 \leq i < j \leq r\}$. By an abuse of notation, denote also by $X_1, \ldots, X_n$, the corresponding left-invariant frame of TG. Let $\theta_1, \ldots, \theta_n$ be the dual left-invariant frame of T^*G. Any covector $\lambda \in T_g^*G$ can be written in the frame as

$$\lambda = \sum_{i=1}^{n} a_i(\lambda) \theta_i(g).$$

The functions $a_i \colon T^*G \to \mathbb{R}$ together with coordinates $g \in G$ define left-trivialized coordinates on T^*G.

Lemma 3.3. *For any vector $v \in V_1$, the Hamiltonian vector field of the normal left-invariant Hamiltonian $h_{v,-1}$ has in left-trivialized coordinates the expression*

$$\vec{h}_{v,-1}(\lambda) = \sum_{1 \leq i \leq r} \lambda((L_{\pi(\lambda)})_*[v, X_i]) \partial a_i + (L_{\pi(\lambda)})_* v \in T_\lambda(T^*G).$$

Proof. Let

$$F(v,\lambda) := \sum_{1\le i\le r} \lambda((L_{\pi(\lambda)})_*[v,X_i])\partial a_i + (L_{\pi(\lambda)})_* v \tag{8}$$

be the expression on the right-hand side of the claimed formula for the normal Hamiltonian $\vec{h}_{v,-1}$. By the definition of a Hamiltonian vector field, it suffices to verify that the vector field $\lambda \mapsto F(v,\lambda)$ satisfies the duality (7), i.e., that

$$\omega(w, F(v,\lambda)) = dh_{v,-1}(w) \quad \forall w \in T_\lambda(T^*G). \tag{9}$$

The differential on the right-hand side of (9) is easily computed. The normal Hamiltonian $h_{v,-1}\colon T^*G \to \mathbb{R}$ defined by (6) is left-invariant and linear on fibers. Therefore in left-trivialized coordinates, the differential has the expression

$$dh_{v,-1} = \sum_{i=1}^{n} v_i da_i. \tag{10}$$

The expression for the symplectic form is more intricate. By definition the canonical symplectic form ω on the cotangent bundle T^*G is the differential of the tautological one-form $\sum_{i=1}^{n} a_i\theta_i$. That is, in left-trivialized coordinates, the symplectic form has the expression (see e.g. [ABB19, Section 4.2] for more details)

$$\omega = \sum_{i=1}^{n} da_i \wedge \theta_i + \sum_{i=1}^{n} a_i d\theta_i. \tag{11}$$

The differentials $d\theta_i$ can be evaluated along vector fields $X, Y \in \Gamma(TG)$ by the classical formula

$$d\theta_i(X,Y) = X\theta_i(Y) - Y\theta_i(X) - \theta_i([X,Y]).$$

For left-invariant vector fields this reduces to $d\theta_i(X,Y) = -\theta_i([X,Y])$.

Let W and Z be vector fields on the cotangent T^*G such that the projections $X := \pi_*W$ and $Y := \pi_*Z$ are left invariant. Writing the vector fields in left-trivialized coordinates as $W = \sum_{i=1}^{n} w_i\partial_{a_i} + X$ and $Z = \sum_{i=1}^{n} z_i\partial_{a_i} + Y$, the identity (11) gives the expression

$$\omega(W,Z) = \sum_{i=1}^{n} w_i\theta_i(Y) - \sum_{i=1}^{n} z_i\theta_i(X) - \sum_{i=1}^{n} a_i\theta_i([X,Y]). \tag{12}$$

The duality (9) will now be deduced by comparing the expressions (10) and (12). For an arbitrary vector $w \in T_\lambda(T^*G)$, let W be the left-invariant extension to a vector field on the cotangent T^*G. That is, the vector field W has a constant coefficient coordinate expression

$$W = \sum_{i=1}^{n} w_i \partial a_i + \sum_{i=1}^{n} x_i X_i. \tag{13}$$

Denote by $X := \pi_* W$ and $Y := \pi_* F(v, \cdot)$ the projection vector fields on the group G of the vector fields W and $\lambda \mapsto F(v, \lambda)$. Substituting the expressions (13) of W and (8) of $F(v, \cdot)$ into the expression (12) for $\omega(W, Z)$ with $Z = F(v, \lambda)$ gives the three sums

$$\sum_{i=1}^{n} w_i \theta_i((L_{\pi(\lambda)})_* v) = \sum_{i=1}^{n} w_i v_i,$$

$$\sum_{i=1}^{n} \lambda((L_{\pi(\lambda)})_* [v, X_i]) \theta_i(X) = \lambda([V, X]) \quad \text{and}$$

$$\sum_{i=1}^{n} a_i \theta_i([X, V]) = \lambda([X, V]).$$

By anti-commutativity of the Lie bracket, the last two sums cancel out, so $\omega(W, F(v, \lambda)) = \sum_{i=1}^{n} w_i v_i$. Since the expression (10) of the differential also gives $dh_{v,-1}(W) = \sum_{i=1}^{n} w_i v_i$, the vector field $\lambda \mapsto F(v, \lambda)$ indeed satisfies the duality (9). □

3.4. **Conclusion of the step 2 sub-Finsler PMP.**

Proof of Proposition 3.1. The curve $a\colon [0, T] \to V_1^*$ will be given by restricting the linear map

$$a(t) := (L_{\gamma(t)})^* \lambda(t) \colon \mathfrak{g} \to \mathbb{R} \tag{14}$$

to V_1. The skew-symmetric bilinear form $B\colon V_1 \times V_1 \to \mathbb{R}$ will be given by

$$B(X, Y) = a(t)[X, Y]. \tag{15}$$

In order to see that the above expressions are well defined and have the desired properties, express the curve $\lambda(t)$ in left-trivialized coordinates as

$$\lambda(t) = \sum_{i=1}^{n} a_i(t) \theta_i(\gamma(t)).$$

The curve $a(t)$ of (14) has the same coefficients as the curve $\lambda(t)$, i.e., it is given by $a(t) = \sum_{i=1}^{n} a_i(t) \theta_i(e)$. Using the explicit expression for the normal Hamiltonian vector field from Lemma 3.3, the normal

Hamiltonian ODE $\dot\lambda = \vec{h}_{u(t),-1}(\lambda)$ implies that the components of the curve a have the derivatives

$$\begin{cases} \dot a_i(t) = a(t)[u(t), X_i], & i \in \{1,\dots,r\} \\ \dot a_i(t) = 0, & i \in \{r+1,\dots,n\} \end{cases} \tag{16}$$

Observe that the vertical coefficients $a_{r+1},\dots,a_n$ are all constant, and that $\theta_i([X,Y]) = 0$ for $i = 1,\dots,r$. Therefore $a(t)[X,Y] = \sum_{i=r+1}^n a_i\theta_i([X,Y])$ is constant in t, so (15) defines a unique bilinear form B independent from t. Moreover, the non-trivial equations of the system (16) are then exactly

$$\dot a_i(t) = a(t)[u(t), X_i] = B(u(t), X_i), \quad i \in \{1,\dots,r\}.$$

Writing an arbitrary vector $Y \in V_1$ in the basis $X_1,\dots,X_r$ as $Y = y_1X_1 + \dots + y_rX_r$, the derivative condition 3.1(i) follows by linearity:

$$\frac{d}{dt}a(t)Y = \frac{d}{dt}\sum_{i=1}^n a_i(t)y_i = \sum_{i=1}^n B(u(t), X_i)y_i = B(u(t), Y).$$

The subdifferential condition 3.1(ii) for the linear functions $a(t)$ follows from rephrasing the maximality condition (5). Namely, expanding out the explicit expressions of the normal Hamiltonians $h_{u(t),-1}$ and $h_{v,-1}$ from (2) and reorganizing terms, the maximality condition (5) is equivalently stated as

$$a(t)v - a(t)u(t) \le \frac{1}{2}\|v\|^2 - \frac{1}{2}\|u(t)\|^2 \quad \forall v \in V_1 \quad \text{a.e. } t \in [0,T].$$

This is exactly Definition 2.16 for the linear function $a(t)$ being a subdifferential of the squared norm $\frac{1}{2}\|\cdot\|^2$ at the point $u(t) \in V_1$. □

4. Asymptotic behavior of controls

In this section, let $u\colon [0,\infty) \to V_1$ be a fixed control satisfying the PMP 3.1. Let $a\colon [0,\infty) \to V_1^*$ be the associated curve of subdifferentials and let $B\colon V_1 \times V_1 \to \mathbb{R}$ be the associated bilinear form.

Lemma 4.1. *For every vector $X \in V_1$,*

$$\lim_{T\to\infty} B\left(⨍_0^T u(t)\,dt, X\right) = 0.$$

Proof. Fix an arbitrary vector $X \in V_1$. Bilinearity of the map B implies that

$$B\left(⨍_0^T u(t)\,dt, X\right) = \frac{1}{T}\int_0^T B\left(u(t), X\right)\,dt. \tag{17}$$

Since the curve a is absolutely continuous, the derivative condition PMP 3.1(i) implies that

$$\int_0^T B(u(t), X) = \int_0^T \frac{d}{dt} a(t)X = a(T)X - a(0)X. \tag{18}$$

For almost every t, the linear map $a(t)$ is a subdifferential of the squared norm $\frac{1}{2}\|\cdot\|^2$ at the point $u(t)$. Since $\|u(t)\| \equiv 1$ is constant, continuity of the curve a and Lemma 2.19 imply the bound $|a(t)X| \leq \|X\|$ for every $t \in [0, T]$. The identities (17) and (18) then imply the desired conclusion that

$$\lim_{T\to\infty} \left| B\left(\fint_0^T u(t)\, dt, X \right) \right| \leq \lim_{T\to\infty} \frac{2}{T} \|X\| = 0. \qquad \square$$

Lemma 4.2. *Let $h_k \to \infty$ be a diverging sequence and let $u_{h_k}(t) = u(h_k t)$ be the corresponding dilated controls. If $u_{h_k} \to \tilde{u}$ in $\mathrm{L}^2_{loc}([0, \infty); V_1)$, then $\tilde{u}(t) \in \ker B$ for almost every $t \in [0, \infty)$.*

Proof. By the Lebesgue differentiation theorem it suffices to prove that $\fint_a^b \tilde{u}(t)\, dt \in \ker B$ for any $0 \leq a < b < \infty$.

Fix $0 \leq a < b < \infty$. By assumption $u_{h_k} \to \tilde{u}$ in $\mathrm{L}^2([a, b]; V_1)$, so there exists some error term $\epsilon \colon \mathbb{N} \to V_1$ with $\lim\limits_{k\to\infty} \epsilon(k) = 0$ such that

$$\fint_a^b \tilde{u}(t)\, dt = \fint_a^b u(h_k t)\, dt + \epsilon(k) = \fint_{ah_k}^{bh_k} u(t)\, dt + \epsilon(k). \tag{19}$$

The right-hand integral average can further be expressed as a difference of integral averages as

$$\fint_{ah_k}^{bh_k} u(t)\, dt = \frac{b}{b-a} \cdot \fint_0^{bh_k} u(t)\, dt - \frac{a}{b-a} \cdot \fint_0^{ah_k} u(t)\, dt. \tag{20}$$

Lemma 4.1 implies that for any $X \in V_1$

$$\lim_{k\to\infty} B\left(\fint_0^{bh_k} u(t)\, dt, X \right) = \lim_{k\to\infty} B\left(\fint_0^{ah_k} u(t)\, dt, X \right) = 0$$

Combining the identities (19) and (20) and using bilinearity of B then implies that $B\left(\fint_a^b \tilde{u}(t)\, dt, X \right) = 0$. Since the vector $X \in V_1$ was arbitrary, this proves the desired claim that $\fint_a^b \tilde{u}(t)\, dt \in \ker B$. $\square$

5. Affinity of infinite geodesics

5.1. **Sub-Finsler Carnot groups.** The proof of Theorem 1.1 will now be concluded. The key ingredients are the sub-Finsler PMP 3.1, the knowledge of asymptotic behavior of blowdown controls from Lemma 4.2, and the convex analysis arguments from Subsection 2.5.

Proof of Theorem 1.1. Let $\gamma\colon [0,\infty) \to G$ be an infinite geodesic and let $u\colon [0,\infty) \to V_1$ be its control. Let $a\colon [0,\infty) \to V_1^*$ be the curve of subdifferentials of the squared norm $\frac{1}{2}\|\cdot\|^2$ and let $B\colon V_1 \times V_1 \to \mathbb{R}$ be the skew-symmetric bilinear form given by the PMP 3.1.

By Lemma 2.15, there exists a sequence $h_k \to \infty$ such that the blowdown $\tilde{\gamma} = \lim_{k\to\infty} \delta_{1/h_k} \circ \gamma \circ \delta_{h_k}\colon [0,\infty) \to G$ is affine. By Lemma 2.14, taking a subsequence if necessary, the dilated controls $u_{h_k}(t) = u(h_k t)$ converge in $\mathrm{L}^2_{\mathrm{loc}}([0,\infty); V_1)$ to the control $\tilde{u}$ of the curve $\tilde{\gamma}$. Since the curve $\tilde{\gamma}$ is affine, the control $\tilde{u}$ is constant. That is, there exists a constant vector $Y \in V_1$, which for almost every $t \in [0,\infty)$ is the limit

$$Y = \tilde{u}(t) = \lim_{k\to\infty} u(h_k t). \tag{21}$$

By Lemma 4.2, $Y \in \ker B$, so the derivative condition PMP 3.1(i) implies that the curve $t \mapsto a(t)Y$ is constant $a(t)Y \equiv: C$.

Fix any $t \in [0,\infty)$ such that the limit (21) holds. By Lemma 2.17, up to taking a further subsequence, the subdifferentials $a(h_k t)$ of the squared norm $\frac{1}{2}\|\cdot\|^2$ at the points $u(h_k t)$ converge to a subdifferential $\tilde{a}\colon V_1 \to \mathbb{R}$ of the squared norm $\frac{1}{2}\|\cdot\|^2$ at the point Y. Moreover, since $a(t)Y \equiv C$ is constant, also the limit evaluates to $\tilde{a}Y = C$. Applying Lemma 2.19 for the subdifferential $\tilde{a}$ shows that $C = \tilde{a}Y = \|Y\|^2$. Similarly applying Lemma 2.19 for the subdifferential $a(t)$ shows that $a(t)u(t) = \|u(t)\|^2$. Since the curves γ and $\tilde{\gamma}$ are both geodesics, $\|u(t)\| = 1 = \|Y\|$, so combining all of the above gives the equality

$$a(t)Y = 1 = a(t)u(t). \tag{22}$$

The norm $\|\cdot\|$ is by assumption strictly convex and the map $x \mapsto \frac{1}{2}x^2$ is strictly increasing and convex on $[0,\infty)$, so also the squared norm $\frac{1}{2}\|\cdot\|^2$ is strictly convex. Then by Lemma 2.18, the point $u(t)$ is the unique maximizer for the linear map $a(t)$ in the corresponding sublevel set, so the equality (22) implies that $u(t) = Y$. Repeating the same argument at all the times t satisfying the limit (21), it follows that $u(t) = Y$ for almost every $t \in [0,\infty)$, so the geodesic γ is itself affine. □

5.2. **Arbitrary homogeneous distances.** The proof of Corollary 1.2 about infinite geodesics for arbitrary homogeneous norms follows from the sub-Finsler case by passing to the induced length metric. The relevant properties are captured in the next lemma.

Lemma 5.1. *Let (G,d) be a stratified group of step 2 equipped with a homogeneous distance d and let d_ℓ be the length metric of d. Then*

(i) (G,d_ℓ) is a sub-Finsler Carnot group.
(ii) All geodesics of (G,d) are also geodesics of (G,d_ℓ).

(iii) The projection norm of d is the sub-Finsler norm of d_ℓ.

Proof. **(i).** In [LD15, Theorem 1.1] sub-Finsler Carnot groups are characterized as the only geodesic metric spaces that are locally compact, isometrically homogeneous, and admit a dilation. Therefore it suffices to verify that the length metric associated to a homogeneous distance satisfies these properties.

The claims of isometric homogeneity and admitting a dilation follow directly from the corresponding properties of the metric d. Namely, since left-translations are isometries of the metric d, they preserve the length of curves, and hence are also isometries of the length metric d_ℓ. Similarly since dilations scale the length of curves linearly, they are dilations for the length metric d_ℓ.

Finiteness of the length metric d_ℓ follows from the stratification assumption: each element $g \in G$ can be written as a product of elements in $\exp(V_1)$ and the horizontal lines $t \mapsto \exp(tX)$ are all geodesics. Therefore concatenation of suitable horizontal line segments defines a finite length curve from the identity e to any desired point g. It follows that the length metric d_ℓ determines a well defined homogeneous distance on G, so by [LD17, Proposition 3.5] it induces the manifold topology of G. In particular (G, d_ℓ) is a boundedly compact length space, so it is a geodesic metric space (see [BBI01, Corollary 2.5.20]). Applying [LD15, Theorem 1.1] shows that (G, d_ℓ) is a sub-Finsler Carnot group.

(ii). The lengths of all rectifiable curves in the original metric d and its associated length metric d_ℓ always agree (see [BBI01, Proposition 2.3.12]). In particular, the claim that the geodesics of (G, d) are geodesics of (G, d_ℓ) follows.

(iii). The horizontal projection $\pi\colon (G, d) \to V_1$ is a submetry both for the sub-Finsler norm $\|\cdot\|_{SF}$ (by definition) and for the projection norm $\|\cdot\|_d$ (by Lemma 2.5). Hence the norms $\|\cdot\|_{SF}$ and $\|\cdot\|_d$ have exactly the same balls, so $\|\cdot\|_{SF} = \|\cdot\|_d$.

□

Proof of Corollary 1.2. Let (G, d) be a stratified group of step 2 equipped with a homogeneous distance d whose projection norm is strictly convex, and let $\gamma\colon [0, \infty) \to (G, d)$ be an infinite geodesic.

Let d_ℓ be the length-metric associated to d. By Lemma 5.1(i) and (ii), the curve γ is also a geodesic of $(G, \|\cdot\|)$, where $\|\cdot\| : V_1 \to \mathbb{R}$ is the sub-Finsler norm of the sub-Finsler metric d_ℓ. Moreover by Lemma 5.1(iii) the norm $\|\cdot\| = \|\cdot\|_d$ is by assumption strictly convex.

Consequently by Theorem 1.1, the geodesic γ is affine. □

The necessity of the strict convexity assumption is an immediate consequence of the classical case of normed spaces by the following simple lifting argument.

Proposition 5.2. *Let G be a stratified group of step 2 equipped with an arbitrary homogeneous distance d. If the projection norm of d is not strictly convex, then there exist an infinite geodesic $\gamma\colon \mathbb{R} \to G$ which is not affine.*

Proof. If the projection norm $\|\cdot\|_d : V_1 \to \mathbb{R}$ is not strictly convex, then there exists a non-linear geodesic $\beta\colon \mathbb{R} \to V_1$. For example, if the norm $\|X + cY\|_d$ is constant for $-\epsilon \le c \le \epsilon$, then the curve $\beta(t) = tX + \epsilon \sin(t)Y$ is an infinite geodesic.

By Lemma 2.5, the projection $\pi\colon (G, d) \to (V_1, \|\cdot\|)$ is a submetry, so the geodesic $\beta\colon \mathbb{R} \to V_1$ lifts to an infinite geodesic $\gamma\colon \mathbb{R} \to G$. Since the projection is a homomorphism and the geodesic β is not affine, neither is the geodesic γ. □

6. Affinity of isometric embeddings

Theorem 1.3 about isometric embeddings being affine follows from Corollary 1.2 by an abstraction of the argument of [BFS18, Theorem 4.1]. The key link between the metric and algebraic properties is the following simple lemma stating that the distance between two lines grows sublinearly if and only if the lines are parallel.

Lemma 6.1. *Let (G, d) be a stratified group with a homogeneous distance. Then for all points $g, h \in G$ and all vectors $X, Y \in V_1$*

$$d(g\exp(tX), h\exp(tY)) = o(t) \text{ as } t \to \infty \iff X = Y.$$

Proof. Consider dilations by $1/t$. Since dilations are homomorphisms, continuity of the distance gives the limit

$$\begin{aligned}\lim_{t\to\infty} \frac{d(g\exp(tX), h\exp(tY))}{t} &= \lim_{t\to\infty} d(\delta_{1/t}(g)\exp(X), \delta_{1/t}(h)\exp(Y)) \\ &= d(\exp(X), \exp(Y)).\end{aligned}$$

□

Proof of Theorem 1.3. Let $\varphi\colon (H, d_H) \hookrightarrow (G, d_G)$ be an isometric embedding. Since left-translations are isometries, it suffices to consider the case when the map φ preserves the identity element, and prove that such an isometric embedding is a homomorphism.

Consider an arbitrary point $h \in H$ and a horizontal vector $X \in V_1^H$. The horizontal line $t \mapsto h\exp(tX)$ is an infinite geodesic with speed $\|X\|_H$ through the point $h \in H$. The image of the line under the isometric embedding φ is an infinite geodesic in the group G through the point $\varphi(h)$ with exactly the same speed. By Corollary 1.2 all infinite geodesics in the group G are horizontal lines, so there exists some vector $Y \in V_1^G$ (a priori depending on the point h and the vector X) with $\|X\|_H = \|Y\|_G$ such that

$$\varphi(h\exp(tX)) = \varphi(h)\exp(tY) \quad \forall t \in \mathbb{R}.$$

Consider then the two parallel infinite geodesics $t \mapsto \exp(tX)$ and $t \mapsto h\exp(tX)$ with speed $\|X\|_H$. Repeating the previous consideration, since the map φ was assumed to preserve the identity, there exists another horizontal direction $Z \in V_1^G$ such that $\varphi(\exp(tX)) = \exp(tZ)$. By Lemma 6.1, the distance between the two lines in the group H grows sublinearly. Since the map φ is an isometric embedding, also the distance between the image lines in the group G grows sublinearly. Hence applying Lemma 6.1 in the converse direction implies that $Y = Z$. That is, the vector $Y \in V_1^G$ does not depend on the point $h \in H$, only on the vector $X \in V_1^H$.

The above shows that there is a well defined map $\varphi_*\colon V_1^H \to V_1^G$ such that $\varphi(h\exp(X)) = \varphi(h)\exp(\varphi_* X)$. In particular,

$$\varphi(h_1h_2) = \varphi(h_1)\varphi(h_2) \quad \forall h_1 \in H\ \forall h_2 \in \exp(V_1^H). \tag{23}$$

Since the group H is stratified, the subset $\exp(V_1^H)$ generates the entire group H. That is, any element $h \in H$ can be written as a finite product of elements in $\exp(V_1^H)$. Applying the identity (23) repeatedly using such decompositions shows that the map φ is a homomorphism. □

Acknowledgements. The author wishes to thank Enrico Le Donne and Yuri Sachkov for helpful discussions on the control viewpoint to infinite geodesics, which paved the way to the conclusion of the main proof. The author also wishes to thank Ville Kivioja for his help during the outset of the project in the study of the known results and their abstractions.

References

[ABB19] Andrei Agrachev, Davide Barilari, and Ugo Boscain, *A comprehensive introduction to sub-riemannian geometry*, 2019, p. 634.

[AS04] Andrei A. Agrachev and Yuri L. Sachkov, *Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences, vol. 87, Springer-Verlag, Berlin, 2004, Control Theory and Optimization, II. MR 2062547

[AS15] A. A. Ardentov and Yu. L. Sachkov, *Cut time in sub-Riemannian problem on Engel group*, ESAIM Control Optim. Calc. Var. **21** (2015), no. 4, 958–988. MR 3395751

[BBI01] Dmitri Burago, Yuri Burago, and Sergei Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001. MR 1835418

[BC18] Z. M. Balogh and A. Calogero, *Infinite geodesics of sub-Finsler distances in the Heisenberg groups*, ArXiv e-prints (2018), arXiv:1807.10369.

[BFS18] Zoltán M. Balogh, Katrin Fässler, and Hernando Sobrino, *Isometric embeddings into Heisenberg groups*, Geom. Dedicata **195** (2018), 163–192. MR 3820500

[HLD18] Eero Hakavuori and Enrico Le Donne, *Blowups and blowdowns of geodesics in Carnot groups*, ArXiv e-prints (2018).

[Kis03] Iwao Kishimoto, *Geodesics and isometries of Carnot groups*, J. Math. Kyoto Univ. **43** (2003), no. 3, 509–522. MR 2028665

[KLD17] Ville Kivioja and Enrico Le Donne, *Isometries of nilpotent metric groups*, J. Éc. polytech. Math. **4** (2017), 473–482. MR 3646026

[LD15] Enrico Le Donne, *A metric characterization of Carnot groups*, Proc. Amer. Math. Soc. **143** (2015), no. 2, 845–849. MR 3283670

[LD17] ———, *A primer on Carnot groups: homogenous groups, Carnot-Carathéodory spaces, and regularity of their isometries*, Anal. Geom. Metr. Spaces **5** (2017), no. 1, 116–137. MR 3742567

[LDO16] Enrico Le Donne and Alessandro Ottazzi, *Isometries of Carnot groups and sub-Finsler homogeneous manifolds*, J. Geom. Anal. **26** (2016), no. 1, 330–345. MR 3441517

[MPV18] Roberto Monti, Alessandro Pigati, and Davide Vittone, *On tangent cones to length minimizers in Carnot-Carathéodory spaces*, SIAM J. Control Optim. **56** (2018), no. 5, 3351–3369. MR 3857882

[Pan89] Pierre Pansu, *Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un*, Ann. of Math. (2) **129** (1989), no. 1, 1–60. MR 979599

[Roc70] R. Tyrrell Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970. MR 0274683

Email address: `eero.j.hakavuori@jyu.fi`

(Hakavuori) Department of Mathematics and Statistics, University of Jyväskylä, 40014 Jyväskylä, Finland