NILPOTENT GROUPS, ASYMPTOTIC CONES AND SUBFINSLER GEOMETRY

EMMANUEL BREUILLARD AND ENRICO LE DONNE

Abstract. We give an estimate of the speed of convergence of Cayley graphs of general finitely generated nilpotent groups towards their asymptotic cone. This yields an error term in the asymptotics for the volume of large balls, namely $|B(n)| = cn^d + O(n^{d-\alpha})$, where the exponent $\alpha > 0$ in the error term depends only on the nilpotency class. Conjecturally this holds for $\alpha = 1$. We relate this conjecture to other well-known conjectures in subRiemannian geometry and show that abnormal geodesics play an important role. We also study in some detail the geometry of the Heisenberg group (equipped with the Pansu metric) and show that our results are sharp for 2-step groups by giving an example for which the speed of convergence to the asymptotic cone is no faster than $n^{-\frac{1}{2}}$.

Contents

1. Introduction

Let Γ be a nilpotent group generated by a finite set of elements S. We will assume that $1 \in S$ and that $S = S^{-1}$. Following earlier results of Wolf [22], Bass [3], and Guivarc'h [12], Pansu [17] established in 1983 that the cardinality of the balls $S^n = S \cdot \ldots \cdot S$ of radius n for the word metric induced by S on the Cayley graph of G is asymptotic to $c_S n^d$, where $c_S > 0$ is a positive constant depending on S and d is an integer independent of S and given by the Bass-Guivarc'h formula:

$$d = \sum_{k \geqslant 1} k d_k, \tag{1.1}$$

where d_k is the rank of the Abelian group $C^k(\Gamma)/C^{k+1}(\Gamma)$, where $C^1(\Gamma) = \Gamma$, $C^{i+1}(\Gamma) = [\Gamma, C^i(\Gamma)]$ for $i \geqslant 1$, is the central descending series of Γ.

Date: December 7, 2012.
1991 *Mathematics Subject Classification.* 53C17,22E25.
Key words and phrases. subRiemannian geometry, polynomial growth, asymptotic cone, nilpotent groups, Gromov-Hausdorff convergence, abnormal geodesics.

Our main result is the following bound on the second term of the asymptotics. Let r be the nilpotency step of Γ, i.e., the smallest integer such that $C^{r+1}(\Gamma) = \{1\}$.

Theorem 1.1. *There is $\alpha_r > 0$ such that*

$$|S^n| = c_S n^d + O_S(n^{d-\alpha_r}), \text{ as } n \to \infty$$

and one can take $\alpha_r = \frac{2}{3r}$.

This improves Pansu's theorem, which gave no error term. When $r = 1$, i.e. in the Abelian case, the result holds with $\alpha_1 = 1$; this is well-known and easy to prove (see [9] for an explicit derivation and for more on the Abelian case). Stoll showed in [19] that if Γ is a 2-step nilpotent group, i.e. when $r = 2$, one can also take $\alpha_2 = 1$. Unfortunately his proof breaks down for groups of nilpotency step 3 and higher (see the remark following [19, Lemma 3.3] and the end of subsection 6.4). Nevertheless we have no example that may rule out the possibility that one could always take $\alpha_r = 1$ for every r. We discuss this conjecture further in Section 6 and its connection with other well-known open problems in subRiemannian geometry.

By way of contrast, the error term obtained in Theorem 1.1 admits no analogue in general for nondiscrete groups of polynomial growth, where the speed can sometimes be arbitrarily slow. A simple example is given by the group $\mathbb{R}^2 \rtimes_\theta \mathbb{Z}$, where $\mathbb{Z}$ acts by a rotation R_θ whose angle θ is very well approximable by rationals multiples of 2π but not in $\pi\mathbb{Q}$. In this example, the error term can be shown to be arbitrarily bad if θ is chosen carefully (see [6, §8.1]).

In [17], Pansu gave a beautiful description of the asymptotic cone of an arbitrary finitely-generated torsion-free nilpotent group Γ. Let us briefly recall his results. The Malcev closure G of Γ is a simply-connected nilpotent Lie group in which Γ embeds as a co-compact discrete subgroup (see [18]). On every simply-connected nilpotent Lie group G, one can modify the Lie product structure in a natural (yet nonunique) way and obtain the so-called *graded group* associated to G. Endowed with this new Lie product $(G, *)$ is a graded nilpotent Lie group (also often called *Carnot* group when endowed with a subRiemannian or subFinsler metric) in the sense that it admits a one-parameter subgroup of $\mathbb{R}$-diagonalisable automorphisms. We refer the reader to Section 2 for this construction.

In his work on groups of polynomial growth [10], Gromov observed that if we renormalize the word metric ρ_S in the Cayley graph of Γ by a factor $\frac{1}{n}$, then there is a subsequence such that the balls of any given radius converge in the Gromov-Hausdorff topology (see [11, chapter 3], [10]). Pansu [17] showed that the entire sequence converges and that the limit is a certain metric space, the asymptotic cone of Γ, which can be described as follows. It is the graded Malcev closure $(G, *)$ of Γ endowed with a certain subFinsler $*$-left-invariant metric d_∞ (the Pansu limit metric), which is induced in the usual way from a certain polyhedral norm on the horizontal subspace of the Lie algebra (which is a transverse to the commutator subalgebra). The unit ball of this norm, which we call hereafter the Pansu limit norm, is defined as the convex hull of the projections of the generating set S to the horizontal subspace of $(G, *)$. The coefficient c_S in the main term of the asymptotics in Theorem 1.1 is the Lebesgue measure of the unit ball of $(G, *)$, where the measure is normalized so that the lattice Γ has co-volume 1. In Section 2, we describe Pansu's limit norm and the associated subFinsler metric (also known as Carnot-Carathéodory metric).

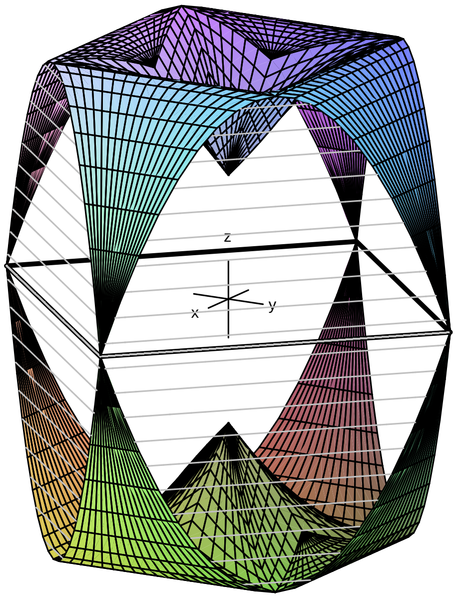


FIGURE 1. The unit ball for the Pansu limit metric d_3 of the Cayley graph of the discrete Heisenberg group $H_3(\mathbb{Z})$ with standard generators (see §5.5 for an explicit formula for d_3).

The Gromov-Hausdorff convergence implies that as metric spaces, all the asymptotic cones of Γ are all isometric to $(G, *, d_\infty)$. Thus $(G, *, d_\infty)$ will be referred to as *the asymptotic cone* of (Γ, ρ_S). In this paper we give an estimate of the speed of convergence towards the asymptotic cone in the Gromov-Hausdorff metric d_{GH}. Let ρ_S be the word metric induced by S on Γ, i.e. $\rho_S(x, y) = \inf\{n \in \mathbb{N}; x^{-1}y \in S^n\}$. For $\gamma \in \Gamma$, we set $|\gamma|_S = \rho_S(\mathrm{id}, \gamma)$ and $|\gamma|_\infty = d_\infty(\mathrm{id}, \gamma)$ and $B_S(n)$ the Cayley ball centered at id and radius $n \in \mathbb{N}$ and $B_\infty(R)$ the ball centered at the origin for d_∞ and radius $R > 0$.

Theorem 1.2. *Let Γ be a torsion free nilpotent group with nilpotency class r generated by a finite subset S (with $S = S^{-1}$, $\mathrm{id} \in S$). Let $B_S(n)$ be the ball of radius n in Γ centered at the identity for the word distance ρ_S. Let $B_\infty(1)$ the unit ball at the origin for the Pansu limit metric d_∞ on the graded Malcev closure $(G, *)$ of Γ. Then the metric spaces $X_n := (B_S(n), \frac{1}{n}\rho_S)$ and $X_\infty := (B_\infty(1), d_\infty)$ satisfy*

$$d_{GH}(X_n, X_\infty) = O(n^{-\alpha_r}), \qquad \text{as } n \to \infty,$$

for some $\alpha_r > 0$. We can take $\alpha_r := \frac{2}{3r}$ if $r > 2$, $\alpha_2 = \frac{1}{2}$ and $\alpha_1 = 1$.

Theorems 1.1 and 1.2 will follow from the following distance comparison theorem, which compares the word metric on Γ with the Pansu limit metric on the asymptotic cone of Γ.

Theorem 1.3. *(Word metric versus asymptotic metric comparison) For each $r \in \mathbb{N}$, there is $\alpha_r > 0$ such that for every finitely-generated torsion-free nilpotent group Γ of nilpotency step r the following property holds. Let S be a finite generating set for Γ with $S = S^{-1}$ and $1 \in S$. Denote by $|\cdot|_S = \rho_S(\mathrm{id}, \cdot)$ the word distance from the identity and by $|\cdot|_\infty = d_\infty(\mathrm{id}, \cdot)$ the distance from the origin in the Pansu* *limit metric, as explained above. Then, as γ tends to ∞ in Γ, we have*

$$||\gamma|_S - |\gamma|_\infty| = O_S(|\gamma|_\infty^{1-\alpha_r}).$$

Moreover, one can take $\alpha_r = \frac{2}{3r}$ if $r > 2$, $\alpha_2 = \frac{1}{2}$ and $\alpha_1 = 1$.

We refer the reader to Sections 5 and 6 for a discussion on the sharpness of this result and some related open questions. Basically it is sharp for step 2 and we conjecture that the $\frac{1}{2}$ exponent holds in general, at least for stratified Lie groups. For step 2 groups, sharpness is shown by using the same example from [6, §8.2], which was used there to disprove a conjecture of Burago and Margulis on asymptotic metrics. In particular we have:

Proposition 1.4. *On the 2-step group $G := \mathbb{Z}\times H_3(\mathbb{Z})$ (where $H_3(\mathbb{Z})$ is the discrete Heisenberg group), one can find two left-invariant word metrics ρ_1 and ρ_2 such that (G, ρ_1) and (G, ρ_2) have isometric asymptotic cones, but are not $(1, C)$-quasi-isometric for any $C > 0$. Moreover, the convergence of $(B_{\rho_1}(n), \frac{1}{n}\rho_1)$ to the asymptotic cone is no faster than $n^{-\frac{1}{2}}$ in the Gromov-Hausdorff metric (i.e., exponent $\alpha_2 = \frac{1}{2}$ is sharp in Theorem 1.2).*

The sharpness of the exponent $\frac{1}{2}$ for 2-step groups is a major difference between the nilpotent case and the abelian case (for which the power saving is 1). The reason for it is also very interesting, because it is related to the existence of *abnormal geodesics* in certain Carnot groups. Abnormal geodesics do not exist in classical Riemannian (or Finsler) geometry and are a distinctive feature of the underlying subRiemannian geometry of nilpotent groups (see Sections 5, 6 and Figure 5).

Our treatment of Theorem 1.3 here differs slightly from Pansu's original work (which gave the convergence with no error term), in particular in our use a single underlying manifold (the Lie algebra of G) on which we consider several Lie group structures. However our proof of Theorem 1.3 is ultimately an effectivization of Pansu's argument, in which we have to replace the compactness arguments used in several places by effective arguments. In fact a mere effectivization of Pansu's proof yields an exponent $\alpha_r = \frac{1}{2r}$. In order to obtain $\alpha_r = \frac{2}{3r}$, we make use of Stoll's result in [19]. Apart from some classical facts about subRiemannian geometry and lattices in nilpotent Lie groups, for which we refer the reader to the books by Montgomery [16] and Raghunathan [18], our treatment is self-contained and the reader need not have read Pansu's paper as a prerequisite.

This effectivization and the error term in the distance comparison theorem above are intimately related to the regularity of the distance function $g \mapsto |g|_\infty$ and the singularities of the sphere $\{g; |g|_\infty = 1\}$. In general, even for subRiemannian metrics on 2-step groups, the sphere is not smooth and singularities of polynomial type can occur (e.g. see Figure 5).

For us, the key to Theorem 1.3 will be the following lemma about the unit sphere and geodesics in the asymptotic cone of Γ.

Lemma 1.5 (Almost geodesic piecewise linear paths)**.** *Let (G, d) be a Carnot group of nilpotency step r. Let d_e be a left-invariant Riemannian metric on G. Let $\partial B(1) := \{x \in G; d(\mathrm{id}, x) = 1\}$ be the sphere of radius 1 for the Carnot metric. Then for every large enough integer n, and every $x \in \partial B(1)$, there exists a continuous piecewise linear path $\{\xi(t)\}_{t\in[0,1]}$ starting at id, such that $\xi(t)$ is horizontal* *linear on each interval* $[\frac{k}{n}, \frac{k+1}{n}]$*, with derivative of norm at most* 1 *almost everywhere, and is such that* $d_e(x, \xi(1)) = O_G(\frac{1}{n})$ *as* $n \to +\infty$*.*

In Stoll's proof of the optimal error term for 2-step groups ([19]), a key lemma consisted in establishing that in 2-step groups endowed with a subFinsler metric induced from a polyhedral norm, one can always connect any point on the unit sphere to the origin by a piecewise linear geodesic path with a uniformly bounded number of breakpoints. So Lemma 1.5, which is in fact not hard to prove, can be seen as an approximation to Stoll's lemma, which says that, allowing n breakpoints, one can find an almost geodesic piecewise linear path (i.e., of total length at most $1 + O(n^{-\frac{1}{r}})$) between any point on the unit sphere and the origin.

This paper is organized as follows. In Section 2, we set some notation and describe Pansu's construction of the asymptotic cone. In Section 3 we prove the main technical lemmas about approximating continuous and discrete geodesic paths including Lemma 1.5. In Section 4, we complete the proofs of Theorem 1.3, 1.2, and 1.1, in this order. In Section 5, we discuss the sharpness of our results, prove Proposition 1.4 and discuss in some detail the underlying geometry of the Heisenberg group equiped with the Pansu limit metric. Finally in Section 6, we discuss some further consequences of our theorems and state some open questions.

Acknowledgments E.B. is grateful to the ERC for its support through grant GADA-208091. Both authors would like to thank the MSRI, Berkeley, for perfect working conditions during the Quantitative Geometry program, when part of this research was conducted. E.B. also thanks Fudan University, Shanghai, for its hospitality.

2. SubFinsler metrics and dilations on nilpotent Lie groups

2.1. **Left-invariant SubFinsler metrics on Lie groups.** Let G be a Lie group. Denote by $\mathfrak{g}$ the Lie algebra of G considered as the tangent space at the identity element. We will often identify the Lie algebra with the space of left-invariant vector fields on G.

Suppose a linear subspace $V_1 \subset \mathfrak{g}$ and a norm $\|\cdot\|$ on V_1 are given. Then V_1 induces a left-invariant subbundle Δ of the tangent bundle of G. Namely, a vector v at a point $p \in G$ is an element of Δ if $(L_p)^* v \in V_1$, where $L_p : G \to G$ is the left multiplication by p and F^* denotes the pull back by a diffeomorphism F. For such a v, we set $\|v\| := \|(L_p)^* v\|$. Such a Δ is called a *horizontal distribution.* The triple $(G, \Delta, \|\cdot\|)$ is an example of a subFinsler manifold, cf. [16, 14].

Any subFinsler manifold has an associated distance function, which is called a subFinsler metric (also known as Carnot-Carathéodory-Finsler metric) and it is defined as follows. One says that an absolutely continuous curve $\gamma : [a, b] \to G$, with $a, b \in \mathbb{R}$, is *horizontal* (with respect to Δ) if the derivative $\dot\gamma(t)$ belongs to Δ, for almost all $t \in [a, b]$. For such a curve, the value $\|\dot\gamma(t)\|$ is almost everywhere defined. Hence each horizontal curve $\gamma : [a, b] \to G$ has an associated length defined as

$$L(\gamma) := \int_a^b \|\dot\gamma(t)\| \, \mathrm{d}t.$$

Then one defines the *subFinsler distance* between two points $p, q \in G$ as

$$d_{sF}(p, q) := \inf\{L(\gamma) \mid \gamma \text{ horizontal, from } p \text{ to } q\}. \tag{2.1}$$

Remark. It is possible to show that one may restrict to piecewise linear horizontal paths in the definition of the length, i.e., to those paths for which $\dot\gamma(t)$ takes only finitely many values. An explicit proof of such a fact can be found in [15, Theorem 1.2].

If the Lie algebra generated by V_1 is the whole of $\mathfrak{g}$, then the function $d_{sF}(\cdot,\cdot)$ is finite and in fact defines a geodesic distance that induces the manifold topology on G. This is a particular case of a theorem by Chow [16, Chapter 2].

The subFinsler metric d_{sF} is said to be subRiemannian if the norm $\|\cdot\|$ is an Euclidean norm. Moreover d_{sF} is said to be Finsler if $V_1 = \mathfrak{g}$.

SubFinsler metrics are *geodesic* distances. This means that every two points can be connected by a geodesic path, i.e., an isometric embedding of an interval. For left-invariant geodesic metrics on Lie groups, it turns out that the converse also holds, and this is a special case of a more general result due to Berestovskiĭ, see [4, Theorem 2, page 892]. Namely:

Theorem 2.2 (Berestovskiĭ)**.** *Let d be a geodesic left-invariant metric on a Lie group that induces the manifold topology. Then there exist a left-invariant subbundle Δ and a left-invariant norm $\|\cdot\|$ on Δ such that the distance d is the subFinsler metric associated to Δ and $\|\cdot\|$.*

Any two left-invariant Finsler metrics are globally bi-Lipschitz equivalent on G. Things are different if we compare a Finsler and a subFinsler metric. On the large scale they remain bi-Lipschitz, because, as it is well-known and easy to check, any two left-invariant geodesic distances on any given locally compact group are quasi-isometric (see Proposition 2.3 below for a proof). But on small scale, they can be drastically different. Clearly Riemannian (or Finsler) metrics on G are dominated by subFinsler metrics up to multiplicative constants. The opposite inequality typically does not hold, but we always have a lower bound of polynomial type as we now describe.

Assume that V_1 together with all brackets of order at most r in the elements of V_1 span $\mathfrak{g}$ linearly. Let d_e be any Riemanian metric on G. The Ball-Box Theorem (e.g. [16, Theorem 2.4.2]) implies that, for any compact set $K \subset G$, there exists a constant C such that

$$\frac{1}{C}(d_{sF}(\mathrm{id},x))^r \leqslant d_e(\mathrm{id},x) \leqslant C d_{sF}(\mathrm{id},x), \qquad \forall x \in K.$$

In particular, we have

$$d_{sF}(\mathrm{id},x) = O(d_e(\mathrm{id},x)^{\frac{1}{r}}), \text{ as } x \to 0. \tag{2.2}$$

We now record the well-known fact that left-invariant quasi-geodesic metrics on a group G are always quasi-isometric. Recall that a metric space (X,d) is said to be *quasi-geodesic* if there exist constants $C>0$ and $L>1$ such that every two points in X can be join with a (L,C)-quasi-arc. In other words, for all $x,x' \in X$, there exist $k \in \mathbb{N}$ and $x_0,x_1,\dots,x_k \in X$ such that $x_0 = x$, $x_k = x'$, $d(x_{i-1},x_i) \leqslant C$, for $i = 1,\dots,k$, and $\sum_{i=1}^k d(x_{i-1},x_i) \leqslant Ld(x,x') + C$.

Proposition 2.3. *Let d and d' be two quasi-geodesic left-invariant distances on a locally compact group. Assume that d and d' are locally bounded (i.e. bounded on compact sets) and proper (i.e. $g \mapsto d(\mathrm{id},g)$ is a proper map). Then there exist constants $c>0$ and $L>1$ such that $L^{-1}d - c < d' < Ld + c$.*

Proof. Since d is quasi-geodesic, there are two constants $C_1 > 0$ and $L_1 \geqslant 1$ with the following property. Given any group element g, we may find $g_1,...,g_n$ such that $d(\mathrm{id},g_i) \leqslant C_1$ for all i and $g = g_1 \cdot \ldots \cdot g_n$, while $\sum_1^n d(\mathrm{id}, g_i) \leqslant L_1 d(\mathrm{id}, g) + C_1$. Grouping some g_i's together if necessary, we may assume that $C_1/2 \leqslant d(\mathrm{id}, g_i)$. Hence $n \leqslant \frac{2L_1}{C_1} d(\mathrm{id}, g) + 2$. But then by our assumptions of local boundedness and properness, all g_i's lie in a fixed compact subset of the group and thus $d'(\mathrm{id}, g_i)$ is uniformly bounded say by some constant $\tilde{C} > 0$. Then $d'(\mathrm{id}, g) \leqslant \sum_1^n d'(\mathrm{id}, g_i) \leqslant \tilde{C} n \leqslant L d(\mathrm{id}, g) + c$ for $L = \frac{2\tilde{C} L_1}{C_1}$ and $c = 2\tilde{C}$. The proposition follows by exchanging the roles of d and d' and using the assumed left invariance. $\square$

The above applies in particular to Finsler and subFinsler left-invariant metrics on Lie groups.

2.4. **Stratified Lie algebras and Carnot groups.** A special role in subFinsler geometry is played by those Lie groups whose Lie algebra admits a stratification. We say that a Lie algebra $\mathfrak{g}$ *admits an s-step stratification*, with $s \in \mathbb{N}$, if there exist vector subspaces $V_1, \cdots, V_s \subseteq \mathfrak{g}$, such that

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_s,$$

$[V_j, V_1] = V_{j+1}$, for $1 \leqslant j \leqslant s-1$, with $V_s \neq \{0\}$, and $[V_s, V_1] = \{0\}$.

Here we are using the following notation. Given two vector subspaces V, W of a Lie algebra $\mathfrak{g}$, the set $[V, W]$ is the vector subspace generated by the commutators of the form $[v, w]$ with $v \in V$ and $w \in W$. The vector subspaces $V_1, \cdots, V_s$ in the definition of a stratification of an algebra are called the *strata* of the stratified Lie algebra $\mathfrak{g}$.

Note that every stratified Lie algebra is nilpotent, with nilpotency class (or step) s. Observe further that the commutator subalgebra $[\mathfrak{g}, \mathfrak{g}]$ coincides with $V_2 \oplus \cdots \oplus V_s$ and thus that V_1 is in bijection with the abelianization $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ via the projection map onto the V_1 component. Similarly, for every i, the i-th term $\mathfrak{g}^{(i)}$ in the central descending series of $\mathfrak{g}$, coincides with $V_i \oplus \ldots \oplus V_s$.

Note that the first stratum V_1 completely determines the other strata. We also remark that every linear subspace in direct sum with $[\mathfrak{g}, \mathfrak{g}]$ generates $\mathfrak{g}$, but that it may not always give rise to a stratification of $\mathfrak{g}$ (exercise). One can show that any two stratifications on a given Lie algebra $\mathfrak{g}$ are isomorphic in the sense that there is a linear automorphism of $\mathfrak{g}$ exchanging them (take the change of coordinates between the two direct sum decompositions).

Let $\mathbf{G}$ be a stratified group, i.e., a connected, simply-connected Lie group with stratified Lie algebra $\mathfrak{g}$. In such a group the first stratum V_1 generates $\mathfrak{g}$. Hence, if V_1 is equipped with a norm, we consider the subFinsler distance $d_{sF}(\cdot, \cdot)$. The metric space $(\mathbf{G}, d_{sF})$ is called *Carnot group*.

One peculiarity of Carnot groups is that they admits dilations in the following sense. For each $\lambda \in \mathbb{R}$, the *algebra-dilation* $\delta_\lambda : \mathfrak{g} \to \mathfrak{g}$ is defined linearly by imposing $\delta_\lambda(X) := \lambda^j X$, for every $X \in V_j$ and every $j = 1, \ldots, s$. If $\lambda \neq 0$, then δ_λ is an automorphism of $\mathfrak{g}$. Since the Lie group $\mathbf{G}$ is simply-connected, the dilation induces a unique automorphism on the group, which we still denote by δ_λ and call the *group-dilation*.

Since $\mathbf{G}$ is nilpotent and simply-connected, the exponential map $\exp : \mathfrak{g} \to \mathbf{G}$ is a diffeomorphism. Thus, group-dilations δ_λ can be equivalently defined as $\delta_\lambda(p) = \exp \circ \delta_\lambda \circ \exp^{-1}(p)$, for all $\lambda \in \mathbb{R}$ and $p \in \mathbf{G}$. Note that $\lambda \mapsto \delta_\lambda$, as a map from the positive reals to the group of automorphisms of $\mathbf{G}$ yields a one-parameter group. Since δ_λ stretches vectors in the horizontal distribution by λ, then $\mathrm{Length}(\delta_\lambda \circ \gamma) = \lambda \mathrm{Length}(\gamma)$, for each horizontal curve γ. Hence, δ_λ is a dilation by a factor of λ with respect to the subFinsler distance, i.e.,

$$d_{sF}(\delta_\lambda(p), \delta_\lambda(q)) = \lambda d_{sF}(p,q), \quad \forall p,q \in \mathbf{G}. \tag{2.3}$$

To every nilpotent Lie algebra $\mathfrak{g}$, one can associate in a canonical way a certain stratified Lie algebra, called the *graded Lie algebra* associated to $\mathfrak{g}$. We give now the abstract definition of the graded algebra. Later, we shall give a more concrete (but noncanonical) construction.

Definition 2.5 (Graded algebra)**.** Let $\mathfrak{g}$ be a Lie algebra that is nilpotent of step s. Let $\mathfrak{g}^{(1)} := \mathfrak{g}$ and $\mathfrak{g}^{(i+1)} := [\mathfrak{g}, \mathfrak{g}^{(i)}]$ be the descending central series of $\mathfrak{g}$. The *graded algebra associated to* $\mathfrak{g}$ (or simply the graded algebra of $\mathfrak{g}$) is the Lie algebra $\mathfrak{g}_\infty$ given by the direct-sum decomposition

$$\mathfrak{g}_\infty := \bigoplus_{i=1}^{s} \mathfrak{g}^{(i)}/\mathfrak{g}^{(i+1)},$$

endowed with the unique Lie bracket $[\cdot,\cdot]_\infty$ that has the property that, if $x \in \mathfrak{g}^{(i)}$ and $y \in \mathfrak{g}^{(j)}$, the bracket is defined, modulo $\mathfrak{g}^{(i+j+1)}$, as

$$[\bar{x}, \bar{y}]_\infty = \overline{[x,y]}.$$

Notice that the graded algebra associated to an algebra is a stratified algebra. Hence there exists a unique connected, simply-connected Lie group G_∞ whose Lie algebra is $\mathfrak{g}_\infty$. We refer to such a group G_∞ as the *graded group* of $\mathfrak{g}$.

There is a natural way to identify the underlying vector spaces of $\mathfrak{g}$ and its graded algebra $\mathfrak{g}_\infty$. This identification depends on the choice of s subspaces of the Lie algebra $\mathfrak{g}$. Namely, for each $j = 1, ..., s$, one chooses a subspace $V_j \subset \mathfrak{g}$ such that

$$\mathfrak{g}^{(j)} = \mathfrak{g}^{(j+1)} \oplus V_j. \tag{2.4}$$

In particular, the subspace V_1 is a complementary of the commutator subalgebra, i.e.,

$$\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] \oplus V_1. \tag{2.5}$$

We have that the projection whose kernel is $\mathfrak{g}^{(i+1)}$ gives a linear isomorphism between V_i and $\mathfrak{g}^{(i)}/\mathfrak{g}^{(i+1)}$ and this induces a linear isomorphisms of vector spaces:

$$\mathfrak{g} \simeq \mathfrak{g}_\infty. \tag{2.6}$$

Under this identification, which depends on the choice of the subspaces V_j's, we can pull back the dilations δ_λ from $\mathfrak{g}_\infty$ onto $\mathfrak{g}$ and thus define the map $\delta_\lambda : \mathfrak{g} \to \mathfrak{g}$ to be the linear map such that $\delta_\lambda(X) := \lambda^j X$, for every $X \in V_j$ and every $j = 1, \dots, s$. Under this identification, we can also pull back the Lie algebra structure from $\mathfrak{g}_\infty$ onto $\mathfrak{g}$ and thus define a new Lie bracket $[\cdot,\cdot]_\infty$ on the underlying vector space $\mathfrak{g}$.

Let us remark that each linear map δ_λ is not necessarily an algebra homomorphism of $\mathfrak{g}$. However, given the choice of the V_i's, the δ_λ's will be automorphisms for the new Lie algebra structure on $\mathfrak{g}$ given by $[\cdot,\cdot]_\infty$.

Lemma 2.6. *The new Lie bracket on* $\mathfrak{g}$ *obtained from the identification* (2.6) *satisfies the formula*

$$[X,Y]_\infty = \lim_{\lambda \to +\infty} \delta_\lambda^{-1}[\delta_\lambda X, \delta_\lambda Y]. \tag{2.7}$$

Consequently $[\delta_\lambda X, \delta_\lambda Y]_\infty = \delta_\lambda [X,Y]_\infty$ *for all* $X, Y \in \mathfrak{g}$.

Proof. Indeed, since the V_j's are a direct decomposition of $\mathfrak{g}$, it suffices to show (2.7) for $X \in V_i$ and $Y \in V_j$, for some i, j. In this case, we have that

$$[X,Y] = Z_{i+j} + Z_{i+j+1} + \ldots + Z_s,$$

for some vectors $Z_k \in V_k$. Hence

$$\begin{aligned} \delta_\lambda^{-1}[\delta_\lambda X, \delta_\lambda Y] &= \delta_\lambda^{-1}[\lambda^i X, \lambda^j Y] \\ &= \lambda^{i+j}\delta_\lambda^{-1}(Z_{i+j} + Z_{i+j+1} + \ldots + Z_s) \\ &= Z_{i+j} + \lambda^{-1} Z_{i+j+1} + \ldots + \lambda^{i+j-s} Z_s, \end{aligned}$$

which goes to Z_{i+j}, as $\lambda \to \infty$. The proof of (2.7) is concluded by observing that Z_{i+j} is a vector that represent $[X,Y]$ modulo $\mathfrak{g}^{(i+j+1)}$. □

Finally note that identifying G with $\mathfrak{g}$ and G_∞ with $\mathfrak{g}_\infty$ via the exponential map, the identification (2.6) also yields an identification at the group level:

$$G \simeq G_\infty, \tag{2.8}$$

which is a diffeomorphism, but *not* a group isomorphism. Under this identification, the underlying manifold of the Lie group G can be given a new Lie group structure by pulling back the Lie group structure of G_∞. This Lie structure is also the one induced by the new Lie bracket $[X,Y]_\infty$ on $\mathfrak{g}$. In order to distinguish it from the original Lie product $x \cdot y$, we will denote the new Lie product on G obtained in this way by $x * y$.

2.7. **The Campbell-Baker-Hausdorff formula.** Suppose G is simply connected nilpotent Lie group with Lie algebra $\mathfrak{g}$ and we have chosen subspaces V_i's as in (2.4). Choose a basis $(e_1, \ldots, e_n)$ of $\mathfrak{g}$ which is adapted to the direct sum decomposition $\mathfrak{g} = \oplus_i V_i$. The Campbell-Baker-Hausdorff formula allows one to express the exponential coordinates of the product of two elements of G in terms of the coordinates of each factor. Given an index $i \in [1,n]$, let d_i be the degree of e_i, namely the integer such that $e_i \in V_{d_i}$. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ a multi-index, we set $x^\alpha := x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$ and $d_\alpha := d_1\alpha_1 + \ldots + d_n\alpha_n$. We have for some constants $C_{\alpha,\beta} \in \mathbb{R}$

Campbell-Baker-Hausdorff formula:

$$(x \cdot y)_i = x_i + y_i + \sum_{\{\alpha,\beta \ | \ d_\alpha \geqslant 1, d_\beta \geqslant 1, d_\alpha + d_\beta \leqslant d_i\}} C_{\alpha,\beta} x^\alpha y^\beta.$$

The new Lie product $x*y$ defined in the previous subsection with the help of the identification (2.8) then takes the following simple form (we have chopped the terms with $d_\alpha + d_\beta < d_i$):

$$(x * y)_i = x_i + y_i + \sum_{\{\alpha,\beta \ | \ d_\alpha \geqslant 1, d_\beta \geqslant 1, d_\alpha + d_\beta = d_i\}} C_{\alpha,\beta} x^\alpha y^\beta.$$

2.8. **Homogeneous quasi-norms and a lemma of Guivarc'h.** Suppose G is a stratified nilpotent Lie group with Lie algebra $\mathfrak{g}$ and stratification $\mathfrak{g} = \oplus_i V_i$. Let δ_λ as above be the one parameter group of dilations. A non-negative continuous function $|\cdot|$ on G is called a *homogeneous quasi-norm* if it satisfies the two axioms: $|x| = 0$ if and only if $x = \mathrm{id}$, and $|\delta_\lambda(x)| = \lambda|x|$ for every $\lambda > 0$ and $x \in G$.

Here are two typical examples:

a) a left-invariant subFinsler metric d_{sF} on G with horizontal subspace V_1,

b) the function $|x| := \max_i c_i|x_i|^{1/d_i}$ for any choice of constants $c_i > 0$ (in the notation of subsection 2.7).

The following is a simple observation:

Lemma 2.9. *Any two homogeneous quasi-norms $|\cdot|_1$ and $|\cdot|_2$ on G are equivalent in the sense that there is a constant $C > 0$ such that $\frac{1}{C}|\cdot|_1 \leqslant |\cdot|_2 \leqslant C|\cdot|_1$.*

When applied to the above two examples of homogeneous quasi-norms, this lemma becomes an instance of the *ball-box* principle for Carnot groups.

When the group G is not stratified, it turns out that this ball-box principle remains true under the identification (2.8), even though the group structures on G and G_∞ are not the same. This was first observed by Guivarc'h in [12]. Using the Campbell-Baker-Hausdorff formula he showed that, even when G is not stratified, given any choice of subspaces V_i's as in (2.4), one can always find constants $c_i > 0$ such that $|x| := \max_i c_i|x_i|^{1/d_i}$ satisfies $|x \cdot y| \leqslant |x| + |y| + 1$. As a consequence, he proved the following comparison theorem:

Proposition 2.10 (Guivarc'h)**.** *Let G be a simply connected nilpotent Lie group with Lie algebra $\mathfrak{g}$ and U a bounded symmetric neighborhood of* id *in G. Let $\rho_U(x,y) := \inf\{n \in \mathbb{N}; x^{-1}y \in U^n\}$ be the word metric induced by U. Let $\mathfrak{g} = \oplus_i V_i$ a decomposition as in* (2.4) *and set $|x| = \max |x_i|^{1/d_i}$. Then there are constants $C, L > 0$ such that*

$$\frac{1}{L}\rho_U(\mathrm{id}, x) - C \leqslant |x| \leqslant L\rho_U(\mathrm{id}, x) + C$$

For the proof, we refer the reader to Guivarc'h thesis [12], or to [6, Theorem 3.7].

2.11. **Geodesic left-invariant distances on nilpotent Lie groups.** Let us now turn to simply-connected nilpotent Lie groups and compare subFinsler metrics on them. Let d_{sF} be a left-invariant subFinsler metric on a simply-connected nilpotent Lie group G. The projection map $\pi : \mathfrak{g} \to \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ can be viewed as a map from G to $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ by precomposing it by the inverse of the exponential map (which we recall is a diffeomorphism).

Let Δ be the horizontal distribution for d_{sF} and let $V_1 := \Delta_{\mathrm{id}}$ be the horizontal space at the identity. Since d_{sF} is finite, we have that V_1 is bracket generating, i.e., V_1 generates the whole Lie algebra. Hence $\pi(V_1) = \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$, since $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is an Abelian algebra. We consider the norm $\|\cdot\|_{ab}$ on $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ whose unit ball is the image in $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ under $\pi|_{V_1}$ of the unit ball of $\|\cdot\|$ in V_1.

Then we have:

Lemma 2.12 (π is a submetry)**.** *The projection map $\pi : G \to \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is a group homomorphism (i.e. $\pi(xy) = \pi(x) + \pi(y)$) and a distance nonincreasing submetry between G metrized with d_{sF} and the* *normed vector space $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ endowed with the norm $\|\cdot\|_{ab}$. This means that the image of any r-ball on (G, d_{sF}) is an r-ball on $(\mathfrak{g}/[\mathfrak{g},\mathfrak{g}], \|\cdot\|_{ab})$.*

Proof. The identity $\pi(xy) = \pi(x) + \pi(y)$ follows from the Campbell-Baker-Hausdorff formula, which tells us that $\exp^{-1}(xy) = \exp^{-1}(x) + \exp^{-1}(y)$ modulo $\mathfrak{g}^{(2)}$.

Given $r > 0$ and $g \in G$, we need to show that $\pi(B_{sF}(g, r)) = B_{\|\cdot\|_{ab}}(\pi(g), r)$. From the left invariance of d_{sF}, we may assume that $g = \mathrm{id}$. By definition of $\|\cdot\|_{ab}$, we have $\|\pi(x)\|_{ab} \leqslant \|x\|$ for every $x \in \mathfrak{g}$. Integrating this inequality along a horizontal path connecting id and h, it follows immediately that $\|\pi(h)\|_{ab} \leqslant d_{sF}(\mathrm{id}, h)$ for every $h \in G$. So π does not increases distances and we have one inclusion $\pi(B_{sF}(\mathrm{id}, r)) \subset B_{\|\cdot\|_{ab}}(0, r)$.

Moreover if $X \in \mathfrak{g}$ satisfies $\|\pi(X)\|_{ab} \leqslant r$, then, by definition of the norm $\|\cdot\|_{ab}$, there exists $X \in V$, such that $\|X\| \leqslant r$. Then $d_{sF}(\mathrm{id}, \exp(X)) \leqslant r$, because $\{\exp(tX)\}_{t\in[0,1]}$ is a horizontal path connecting id and $\exp(X)$ with length at most r. Finally $\pi(\exp(X)) = \pi(X)$, so we have proved the opposite inclusion $B_{\|\cdot\|_{ab}}(0, r) \subset \pi(B_{sF}(\mathrm{id}, r))$. □

Remark. Under the identification (2.8) the group G is endowed with a new Lie product $*$. The two projection maps on G and G_∞ agree, because they do so at the Lie algebra level. It follows that π is also a group homomorphism for the $*$ product, namely $\pi(x * y) = \pi(x) + \pi(y)$.

We will also want to consider $*$-left-invariant subFinsler metrics on G. Although the two Lie products are typically different (and sometimes not even isomorphic), the associated subFinsler metrics are comparable, namely:

Proposition 2.13 (left and $*$-left-invariant metrics are comparable)**.** *Let G be a simply-connected nilpotent Lie group and d_1 and d_2 be two subFinsler metrics on G and suppose that, for each $i = 1, 2$, d_i is either left invariant or $*$-left invariant. Then there are constants $C > 0$ and $L \geqslant 1$ such that $L^{-1}d_1(\mathrm{id}, g) - C < d_2(\mathrm{id}, g) < Ld_1(\mathrm{id}, g) + C$ for all $g \in G$.*

Proof. If both d_i are left invariant, or both d_i are $*$-left invariant, then Proposition 2.3 applies. So we can assume that say d_2 is the $*$-left-invariant subFinsler metric with horizontal subspace the subspace V_1 used in the construction of the identification (2.6) and that d_1 is a left-invariant word metric on G. Then this is precisely the result of Guivarc'h quoted in Proposition 2.10. □

A natural question is to ask for a finer comparison between subFinsler metrics. We will say that d_1 and d_2 are *asymptotic* if $\frac{d_1(\mathrm{id},g)}{d_2(\mathrm{id},g)}$ tends to 1 as g tends to infinity in the group. The following gives a simple criterion for when two subFinsler metric (be them left invariant or $*$-left invariant) are asymptotic.

Proposition 2.14 (Asymptotic metrics)**.** *Two left-invariant subFinsler metrics d_1 and d_2 are asymptotic if and only if the respective projections to $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ of the unit balls of the norms coincide.*

Proof. The 'only if' part is the easier half of the statement and follows easily from Lemma 2.12, which reduces the question to when two norms on a vector space are asymptotic, and this happens if and only if they are identical of course. The 'if' part is harder and we will in fact prove a stronger statement with error term in Proposition 4.1 below. □

2.15. **The asymptotic cone of a discrete nilpotent group.** We now pass to discrete nilpotent groups and explain Pansu's description of their asymptotic cone.

According to a well-known theorem of Malcev ([18, chapter 2]), every torsion-free finitely-generated nilpotent group Γ is isomorphic to a discrete cocompact subgroup in a connected, simply-connected, and nilpotent Lie group G. The Lie group G is uniquely determined by Γ and is called the Malcev closure of Γ.

Let $\pi : G \to \mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ be the projection homomorphism (see §2.11). Recall that π is a group homomorphism (cf. Lemma 2.12), and so in particular $\pi(\Gamma)$ is a subgroup of $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$. Since Γ is co-compact in G, $\pi(\Gamma)$ must also be co-compact in $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ and thus be a discrete lattice of full rank there.

Let S be a finite and symmetric generating set for Γ. Consider the image of S under the projection map π, then take its convex hull B in $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$. So

$$B = ConvexHull\{\pi(S)\}.$$

It is clear that B has nonempty interior. Indeed, otherwise B would be contained in a proper subspace of $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$, but $\pi(S)$ generates the lattice of full rank $\pi(\Gamma)$ in $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$, hence cannot be contained in a proper vector subspace. Moreover B is symmetric with respect to the origin, since we assumed $S = S^{-1}$. Therefore B is a symmetric convex body with nonempty interior and we can define $\|\cdot\|$ to be the norm on $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ for which the set B is the unit ball. We will call this norm the *Pansu limit norm.*

Recall that $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ is the first stratum of the graded algebra $\mathfrak{g}_\infty$ of $\mathfrak{g}$. Let G_∞ be the graded group of $\mathfrak{g}$. We call G_∞ the *graded Malcev closure* of Γ. Hence the triple $(G_\infty, \mathfrak{g}/[\mathfrak{g},\mathfrak{g}], \|\cdot\|)$ induces a subFinsler distance d_∞ on the group G_∞ and the metric space (G_∞, d_∞) is a Carnot group. We will call d_∞ the *Pansu limit metric.*

Recall that S induces on Γ a (left-invariant) word metric ρ_S defined by setting $\rho_S(\mathrm{id},\gamma) := \inf\{n \in \mathbb{N}; \gamma \in S^n\}$. We can now state:

Theorem 2.16 (Pansu [17])**.** *The sequence of pointed metric spaces* $(\Gamma, \frac{1}{n}\rho_S, \mathrm{id})$ *converges in the Gromov-Hausdorff topology to the Carnot group* $(G_\infty, d_\infty, \mathrm{id})$*. In particular all asymptotic cones of* (Γ, ρ_S) *are isometric to* (G_∞, d_∞)*.*

For the Gromov-Hausdorff topology, we refer the reader to Gromov's book [11] and to his paper [10]. Let us only recall the definition. A sequence of pointed metric spaces (X_n, d_n, x_n) is said to converge to (X, d, x) if there is $\varepsilon_n \to 0$ such that, for every $R \geqslant 1$, the sequence of bounded metric spaces $(B_{X_n}(x_n, R+\varepsilon_n), d_n)$ converges to $(B_X(x,R), d)$ in the Gromov-Hausdorff topology. Now the Gromov-Hausdorff metric is a distance on the set of (bounded) metric spaces (X, d_X) and (Y, d_Y) defined as follows:

$$d_{GH}(X,Y) = \inf\{d_{H,Z}(X,Y); Z = X \sqcup Y, d_Z \text{ admissible}\},$$

where Z is the disjoint union of X and Y and d_Z is an admissible metric on Z, namely a distance function, which restricts to d_X on X and to d_Y on Y. Here $d_{H,Z}$ is the Hausdorff distance on compact subsets of Z, namely the smallest $r > 0$ such that X lies in the r-neighborhood of Y and Y lies in the r-neighborhood of X.

We will not give here the definition of asymptotic cones, and will refer the reader to standard expositions, such as in [21], and also to [11, chapter 3]. The asymptotic cone of a metric space (X, d) is another metric space (Y, d_∞), which roughly speaking is the limit of the original metric space "viewed from very far". The construction of the limit typically depends on a choice and is not canonical (a choice of a non principal ultrafilter). However, if the sequence of pointed metric spaces $(X, x, \frac{1}{n}d)$ converges in the Gromov-Hausdorff topology to (Y, d_∞), then all limits obtained in the asymptotic cone construction are isometric to (Y, d_∞) and then one can speak of *the* asymptotic cone of (X, d).

We will give now a reformulation of Pansu's theorem, which makes no mention of asymptotic cones, but takes the form of a distance comparison theorem between ρ_S and the subFinsler distance d_∞, when viewed on G after identifying G and G_∞ via (2.8). Recall (see the discussion after Definition 2.5) that any choice of supplementary subspaces V_i's of $\mathfrak{g}^{(i+1)}$ inside $\mathfrak{g}^{(i)}$ gives rise to a natural identification between $\mathfrak{g}$ and $\mathfrak{g}_\infty$ and thus between G and G_∞. In particular G is then endowed with a new Lie product $*$ and d_∞ becomes a $*$-left-invariant subFinsler metric on G. Theorem 2.16 can be deduced easily from the following result:

Theorem 2.17 (Pansu [17])**.** *As $\gamma \in \Gamma$ tends to infinity, we have:*

$$\frac{\rho_S(\mathrm{id}, \gamma)}{d_\infty(\mathrm{id}, \gamma)} \to 1$$

Remark. In Theorem 2.17 we may have replaced d_∞ (which is a $*$-left-invariant metric) by the associated left-invariant metric with the same norm on the same horizontal subspace V_1. Indeed this is an instance of Proposition 2.14 above. Also the theorem holds regardless of the choice of the V_i's used to identify G with G_∞.

Remark. We will give a full proof below of Theorem 2.17. In fact our proof will give an error term and yield Theorem 1.3, our main technical result.

Another interesting metric associated to the word metric on Γ is the following one, which we will call the Stoll metric relative to (Γ, ρ_S), because Stoll proved in [19] that in 2-step nilpotent groups, it lies at a bounded distance from the word metric. The Stoll distance of $g \in G$ from the identity is defines as

$$d(\mathrm{id}, g) := \inf\{|t_1| + ... + |t_n|; g = s_1^{t_1} \cdot \ldots \cdot s_n^{t_n}, n \in \mathbb{N}, s_1, \ldots, s_n \in S, t_1, \ldots, t_n \in \mathbb{R}\}$$

The following is clear (either directly or by invoking Berestowski's theorem 2.2):

Lemma 2.18. *The Stoll metric $d(\cdot, \cdot)$ coincides with the left-invariant subFinsler metric induced by the norm whose unit ball is the convex hull of S in the Lie algebra $Lie(G)$ and with left-invariant distribution induced by the subspace spanned by S in $Lie(G)$. In particular (by Proposition 2.14 and Theorem 2.17) we also have $\frac{\rho_S(\mathrm{id},\gamma)}{d(\mathrm{id},\gamma)} \to 1$ as $\gamma \in \Gamma$ tends to ∞.*

In fact in [19, Theorem 4.5], Stoll proved the following:

Theorem 2.19 (Stoll)**.** *Suppose G is a 2-step simply-connected nilpotent Lie group and Γ a lattice in it generated by a symmetric generating set S. Let ρ_S be the word metric on Γ and d be the Stoll metric on G. Then there is $C > 0$ such that for all $\gamma \in \Gamma$*

$$|d(\mathrm{id}, \gamma) - \rho_S(\mathrm{id}, \gamma)| \leqslant C.$$

Whether this continues to hold in arbitrary step remains an open problem.

3. Approximation by horizontal paths and discretization of continuous geodesics

We will need a well-known and simple lemma about perturbations of controlled paths in Lie groups. If the derivatives of two paths in V_1 are close in L^2, then their horizontal lifts are also close.

Lemma 3.1. *Let G be a Lie group, let $\|\cdot\|$ be some norm on the Lie algebra of G and let $d_e(\cdot,\cdot)$ be a Riemannian metric on G. Then for every $L>0$ there is a constant $C=C(d_e,\|\cdot\|,L)>0$ with the following property. Assume $\xi_1,\xi_2:[0,1]\to G$ are two piecewise smooth paths in the Lie group G with $\xi_1(0)=\xi_2(0)=\mathrm{id}$. Let $\xi_i'\in Lie(G)$ be the tangent vector pulled back at the identity by a left translation of G. Assume that $\sup_{t\in[0,1]}\|\xi_i'(t)\|\leqslant L$, and that $\|\xi_1'(t)-\xi_2'(t)\|_{L^2([0,1])}\leqslant\varepsilon$. Then*

$$d_e(\xi_1(1),\xi_2(1))\leqslant C\varepsilon.$$

Proof. Note that the paths ξ_1 and ξ_2 live in a bounded region of G determined by L. For simplicity of exposition we may assume that such a region is contained in a single coordinate chart. On it the Riemannian distance is bi-Lipschitz to the Euclidean norm $\|\cdot\|_e$ on the coordinates, thus we may as well consider the distance associated to this norm instead of d_e.

We can differentiate and have, for suitable constants C_1, C_2, K_1, and K_2,

$$\frac{d}{dt}\|\xi_1(t)-\xi_2(t)\|_e^2 = 2\langle\frac{d}{dt}(\xi_1(t)-\xi_2(t)),\xi_1(t)-\xi_2(t)\rangle \tag{3.1}$$

$$\leqslant 2\|\xi_1(t)-\xi_2(t)\|_e\left\|\frac{d}{dt}(\xi_1(t)-\xi_2(t))\right\|_e \tag{3.2}$$

$$= 2\|\xi_1(t)-\xi_2(t)\|_e\|\xi_1(t)\cdot\xi_1'(t)-\xi_2(t)\cdot\xi_2'(t)\|_e \tag{3.3}$$

$$\leqslant 2\|\xi_1(t)-\xi_2(t)\|_e(\|\xi_1(t)\cdot\xi_1'(t)-\xi_1(t)\cdot\xi_2'(t)\|_e+\|\xi_1(t)\cdot\xi_2'(t)-\xi_2(t)\cdot\xi_2'(t)\|_e) \tag{3.4}$$

$$\leqslant 2\|\xi_1(t)-\xi_2(t)\|_e(C_1\|\xi_1'(t)-\xi_2'(t)\|_e+C_2\|\xi_1(t)-\xi_2(t)\|_e) \tag{3.5}$$

$$\leqslant K_1\|\xi_1(t)-\xi_2(t)\|_e^2+K_2\|\xi_1'(t)-\xi_2'(t)\|_e^2 \tag{3.6}$$

where, after the triangle inequality, we have used the fact that the map $(g,v)\to g\cdot v$ is locally Lipschitz in both variables and then the inequality $2ab\leqslant a^2+b^2$.

Thus

$$\frac{df}{dt}\leqslant K_1f(t)+\alpha(t),$$

where $f(t)=\|\xi_1(t)-\xi_2(t)\|_e^2$ and $\alpha(t)=K_2\|\xi_1'(t)-\xi_2'(t)\|_e^2$. Applying Gronwall's lemma (or just noting that the derivative of $e^{-K_1t}f(t)$ is at most $e^{-K_1t}\alpha(t)$), we get for $t\in[0,1]$

$$f(t)\leqslant e^{-K_1t}\int_0^t e^{-K_1s}\alpha(s)ds\leqslant K_2e^{K_1}\|\xi_1'(t)-\xi_2'(t)\|_{L^2([0,1])}^2,$$

and the result follows. ☐

Let G be a simply connected nilpotent Lie group, $(\delta_t)_t$ a one-parameter subgroup of dilations as described in Section 2 and let $*$ be the new Lie product on G associated to the Lie bracket $[\cdot,\cdot]_\infty$ obtained from the original one by setting $[x,y]_\infty:=\lim_{t\to+\infty}\delta_{\frac{1}{t}}[\delta_t(x),\delta_t(y)]$. Let V_1 be the first eigenspace of δ_t, which is a subspace transverse to $[G,G]$. We set V_1 as our horizontal subspace. Let

d_∞ be some $*$-left-invariant subFinsler metric on G, with horizontal subspace V_1. Finally we take a fixed $*$-left-invariant Riemannian metric d_e on G, we also fix a norm $\|\cdot\|$ on the Lie algebra $\mathfrak{g}$ and we set $|g|_\infty = d_\infty(\mathrm{id}, g)$.

Lemma 3.2. *Given $C > 1$, there is $D = D(C, G) > 0$ such that the following holds. Let $n \in \mathbb{N}$ and $s, t > 0$ such that $ns \leqslant Ct$. Let $x_1, ..., x_n$ be n elements of G with $|x_i|_\infty \leqslant s$, then*

$$d_e(\delta_{\frac{1}{t}}(x_1 * \ldots * x_n), \delta_{\frac{1}{t}}(\pi_1(x_1) * \ldots * \pi_1(x_n)) \leqslant D\frac{ns^2}{t^2},$$

and

$$d_e(\delta_{\frac{1}{t}}(x_1 \cdot \ldots \cdot x_n), \delta_{\frac{1}{t}}(x_1 * \ldots * x_n)) \leqslant D\frac{ns}{t^2},$$

Proof. See [6, Lemma 6.12]. The first estimate is a simple application of Lemma 3.1, where the two paths ξ_1 and ξ_2 are taken to be piecewise linear with derivative $n\delta_{\frac{1}{t}}(x_i)$ and $n\delta_{\frac{1}{t}}(\pi_1(x_i))$ respectively on each interval $[\frac{i-1}{n}, \frac{i}{n}]$. Indeed, let (e_j) be a basis of $Lie(G)$ adapted to the direct sum $V_1 \oplus \ldots \oplus V_r$. In particular, for each j there exists d_j such that $e_j \in V_{d_j}$. For $x \in G$, denote by $(x)_j$ the j-th component of x with respect to the basis. Notice that there exists a constant $K > 0$ such that $|(x)_j| \leqslant K(|x|_\infty)^{d_j}$. This follows from the equivalence of homogeneous quasi-norms (Lemma 2.9). Then, when $d_j \geqslant 2$, we have the following bound

$$\frac{(x)_j}{t^{d_j}} \leqslant K\left(\frac{|x|_\infty}{t}\right)^{d_j} \leqslant K\left(\frac{s}{t}\right)^{d_j} \leqslant K\left(\frac{s}{t}\right)^2\left(\frac{s}{t}\right)^{d_j-2} \leqslant KC^{r-2}\left(\frac{s}{t}\right)^2.$$

Therefore, we have

$$\left\|n\delta_{\frac{1}{t}}(x_i) - n\delta_{\frac{1}{t}}(\pi_1(x_i))\right\|_e = \left\|n\sum_{j:d_j\geqslant 2}\frac{(x_i)_j}{t^{d_j}}e_j\right\|_e \leqslant KC^{r-2}\dim(G)\frac{ns^2}{t^2}.$$

Hence, for some constant $D > 0$ depending only on C and G, we have

$$d_e(n\delta_{\frac{1}{t}}(x_i), n\delta_{\frac{1}{t}}(\pi_1(x_i))) \leqslant D\frac{ns^2}{t^2}.$$

The conclusion follows from Lemma 3.1.

For the second estimate, we use the Campbell-Baker-Hausdorff formula recalled in Subsection 2.7. First, letting $z_k = x_{k+1} \cdot \ldots \cdot x_n$ and $y_k = x_1 * \ldots * x_{k-1}$, we may write by the triangle inequality:

$$\begin{aligned} d_e(\delta_{\frac{1}{t}}(x_1 \cdot \ldots \cdot x_n), \delta_{\frac{1}{t}}(x_1 * \ldots * x_n)) &\leqslant \sum_{k=2}^{n} d_e(\delta_{\frac{1}{t}}(y_k * x_k * z_k), \delta_{\frac{1}{t}}(y_k * (x_k \cdot z_k))) \\ &\leqslant \sum_{k=2}^{n} d_e(\delta_{\frac{1}{t}}(x_k * z_k), \delta_{\frac{1}{t}}(x_k \cdot z_k)), \end{aligned}$$

where we used the $*$-left invariance of d_e. Note that $|z_k|_\infty = O(ns)$. From the Campbell-Baker-Hausdorff formula we know that for every $x, y \in G$, $xy - x * y = \sum_i (xy - x * y)_i e_i$ and

$$|(xy - x * y)_i| \leqslant \sum_{d_\alpha \geqslant 1, d_\beta \geqslant 1, d_\alpha + d_\beta < d_i} C_{\alpha,\beta}|x^\alpha y^\beta|,$$

and thus, noting that $|x_i| = O(|x|_\infty^{d_i})$ and $|x^\alpha| = O(|x|_\infty^{d_\alpha})$, if $|x|_\infty \leqslant Cs$ and $|y|_\infty \leqslant Ct$,

$$|(xy - x * y)_i| = \sum_{d_\alpha \geqslant 1, d_\beta \geqslant 1, d_\alpha + d_\beta < d_i} O(|x|_\infty^{d_\alpha} |y|_\infty^{d_\beta}) = O_C(st^{d_i - 2})$$

Hence we have shown that if $|x|_\infty \leqslant Cs$ and $|y|_\infty \leqslant Ct$, then

$$d_e(\delta_{\frac{1}{t}}(xy), \delta_{\frac{1}{t}}(x * y)) = O_C(\frac{s}{t^2}) \tag{3.7}$$

Applying this to $x = x_k$ and $y = z_k$ and summing, we finally obtain

$$d_e(\delta_{\frac{1}{t}}(x_1 \cdot \ldots \cdot x_n), \delta_{\frac{1}{t}}(x_1 * \ldots * x_n)) = O_C(\frac{ns}{t^2}),$$

as desired. □

We will also need the following lemma:

Lemma 3.3. *Let S be a generating set of Γ and $\|\cdot\|$ be the Pansu limit norm on V_1. Then for every $x \in V_1$, there is $\gamma \in \Gamma$ such that $|\gamma|_S \leqslant \|x\|$ and*

$$\|x - \pi_1(\gamma)\| \leqslant \dim V_1.$$

Proof. Recall that the unit ball of the Pansu limit norm $\|\cdot\|$ is defined as the convex hull of the linear projections onto m_1 (modulo the commutator subgroup) of the generating set S. If $\|x\| \leqslant 1$ we can take $\gamma = \mathrm{id}$ and there is nothing to prove. If $\|x\| > 1$, we let $s = \|x\|$ and $y = \frac{1}{s}x$. Since we have assumed S to be finite, the unit ball of $\|\cdot\|$ is a polyhedron and any point y on the unit sphere lies in some codimension one face. Thus one may find $d = \dim m_1$ vertices of the form $\pi_1(s)$, $s \in S$ such that y lies in their convex hull. That is $y = \sum_{i=1}^d y_i \pi_1(s_i)$ for some $y_i \geqslant 0$, $\sum y_i = 1$. Let n_i be the largest integer smaller or equal to sy_i. We have $n_i \leqslant sy_i < n_i + 1$. Let $\gamma = s_1^{n_1} \cdot \ldots \cdot s_d^{n_d}$. Since π_1 is a homomorphism, we have $\pi_1(\gamma) = \sum n_i \pi_1(s_i)$, and moreover $\|\pi_1(s_i)\| \leqslant 1$ by definition of the norm. Hence

$$\|y - \frac{\pi_1(\gamma)}{s}\| \leqslant \frac{d}{s}.$$

Finally $|\gamma|_S \leqslant \sum n_i \leqslant \sum sy_i = s$ and we are done. □

Using Stoll's theorem 2.19, one can improve the above to projections modulo $G^{(3)}$ rather than $G^{(2)}$. Namely, let π_2 be the quotient homomorphism $G \to G/G^{(3)}$ and let d be the Stoll metric in $G/G^{(3)}$ induced by $\pi_2(S)$ (see the end of §2.15).

Lemma 3.4. *There is a constant $C = C(S) > 0$ such that for every $u \in G/G^{(3)}$, there is $\gamma \in \Gamma$ with $|\gamma|_S \leqslant d(\mathrm{id}, u)$ such that $d(\pi_2(\gamma), u) \leqslant C$.*

Proof. Stoll's theorem (i.e. Theorem 2.19) tells us that for every $\gamma \in \Gamma$ we have $||\pi_2(\gamma)|_{\pi_2(S)} - d(\mathrm{id}, \pi_2(\gamma))| \leqslant C$ for some constant $C > 0$. Now, enlarging C if necessary, and since $\pi_2(\Gamma)$ is co-compact in $G/G^{(3)}$, there exists $\gamma \in \Gamma$ such that $||\pi_2(\gamma)|_{\pi_2(S)} - d(\mathrm{id}, u)| \leqslant 2C$. By choosing v on a d-geodesic connecting u to id such that $d(\mathrm{id}, v) = d(\mathrm{id}, u) - 2C$, and choosing $\gamma \in \Gamma$ as above but for v instead of u (namely $||\pi_2(\gamma)|_{\pi_2(S)} - d(\mathrm{id}, v)| \leqslant 2C$), we may ensure that $|\pi_2(\gamma)|_{\pi_2(S)} \leqslant d(\mathrm{id}, u)$ while $d(\pi_2(\gamma), u) \leqslant 3C$. Choosing $s_1, ..., s_k \in S$ such that $\pi_2(\gamma) = \pi_2(s_1) \cdot \ldots \cdot \pi_2(s_k)$ and $k = |\pi_2(\gamma)|_{\pi_2(S)}$, we may replace γ by $s_1 \cdot \ldots \cdot s_k$ and assume $|\gamma|_S = k \leqslant d(\mathrm{id}, u)$ while $d(\pi_2(\gamma), u) \leqslant 3C$. □

The above lemma will be useful in our main theorems in order to obtain the exponent $\frac{2}{3r}$ instead of the exponent $\frac{1}{2r}$, which is what one gets by using Lemma 3.3.

We complete this section with the proof of Lemma 1.5 from the Introduction.

Proof of Lemma 1.5. Let π_1 be the projection to the first stratum of G, and let $\{\gamma(t)\}_{t\in[0,1]}$ be a geodesic path connecting the origin to x. We set

$$x_i = \delta_n(\gamma(\frac{i-1}{n})^{-1}\gamma(\frac{i}{n})),$$

for $i = 1, \ldots, n$. Let $\{\xi(t)\}_{t\in[0,1]}$ be the path whose derivative is piecewise constant equal $\pi_1(x_i)$ on each interval $[\frac{i-1}{n}, \frac{i}{n}]$. Note that $d(\mathrm{id}, x_i) = 1$ and hence $\|\pi_1(x_i)\| \leqslant 1$ by Lemma 2.12. Then $\xi(1) = \delta_{\frac{1}{n}}(\pi_1(x_1) \cdot \ldots \cdot \pi_1(x_n))$ and Lemma 3.2 applies, with $s = 1$ and $t = n$, and yields $d_e(\xi(1), x) = O(n^{-1})$. $\square$

4. Proofs of the main results

In this section we prove our main theorems.

The following proposition is a simple consequence of Lemma 3.2. Let G be a simply connected nilpotent Lie group, $(\delta_t)_t$ a one-parameter subgroup of dilations as described in Section 2 and let $*$ be the new Lie product on G associated to the Lie bracket $[\cdot,\cdot]_\infty$ obtained from the original one by setting $[x,y]_\infty := \lim_{t\to+\infty} \delta_{\frac{1}{t}}[\delta_t(x), \delta_t(y)]$. Let d be some left-invariant geodesic metric on G, let V_1 be the first eigenspace of δ_t, which is a subspace transverse to $[G,G]$. By Berestovski's theorem (see Section 2) d is a subFinsler metric associated to some norm on a subspace of the Lie algebra of G. Observe that this subspace projects surjectively onto V_1 modulo $[\mathfrak{g},\mathfrak{g}]$ (because the horizontal subspace at the origin for d generates the Lie algebra) and that the projection of the unit ball of this norm on V_1 modulo $[\mathfrak{g},\mathfrak{g}]$ defines a norm $\|\cdot\|$ on V_1. This norm induces a $*$-left-invariant subFinsler metric d_∞ on G with horizontal subspace V_1. We set the following notation $|g| = d(\mathrm{id}, g)$ and $|g|_\infty = d_\infty(\mathrm{id}, g)$.

Proposition 4.1 (Comparison of subFinsler metrics)**.** *We have:*

$$||g|_\infty - |g|| = O(|g|^{1-\frac{1}{r}}), \qquad \textit{as } |g| \to \infty.$$

Proof. Recall that by Proposition 2.13 we have $|g|_\infty = O(|g|)$ and $|g| = O(|g|_\infty)$, as $|g|$ or $|g|_\infty \to \infty$. We first prove one side of the inequality, namely $|g|_\infty \leqslant |g| + O(|g|^{1-\frac{1}{r}})$. Let $t = |g|_\infty$. Since d is left invariant and geodesic, we may write $g = g_1 \cdot \ldots \cdot g_n$ for some g_i in a fixed compact set and $n \simeq t$, so that $|g| = \sum |g_i|$. Indeed simply take $g_i = \gamma(i-1)^{-1}\gamma(i)$, where $\{\gamma(t)\}_{t\in[0,|g|]}$ is a geodesic path connecting the origin to g. By Lemma 3.2 we may write,

$$d_e(\delta_{\frac{1}{t}}(g), \delta_{\frac{1}{t}}(\pi_1(g_1) * \ldots * \pi_1(g_n))) = O(\frac{1}{n}),$$

where d_e is a $*$-left-invariant Riemannian metric on G. Hence recalling (2.2)

$$d_\infty(g, \pi_1(g_1) * \ldots * \pi_1(g_n)) = O(t^{1-\frac{1}{r}})$$

and $|g|_\infty \leqslant |\pi_1(g_1) * \ldots * \pi_1(g_n)|_\infty + O(t^{1-\frac{1}{r}})$. On the other hand $|\pi_1(g_i)|_\infty = \|\pi_1(g_i)\| \leqslant |g_i|$ by definition of the norm $\|\cdot\|$, hence

$$|g|_\infty \leqslant \sum |g_i| + O(t^{1-\frac{1}{r}}) = |g| + O(|g|^{1-\frac{1}{r}})$$

as desired.

We now turn to the reverse inequality. We set this time $t = |g|$ and consider a d_∞-geodesic $g = g_1 * \ldots * g_n$, where the g_i's are in a fixed compact set, $n \simeq t$, and $\sum |g_i|_\infty = |g|_\infty$. By definition of the norm $\|\cdot\|$ we may find $y_1, \ldots, y_n$ in G such that $\pi_1(y_i) = \pi_1(g_i)$ and $|y_i| \leqslant \|\pi_1(g_i)\|$. Now we set $h := \pi_1(g_1) * \ldots * \pi_1(g_n)$, and $y = y_1 \cdot \ldots \cdot y_n$ and observe that $|y| \leqslant \sum |y_i| \leqslant \sum \|\pi_1(g_i)\| \leqslant \sum |g_i|_\infty = |g|_\infty$. Then applying Lemma 3.2 we get

$$d_e(\delta_{\frac{1}{t}}(y), \delta_{\frac{1}{t}}(h)) = O(\frac{1}{t}),$$

and

$$d_e(\delta_{\frac{1}{t}}(g), \delta_{\frac{1}{t}}(h)) = O(\frac{1}{t}).$$

Hence by the triangle inequality $d_e(\delta_{\frac{1}{t}}(g), \delta_{\frac{1}{t}}(y)) = O(\frac{1}{t})$. Recalling (3.7), we have $d_e(\delta_{\frac{1}{t}}(g^{-1}y), \delta_{\frac{1}{t}}(g^{-1} * y)) = O(\frac{1}{t})$. Then by the triangle inequality $d_e(\mathrm{id}, \delta_{\frac{1}{t}}(g^{-1}y)) = O(\frac{1}{t})$ and by (2.2) we get

$$|\delta_{\frac{1}{t}}(g^{-1}y)|_\infty = O(t^{-\frac{1}{r}}).$$

Finally we obtain the desired bound using Proposition 2.13:

$$|g| \leqslant |y| + |g^{-1}y| \leqslant |g|_\infty + O(|g^{-1}y|_\infty) \leqslant |g|_\infty + O(|g|_\infty^{1-\frac{1}{r}}).$$

□

4.2. **Proof of Theorem 1.3.** In the proof of the previous proposition, we considered in each case a geodesic connecting g to the identity and we split it into $n \simeq |g|$ pieces and approximate it by piecewise linear horizontal paths using Lemma 3.2. In the forthcoming proof of Theorem 1.3 below, we will follow a similar strategy, except that we will split a d_∞-geodesic of length t into roughly $\sqrt{t}$ pieces of equal length, and then take advantage of Lemma 3.3 to find a nearby path inside the discrete group Γ. This will lead to an error term with $\alpha_r = \frac{1}{2r}$, which is slightly worse than the $\alpha_r = \frac{2}{3r}$ that we claimed in the introduction. We explain how to modify the argument to get the $\frac{2}{3r}$ exponent afterwards.

Proof of Theorem 1.3 with exponent $\frac{1}{2r}$. There are two distinct arguments for the lower bound and for the upper bound. The easier of the two, namely showing the lower bound, follows from Proposition 4.1 above. Indeed, we let d be the Stoll metric (see Lemma 2.18) associated to S on G, namely the left-invariant subFinsler metric on G induced by the set S by setting $d(\mathrm{id}, g) = \inf\{|t_1| + \ldots + |t_n|; g = s_1^{t_1} \cdot \ldots \cdot s_n^{t_n}, t_j \in \mathbb{R}\}$. For $\gamma \in \Gamma$, set $|\gamma| = d(\mathrm{id}, \gamma)$. Clearly $|\gamma| \leqslant |\gamma|_S$ for every $\gamma \in \Gamma$. Moreover the associated $*$-left-invariant subFinsler metric on G induced by the norm on V_1 obtained by projecting the convex hull of S coincides with the asymptotic cone metric d_∞. Thus by Proposition 4.1 $|\gamma|_\infty \leqslant |\gamma| + O(|\gamma|_\infty^{1-\frac{1}{r}}) \leqslant |\gamma|_S + O(|\gamma|_\infty^{1-\frac{1}{r}})$ as desired (note that we get the better exponent $\frac{1}{r}$ here for the lower bound).

We now pass to the upper bound. This will require Lemma 1.5 which we proved in Section 3, and is about approximating geodesic paths by piecewise linear paths. Let $x \in G$ such that $|x|_\infty = 1$ and let $t \geqslant 1$. From Lemma 1.5, we can find, for every $n \geqslant 1$ (large but much smaller than t, to be determined later), a continuous unit speed piecewise horizontal path $\xi(s)_{s\in[0,1]}$ with $\xi(0) = \mathrm{id}$ and $d_e(x, \xi(1)) = O(\frac{1}{n})$, such that on each $[\frac{k}{n}, \frac{k+1}{n}]$ the curve $\xi(s)$ is horizontal (recall that d_e is a fixed $*$-left-invariant Riemannian metric on G). Note that 'horizontal' here is to be understood with respect to the $*$-left-invariant bundle induced by V_1. This means that there are elements $y_1, ..., y_n$ in V_1 on the unit ball of the Pansu limit norm such that $\xi(1) = (\frac{1}{n}y_1) * \ldots * (\frac{1}{n}y_n)$. We may then apply Lemma 3.3 and find group elements $\gamma_1, \ldots, \gamma_n \in \Gamma$ such that $|\gamma_i|_S \leqslant \frac{t}{n}$ and

$$\|y_i - \frac{\pi_1(\gamma_i)}{t/n}\| \leqslant \frac{\dim G}{t/n}.$$

We let $z := (\frac{1}{t}\pi_1(\gamma_1)) * \ldots * (\frac{1}{t}\pi_1(\gamma_n))$ and observe that applying Lemma 3.1 yields $d_e(z, \xi(1)) = O_G(\frac{1}{t/n})$. Now by Lemma 3.2 applied to $x_i = \gamma_i$, we obtain

$$d_e(z, \delta_{\frac{1}{t}}(g_n)) = O(\frac{n(t/n)^2}{t^2}) = O(\frac{1}{n}),$$

where $g_n := \gamma_1 \cdot \ldots \cdot \gamma_n$. Finally

$$d_e(x, \delta_{\frac{1}{t}}(g_n)) = O(\frac{1}{t/n}) + O(\frac{1}{n}).$$

Then setting $n \simeq t^{\frac{1}{2}}$, we obtain the bound $d_e(x, \delta_{\frac{1}{t}}(\gamma_1 \cdot \ldots \cdot \gamma_n)) = O(t^{-\frac{1}{2}})$. Now suppose $\gamma \in \Gamma$ and set $t = |\gamma|_\infty$ and $x = \delta_{\frac{1}{t}}(\gamma)$. Then both $|\gamma|_\infty$ and $|\gamma_1 \cdot \ldots \cdot \gamma_n|_\infty$ are at most $O(t)$, and so we know by (3.7) that $d_e(\delta_{\frac{1}{t}}(\gamma^{-1} * g_n), \delta_{\frac{1}{t}}(\gamma^{-1} \cdot g_n)) = O(\frac{1}{t})$. We conclude from the triangle inequality that $d_e(\mathrm{id}, \delta_{\frac{1}{t}}(\gamma^{-1} \cdot g_n)) = O(t^{-\frac{1}{2}})$.

By (2.2), we have that

$$|\delta_{\frac{1}{t}}(\gamma^{-1} \cdot g_n)|_\infty = O(t^{-\frac{1}{2r}}).$$

Finally, using Proposition 2.13,

$$\begin{aligned} |\gamma|_S &\leqslant |g_n|_S + |\gamma^{-1} \cdot g_n|_S \\ &\leqslant n \cdot \frac{t}{n} + O(|\gamma^{-1} \cdot g_n|_\infty) \\ &\leqslant t + O(t^{1-\frac{1}{2r}}) = |\gamma|_\infty + O(|\gamma|_\infty^{1-\frac{1}{2r}}) \end{aligned}$$

and we are done. □

Note that the proof actually gave an exponent $\frac{1}{r}$ in the lower bound, and an exponent $\alpha_r = \frac{1}{2r}$ for the upper bound. However if we repeat the same proof and replace d_∞ with the left-invariant subFinsler metric induced by the projection of S on the first two strata and use Theorem 2.19 (Stoll's theorem) and Lemma 3.4 in place of Lemma 3.3, this will allow us to subdivide the geodesic into $n \simeq t^{\frac{1}{3}}$ intervals of equal length (instead of $\sqrt{t}$ intervals). Ultimately this will give an element $g_n \in \Gamma$ with $|g_n|_S \leqslant t$ and $d_e(x, \delta_{1/t}(g_n)) = O(t^{-\frac{2}{3}})$ and thus an exponent $\alpha_r = \frac{2}{3r}$ in the upper bound. We now pass to the details.

Let $\pi_2 : \mathfrak{g} \to V_1 \oplus V_2$ be the linear projection modulo $\mathfrak{g}^{(3)}$. Note that the 2-step Lie algebras $\mathfrak{g}/\mathfrak{g}^{(3)}$ and $\mathfrak{g}_\infty/\mathfrak{g}_\infty^{(3)}$ are isomorphic under the identification of $\mathfrak{g}$ and $\mathfrak{g}_\infty$ given by *any* choice of V_i's. Thus $V_1 \oplus V_2$ is given this 2-step Lie algebra structure and then π_2 becomes a Lie algebra homomorphism. We can also view $V_1 \oplus V_2$ as endowed with the associated Lie product. Then π_2 becomes a homomorphism defined on G. Finally we note that π_2 respects both Lie products on G, namely $\pi_2(xy) = \pi_2(x * y) = \pi_2(x)\pi_2(y)$.

Now let d be the $*$-left-invariant subFinsler metric on G which is defined exactly as d_∞, except that we consider the linear projection π_2 of S to the first two strata $V_1 \oplus V_2$, then take the left-invariant distribution induced by $V_1 \oplus V_2$. If we endow $V_1 \oplus V_2 = \mathfrak{g}/\mathfrak{g}^{(3)}$ with the left-invariant metric induced by the same norm, then π_2 becomes a distance non-increasing map (and in fact a submetry) from (G, d) to $V_1 \oplus V_2$. Abusing notation, we again denote by d this subFinsler metric on $V_1 \oplus V_2$. We will also abuse notation similarly and continue to denote by d_∞ the subFinsler metric on $V_1 \oplus V_2$ which is the projection of d_∞ from G to $G/G^{(3)} \simeq V_1 \oplus V_2$.

Now pick $g \in G$ and let $t = d(\mathrm{id}, g)$. Let $n \leqslant t$ to be determined later. Now connect g to id by a d-geodesic, so that we have $g = x_1 * \ldots * x_n$ with $d(\mathrm{id}, x_i) \leqslant \frac{t}{n}$. Let $x = \delta_{\frac{1}{t}}(g)$. Let $\xi = \delta_{\frac{1}{t}}(\pi_2(x_1)) * \ldots * \delta_{\frac{1}{t}}(\pi_2(x_n))$. Then because of the π_2 projection, we gain a little more in the approximation of x by ξ. Namely (taking any norm on the vector space $\mathfrak{g}$) :

$$||n\delta_{\frac{1}{t}}(\pi_2(x_i)) - n\delta_{\frac{1}{t}}(x_i)|| = O(\frac{1}{n^2}) \tag{4.1}$$

and thus, applying Lemma 3.1, we obtain

$$d_e(x, \xi) = O(\frac{1}{n^2}).$$

Now the projection of Γ on $V_1 \oplus V_2 \simeq G/G^{(3)}$ is a discrete co-compact subgroup of $G/G^{(3)}$ and thus every element u in $V_1 \oplus V_2$ admits an element $\gamma \in \Gamma$ such that $d(u, \pi_2(\gamma)) \leqslant C$ for some constant C. Recall Lemma 3.4, which is the 2-step analogue to Lemma 3.3 and follows from Stoll's theorem (Theorem 2.19), namely:

Lemma 4.3. *There is a constant $C = C(S) > 0$ such that for every $u \in G/G^{(3)}$, there is $\gamma \in \Gamma$ with $|\gamma|_S \leqslant d(\mathrm{id}, u)$ such that $d(\pi_2(\gamma), u) \leqslant C$.*

Apply the lemma to each $u = \pi_2(x_i)$ and obtain $\gamma_i \in \Gamma$ such that $|\gamma_i|_S \leqslant \frac{t}{n}$ and $d(\pi_2(\gamma_i), \pi_2(x_i)) \leqslant C$. By left invariance we get $d_\infty(\pi_2(\gamma_i), \pi_2(x_i)) = O(1)$ and hence $d_\infty(\delta_{\frac{1}{t}}(\pi_2(\gamma_i)), \delta_{\frac{1}{t}}(\pi_2(x_i))) = O(\frac{1}{t})$. Since $d_e \ll d_\infty$, we deduce that

$$||n\delta_{\frac{1}{t}}(\pi_2(x_i)) - n\delta_{\frac{1}{t}}(\pi_2(\gamma_i))|| = O(\frac{1}{t/n})$$

Then consider $z := \delta_{\frac{1}{t}}(\pi_2(\gamma_1)) * \ldots * \delta_{\frac{1}{t}}(\pi_2(\gamma_n))$ and apply Lemma 3.1 again to obtain:

$$d_e(\xi, z) = O(\frac{1}{t/n})$$

Now set $g_n := \gamma_1 \cdot \ldots \cdot \gamma_n$. We have $|g_n|_S \leqslant n\frac{t}{n} = t$. On the other hand the second estimate of Lemma 3.2 applied to the γ_i's with $s = t/n$ shows that

$$d_e(\delta_{\frac{1}{t}}(g_n), \delta_{\frac{1}{t}}(\gamma_1) * \ldots * \delta_{\frac{1}{t}}(\gamma_n)) = O(\frac{1}{t}).$$

Now applying Lemma 3.1 again exactly as we did in (4.1) yields

$$d_e(z, \delta_{\frac{1}{t}}(\gamma_1) * \ldots * \delta_{\frac{1}{t}}(\gamma_n)) = O(\frac{1}{n^2})$$

It follows from the triangle inequality that

$$d_e(x, \delta_{\frac{1}{t}}(g_n)) = O(\frac{1}{n^2}) + O(\frac{1}{t/n}).$$

The remainder of the proof is then exactly as before, except we set the value of n to be roughly $n \simeq t^{\frac{1}{3}}$ instead of $t^{\frac{1}{2}}$ as before. So we have

$$d_e(\delta_{\frac{1}{t}}(g), \delta_{\frac{1}{t}}(g_n)) = O(t^{-\frac{2}{3}}),$$

while $t = d(\mathrm{id}, g)$ and $g_n \in \Gamma$ with $|g_n|_S \leqslant t$. We conclude as before using (3.7) that $d_e(\mathrm{id}, \delta_{\frac{1}{t}}(g^{-1}g_n)) = O(t^{-\frac{2}{3}})$. Then if $g \in \Gamma$, $|g^{-1}g_n|_S = O(|g^{-1}g_n|_\infty) = O(t^{-\frac{2}{3r}})$ so finally $|g|_S \leqslant |g_n|_S + O(t^{-\frac{2}{3r}}) \leqslant d(\mathrm{id}, g) + O(t^{-\frac{2}{3r}})$. Finally Proposition 4.1 shows that d and d_∞ are $O(d^{1-\frac{1}{r}})$ away from each other, hence $||g|_S - |g|_\infty|$ is also $O(|g|_\infty^{1-\frac{2}{3r}})$, and we are done. This ends the proof of Theorem 1.3 with the exponent $\frac{2}{3r}$.

We remark that the above proof shows that the renormalized Cayley ball $\delta_{\frac{1}{n}}(B_S(n))$ converges (in the Hausdorff metric of compact subsets of G) towards the Pansu limit ball $B_\infty(1)$ (i.e. the unit ball for the d_∞ metric). We are going to prove that it also converges as metric spaces in the Gromov-Hausdorff metric and prove Theorem 1.2, which gives an estimate on the speed of convergence.

To this end, observe that given $x \in B_\infty(1)$ and $n \in \mathbb{N}$, the proof above builds an element $\gamma_x \in \Gamma$ (denoted g_n above) such that $|\gamma_x|_S \leqslant n$ and $d_\infty(x, \delta_{\frac{1}{n}}(\gamma_x)) = O(n^{-\alpha_r})$. Similarly Proposition 3.2 shows that given $x = \delta_{\frac{1}{n}}(\gamma)$ for $\gamma \in B_S(n)$, there exists an element $y_\gamma \in B_\infty(1)$ with $d_\infty(x, y_\gamma) = O(n^{-\frac{1}{r}})$ and $y_\gamma = \delta_{\frac{1}{n}}(\pi_1(s_1) * \ldots * \pi_1(s_n))$ for some $s_1, \ldots, s_n$ in S such that $\gamma = s_1 \cdot \ldots \cdot s_n$. We are now ready for the proof of Theorem 1.2.

Proof of Theorem 1.2. Let $d_n = \frac{1}{n}\rho_S$ be the distance on X_n. Let $Z = X_n \sqcup X$ be the disjoint union of X and X_n and let $\phi : Z \to X_n$ and $\psi : Z \to X$ be the surjective maps defined as follows. If $\alpha = x \in X$ we let $\psi(\alpha) = \alpha$ and set $\phi(\alpha)$ to be the element $\delta_{\frac{1}{n}}(\gamma_x)$ defined above. While if $\alpha = \delta_{\frac{1}{n}}(\gamma) \in X_n$ for $\gamma \in B_S(n)$, then we set $\phi(\alpha) = \alpha$ and set $\psi(\alpha)$ to be the element $y_\gamma \in X$ defined above.

Next we define a distance on $Z = X_n \sqcup X$ which restricts to $d_n = \frac{1}{n}\rho_S$ on X_n and to d_∞ on X and is defined for $x \in X$ and $x' = \delta_{\frac{1}{n}}(\gamma) \in X_n$ by

$$d(x, x') := \inf_{\alpha \in Z}\{d_n(x', \phi(\alpha)) + d_\infty(\psi(\alpha), x)\} + \varepsilon_n,$$

where ε_n is soon to be determined. It is easy to check that this is indeed a distance on Z provided

$$\sup_{\alpha,\beta \in Z} |d_n(\phi(\alpha), \phi(\beta)) - d_\infty(\psi(\alpha), \psi(\beta))| \leqslant 2\varepsilon_n. \tag{4.2}$$

Moreover, by construction, $d(\phi(\alpha), \psi(\alpha)) \leqslant \varepsilon_n$ and hence $d_{GH}(X_n, X) \leqslant \varepsilon_n$ by definition of d_{GH} (see Section 2.15). We now show that (4.2) holds with $\varepsilon_n = O(n^{-\alpha_r})$ and this will finish the proof of Theorem 1.2.

We have three kinds of quantities to estimate depending on whether α and β belong to X or to X_n. However it will be enough to prove the estimate say for $\alpha, \beta \in X$, provided we show that $d_n(\alpha, \phi \circ \psi(\alpha)) = O(n^{-\alpha_r})$ when $\alpha \in X_n$. Both estimates are easy to prove given the definitions with the help of (3.7).

For $d_n(\alpha, \phi \circ \psi(\alpha)) = O(n^{-\alpha_r})$, we recall that given $\gamma \in B_S(n)$, $d_\infty(y_\gamma, \delta_{\frac{1}{n}}(\gamma_{y_\gamma})) = O(n^{-\alpha_r})$ and that $d_\infty(\delta_{\frac{1}{n}}(\gamma), y_\gamma) = O(n^{-\frac{1}{r}})$. Hence by the triangle inequality $d_\infty(\delta_{\frac{1}{n}}(\gamma_{y_\gamma}), \delta_{\frac{1}{n}}(\gamma)) = O(n^{-\alpha_r})$. Now applying (3.7) we get $d_\infty(\mathrm{id}, \delta_{\frac{1}{n}}(\gamma_{y_\gamma}^{-1}\gamma)) = O(n^{-\alpha_r})$, and since d_∞ and ρ_S are comparable by Proposition 2.13 we finally obtain $d_n(\mathrm{id}, \delta_{\frac{1}{n}}(\gamma_{y_\gamma})^{-1}\gamma) = O(n^{-\alpha_r})$ as desired.

The estimate $|d_\infty(x,y) - d_n(\delta_{\frac{1}{n}}(\gamma_x), \delta_{\frac{1}{n}}(\gamma_y))| = O(n^{-\alpha_r})$ is dealt with in a similar fashion. Namely recall that $d_\infty(x, \delta_{\frac{1}{n}}(\gamma_x)) = O(n^{-\alpha_r})$. We thus get $|d_\infty(x,y) - d_\infty(\delta_{\frac{1}{n}}(\gamma_x), \delta_{\frac{1}{n}}(\gamma_y))| = O(n^{-\alpha_r})$. On the other hand, $|d_\infty(\delta_{\frac{1}{n}}(\gamma_x), \delta_{\frac{1}{n}}(\gamma_y)) - d_\infty(\mathrm{id}, \delta_{\frac{1}{n}}(\gamma_x^{-1}\gamma_y))| \leqslant d_\infty(\delta_{\frac{1}{n}}(\gamma_x^{-1} * \gamma_y), \delta_{\frac{1}{n}}(\gamma_x^{-1}\gamma_y))$ which is $\leqslant$ $d_e(\delta_{\frac{1}{n}}(\gamma_x^{-1} * \gamma_y), \delta_{\frac{1}{n}}(\gamma_x^{-1}\gamma_y))^{\frac{1}{r}}$ by (2.2), which in turn is a $O(n^{-\frac{1}{r}})$ by (3.7). Finally $|d_\infty(\mathrm{id}, \delta_{\frac{1}{n}}(\gamma_x^{-1}\gamma_y)) - d_n(\mathrm{id}, \delta_{\frac{1}{n}}(\gamma_x^{-1}\gamma_y))| = O(n^{-\alpha_r})$ by Theorem 1.3, and we are done. □

Proof of Theorem 1.1. We first assume that Γ is torsion free and thus embeds in its Malcev closure G as a discrete co-compact subgroup. The result will follow easily from Theorem 1.3 and Proposition 4.1 above. First, normalize the Haar measure on G so that G/Γ has total volume 1, and let F be a compact fundamental domain for the action of Γ on G. Let d and d_∞ be the left-invariant and $*$-left-invariant subFinsler metrics on G respectively, which are induced from the Pansu limit norm on V_1 associated to S as described in Section 2. Combining Theorem 1.3 and Proposition 4.1, we see that $||\gamma|_S - |\gamma|| = O_S(|\gamma|^{1-\alpha_r})$ for every $\gamma \in \Gamma$. Given that F is compact, there must be a constant $C > 0$ such that $S^nF \subset B_d(n + O(n^{1-\alpha_r}))$ for all $n \geqslant 1$. And conversely, if $g \in B_d(n)$, then there is $f \in F$ such that $gf^{-1} \in \Gamma$ and we thus get $B(n) \subset S^{n+n^{1-\alpha_r}}F$ for every $n \geqslant 1$. Taking the Haar volume of S^nF, we obtain:

$$\operatorname{vol} B_d\big(n - O(n^{1-\alpha_r})\big) \leqslant \operatorname{vol}(S^nF) \leqslant \operatorname{vol} B_d\big(n + O(n^{1-\alpha_r})\big).$$

On the other hand we can compare B_d and B_{d_∞} using Proposition 3.2 and conclude that the above inequality also holds with B_{d_∞} in place of B_d. However d_∞ is admits the scaling property $|\delta_t(g)|_\infty = t|g|_\infty$, and thus $\operatorname{vol} B_{d_\infty}(t) = t^d \operatorname{vol} B_{d_\infty}(1)$. Observing that with our choice of normalization for the Haar measure $\operatorname{vol}(S^nF) = |S^n|$, we obtain the desired result with $c_S = \operatorname{vol} B_{d_\infty}(1)$.

Now a word about the torsion case. As is well-known (see [18, chapter 2]), the set T of torsion elements in Γ is a finite normal subgroup of Γ and Γ/T is torsion free. In particular, there exists $n_0 \geqslant 1$ such that $T \subset S^{n_0}$. If $\overline{S}$ is the projection of S in Γ/T, then $|\overline{S}^n| = c_{\overline{S}}n^d + O(n^{d-\alpha_r})$. However, on the one hand $|S^n| \leqslant |S^nT| = |\overline{S}^n||T|$ and on the other hand $|\overline{S}^n||T| = |S^nT| \leqslant |S^{n+n_0}|$. It follows that $|S^n| = c_{\overline{S}}|T|n^d + O(n^{d-\alpha_r})$ as desired. □

5. Sharpness of the error terms for step-2 groups and the Burago-Margulis conjecture

In [1] D. Burago and G.A. Margulis conjectured that any two left-invariant word distances on a group ρ_1 and ρ_2 which are asymptotic in the sense that $\frac{\rho_1(\mathrm{id},\gamma)}{\rho_2(\mathrm{id},\gamma)} \to 1$ as $\gamma \to \infty$, must be at a bounded distance from each other, i.e. $|\rho_1(\mathrm{id},\gamma) - \rho_2(\mathrm{id},\gamma)| = O(1)$. Burago proved the conjecture for $\mathbb{R}^n$ and $\mathbb{Z}^n$ in [7] and Abels and Margulis [2] did so for word metrics in reductive real Lie groups. Krat proved it for the discrete Heisenberg group in [13] and also for word hyperbolic groups. However it turned out that their conjecture failed for general discrete nilpotent groups, and the first author gave a counter-example in [6, §8.2]. In this counter-example the difference $|\rho_1(\mathrm{id},\gamma) - \rho_2(\mathrm{id},\gamma)|$ can be of order $\sqrt{\rho_2(\gamma)}$.

In this section we recall the counter-example from [6, §8.2] and proceed by proving Proposition 1.4, which is the much stronger statement that although ρ_1 and ρ_2 are asymptotic and have isometric asymptotic cones, they are not $(1, C)$-quasi-isometric for any $C > 0$ (note that the failure of the

Burago-Margulis conjecture boils down to the weaker fact that the identity map is not a $(1,C)$-quasi-isometry). Recall that a map $\phi : X \to Y$ between two metric spaces (X, d_X) and (Y, d_Y) is called a $(1,C)$-quasi-isometry if $d_X(a,b) - C \leqslant d_Y(\phi(a), \phi(b)) \leqslant d_X(a,b) + C$, for all $a, b \in X$, and every $y \in Y$ is at distance at most C from some element of $\phi(X)$.

This will be done by proving that the renormalized Cayley ball are at least $n^{-\frac{1}{2}}$ away (instead of the expected n^{-1}) from the asymptotic cone metric. Hence the sharpness of the exponent $\frac{1}{2}$ in our main theorem, Theorem 1.3, and in the Gromov-Hausdorff convergence of Theorem 1.2 for step-2 groups.

The counter-example from [6, §8.2] was built as follows. Let $G = G(\mathbb{Z}) = \mathbb{Z} \times H_3(\mathbb{Z})$, where $H_3(\mathbb{Z})$ is the 3-dimensional discrete Heisenberg group. We will use exponential coordinates $g = (v; x, y, z)$ for the element $g = \exp(vV + xX + yY + zZ)$ in G, where V, X, Y, Z are the generators of the (4-dimensional) Lie algebra of $G(\mathbb{R})$ defined by the relations $[X, Y] = Z$ and all other brackets are zero.

We consider the two left-invariant word metrics ρ_1 and ρ_2 on G induced by the following choice of generating set:

$$\Omega_1 := \{(1; 0, 0, 1)^{\pm 1}, (1; 0, 0, -1)^{\pm 1}, (0; 1, 0, 0)^{\pm 1}, (0; 0, 1, 0)^{\pm 1}\}$$

gives rise to ρ_1, and

$$\Omega_2 := \{(1; 0, 0, 0)^{\pm 1}, (0; 1, 0, 0)^{\pm 1}, (0; 0, 1, 0)^{\pm 1}\}$$

gives rise to ρ_2.

The abelianization of the Lie algebra $\mathfrak{g}$ of $G(\mathbb{R})$ is 3-dimensional. We will set the first stratum of $\mathfrak{g}$ to be the linear span of V, X and Y. The projections of Ω_1 and Ω_2 on the abelianization $\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ coincide. It thus follows from Pansu's theorem that ρ_1 and ρ_2 are asymptotic, i.e. $\frac{\rho_1(\mathrm{id},\gamma)}{\rho_2(\mathrm{id},\gamma)}$ converges to 1 as γ tends to ∞ in G. In particular their asymptotic cones are isometric and coincide with the Carnot group $G(\mathbb{R})$ endowed with the left-invariant subFinsler metric induced by the convex hull of $\pi(\Omega_1) = \pi(\Omega_2)$. Let $d_\infty(\cdot, \cdot)$ be this Pansu limit metric on $G(\mathbb{R})$.

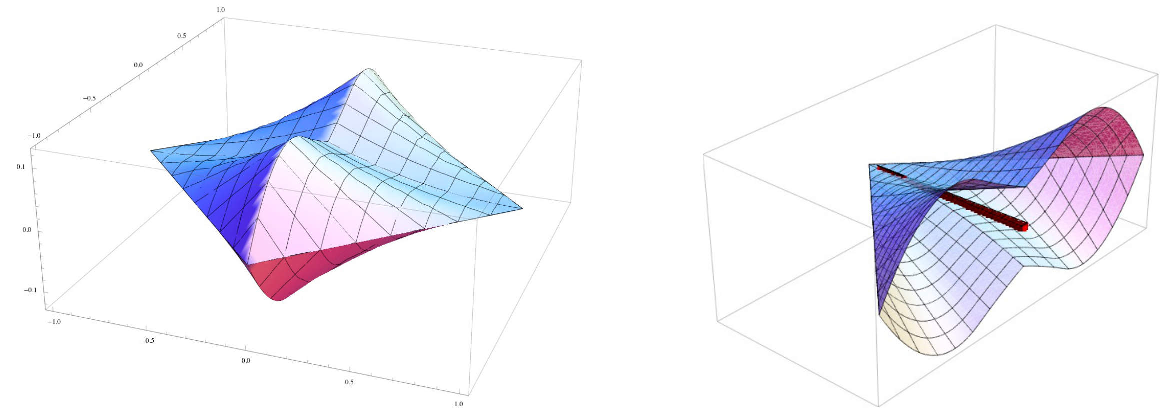
Observe that although the Pansu limit metrics associated to ρ_1 and ρ_2 coincide, the corresponding Stoll limit metrics are different. For ρ_2 the Stoll limit metric is precisely d_∞, but this is not the case for ρ_1. In view of Stoll's theorem (see Theorem 2.19), ρ_2 and d_∞ are at a bounded distance from each other. This implies that the convergence of the renormalized Cayley balls $(B_{\rho_2}(\mathrm{id}, n), \frac{1}{n}\rho_2)$ towards the asymptotic cone (i.e. $(B_{d_\infty}(\mathrm{id}, 1), d_\infty)$ is best possible, that is the Gromov-Hausdorff distance is at most $O(\frac{1}{n})$). However this is not the case for ρ_1 and we show:

Proposition 5.1. *Let $X_n^1 = (B_{\rho_1}(\mathrm{id}, n), \frac{1}{n}\rho_1)$ and $X_\infty := (B_{d_\infty}(\mathrm{id}, 1), d_\infty)$. Then*

$$d_{GH}(X_n^1, X_\infty) > \frac{c}{\sqrt{n}},$$

for some $c > 0$. In particular (G, ρ_2) and (G, ρ_1) are not $(1, C)$ quasi-isometric for any $C > 0$ even though they have isometric asymptotic cones.

To see how the statement about the absence of $(1,C)$-quasi-isometries follows from the lower bound on the Gromov-Hausdorff distance, simply observe that if (G, ρ_2) and (G, ρ_1) where $(1,C)$-quasi-isometric for some $C > 0$, then so would be (G, ρ_2) and $(G(\mathbb{R}), d_\infty)$ and we would then have



(a) The section of the ball in the plane $y = 0$; note the cusps where the vertical direction is squashed.

(b) Half of the section of the ball with the abnormal geodesic in red.

FIGURE 2. A hyperplane section of the unit ball for the Pansu limit metric d_∞ on the asympotic cone $\mathbb{R} \times H_3(\mathbb{R})$ of $\mathbb{Z} \times H_3(\mathbb{Z})$ endowed with the word metric ρ_2.

$d_{GH}(X_n^1, X_\infty) = O(\frac{1}{n})$, which contradicts the above lower bound.

We now make the following simple remark: given two Lie groups G_1 and G_2 endowed with left-invariant subFinsler metrics d_1 (associated to a norm $\|\cdot\|_1$ on the Lie algebra $\mathfrak{g}_1$) and d_2 (associated to $\|\cdot\|_2$ on $\mathfrak{g}_2$), we can build on the direct product $G_1 \times G_2$ a left-invariant product metric d defined by $d(\mathrm{id}, (g_1, g_2)) := d_1(\mathrm{id}, g_1) + d_2(\mathrm{id}, g_2)$. Since G_1 and G_2 commute, we see that this product metric d is precisely the left-invariant subFinsler metric associated to the norm $\|\cdot\| := \|\cdot\|_1 + \|\cdot\|_2$ on $Lie(G_1 \times G_2)$. Moreover the unit ball of $\|\cdot\|$ is the convex hull of the unit spheres of $\|\cdot\|_1$ in $\mathfrak{g}_1$ and $\|\cdot\|_2$ in $\mathfrak{g}_2$ viewed in the product $\mathfrak{g}_1 \times \mathfrak{g}_2$.

The Pansu limit metric d_∞ is easy to describe in terms of the Pansu limit metric of $H_3(\mathbb{Z})$ with standard generators (whose unit ball is drawn in Figure 1). Indeed it is just a product metric by the above remark. Let d_3 be the Pansu limit metric of $H_3(\mathbb{R})$ associated to the standard generators $\{(1,0,0)^{\pm 1}, (0,1,0)^{\pm 1}\}$. The metric d_3 was described in detail in the Appendix to [6]. In particular the geodesics in the d_3 metric are completely known. This will be crucial in the proof of Proposition 5.1.

We thus have:

$$d_\infty(\mathrm{id}, (v; x, y, z)) = |v| + d_3(\mathrm{id}, (x, y, z)) \tag{5.1}$$

Pictures of a 3-dimensional hyperplane section of the unit ball for d_∞ are given in Figure 5.

The proof of Proposition 5.1 relies on some elementary geometric considerations involving the precise form of the distances d_∞ and d_3 and in particular the knowledge of their geodesics. The key to it is the fact the curve $t \to (t; 0, 0, 0)$ is an abnormal geodesic in the Carnot group $(G(\mathbb{R}), d_\infty)$, so points lying above $(1; 0, 0, 0)$ of the form $(1 + \varepsilon; 0, 0, 0)$ are much further from id than would have been expected should this geodesic been normal (namely they are $O(\sqrt{\varepsilon})$ away instead of $O(\varepsilon)$ away). This can be seen geometrically on the pictures of Figure 5, where the abnormal geodesic is drawn in red and we see that it causes the unit sphere to have a cusp at the point $(1;0,0,0)$.

For the proof, we will need the concept of ε-submetry. Given $\varepsilon > 0$, an ε-submetry between two metric spaces (X, d_X) and (Y, d_Y) is a map $\phi : X \to Y$ such that every point of Y is at most ε away from $\phi(X)$ and such that the image under ϕ of every ball of radius r in X is within ε Hausdorff distance from a ball of radius r in Y. If $\varepsilon = 0$ we recover the ordinary concept of submetry.

Let B_{ℓ^1} be the unit ball for the ℓ^1 norm $\|\cdot\|_1$ in the abelianization $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ given by $\|(v;x,y)\|_1 := |v| + |x| + |y|$. This unit ball is a regular octahedron with 6 vertices. We prove:

Lemma 5.2. *Let ϕ be an isometry of $X := (B_{d_\infty}(\mathrm{id},1), d_\infty)$ onto itself. Then ϕ permutes the points $\{(-1;0,0,0),(1;0,0,0)\}$.*

Lemma 5.3. *Suppose we are given 5 points $g_1,\ldots,g_5$ in $B_{d_\infty}(\mathrm{id},1)$ in $G(\mathbb{R})$ such that $d_\infty(g_i,g_j) \geqslant 2-\varepsilon$ for all $i \neq j$. Let $g = (v;x,y,z) \in B_{d_\infty}(\mathrm{id},1)$ be such that $d_\infty(g,g_i) \geqslant 2-\varepsilon$ for every $i = 1,\ldots,5$. Then as $\varepsilon \to 0$, either $|v| = O(\varepsilon)$ or $|v-1| = O(\varepsilon)$ or $|v+1| = O(\varepsilon)$.*

Proposition 5.1 follows easily from these two lemmas and we now explain how.

Proof of Proposition 5.1. Let $\varepsilon_n := d_{GH}(X_n^1, X_\infty)$. Then by definition of the Gromov-Hausdorff metric, there exists a $(1,4\varepsilon_n)$-quasi-isometry $\phi_n : X_\infty \to X_n^1$. Moreover, the projection map π between $(G(\mathbb{Z}),\rho_1)$ and $(\mathfrak{g}/[\mathfrak{g},\mathfrak{g}], \|\cdot\|_1)$ is a 1-submetry. Indeed π is a group homomorphism such that $\|\pi(\gamma)\|_1 \leqslant \rho_1(\mathrm{id},\gamma)$, while the image of a ball of radius $n \in \mathbb{N}$ centered at id under π is precisely the integer points lying in the ℓ^1 ball of radius n in $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$, namely the ball of radius n for the word metric on $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$ induced by $\pi(\Omega_1)$. Renormalizing, it follows that π is a $\frac{1}{n}$-submetry between X_n^1 and the unit ball for the ℓ^1 metric on $\mathfrak{g}/[\mathfrak{g},\mathfrak{g}]$. Hence $\pi \circ \phi_n$ is a $(4\varepsilon_n + \frac{1}{n})$-submetry between X_∞ and the ℓ^1 unit ball in $\mathbb{R}^3$.

The two points with coordinates $(n;0,0,n)$ and $(n;0,0,-n)$ are both at ρ_1-distance n from id. After renormalization, they give rise to the two points $(1;0,0,\frac{1}{n})$ and $(1;0,0,-\frac{1}{n})$ in X_n^1. According to Pansu's theorem we have $\rho_1((n;0,0,n),(n;0,0,-n)) \sim d_\infty((n;0,0,n),(n;0,0,-n))$ (i.e. the ratio tends to 1). But $d_\infty((n;0,0,n),(n;0,0,-n)) = n d_\infty((1;0,0,\frac{1}{n}),(1;0,0,-\frac{1}{n}))$. However the d_∞-distance between $(1;0,0,\frac{1}{n})$ and $(1;0,0,-\frac{1}{n})$ is the d_3 distance between the origin and a point in the center of $H_3(\mathbb{R})$ with coordinates $(0,0,\frac{2}{n})$. Hence this distance is at least $c\frac{1}{\sqrt{n}}$ for some $c > 0$ (see the formula for the distance in subsection 5.5).

Now there are x_n and y_n in X_∞ such that $\phi_n(x_n)$ and $\phi_n(y_n)$ lie at distance at most $4\varepsilon_n$ from $(1;0,0,\frac{1}{n})$ and $(1;0,0,-\frac{1}{n})$ respectively. We conclude that

$$d_\infty(x_n,y_n) \geqslant \frac{c}{\sqrt{n}} - 4\varepsilon_n, \tag{5.2}$$

for some $c > 0$. Also $\pi \circ \phi_n(x_n)$ and $\pi \circ \phi_n(y_n)$ are within $O(\varepsilon_n)$ from $(1;0,0)$.

Now since X_n^1 Gromov-Hausdorff converges to X_∞ by Pansu's theorem, the sequence of maps ϕ_n converges to an isometry $\phi : X_\infty \to X_\infty$. By Lemma 5.2 ϕ preserves the pair of points $(\pm 1;0,0,0)$. Hence after possibly precomposing all maps ϕ_n by the symmetry $v \to -v$, we may assume that ϕ fixes both points. Therefore x_n and y_n both converge to $(1;0,0,0)$.

Now since $\pi \circ \phi_n$ is an η_n-submetry to the ℓ^1 unit ball, where $\eta_n := (4\varepsilon_n + \frac{1}{n})$, taking preimages of the 5 remaining vertices of the ℓ^1 unit ball in $\mathbb{R}^3$ (apart from $(1;0,0)$), we can find 5 points in $B_\infty(\mathrm{id},1)$, say $g_1,\dots,g_5$ such that $d_\infty(g_i,g_j) \geqslant 2-\eta_n$ and $d_\infty(x_n,g_i) \geqslant 2-\eta_n$ for all $i \neq j$. Then Lemma 5.3 tells us that the v-component of x_n must be $O(\eta_n)$-close to either -1,0, or 1. However x_n converges to $(1;0,0,0)$. We conclude that the v-component of x_n is $O(\eta_n)$-close to 1 and hence x_n itself is $O(\eta_n)$-close to $(1;0,0,0)$. The same applies to y_n. Therefore $d_\infty(x_n,y_n) = O(\eta_n)$.

Combining this with (5.2) we get $\eta_n \gg \frac{1}{\sqrt{n}}$ and hence $\varepsilon_n \gg \frac{1}{\sqrt{n}}$ as desired. □

For the proof of Lemma 5.3, we will need the following fact about the geometry of the Heisenberg group $H_3(\mathbb{R})$ equiped with the Pansu metric d_3.

Lemma 5.4. *Suppose $h_1,\dots,h_4$ are 4 points in $B_{d_3}(\mathrm{id},1)$ such that $d_3(h_i,h_j) \geqslant 2-\varepsilon$ for every $i \neq j$. Let $p \in B_{d_3}(\mathrm{id},1)$ be such that $d_3(\mathrm{id},p) + d_3(\mathrm{id},h_i) \leqslant d_3(p,h_i) + \varepsilon$ for every $i = 1,\dots,4$. Then $d_3(\mathrm{id},p) = O(\varepsilon)$ as $\varepsilon \to 0$.*

Proof of Lemma 5.3. Let $g_1,\dots,g_6$ be 6 points in $B_\infty(\mathrm{id},1)$ such that $d_\infty(g_i,g_j) \geqslant 2-\varepsilon$ for all $i \neq j$. Write $g_i = (v_i;h_i)$ the coordinates of g_i in $\mathbb{R} \times H_3(\mathbb{R})$. From (5.1) we have $d_\infty(g_i,g_j) = |v_i - v_j| + d_3(h_i,h_j)$. Since $|v_i| + d_3(\mathrm{id},h_i) \leqslant 1$ and $d_\infty(g_i,g_j) \geqslant 2-\varepsilon$, it follows from the triangle inequality that $1-\varepsilon \leqslant |v_i| + d_3(\mathrm{id},h_i) \leqslant 1$ and $|v_i| + |v_j| \leqslant |v_i - v_j| + \varepsilon$ and $d_3(\mathrm{id},h_i) + d_3(\mathrm{id},h_j) \leqslant d_3(h_i,h_j) + \varepsilon$. From this we conclude that for any $i \neq j$:

a) either v_i or v_j have opposite signs or one of them has absolute value at most $\varepsilon/2$, and

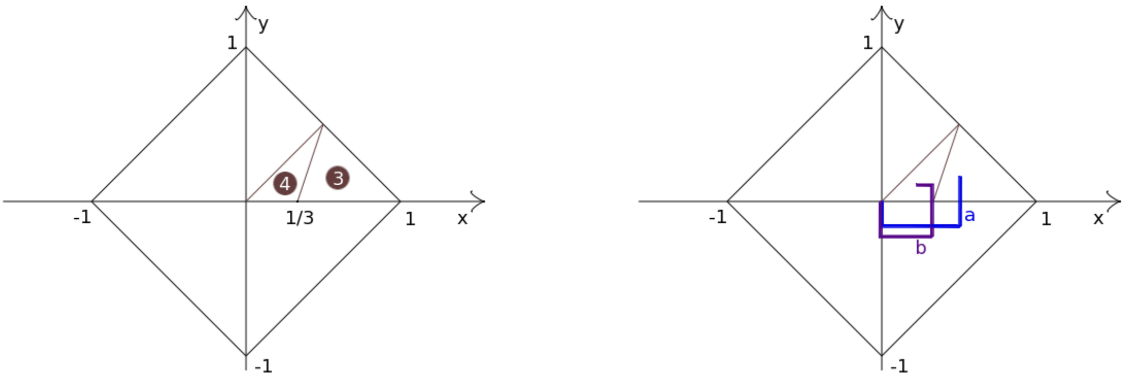
b) the piecewise geodesic path in $H_3(\mathbb{R})$ joining h_i to id and id to h_j is an ε-geodesic.

From a) we see that at most two v_i's might be $> \varepsilon/2$ in absolute value. Hence at least 4 of the 6 points g_i, say $g_1,\dots,g_4$, must have $|v_i| \leqslant \varepsilon/2$. For these points, $d_3(\mathrm{id},h_i) \geqslant 1 - 3\varepsilon/2$ and $d_3(h_i,h_j) \geqslant 2-2\varepsilon$. We may thus apply Lemma 5.4 to $p = h_1$ and $p = h_2$ successively and get $d_3(\mathrm{id},h_i) = O(\varepsilon)$ for $i = 1,2$. Given a) this means that $|v_1 - 1| = O(\varepsilon)$ and $|v_2 + 1| = O(\varepsilon)$ or vice-versa. This establishes Lemma 5.3. □

We now focus on the proofs of Lemma 5.4 and Lemma 5.2. The proof of Lemma 5.2 is postponed to the end, because it will use Lemmas 5.3 and 5.4. Since we will use the precise form of the subFinsler metric d_3 and the knowledge of its geodesics in the proof, we first devote some time to recall some facts about this metric.

5.5. **The Pansu limit metric for the Heisenberg group with standard generators and its geodesics.** Let $H_3(\mathbb{Z})$ be the discrete Heisenberg group. Write the standard generators in exponential coordinates $(x,y,z) = \exp(xX + yY + zZ)$ as $\Omega := \{(1,0,0)^{\pm 1}, (0,1,0)^{\pm 1}\}$. By Pansu's theorem, the left-invariant word metric on $H_3(\mathbb{Z})$ induced by Ω is asymptotic to its Pansu limit metric d_3, which is the left-invariant subFinsler metric on $H_3(\mathbb{R})$ induced by the ℓ^2 do we mean ℓ^1 ? norm on the horizontal subspace $span\{X,Y\}$ and defined as follows: $\|xX + yY\|_1 = |x| + |y|$.

In the Appendix to [6], we computed the geodesics of the Pansu limit metric d_3 and drew the picture of Figure 1 of the unit ball for d_3. This is done via solving the Dido isoperimetric problem for the ℓ^1 norm, that is finding the curve between two points in the plane which maximizes the enclosed area between the curve and the chord. The solution to Dido's problem is a piece of the boundary of a square with sides parallel to either X or Y.



(a) In region (3), geodesics have 3 sides. In region (4), geodesics have 4 sides.

(b) Examples of projections of geodesics.

FIGURE 3. Geodesics in the Pansu metric d_3 on the Heisenberg group $H_3(\mathbb{R})$.

We now recall the conclusion of [6, Appendix] describing the unit ball of d_3. It is of the form $\{(x,y,z) \mid |x|+|y| \leqslant 1 \text{ and } |z| \leqslant z(x,y)\}$ for a certain function $z(x,y)$ which we now describe. First one notes that there are symmetries involved, namely $d_3(\mathrm{id},(x,y,z))$ is invariant under the following operations: $z \to -z$, $x \to -x$,$y \to -y$ and $(x,y) \to (y,x)$. This means that in order to compute $z(x,y)$, it is enough to deal with the case when $0 \leqslant y \leqslant x \leqslant 1$. Then we have:

(i) If $y \leqslant 3x-1$, then $z(x,y) = \frac{1}{2}x(1-x)$
(ii) If $y \geqslant 3x-1$, then $z(x,y) = \frac{1}{16}(1+x+y)^2 - \frac{xy}{2}$

Case (i) corresponds to points (x,y,z) which can be joined to id by a geodesic which is piecewise linear with at most 3 linear pieces (all parallel to either X or Y). In case (ii) we require 4 linear pieces. See Figure 3(a).

Deciding the uniqueness of geodesics is also easy in this case. There is a unique geodesic between id and (x,y,z) on the unit sphere (assuming $0 \leqslant y \leqslant x \leqslant 1$) except in the following two cases: a) $x+y=1$ and $|z| < \frac{xy}{2}$, and b) $y=0$ and $0 \leqslant x < \frac{1}{3}$.

Remark 5.6. The unit ball of d_3 is not convex. The reader will verify easily that there is a unique geodesic between the following points on the unit sphere $(\frac{1}{2},\frac{1}{2},\frac{1}{8})$ and $(\frac{1}{2},-\frac{1}{2},\frac{1}{8})$ and that it pops out of the closed unit ball. A simple consequence of this fact is that the discrete group $H_3(\mathbb{Z})$ is not quasi-convex in the sense that there are points in the word ball of radius N which cannot be joined by a $O(1)$-coarse geodesic that lives in the ball of radius $N+O(1)$.

It is a simple matter from the above considerations to compute the exact form of the distance function d_3. We find for every $(x,y,z) \in H_3(\mathbb{R})$:

(i) If $|z| \leqslant \frac{|xy|}{2}$, then $d_3(\mathrm{id},(x,y,z)) = |x|+|y|$,
(ii) If $\frac{|xy|}{2} \leqslant |z| \leqslant \max\{|x|,|y|\}^2 - \frac{|xy|}{2}$, then $d_3(\mathrm{id},(x,y,z)) = \max\{|x|,|y|\} + \frac{2|z|}{\max\{|x|,|y|\}}$,
(iii) If $\max\{|x|,|y|\}^2 - \frac{|xy|}{2} \leqslant |z|$, then $d_3(\mathrm{id},(x,y,z)) = 4\sqrt{|z|+\frac{|xy|}{2}} - |x| - |y|$.

In particular $d_3(\mathrm{id},(0,0,z)) = 4\sqrt{|z|}$. If we are not in case (i), then every geodesic joining id to (x,y,z) is (in projection to the (x,y)-plane) an arc of square of side length $|x|$ when in case (ii) and $\frac{1}{4} + \frac{|x|+|y|}{4d_3(\mathrm{id},(x,y,z))}$ when in case (iii). Again, when not in case (i), this geodesic is unique except when $y = 0$ and $|x| \leqslant \frac{d_3(\mathrm{id},(x,y,z))}{3}$ (resp. when $x = 0$ and $|y| \leqslant \frac{d_3(\mathrm{id},(x,y,z))}{3}$). In that case the geodesics are arcs of the same square, but the position of the square can vary a little in the x-coordinate (resp. the y-coordinate).

We show in Figure 3(b) the regions where the geodesic(s) connecting id to a point $g = (x,y,z)$ is an arc of square with 3 sides (case (ii) above), or with 4 sides (case (iii) above). Figure 4(a) shows a geodesic of staircase type (case (i) above).

Note also that we always have $d_3(\mathrm{id},(x,y,z)) \geqslant |x| + |y|$, and $d_3(\mathrm{id},(x,y,z)) \geqslant 3\sqrt{|z|}$. We also make the following observation:

Observation: Given a point $g \in H_3(\mathbb{R})$ and a geodesic connecting id to g, the mid-point of the geodesic $m = (x_m, y_m, z_m)$ satisfies $d_3(\mathrm{id},g) \leqslant 4(|x_m| + |y_m|) \leqslant 2d_3(\mathrm{id},g)$ and $\sqrt{|z_m|} \leqslant \frac{1}{6}d_3(\mathrm{id},g)$.

This can be checked easily given the above description of the geodesics after one observes that the side-length of the square the geodesic is an arc of is at least a quarter of the length of the geodesic. The isometries of $(H_3(\mathbb{R}), d_3)$ can also be computed, we have:

Isometries of $(H_3(\mathbb{R}), d_3)$*:* The group of isometries fixing the origin is the dihedral group of order 8 generated by the rotation of 90 degrees $(x,y,z) \to (y,-x,z)$ and the flip around the x-axis $(x,y,z) \to (x,-y,-z)$. Note that the unit ball has an additional symmetry, namely it is centrally symmetric around the origin; this is not an isometry of $H_3(\mathbb{R})$ however.

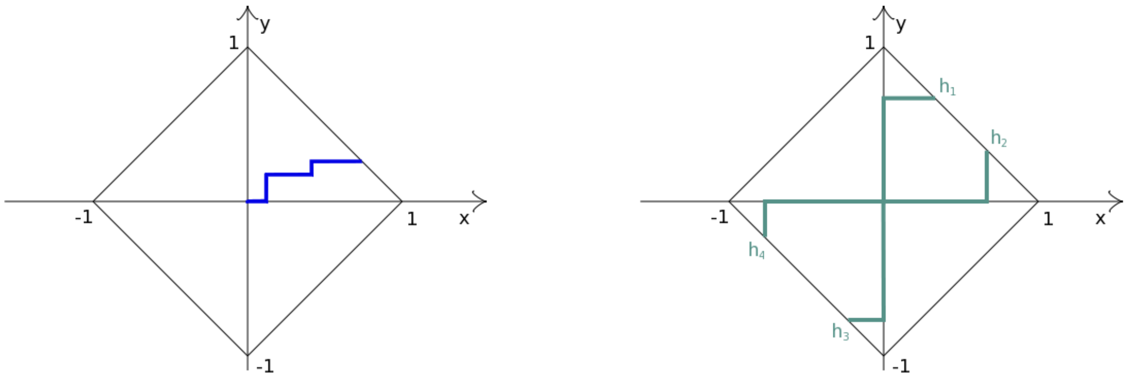
5.7. **Proof of Lemma 5.4.** First we treat the case $\varepsilon = 0$ and determine all possible configurations of 4 points $h_1, \ldots, h_4$ in the unit ball of $(H_3(\mathbb{R}), d_3)$ which satisfy $d_3(h_i, h_j) \geqslant 2$ for all $i \neq j$.

Lemma 5.8. *Let $h_1, \ldots, h_4$ in $(H_3(\mathbb{R}), d_3)$ which satisfy $d_3(h_i, h_j) = 2$ for all $i \neq j$ and $d_3(\mathrm{id}, h_i) \leqslant 1$ for every $i = 1, \ldots, 4$. Then the set of four points is of the form $\{(a, 1-a, a(1-a)), (1-a, a, -a(1-a)), (-b, -(1-b), b(1-b)), (-(1-b), -b, -b(1-b))\}$ for some $a, b \in [\frac{1}{2}, 1]$, or its image under the rotation $(x,y,z) \to (y,-x,z)$.*

For the proof of Lemma 5.4 we will only need the following consequence of this lemma: that the mid-points of the four geodesics connecting the h_i's to id are the points $a := (\frac{1}{2}, 0, 0)$,$b := (0, \frac{1}{2}, 0)$,$c := (-\frac{1}{2}, 0, 0)$ and $d := (0, -\frac{1}{2}, 0)\}$. We postpone the proof of Lemma 5.8 until after we finish the proof of Lemma 5.4.

The configurations are shown in Figure 4(b).

Claim: Suppose $p, h_1, \ldots, h_4 \in B_{d_3}(\mathrm{id}, 1)$ are points such that $d_3(h_i, h_j) = 2$ for $i \neq j$ and $d_3(p, h_i) = d_3(p, \mathrm{id}) + d_3(\mathrm{id}, h_i)$ for each $i = 1, \ldots, 4$. Then $p = \mathrm{id}$.



(a) The projection of an example of geodesic of staircase type. It connects $(0,0)$ to a point $(x, 1-x)$ with $x \in (0,1)$.

(b) A collection of four points in $B_{d_3}(\mathrm{id}, 1)$ at distance 2 from one another and the projection of the geodesics connecting them.

FIGURE 4. Geodesics in the Pansu metric d_3 on the Heisenberg group $H_3(\mathbb{R})$.

To see it, first observe that by Lemma 5.8 the four points $a := (\frac{1}{2},0,0)$,$b := (0,\frac{1}{2},0)$,$c := (-\frac{1}{2},0,0)$ and $d := (0,-\frac{1}{2},0)\}$ lie at distance $\frac{1}{2}$ from id on each one of the four geodesics connecting id to each one the points $h_1,\dots,h_4$. Consequently $d(a,p) = d(a,\mathrm{id}) + d(\mathrm{id},p) = \frac{1}{2} + d(\mathrm{id},p)$ and similarly for b,c and d. Hence without loss of generality we may assume that $p = (x,y,z)$ satisfies $x \geqslant 0$ and $|y| \leqslant x$. Also using the isometry $(x,y,z) \to (x,-y,-z)$ we can assume without loss of generality that $y \geqslant 0$. The proof of the claim is then a rather simple case by case analysis according to the shape of the geodesics between p and id. Recall that geodesics are of three possible types: the staircase type (a succession of up and right moves, say), the 3-side type (an arc of square with 3 sides), and the 4-side type (an arc of square with 4 sides). We need the concatenation of the geodesic from p to id with the geodesic from id to $a = (\frac{1}{2},0,0)$ to be a geodesic. Clearly this geodesic must be an arc of square and the side length of the square must be at least $\frac{1}{2}$ (the geodesic between id and a is a horizontal straight line of length $\frac{1}{2}$). The only way this can happen (apart from $p = \mathrm{id}$) is if $p = (\frac{1}{2},\frac{1}{2},-\frac{1}{8})$ and the geodesic from p to a is an arc of square with side length $\frac{1}{2}$ and three sides, which in projection to the (x,y)-plane are $(\frac{1}{2},\frac{1}{2}) \to (0,\frac{1}{2}) \to (0,0) \to (\frac{1}{2},0)$. However this point p is at distance 1 from id, but at distance $\frac{1}{2}$ only from the other point b. So it cannot satisfy the condition with respect to b. This proves the claim.

From the claim we conclude that as $\varepsilon \to 0$ in Lemma 5.4, any point p satisfying the conditions of the lemma must converge to id. Indeed, for any sequence of h_i's and p's satisfying the conditions of the lemma for $\varepsilon = \varepsilon_n \to 0$, we can pick a subsequence that converges. The limit points will then satisfy the conditions of the claim, and thus the limit of the p's will be id. Similarly the mid-points of the geodesics between id and the four points h_i's are converging to the four points a, b, c and d defined above. This fact will be useful in order to restrict the analysis needed to establish Lemma 5.4. We now proceed with the proof.

Our standing assumption is that $d_3(\mathrm{id},p)+d_3(\mathrm{id},A)\leqslant d_3(p,A)+\varepsilon$ and we want to deduce that $d_3(\mathrm{id},p)=O(\varepsilon)$. Observe that this inequality also holds with p replaced by any point q lying on a geodesic joining p to id. In particular when $q=(x_q,y_q,z_q)$ is the mid-point of that geodesic. Recall the observation made at the end of subsection 5.5, namely that $d_3(\mathrm{id},p)\leqslant 4(|x_q|+|y_q|)\leqslant 2d_3(\mathrm{id},p)$. It follows from this that we may assume that $p:=(x,y,z)$ satisfies the extra condition $d_3(\mathrm{id},p)\leqslant 2(|x|+|y|)$.

As in the proof of the claim above, by symmetry, we may assume that $p=(x,y,z)$ with $0\leqslant y\leqslant x\leqslant 1$. From the observation made above after the proof of the claim, we may assume that x,y and z are small. We need to show that $x=O(\varepsilon)$. Let $A=(X,Y,Z)$ be the mid-point of the geodesic between id and the point h_i for which A is close to $a=(\frac{1}{2},0,0)$. So X is close to $\frac{1}{2}$ and Y and Z are close to zero.

We have $p^{-1}A=(X-x,Y-y,Z-z+\frac{1}{2}(yX-xY))$. Note that the third coordinate is small (because x,y,z and Y,Z are small). From the formula for the distance, which we recalled in subsection 5.5, we see that the geodesic connecting p to A is either of type (i) (staircase) or of type (ii) (3-side arc of square). But not of type (iii) (4-side arc of square) because the third coordinate is negligible compared to $(X-x)^2-\frac{(X-x)(Y-y)}{2}$ (itself close to $\frac{1}{4}$). The same holds for the geodesic connecting id to A (in particular $d_3(\mathrm{id},A)\geqslant X+\frac{2|Z|}{X}$).

If we are in case (i), then the formula for the distance gives: $d_3(p,A)=X-x+|Y-y|$. On the other hand, $d_3(\mathrm{id},p)\geqslant x+y$ and $d_3(\mathrm{id},A)\geqslant X+|Y|$. Therefore $x+y+X+|Y|\leqslant X-x+|Y-y|+\varepsilon$, and it follows that $x\leqslant\frac{1}{2}\varepsilon$. Since $0\leqslant y\leqslant x$, we also have $y\leqslant\frac{1}{2}\varepsilon$ and $d_3(\mathrm{id},p)\leqslant 2(x+y)\leqslant 2\varepsilon$.

If we are in case (ii), then the formula for the distance gives: $d_3(p,A)=X-x+2\frac{|Z-z+\frac{1}{2}(yX-xY)|}{X-x}$. We write $\frac{1}{X-x}=\frac{1}{X}(1+\frac{x}{X}+o(x))$ as x nears 0. Hence recalling that Y is small and $0\leqslant y\leqslant x$, we get

$$2\frac{Z-z+\frac{1}{2}(yX-xY)}{X-x}=\frac{2Z}{X}-\frac{2z}{X}+y+o(x),$$

hence, since $\sqrt{|z|}\leqslant\frac{1}{3}d_3(\mathrm{id},p)\leqslant x+y\leqslant 2x$,

$$d_3(p,A)\leqslant X-x+\frac{2|Z|}{X}+y+o(x).$$

On the other hand, $d_3(\mathrm{id},p)\geqslant x+y$ and $d_3(\mathrm{id},A)\geqslant X+\frac{2|Z|}{X}$. Hence

$$x+y+X+\frac{2|Z|}{X}\leqslant d_3(p,A)+\varepsilon\leqslant X-x+\frac{2|Z|}{X}+y+o(x)+\varepsilon$$

and finally $2x\leqslant\varepsilon+o(x)$. Hence $x=O(\varepsilon)$ as desired and this ends the proof of Lemma 5.4.

5.9. **Proof of Lemma 5.8.** First, using the formula for the distance given in subsection 5.5 (for instance), it is easy to verify that the proposed configurations of h_i's satisfy indeed the conditions of the lemma. We thus turn to the task of proving that these are the only such configurations.

Connect id to each h_i by a geodesic. From the description of geodesics given in subsection 5.5, we know that we can pick a geodesic which (in projection to the (x,y)-plane) is a concatenation of horizontal and vertical moves. In particular the initial segment of each of the four geodesics must leave the origin $(0,0)$ in the (x,y)-plane by following the x-axis or the y-axis. Now the conditions $d_3(h_i,h_j)=2$ for $i\neq j$ and $h_i\in B_{d_3}(\mathrm{id},1)$ for all i's imply that the concatenation of a geodesic from

h_i to id and from id to h_j (for $i \neq j$) is itself a geodesic. Hence these paths do not backtrack and we conclude that the four initial segments of our four geodesics must be the four directions: the positive x-axis, the negative x-axis, the positive y-axis and the negative y-axis.

Now consider the largest $a > 0$ such that the initial segment of all four geodesics (in projection to the (x, y)-plane) start at $(0, 0)$ and end at $(a, 0)$, $(-a, 0)$, $(0, a)$ and $(0, -a)$. If $a = 1$ we are done (it is one of the proposed configurations). If $a < 1$, then one of the four geodesics, say the one moving along the positive x-axis an connecting id to h_1, must turn when it is at distance a from id. Without loss of generality, we may assume that it turns in the direction of the positive y-axis (otherwise apply the isometry $(x, y, z) \to (x, -y, -z)$) and arrive say in (a, t) for some $t > 0$. Now if we follow the path backwards from (a, t) to $(a, 0)$ then to $(0, 0)$ and to $(0, a)$, we have a geodesic. From the description of geodesics, this geodesic must be an arc of square with side length equal to a. This means that the other geodesic, which connects id to say h_2 and starts along the positive y-axis, must turn at the point $(0, a)$ in the direction of the positive x-axis. It also implies that the concatenation of the geodesic from h_1 to id and from id to h_2 must be an arc of the same square and that the total length of this path is at most $4a$. Hence $a \geqslant \frac{1}{2}$ and we must have $h_1 = (a, 1 - a, \frac{1}{2}a(1 - a))$ and $h_2 = (1 - a, a, -\frac{1}{2}a(1 - a))$. Finally the same applies to h_3 and h_4, we only need to argue that the turn occurs in the opposite corner and this is again forced by the form of the geodesics. This ends the proof of Lemma 5.8.

5.10. **Proof of Lemma 5.2.** Combining the case $\varepsilon = 0$ of Lemma 5.3 with Lemma 5.4, we obtain a classification of 6-tuples of points in $B_{d_\infty}(\mathrm{id}, 1)$ which are at distance 2 from one another: they must consist of the two points $(1; 0, 0, 0)$ and $(-1; 0, 0, 0)$ and 4 points in the Heisenberg subgroup $\{v = 0\}$ at distance 2 from one another and hence in one of the configurations described in Lemma 5.4.

For any of the four points that lie in the Heisenberg subgroup, there is a unique geodesic joining it to one of the remaining three (this follows from Lemma 5.4, the known shape of geodesics and the criterion for uniqueness recalled in subsection 5.5). However there are many geodesics joining any of the two points $(\pm 1; 0, 0, 0)$ to any of the four points in the Heisenberg subgroup $\{v = 0\}$, because the v-direction commutes with the Heisenberg subgroup. Therefore in any collection of 6 points at distance 2 from one another, the four points lying in the Heisenberg subgroup are the only subcollection of 4 points with the property that any of the four points is joined to one of the remaining 3 points by a unique geodesic. This determines these four points purely in metric terms and we conclude that any isometry of $B_{d_\infty}(\mathrm{id}, 1)$ must perserve this set of four points, and hence also its complement, namely $\{(\pm 1; 0, 0, 0)\}$. This establishes Lemma 5.2.

Remark 5.11. Given Lemma 5.2 it is easy to determine all isometries of $B_{d_\infty}(\mathrm{id}, 1)$. They respect the product structure and are of the form $(g, \pm 1)$, where g is an isometry of $(H_3(\mathbb{R}), d_3)$. The group is isomorphic to $D_8 \times \mathbb{Z}/2\mathbb{Z}$, where D_8 is the dihedral group of order 8.

6. Volume of Cayley spheres, regularity of subFinsler spheres and other open problems

6.1. **Volume of Cayley spheres.** The error term in the volume asymptotics for balls $B_S(n)$ in the Cayley graph of Γ is of course related to the volume of spheres $S_S(n) = B_S(n) \backslash B_S(n-1)$. Clearly, if one has the asymptotics $|B_S(n)| = c_S n^d + O(n^{d-\alpha})$ for some $\alpha \leqslant 1$, then one also have $|S_S(n)| = O(n^{d-\alpha})$.

Note however knowledge of an upper bound on the size of the spheres does not seem to give any information on the error terms in the volume of balls.

Colding and Minicozzi gave in [8, Lemma 3.3.] a simple argument yielding an upper bound for the volume of spheres in doubling metric spaces. Their argument was rediscovered by Tessera in [20]. This applies to our situation since the polynomial growth of nilpotent groups implies that they are doubling[1]. For nilpotent groups the argument gives an upper bound of the form:

$$\frac{|S_S(n)|}{|B_S(n)|} = O(n^{-K^{-d}}),$$

for any $K > 4$, where d is the growth exponent so that $|B_S(n)| \simeq c_S n^d$. So our Theorem 1.1 improves this bound quite a bit by giving $\frac{|S_S(n)|}{|B_S(n)|} = O(n^{-\frac{2}{3r}})$, where r is the nilpotency class.

Conjecture 6.2. *We have $|B_S(n)| = c_S n^d + O_S(n^{d-1})$ and thus $|S_S(n)| = O_S(n^{d-1})$ for all finitely generated nilpotent groups.*

This discussion was about upper bounds on the size of spheres. We conclude this subsection by recalling a simple argument giving a lower bound, which we learned from V.I. Trofimov. In any Cayley graph Γ

$$|B_S(n)| \leqslant 2n|S_S(n)|.$$

Indeed, pick an element g at distance $2R$ from the identity. Let $\gamma : \{0, \dots, 2R\} \to \Gamma$ be a (discrete) geodesic from the identity to g. Let x be the midpoint of γ. Hence $x^{-1}\gamma$ consists of $2R$ points in the ball $B_S(n)$. If p is any point in $B_S(n)$, then either $d(x^{-1}\gamma(0), p)$ or $d(x^{-1}\gamma(2R), p)$ must be at least R, while $d(id, p)$ is at most R, so there must be a point on $x^{-1}\gamma$ that is at distance exactly R from p. By homogeneity all spheres of given radius have the same size and consequently $|B_S(n)| \leqslant 2n|S_S(n)|$.

Combining the above bound with our result, we thus get:

Corollary 6.3. *There are constants C_1, C_2 depending on S such that, for all $n \in \mathbb{N}$ we have*

$$C_1 n^{d-1} \leqslant |S_S(n)| \leqslant C_2 n^{d-\alpha_r},$$

where α_r is as in Theorem 1.1, namely $\alpha_r = \frac{2}{3r}$

6.4. **The regularity of subFinsler spheres and the error term.** In [19] M. Stoll established the optimal error term on the volume of Cayley balls of 2-step nilpotent groups, namely he showed that $|B_S(n)| = c_S n^d + O(n^{d-1})$. His proofs relied on two key ingredients. First he proved that the Stoll metric d (as defined before Lemma 2.18), which is a subFinsler metric on the Malcev closure G of the finitely generated nilpotent group Γ, lies at a bounded distance from the word metric, namely $|d(\mathrm{id}, \gamma) - \rho_S(\mathrm{id}, \gamma)| = O(1)$. And then he proved that the unit sphere of the Stoll metric is a rectifiable set by proving that it is the image of a polyhedron by a polynomial map. Combined together, these two facts easily yield the error term $O(n^{d-1})$. It is thus natural to conjecture:

[1]In fact Pansu's theorem (i.e. $|B_S(n)| \simeq c_S n^d$) implies that the doubling constant is $\leqslant (1+\varepsilon)2^d$ for all balls of radius $\geqslant r(S, \varepsilon)$, while our Theorem 1.1 implies that this holds already for balls of radius $\geqslant r(S)\varepsilon^{-1/\alpha_r}$.

Conjecture 6.5. *The Stoll metric d lies at a bounded distance from the word metric ρ_S, that is there is $C = C(S) > 0$ such that for all $\gamma \in \Gamma$,*

$$|d(\mathrm{id}, \gamma) - \rho_S(\mathrm{id}, \gamma)| \leqslant C$$

and

Conjecture 6.6. *The unit sphere of the Pansu metric is rectifiable with respect to any Riemannian distance. In particular, if the group has topological dimension n, the sphere has finite $n-1$-dimensional Lebesgue measure.*

Recall that in the proof of Proposition 4.1, we showed that any two subFinsler metrics d_1, d_2 that are asymptotic (i.e. have the same projection on V_1) are such that the Hausdorff distance (for any Riemannian metric) between their renormalized balls $\delta_{1/R}(B_{d_1}(\mathrm{id}, R))$ and $\delta_{1/R}(B_{d_2}(\mathrm{id}, R))$ is at most $O(\frac{1}{R})$. In view of this fact, Conjecture 6.6, which in substance says that a Pansu metric sphere is not a fractal set, implies that $\mathrm{vol}(B_{d_1}(r)) = \mathrm{vol}(B_{d_2}(r)) + O(r^{d-1})$. Combining this with Conjecture 6.5 yields the desired error term in $O(n^{d-1})$ in the asymptotics of the Cayley ball $B_S(n)$.

Finally we note that Stoll's own proof of Conjecture 6.5 for 2-step groups is based on the proof of a statement on geodesics. He shows that every point on the r-sphere for the d metric can be joined to the origin by a piecewise linear horizontal geodesics with at most a bounded number (his proof gives $|S|^2$) of pieces. It is still an open question whether this can hold in general in higher step. Here we do not require that the horizontal pieces be in the direction of one of the generators from S. And indeed, while this can be guaranteed for 2-step groups as shown in [19, Lemma 3.3.], Stoll showed in an example that this fails for 3-step groups.

6.7. **Sharpness of our results and a modified Burago-Margulis conjecture.** The error terms in this paper are not sharp, except when the nilpotent group has step 2, in which case they are all sharp (by Proposition 1.4). It turns out that Proposition 4.1 is also sharp for general nilpotent groups as we show below in Example 1. In this paragraph we explain what we expect the optimal error terms should be in each of our theorems.

Proposition 4.1 compares asymptotic subFinsler metrics on G endowed with possibly different Lie group structure. Example 1 below shows that when the metrics are left invariant with respect to two different Lie group structures associated to different choices of the supplementary subspaces V_i, then the estimate in $O(d^{1-\frac{1}{r}})$ given in that proposition is sharp.

Example 1. Let N be a stratified simply connected nilpotent Lie group of step r with Lie algebra $\mathfrak{n}$. Let $\mathfrak{n} = \mathfrak{n}_1 \oplus \ldots \oplus \mathfrak{n}_r$ be a stratification. Let d be a left invariant subFinsler metric on N induced by a norm on $\mathfrak{n}_1$. Let $G = N^{(1)} \times N^{(2)}$ be the direct product of two copies of N labeled $N^{(1)}$ and $N^{(2)}$ and endow G with the product subFinsler metric. In particular note that if $g = (n_1, n_2)$, then $d(\mathrm{id}, g) = d(\mathrm{id}, n_1) + d(\mathrm{id}, n_2)$. Let $\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_r$ be the direct sum stratification, where $\mathfrak{g}_i = \mathfrak{n}_i^{(1)} \oplus \mathfrak{n}_i^{(2)}$. Now let ϕ be the linear map $\mathfrak{g} \to \mathfrak{g}$ defined by $\phi(x) = x$ if $x \in \mathfrak{g}_i$ for $i < r-1$ or $i = r$ and set $\phi(x) = x$ if $x \in \mathfrak{n}_{r-1}^{(1)}$ while $\phi(x) = x + \ell(x)e_r$ if $x \in \mathfrak{n}_{r-1}^{(2)}$, where $\ell : \mathfrak{n}_{r-1} \to \mathbb{R}$ is a non zero linear form and e_r a non-zero vector in $n_r^{(1)}$. Now note that $\phi(\mathfrak{g}^{(i)}) = \mathfrak{g}^{(i)}$ for all $i = 1, ..., r$, while $\phi(\mathfrak{g}_i) = \mathfrak{g}_i$ for all $i \neq r-1$ and $\phi(\mathfrak{g}_{r-1}) \neq \mathfrak{g}_{r-1}$. And $\mathfrak{g}^{(i)} = \phi(\mathfrak{g}_i) \oplus \mathfrak{g}^{(i+1)}$ for $i = 1, ..., r-1$. So the $\phi(\mathfrak{g}_i)$'s are a new choice of supplementary subspaces out of which we may define a one parameter group of dilations $\delta_t^\phi := \phi \circ \delta_t \circ \phi^{-1}$ which preserves the direct sum $\mathfrak{g} = \phi(\mathfrak{g}_1) \oplus \ldots \oplus \phi(\mathfrak{g}_r)$, where δ_t is the one parameter group of dilations associated to the original stratification $\mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_r$. The new Lie bracket is $[X,Y]^\phi := \phi([\phi^{-1}(X), \phi^{-1}(Y)]$, which in our case (because $\mathfrak{g}^{(r)}$ lies in the center of $\mathfrak{g}$) is equal to $\phi([X,Y])$. Denote by $*$ the new Lie product structure on G thus defined. Note that $\phi(xy) = \phi(x) * \phi(y)$. Let d^ϕ be the $*$-left invariant subFinsler metric on $(G,*)$ induced by the norm $\|\phi(x)\|^\phi := \|x\|$ (where $\|\cdot\|$ is the norm used to define d on G). Note that if c is a geodesic for d, then $\phi(c)$ is a d^ϕ-geodesic and $d^\phi(\mathrm{id}, \phi(g)) = d(\mathrm{id}, g)$. Also d and d^ϕ are asymptotic to each other because the projection of $\|\cdot\|$ and $\|\cdot\|^\phi$ to $\mathfrak{g}_1$ coincide.

Now pick $x \in \mathfrak{n}_{r-1}^{(2)}$ with $\ell(x) \neq 0$ and $d(\mathrm{id}, x) = 1$. Let $y := \phi(x)$ and for $t \geqslant 1$ and $y_t = \delta_t^\phi(y)$. We have $d^\phi(\mathrm{id}, y_t) = d^\phi(\mathrm{id}, \delta_t^\phi(y)) = td^\phi(\mathrm{id}, \phi(x)) = td(\mathrm{id}, x) = t$. On the other hand $d(\mathrm{id}, y_t) = d(\mathrm{id}, \delta_t^\phi(y)) = d(\mathrm{id}, \phi(\delta_t(x))) = d(\mathrm{id}, \phi(t^{r-1}x)) = d(\mathrm{id}, t^{r-1}(x + \ell(x)e_r))$. Hence $\frac{1}{t}d(\mathrm{id}, y_t) = d(\mathrm{id}, x + \ell(x)\frac{1}{t}e_r)$. However $x \in \mathfrak{n}^{(2)}$, while $e_r \in \mathfrak{n}^{(1)}$ and the metric d is the product metric, so $d(\mathrm{id}, x + \ell(x)\frac{1}{t}e_r) = d(\mathrm{id}, x) + d(\mathrm{id}, \frac{\ell(x)}{t}e_r)$. However $d(\mathrm{id}, \frac{\ell(x)}{t}e_r)$ is of order $\frac{1}{t^{\frac{1}{r}}}$. And thus $|d(\mathrm{id}, y_t) - d^\phi(\mathrm{id}, y_t)|$ is of order $t^{1-\frac{1}{r}}$, and we have shown that the error term in Proposition 4.1 is sharp.

If G is a simply connected nilpotent Lie group which is not stratifiable, then there is no preferred choice for the V_i's and the Pansu limit metric is not left invariant for the original Lie structure. In that case we don't expect any improvement on the $O(d^{\frac{1}{r}})$ error term.

However if we start with a stratified nilpotent Lie group and consider subFinsler metrics which are left invariant for that same Lie structure, then we believe that the error term can be improved all the way to $O(d^{\frac{1}{2}})$. Combined with Conjecture 6.5, this would also give a square root error term in Theorems 1.3 and in Theorem 1.2 for stratified nilpotent Lie groups and their lattices.

As we have shown in Section 5 this square root error term cannot be improved already for step-2 groups and it is connected to the failure of the Burago-Margulis conjecture. However this suggests that in the case of stratified nilpotent groups and their lattices, the Burago-Margulis conjecture ought to be reformulated as follows.

Conjecture 6.8 (Modified Burago-Margulis conjecture)**.** *Suppose G is a stratified Lie group and Γ a lattice in it. Let d_1 and d_2 be two left invariant word metrics on Γ such that $\frac{d_1(\mathrm{id},\gamma)}{d_2(\mathrm{id},\gamma)} \to 1$ at infinity. Then*

$$|d_1(\mathrm{id}, \gamma) - d_2(\mathrm{id}, \gamma)| = O(d_1(\mathrm{id}, \gamma)^{\frac{1}{2}}).$$

6.9. **Abnormal geodesics and the relation between Conjecture 6.2 and other well-known conjectures in subRiemannian geometry.** In subRiemannian geometry a horizontal curve from x to y is called abnormal or singular it is a critical point of the end point map. This means that if we perturb the derivative of the curve by an ε amount in L^2-norm, and consider the endpoints of the corresponding perturbated horizontal curves, then we cannot cover a full round ball of radius $C\varepsilon$ around y. Rather we can cover a $C\varepsilon$ ball in some proper subspace of the tangent space, the range of the differential of the endpoint map. Abnormal geodesics are problematic in many respects and are a key difference between subRiemannian and Riemannian geometries. For example, they do not necessarily satisfy the geodesic equations and therefore their smoothness is not guaranteed (and still an open problem!).

Abnormal geodesics exist in most Carnot groups. For example, pieces of horizontal one-parameter subgroups are abnormal geodesics in the free nilpotent Lie group of rank at least 3 (in every step $\geqslant 2$). In the Heisenberg groups however, there are no abnormal geodesics.

It turns out that the presence of abnormal geodesics is precisely the reason why the Burago-Margulis fails. The counter-example from [6, §8.2] and Section 5 of the present paper are based on the idea that given two asymptotic metrics, an abnormal curve can be a geodesic for one metric, but be far from being a geodesic for the other. The absence of abnormal geodesics in the Heisenberg group is precisely what is responsible for the fact that asymptotic left invariant metrics are always at a bounded distance from each other on this group (Krat's theorem [13]). However this is an exceptional case. One can see the absence of abnormal geodesics in the Heisenberg group by looking at Figure 1 and noting that the sphere has no cusps (near every point, the sphere looks like a standard full dimensional cone). In the counter-example to the Burago-Margulis conjecture however (see Figure 5), the presence of the abnormal curve (in red in the picture) produces a cusp on the unit sphere (the z-direction gets squashed).

Abnormal geodesics are also behind Conjecture 6.6 above, in fact they are the reason why this conjecture is not obvious and hence neither is the $O(r^{d-1})$ error term in the volume asymptotics of r-balls for subFinsler metrics. Indeed if there were no abnormal geodesic, the distance function $g \mapsto d(\mathrm{id}, g)$ would be smooth and its level sets (the spheres) would be rectifiable. Note that it is known that for certain Carnot-Caratheodory manifolds, the distance function and the spheres are not subanalytic (see [5]).

Even if abnormal curves exist in most Carnot groups, they are conjectured to be sparse. According to Montgomery [16, chapter 10.2] there ought to be a Sard theorem for the endpoint map, implying in particular that the set of points in G which can be reached by a singular curve of length at most 1, say, must be a nowhere dense set of zero Lebesgue measure. This is still an open problem for general Carnot groups. Should the answer be yes, it would then be possible to prove that subFinsler spheres are not fractal objects and that the $d-1$-dimensional Lebesgue measure of subFinsler spheres is finite. This would be enough to establish the $O(r^{d-1})$ error term in the volume asymptotics of left invariant subFinsler metrics on Carnot groups.

References

[1] Geometric group theory, hyperbolic dynamics and symplectic geometry. *Oberwolfach Rep.*, 3(3):1991–2057, 2006. Abstracts from the workshop held July 23–29, 2006, Organized by Gerhard Knieper, Leonid Polterovich and Leonid Potyagailo, Oberwolfach Reports. Vol. 3, no. 3.

[2] H. Abels and G. Margulis. Coarsely geodesic metrics on reductive groups. In *Modern dynamical systems and applications*, pages 163–183. Cambridge Univ. Press, Cambridge, 2004.

[3] H. Bass. The degree of polynomial growth of finitely generated nilpotent groups. *Proc. London Math. Soc. (3)*, 25:603–614, 1972.

[4] V. N. Berestovskiĭ. Homogeneous manifolds with an intrinsic metric. I. *Sibirsk. Mat. Zh.*, 29(6):17–29, 1988.

[5] B. Bonnard, M. Chyba, and I. Kupka. Nonintegrable geodesics in SR-Martinet geometry. In *Differential geometry and control (Boulder, CO, 1997)*, volume 64 of *Proc. Sympos. Pure Math.*, pages 119–134. Amer. Math. Soc., Providence, RI, 1999.

[6] E. Breuillard. Geometry of groups of polynomial growth and shape of large balls. *arXiv:0704.0095*, 2007. Preprint.

[7] D. Y. Burago. Periodic metrics. In *Representation theory and dynamical systems*, volume 9 of *Adv. Soviet Math.*, pages 205–210. Amer. Math. Soc., Providence, RI, 1992.

[8] T. H. Colding and W. P. Minicozzi, II. Liouville theorems for harmonic sections and applications. *Comm. Pure Appl. Math.*, 51(2):113–138, 1998.

[9] M. Duchin, S. Lelievre, and C. Mooney. The geometry of spheres in free abelian groups. *arXiv:1004.0053*, 2009. Preprint.

[10] M. Gromov. Groups of polynomial growth and expanding maps. *Inst. Hautes Études Sci. Publ. Math.*, (53):53–73, 1981.

[11] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. 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 by Sean Michael Bates.

[12] Y. Guivarc'h. Croissance polynomiale et périodes des fonctions harmoniques. *Bull. Soc. Math. France*, 101:333–379, 1973.

[13] S. A. Krat. Asymptotic properties of the Heisenberg group. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 261(Geom. i Topol. 4):125–154, 268, 1999.

[14] E. Le Donne. Lecture notes on sub-Riemannian geometry. *Preprint*, 2010.

[15] E. Le Donne. Geodesic manifolds with a transitive subset of smooth biLipschitz maps. *Groups Geom. Dyn.*, 5(3):567–602, 2011.

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

[17] P. Pansu. Croissance des boules et des géodésiques fermées dans les nilvariétés. *Ergodic Theory Dynam. Systems*, 3(3):415–445, 1983.

[18] M. S. Raghunathan. *Discrete subgroups of Lie groups*. Springer-Verlag, New York, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68.

[19] M. Stoll. On the asymptotics of the growth of 2-step nilpotent groups. *J. London Math. Soc. (2)*, 58(1):38–48, 1998.

[20] R. Tessera. Volume of spheres in doubling metric measured spaces and in groups of polynomial growth. *Bull. Soc. Math. France*, 135(1):47–64, 2007.

[21] L. van den Dries and A. J. Wilkie. Gromov's theorem on groups of polynomial growth and elementary logic. *J. Algebra*, 89(2):349–374, 1984.

[22] J. A. Wolf. Growth of finitely generated solvable groups and curvature of Riemanniann manifolds. *J. Differential Geometry*, 2:421–446, 1968.

Laboratoire de Mathématiques, Bâtiment 425, Université Paris Sud 11, 91405 Orsay, FRANCE
E-mail address: `emmanuel.breuillard@math.u-psud.fr`

Laboratoire de Mathématiques, Bâtiment 425, Université Paris Sud 11, 91405 Orsay, FRANCE
E-mail address: `ledonne@msri.org, enrico.ledonne@math.u-psud.fr`