THE TRAVELING SALESMAN THEOREM IN CARNOT GROUPS

VASILIS CHOUSIONIS, SEAN LI, AND SCOTT ZIMMERMAN

Abstract. Let $\mathbb{G}$ be any Carnot group. We prove that, if a subset of $\mathbb{G}$ is contained in a rectifiable curve, then it satisfies Peter Jones' geometric lemma with some natural modifications. We thus prove one direction of the Traveling Salesman Theorem in $\mathbb{G}$. Our proof depends on new Alexandrov-type curvature inequalities for the Hebisch-Sikora metrics. We also apply the geometric lemma to prove that, in every Carnot group, there exist -1-homogeneous Calderón-Zygmund kernels such that, if a set $E \subset \mathbb{G}$ is contained in a 1-regular curve, then the corresponding singular integral operators are bounded in $L^2(E)$. In contrast to the Euclidean setting, these kernels are nonnegative and symmetric.

1. Introduction

Let X be a metric space. A set $\Gamma \subset X$ is called a *rectfiable curve* if it is the Lipschitz image of a finite interval. The Analyst's Traveling Salesman Problem asks the following: given a set $E \subset X$, is there a finite length rectifiable curve $\Gamma \subset X$ so that $E \subseteq \Gamma$? This would mean that it is possible to visit the set E in finite time. In the case when such curves Γ exist, one can also ask for the smallest length of Γ.

When $X = \mathbb{R}^2$, Jones gave a complete answer to the first question using the notion of β*-numbers* [19]. For $E \subset X$, $x \in \mathbb{R}^2$, and $r > 0$ we define

$$\beta_E(x,r) := \inf_L \sup_{z \in B(x,r) \cap E} \frac{d(z,L)}{r},$$

where the infimum is taken over the set of all affine lines L. Thus, $\beta_E(x,r)$ is a scale-invariant measure of how close the set E lies to some line. He also developed upper and lower bounds for the infimal length of rectifiable curves containing E by using these β-numbers. Okikiolu later generalized his result to Euclidean spaces of all dimensions [24]. The following theorem is now known as the Traveling Salesman Theorem:

Theorem 1.1 (Euclidean Traveling Salesman Theorem (TST) [19, 24])**.** *Let $E \subset \mathbb{R}^n$. Then E lies on a finite length rectifiable curve if and only if*

$$\gamma(E) := \mathrm{diam}(E) + \int_{\mathbb{R}^n} \int_0^\infty \beta_E(x,r)^2 \, \frac{dr}{r^n} dx < \infty. \tag{1.1}$$

Furthermore, if $\gamma(E) < \infty$, then we have an estimate on the infimal length of such curves:

$$\frac{1}{C}\gamma(E) \le \inf_{\Gamma \supset E} \ell(\Gamma) \le C\gamma(E),$$

where C is some constant depending only on n.

It is well known from Rademacher's theorem that rectifiable curves in $\mathbb{R}^n$ infinitesimally resemble lines. However, to answer questions about the boundedness of singular integrals and

V.C. is supported by the Simons Collaboration grant no. 521845. S.L. is supported by NSF grant DMS-1600804.

other problems of a global nature, Rademacher's theorem does not provide enough quantitative information. Stated informally, one would like to know that rectifiable curves admit good affine approximations "at most places and scales". This is typically quantified via integrals over space and scale as in (1.1). Such *Carleson integrals* convey the right quantitative information required for the study of certain well known singular integrals. Jones was the first to realize this connection [18]. He used β-numbers to control the Cauchy singular integral on 1-dimensional Lipschitz graphs. Since Jones' pioneering work, β-numbers have become crucial tools in harmonic analysis, geometric measure theory, and their connections. In fact, the introduction of β-numbers may be viewed as a point of departure for the theory of *quantitative rectifiability* which was developed in the 90's by David and Semmes [8, 9, 10]. The study of quantitative rectifiability led to a rich geometric framework for singular integrals acting on lower dimensional subsets of $\mathbb{R}^n$. For more information, we refer the reader to the books [9, 25, 31]

There have been numerous generalizations and variants of Theorem 1.1 beyond Euclidean spaces. Schul [27] extended Theorem 1.1 to Hilbert spaces, David and Schul [11] recently considered the theorem in the graph inverse limits of Cheeger-Kleiner, and Hahlomaa and Schul (independently)[15, 16, 26] obtained variants of Theorem 1.1 in general metric spaces. In the last case, however, there is no natural notion of lines over which one may infimize in the definition of β, so curvature-type quantities other than β-numbers must be considered.

A natural class of metric spaces in which to study the Analyst's Traveling Salesman Problem (TSP) is the class of Carnot groups (introduced in more detail in Section 2). This is a special subclass of nilpotent Lie groups whose abelian members are precisely the Euclidean spaces. Thus, these groups can be viewed as nonabelian generalizations of Euclidean spaces. Moreover, Carnot groups are locally compact geodesic spaces which admit dilations, and they are isometrically homogeneous. In fact, by a recent observation of Le Donne [20], Carnot groups are the *only* metric spaces with these properties. Developing aspects of quantitative rectifiability (such as the TST) in Carnot groups contributes to the systematic effort which started about 15 years ago to develop Geometric Measure Theory (GMT) on these sub-Riemannian spaces. Rather than providing a long list of highlights in sub-Riemannian GMT, we refer the reader to the recent lecture notes of Serra Cassano [28] which provide a nice overview of the field with ample references to the continuously growing literature.

Like Euclidean spaces, Carnot groups are Ahlfors regular and contain a rich family of lines (which are cosets of 1-dimensional subgroups isometric to $\mathbb{R}$). These are the so-called *horizontal lines*. Hence the definition of β-numbers readily generalizes in this case. Indeed, in the definition of $\beta_E(x,r)$, we instead take the infimum $\inf_L$ over all horizontal lines that intersect $B(x,r)$, and we use the sub-Riemannian metric to measure distance. Ferrari, Franchi and Pajot [13] initialized the study of the TSP in the simplest nonabelian Carnot group; the Heisenberg group $\mathbb{H}$. They proved that, if the Carleson integral of β_E^2 is bounded, then E lies on a rectifiable curve. Schul and the second named author [22, 23] improved the aforementioned result, and they obtained an almost sharp Traveling Salesman Theorem in $\mathbb{H}$:

Theorem 1.2 ([22]). *There exists a universal constant $C > 0$ so that if $\Gamma \subset \mathbb{H}$ is a finite length rectifiable curve, then*

$$\operatorname{diam}(\Gamma) + \int_{\mathbb{H}} \int_0^\infty \beta_\Gamma(x,r)^4 \ \frac{dr}{r^4} \ dx \le C\ell(\Gamma).$$

Theorem 1.3 ([23]). *For any $p < 4$, there exists $C(p) > 0$ so that for any $E \subset \mathbb{H}$ for which*

$$\gamma_p(E) := \operatorname{diam}(E) + \int_{\mathbb{H}} \int_0^\infty \beta_E(x,r)^p \ \frac{dr}{r^4} \ dx < \infty,$$

then there is a finite length rectifiable curve Γ *that contains* E *and*

$$\ell(\Gamma) \leq C_p \gamma_p(E).$$

It is currently unknown whether Theorem 1.3 holds for $p = 4$. This would give a sharp converse to Theorem 1.2. Note that the 4 in the exponent of $\frac{dr}{r^4}$ is an obvious modification resulting from the Hausdorff dimension of the Heisenberg group. However, the exponent 4 of β_E in Theorem 1.2 is a consequence of an Alexandrov-type curvature inequality in $\mathbb{H}$ whose rather delicate proof depends crucially upon the Koranýi metric in $\mathbb{H}$. Note, however, that Theorem 1.2 holds for any homogeneous metric in $\mathbb{H}$ including the sub-Riemannian metric.

We cannot use the Koranýi metric in the general setting since it does not generalize to arbitrary Carnot groups. Instead, we use another family of metrics – the Hebisch-Sikora metrics [17] – for which we will establish a similar curvature inequality (Theorem 3.1). With the new curvature inequality, we can then use the proof of [22] to obtain the following theorem which holds for all homogenous metrics in any Carnot group $\mathbb{G}$.

Theorem 1.4. *Let* $\mathbb{G}$ *be a step* r *Carnot group with Hausdorff dimension* Q*. There is a constant* $C = C(\mathbb{G}) > 0$ *such that, for any rectifiable curve* $\Gamma \subset \mathbb{G}$*, we have*

$$\int_{\mathbb{G}} \int_0^\infty \beta_\Gamma(x,t)^{2r^2} \frac{dt}{t^Q} d\mathcal{H}^Q(x) \leq C\mathcal{H}^1(\Gamma).$$

In the case of step 2 Carnot groups (of which the Heisenberg group is an example), this theorem provides a bound on the Carleson integral involving $\beta^{2\cdot 2^2} = \beta^8$. This is weaker than the bound on the Carleson integral of Theorem 1.2 which involves β^4. We will prove in Section 5 that, in the special case of step 2 Carnot groups, the curvature inequality can be improved so that Theorem 1.4 holds with an exponent 4 on β. Therefore we obtain a genuine generalization of Theorem 1.2 to any step 2 Carnot group.

Theorem 1.5. *Let* $\mathbb{G}$ *be a step 2 Carnot group with Hausdorff dimension* Q*. There is a constant* $C = C(\mathbb{G}) > 0$ *such that, for any rectifiable curve* $\Gamma \subset \mathbb{G}$*, we have*

$$\int_{\mathbb{G}} \int_0^\infty \beta_\Gamma(x,t)^4 \frac{dt}{t^Q} d\mathcal{H}^Q(x) \leq C\mathcal{H}^1(\Gamma).$$

As mentioned earlier, there are deep connections between quantitative rectifiability and singular integral operators (SIO) in Euclidean spaces. In particular, the boundedness of SIOs on Lipschitz graphs (and beyond) is a classical topic developed by Calderón [1], Coifman-McIntosh-Meyer [5], David [6], David-Semmes [8, 9], Tolsa [30], and many others. In all of these contributions, the kernels defining the SIO are odd functions. This is very reasonable since, in order to define a SIO which makes sense on lines and other "nice" 1-dimensional objects, one heavily relies on the cancellation properties of the kernel, see e.g. [29, Proposition 1, pp 289]. Surprisingly, the situation is very different in Carnot groups, and this was first observed in the first Heiseinberg group in [3]. Using Theorem 1.4, we will prove the following theorem.

Theorem 1.6. *Let* $(\mathbb{G}, d)$ *be Carnot group of step* $r \geq 2$ *equipped with a homogeneous metric* d*. There exists a nonnegative, symmetric,* -1 *homogeneous, Calderón-Zygmund kernel* $K : \mathbb{G}\backslash\{0\} \to (0,\infty)$ *such that the corresponding truncated singular integrals*

$$T^\varepsilon f\,(p) = \int_{E\backslash B_d(p,\varepsilon)} K(q^{-1}\cdot p) f(q)\, d\mathcal{H}^1(q)$$

are uniformly bounded in $L^2(\mathcal{H}^1|_E)$ *for every* 1*-regular set* E *which is contained in a* 1*-regular curve.*

The paper is organized as follows. In Section 2, we will introduce the basic properties of Carnot groups that will be needed for our purposes, and we will define the Hebisch-Sikora metric used throughout the paper. We will introduce and prove the curvature estimate Theorem 3.1 in Section 3. This curvature bound will be used to prove Theorem 1.4 in Section 4. This section follows the example set forth in [22]. The case of step 2 groups will be addressed in Section 5. Finally in Section 6 we will prove Theorem 1.6.

Acknowledgments. We extend our sincere thanks to the referee for reading our manuscript extremely carefully and for pointing out to us several mistakes and inconsistencies.

2. Carnot preliminaries

A step r *Carnot group* is a connected, simply connected Lie group $\mathbb{G}$ whose Lie algebra $\mathfrak{g}$ is *stratified* in the following sense:

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_r, \quad [V_1, V_i] = V_{i+1} \text{ for } i = 1, \ldots, r-1, \quad [V_1, V_r] = \{0\}$$

where $V_1, \ldots, V_r$ are non-zero subspaces of the Lie algebra. Write $v_i = \dim V_i$. Choose a basis $\{X_1, \ldots, X_N\}$ of $\mathfrak{g}$ adapted to the stratification in the following sense:

$$\left\{X_{\sum_{j=1}^{i-1} v_j + 1}, \ldots, X_{\sum_{j=1}^{i} v_j}\right\} \text{ is a basis of } V_i \text{ for each } i \in \{1, \ldots, r\}.$$

For any $x \in \mathbb{G}$, we can uniquely write $x = \exp(x^1 X_1 + \cdots + x^N X_N)$ for some $(x^1, \ldots, x^N) \in \mathbb{R}^N$ via the exponential map $\exp : \mathfrak{g} \to \mathbb{G}$. Thus we may identify $\mathbb{G}$ with $\mathbb{R}^N$ using the relationship $x \leftrightarrow (x^1, \ldots, x^N)$. Denote by $|\cdot|$ the Euclidean norm in $\mathbb{G} = \mathbb{R}^N$ (depending on the above choice of basis). Say Q is the *homogeneous dimension* of $\mathbb{G}$ i.e. $Q = \sum_{i=1}^r i \dim V_i$. There is a natural family of automorphisms known as *dilations* on $\mathbb{G}$. If, for any $p \in \mathbb{G}$, we write $p = (p_1, \ldots, p_r)$ where $p_i \in \mathbb{R}^{v_i}$, then for any $s > 0$ define the dilation

$$\delta_s(p) = \left(sp_1, s^2 p_2, \ldots, s^r p_r\right).$$

It follows that $\{\delta_s\}_{s>0}$ is a one parameter family i.e. $\delta_s \circ \delta_t = \delta_{st}$. Given $p \in \mathbb{G}$, we will also often write $p = (p_1, p_2) \in \mathbb{R}^n \times \mathbb{R}^{N-n}$ for $n := v_1$. We may then think of p_1 as the "horizontal part" of p. Define the *non-horizontal part* of $p \in \mathbb{G}$ as

$$NH(p) := \tilde{\pi}(p)^{-1} p$$

where $\tilde{\pi} : \mathbb{G} \to \mathbb{G}$ is the map $\tilde{\pi}(p_1, p_2) = (p_1, 0)$ (note that this is not a group homomorphism!).

We will now endow $\mathbb{G}$ with a metric space structure.

Theorem 2.1 (Hebisch and Sikora, 1990). *There exists $\varepsilon_0 > 0$ so that, for every $\eta < \varepsilon_0$,*

$$\|x\| = \inf\{t : |\delta_{1/t}(x)| < \eta\} \quad \textit{for all } x \in \mathbb{G}$$

is a homogeneous, subadditive norm on $\mathbb{G}$ i.e. for every $s > 0$ and $x, y \in \mathbb{G}$, $\|\delta_s(x)\| = s\|x\|$ and $\|xy\| \leq \|x\| + \|y\|$. In particular, the unit ball in $\|\cdot\|$ coincides with the Euclidean ball $B_{\mathbb{R}^N}(0, \eta)$.

For any $\eta < \varepsilon_0$, call this norm the *Hebisch-Sikora (HS) norm* on $\mathbb{G}$ as introduced in [17]. Define the induced metric d on $\mathbb{G}$ as $d(x, y) = \|y^{-1}x\|$. The continuity of the Carnot dilations implies in particular that $|\delta_{1/\|x\|}(x)| = \eta$. For any horizontal point (that is, $p \in \mathbb{R}^n \times \{0\}$), we have

$$\|p\| = \inf\{t : \tfrac{1}{t}|p| < \eta\} = \inf\{t : \tfrac{1}{\eta}|p| < t\} = \tfrac{1}{\eta}|p|. \tag{2.1}$$

Moreover, we have $\|\tilde{\pi}(p)\| \leq \|p\|$ for any $p \in \mathbb{G}$. Indeed, if there was some $t > 0$ satisfying both $\|\tilde{\pi}(p)\| > t$ and $|\delta_{1/t}(p)| < \eta$, we would have $\eta^2 > \frac{p_1^2}{t^2} + \cdots + \frac{p_r^2}{t^{2r}} \geq \frac{p_1^2}{t^2} = \frac{\|\tilde{\pi}(p)\|^2}{t^2}\eta^2 > \eta^2$ which is impossible. We also record that for any compact $K \subset \mathbb{G}$, there is a constant $C > 0$ (depending on K) so that

$$d(x,y)^r \leq C|x-y| \quad \text{for any } x, y \in K. \tag{2.2}$$

A metric d on $\mathbb{G}$ is said to be *homogeneous* if $d : \mathbb{R}^N \times \mathbb{R}^N \to [0,\infty)$ is continuous with respect to the Euclidean topology, is left invariant, and is 1-homogeneous with respect to the dilations $\{\delta_r\}_{r>0}$. The 1-homogeneity of d means that

$$d(\delta_r(p), \delta_r(q)) = r\, d(p,q)$$

for all $p, q \in \mathbb{G}$ and all $r > 0$. We note in particular that the Hebisch-Sikora and the Carnot–Carathéodory metrics are homogeneous. Any two homogeneous metrics d_1 and d_2 on a given Carnot group $\mathbb{G}$ are equivalent in the sense that there exists a constant $L > 0$ so that

$$L^{-1} d_1(p,q) \leq d_2(p,q) \leq L d_1(p,q) \tag{2.3}$$

for all $p, q \in \mathbb{G}$; this is an easy consequence of the assumptions.

We define the Jones β-numbers for a set $K \subset \mathbb{G}$ as follows: for any $x \in \mathbb{G}$ and $r > 0$,

$$\beta_K(B(x,r)) = \inf_L \sup_{z \in K \cap B(x,r)} \frac{d(z,L)}{r}.$$

Here, $B(x,r)$ is the closed ball, and the infimum is taken over all possible *horizontal lines*

$$L = \{p\delta_s(\tilde{\pi}(p^{-1}q)) \,:\, s \in \mathbb{R}\} \quad \text{where } p, q \in \mathbb{G}.$$

The following is the famous Baker-Campbell-Hausdorff formula.

Theorem 2.2 (Dynkin, '47)**.** *Suppose $e^{(\cdot)} : \mathfrak{g} \to \mathbb{G}$ is the exponential map. Given $X, Y \in \mathfrak{g}$, choose Z such that $e^Z = e^X e^Y$. Then*

$$Z = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \sum P(r_1, s_1, \ldots, r_k, s_k) [\underbrace{X, \cdots [X}_{r_1}, [\underbrace{Y, \cdots [Y}_{s_1}, \cdots [\underbrace{X, \cdots [X}_{r_k}, [\underbrace{Y, \cdots Y}_{s_k}]] \cdots]$$

where the second sum is taken over all $\{r_1, s_1, \ldots, r_k, s_k\} \in \mathbb{N}^{2k}$ satisfying $r_i + s_i > 0$ for $i = 1, \ldots, k$, and

$$P(r_1, s_1, \ldots, r_k, s_k) = \frac{1}{\sum_{i=1}^k (r_i + s_i) \prod_{i=1}^k r_i! s_i!}.$$

Notice that the nested commutators vanish if $s_k > 1$ or if $s_k = 0$ and $r_k > 1$. Also, since $\mathbb{G}$ is nilpotent, the first sum terminates after finitely many terms, and the length of the nested brackets is bounded from above. That is, there are only finitely many summed nested bracket terms in the BCH formula for a Carnot group $\mathbb{G}$.

We will later make use of the following estimates established in [17] (for a verification, see the proof of Lemma 3.2). Choose $a, b \in \mathbb{G}$ with $|a| < 1$ and $|b| < 1$. Write $a = (a_1, a_2)$ and $b = (b_1, b_2)$ as above. Then

$$ab = (a_1 + b_1, a_2 + b_2 + R(a,b))$$

for some polynomial R given by the BCH formula (Theorem 2.2). Write $R_1(a,b) = R((a_1, 0), (b_1, 0))$ and $R_2 = R - R_1$. Then the BCH formula gives

$$|R_2(a,b)| \leq C_2(|a_1||b_2| + |a_2||b_1| + |a_2||b_2|) \tag{2.4}$$

and

$$|R_1(a,b)| \leq C_1 |a_1||b_1| \left| \frac{a_1}{|a_1|} - \frac{b_1}{|b_1|} \right|. \tag{2.5}$$

for some constants C_1 and C_2 depending only on the group structure of $\mathbb{G}$.

For the remainder of the paper, fix a positive constant $\eta < \min\{\varepsilon_0, \frac{1}{2}\}$ (where ε_0 is as in Theorem 2.1) such that, if $|a| \le \eta$ and $|b| \le \eta$, then

$$(C_2+1)(|a_1|+|a_2|+|b_1|+|b_2|) \le \frac{1}{8} \tag{2.6}$$

and

$$(5C_1^2+1)|a_1||b_1| \le \frac{1}{2}. \tag{2.7}$$

Throughout the paper, we will write $a \lesssim b$ to indicate that there is a constant $C > 0$ depending only on the metric space $(\mathbb{G}, d)$ satisfying $a \le Cb$. Similarly, we will write $a \lesssim_\xi b$ if the constant depends also on some other parameter ξ.

3. Curvature bound in a Carnot group

For $p, q \in \mathbb{G}$, denote the *horizontal segment* between them as

$$L_{pq} := \{p\, \delta_t(\tilde{\pi}(p^{-1}q)) \,:\, t \in [0,1]\}.$$

While this segment will always originate at p, it will not intersect q in general. Note also that horizontal segments do not necessarily coincide with Euclidean segments if the step of $\mathbb{G}$ is $r > 2$. For each $t \in [0,1]$, write $L_{pq}(t) = p\, \delta_t(\tilde{\pi}(p^{-1}q))$. If $p = 0$, then $L_q := L_{0q} \subset \mathbb{R}^n \times \{0\}$ is the segment

$$L_q = \{\delta_t(\tilde{\pi}(q)) \,:\, t \in [0,1]\} = \{(tq_1, 0) \,:\, t \in [0,1]\}.$$

That is, L_q is simply the Euclidean line segment from the origin to $\tilde{\pi}(q) = (q_1, 0)$. Our goal in this section will be to prove the following curvature estimate in $\mathbb{G}$, which should be thought of as a lower bound on the excess of the triangle inequality. We remark that the use of the term *curvature bound* is partly motivated by the Menger curvature which turned out to be a very powerful tool in the study of the Cauchy transform, see e.g. [31] and the references therein.

Here, we fix the value $m = 2^{-217}$. (This value will be important in Section 4. The theorem actually holds for any $0 < m < 1$, but then the constant C_0 would depend also on m.)

Theorem 3.1. *Suppose $a, z, v, w \in \mathbb{G}$ satisfy*

$$m\rho \le \min\{d(a,z), d(a,v), d(z,v), d(v,w)\}$$

and

$$\max\{d(a,z), d(a,v), d(z,v), d(v,w), d(a,w)\} \le \rho$$

for some $\rho > 0$. Then there is a constant $C_0 = C_0(\mathbb{G}) > 0$ so that

$$\sup_{t\in[0,1]} d(L_{av}(t), L_{aw})^{2r^2} + \sup_{t\in[0,1]} d(L_{vw}(t), L_{aw})^{2r^2} \le C_0 \rho^{2r^2-1} \Delta$$

where $\Delta := d(a,z) + d(z,v) + d(v,w) - d(a,w)$.

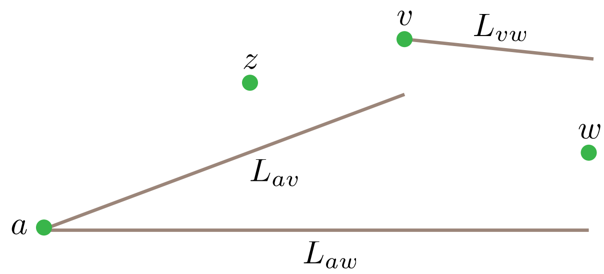

FIGURE 1.

3.1. **Preliminary lemmas.** We will need the estimates from the following two lemmas in the proof of Lemma 3.4. Recall that we have set $m = 2^{-217}$.

Lemma 3.2. *For any $a = (a_1, a_2)$ and $b = (b_1, b_2)$ in $\mathbb{G}$ with $|a|, |b| \in \big(\eta(\frac{m}{2})^r, 1\big)$, we have*

$$\|NH(ab)\|^r \lesssim |a_1||b_1| \left|\frac{a_1}{|a_1|} - \frac{b_1}{|b_1|}\right| + |a_2|(|b_1| + |b_2|) + |b_2|(|a_1| + |a_2|).$$

In particular, we will use the fact that

$$\alpha\|NH(ab)\|^{2r} \leq \left(|a_1||b_1| \left|\frac{a_1}{|a_1|} - \frac{b_1}{|b_1|}\right|\right)^2 + (|a_2|(|b_1| + |b_2|) + |b_2|(|a_1| + |a_2|))^2 \tag{3.1}$$

for some $0 < \alpha < 1$ depending only on the metric and group structure of $\mathbb{G}$.

Proof. We may write $a = e^A$ and $b = e^B$ for $A, B \in \mathfrak{g}$. In other words, $A = \log a$ and $B = \log b$. Write $A = A_1 + A_2$ and $B = B_1 + B_2$ where $A_1, B_1 \in V_1$ lie in the horizontal (first) layer of $\mathfrak{g}$. According to the Baker-Campbell-Hausdorff formula (Theorem 2.2) and the bilinearity of the Lie bracket, $\log(ab)$ is a finite sum of constant multiples of

$$[Z_1, [Z_2, \cdots, [Z_{k-2}, [Z_{k-1}, Z_k]] \cdots]] \tag{3.2}$$

for some positive integer $k \leq r$ where each Z_i is one of A_1, B_1, A_2, or B_2. In particular, $[Z_{k-1}, Z_k]$ must have the form

$$[A_1, B_1], \quad [A_1, B_2], \quad [A_2, B_1], \quad \text{or} \quad [A_2, B_2].$$

Indeed, we must have $s_n = 1$ or we have $s_n = 0$ and $r_n = 1$ (for otherwise the brackets vanish), so the nested brackets (3.2) must have the form

$$[\cdot, [\cdot, \cdots, [A, B] \cdots]] = [\cdot, [\cdot, \cdots, [A_1 + A_2, B_1 + B_2] \cdots]] = \sum_{i=1}^{2}\sum_{j=1}^{2} [\cdot, [\cdot, \cdots, [A_i, B_j] \cdots]]$$

(since $[A, B] = -[B, A]$).

By definition, we have

$$|NH(ab)| = |\tilde{\pi}(ab)^{-1}(ab)| = |(-a_1 - b_1, 0)(a_1 + b_1, a_2 + b_2 + P(a, b))|$$

for some polynomial P (given by the BCH formula). Thus by a similar argument as above, $\log(NH(ab))$ equals $A_2 + B_2$ plus a finite sum of constant multiples of nested brackets (3.2) each of which ends with a term of the form

$$[A_1, B_1], \quad [A_1, B_2], \quad [A_2, B_1], \quad [A_2, B_2], \quad [A_1, A_2], \quad \text{or} \quad [B_1, B_2].$$

Consider the norm $|\cdot|$ on $\mathfrak{g}$ induced by the Euclidean norm on the exponential coordinates $\mathbb{R}^N$, i.e. for $X \in \mathfrak{g}$ with $e^X = x \in \mathbb{G}$, we have $|X| = |x|$. Since

$$[A_1, B_1] = |A_1||B_1| \left[\frac{A_1}{|A_1|}, \frac{B_1}{|B_1|}\right] = |A_1||B_1| \left[\frac{A_1}{|A_1|} - \frac{B_1}{|B_1|}, \frac{B_1}{|B_1|}\right],$$

the bilinearity of the Lie bracket gives the following bound:

$$|[A_1, B_1]| \lesssim |A_1||B_1| \left|\frac{A_1}{|A_1|} - \frac{B_1}{|B_1|}\right|.$$

Thus, for those brackets (3.2) ending with $[A_1, B_1]$, we have

$$|[Z_1, [Z_2, \cdots, [Z_{k-2}, [Z_{k-1}, Z_k]] \cdots]]| \lesssim \left(\prod_{i=1}^{k-2} |Z_i|\right) |[A_1, B_1]| \lesssim |a_1||b_1| \left|\frac{a_1}{|a_1|} - \frac{b_1}{|b_1|}\right|$$

since $|a| < 1$ and $|b| < 1$.

Now write $c = \eta(\frac{m}{2})^r$. Since $|a| > c$, it follows that $|a_1| + |a_2| > c$. Similarly, $|b_1| + |b_2| > c$. Therefore, all other nested brackets (3.2) which do not end with $[A_1, B_1]$ satisfy

$$|[Z_1, [Z_2, \cdots, [Z_{k-2}, [Z_{k-1}, Z_k]] \cdots]]| \lesssim \prod_{i=1}^{k} |Z_i| \lesssim |a_2|(|b_1| + |b_2|) + |b_2|(|a_1| + |a_2|)$$

since, for example,

$$|a_1||a_2| \leq |a_2| \leq c^{-1}|a_2|(|b_1| + |b_2|).$$

Since we may estimate $|NH(ab)| = |\log(NH(ab))|$ by a finite sum of constant multiples of the nested brackets (3.2) plus $|a_2| + |b_2|$, we have proven

$$|NH(ab)| \lesssim |a_1||b_1| \left|\frac{a_1}{|a_1|} - \frac{b_1}{|b_1|}\right| + |a_2|(|b_1| + |b_2|) + |b_2|(|a_1| + |a_2|).$$

Hence there is some compact set $K \subset \mathbb{G}$ (depending only on the group structure and metric) so that $NH(ab) \in K$ for any a and b in the Euclidean unit ball. That is, we may apply (2.2) to conclude $\|NH(ab)\|^r \lesssim |NH(ab)|$. This completes the proof. $\square$

The following lemma is entirely Euclidean in nature and elementary. The details of the proof are included for completeness. Here and in what follows, we will assume that the quantity $|c||d|\left|\frac{c}{|c|} - \frac{d}{|d|}\right|$ vanishes whenever $c = 0$ or $d = 0$.

Lemma 3.3. *Fix $c, d \in \mathbb{R}^n$. Let ℓ_{c+d} denote the segment from the origin to $c + d$. Then*

$$d_{\mathbb{R}^n}(c, \ell_{c+d})^2 \leq \tfrac{1}{2}|c||d| \left|\frac{c}{|c|} - \frac{d}{|d|}\right|^2.$$

Proof. If $c + d = 0$, then $d = -c$ and the result follows. Thus we may assume $c + d \neq 0$. We will make frequent use of the following consequence of the polarization identity:

$$|c||d| \left|\frac{c}{|c|} - \frac{d}{|d|}\right|^2 = |c|^2 + 2|c||d| + |d|^2 - |c + d|^2. \tag{3.3}$$

Let u denote the scalar projection of c along $c+d$. That is, $u = \frac{\langle c, c+d\rangle}{|c+d|}$. If $u \leq 0$, then[1] the closest point in ℓ_{c+d} to c is the origin, so $d_{\mathbb{R}^n}(c, \ell_{c+d}) = |c|$. We then have[2]

$$2d_{\mathbb{R}^n}(c, \ell_{c+d})^2 = 2|c|^2 \leq |c|^2 + |d|^2 - |c+d|^2 \leq |c||d| \left| \frac{c}{|c|} - \frac{d}{|d|} \right|^2.$$

If $u \geq |c+d|$, then the closest point in ℓ_{c+d} to c is $c+d$. Since[3] $|d|^2 \leq |c|^2 - |c+d|^2$, we have

$$2d_{\mathbb{R}^n}(c, \ell_{c+d})^2 = 2|c-(c+d)|^2 = 2|d|^2 = |d|^2 + |d|^2 \leq |c|^2 + |d|^2 - |c+d|^2 \leq |c||d| \left| \frac{c}{|c|} - \frac{d}{|d|} \right|^2.$$

Now suppose $0 < u < |c+d|$. That is, the projection of c to the line containing ℓ_{c+d} actually lies in ℓ_{c+d}. Since this projection divides ℓ_{c+d} into segments of length u and $|c+d| - u$, the Pythagorean Theorem gives

$$\begin{aligned}
|c||d| \left| \frac{c}{|c|} - \frac{d}{|d|} \right|^2 &= |c|^2 + 2|c||d| + |d|^2 - |c+d|^2 \\
&= \left[|c|^2 - u^2\right] + \left[|d|^2 - (|c+d| - u)^2\right] + 2|c||d| + 2u^2 - 2u|c+d| \\
&= d_{\mathbb{R}^n}(c, \ell_{c+d})^2 + d_{\mathbb{R}^n}(c, \ell_{c+d})^2 + 2|c||d| + 2u(u - |c+d|) \\
&= 2d_{\mathbb{R}^n}(c, \ell_{c+d})^2 + 2(|c||d| - u|d| + u|d| + u(u - |c+d|)) \\
&= 2d_{\mathbb{R}^n}(c, \ell_{c+d})^2 + 2|d|(|c| - u) + 2u(|d| - (|c+d| - u)) \\
&\geq 2d_{\mathbb{R}^n}(c, \ell_{c+d})^2
\end{aligned}$$

since $(|c+d| - u)^2 \leq (|c+d| - u)^2 + d_{\mathbb{R}^n}(c, \ell_{c+d})^2 = |d|^2$ and $u \leq |c|$ by the Cauchy-Schwartz inequality. □

The technical proof of this next lemma follows the example of the proof from [17] that the HS-norm is subadditive. By tightening some of the bounds from [17, Theorem 2], we are able to estimate the error in the subadditivity of the norm. Again, we have set $m = 2^{-217}$, but this lemma actually holds for any $m \in [0, 1]$.

Lemma 3.4. *Suppose $x, y \in \mathbb{G}$ satisfy $\|x\|, \|y\| \in [m\rho, \rho]$ for some $\rho > 0$. Then*

$$\frac{m^{2r}}{2^{2r+4}} \left(\frac{\alpha \|NH(xy)\|^{2r}}{(\|x\| + \|y\|)^{2r-1}} + \frac{d_{\mathbb{R}^n}(x_1, \ell_{x_1+y_1})^2}{\|x\| + \|y\|} \right) \leq \|x\| + \|y\| - \|xy\|. \tag{3.4}$$

Proof. We will write A to respresent the left hand side of (3.4). It will be important throughout the proof that we have $4A \leq \min\{\|x\|, \|y\|\}$. This is indeed true since

$$\|NH(xy)\| = d(xy, \tilde{\pi}(xy)) \leq \|xy\| + \|\tilde{\pi}(xy)\| \leq \|xy\| + \|xy\| \leq 2(\|x\| + \|y\|) \leq 4\rho$$

and since

$$d_{\mathbb{R}^n}(x_1, \ell_{x_1+y_1}) \leq |x_1 - (x_1 + y_1)| = |y_1| \leq \rho,$$

and thus we have

$$4A \leq \frac{m^{2r}}{2^{2r+2}} \left(\frac{(4\rho)^{2r}}{(m\rho + m\rho)^{2r-1}} + \frac{\rho^2}{m\rho + m\rho} \right) = \frac{m\rho}{2} + \frac{m^{2r-1}\rho}{2^{2r+3}} \leq m\rho \leq \min\{\|x\|, \|y\|\}$$

[1] Indeed, if $u \leq 0$, then the angle between the vector c and ℓ_{c+d} is between $\pi/2$ and $3\pi/2$.

[2] The assumption $u \leq 0$ implies $|c|^2 + \langle c, d\rangle = \langle c, c\rangle + \langle c, d\rangle = \langle c, c+d\rangle \leq 0$. Hence the polarization identity yields $2|c|^2 \leq -2\langle c, d\rangle = |c|^2 + |d|^2 - |c+d|^2$.

[3] The polarization identity gives $|d|^2 = |c|^2 + |c+d|^2 - 2\langle c, c+d\rangle \leq |c|^2 + |c+d|^2 - 2|c+d|^2$ since the assumption $u \geq |c+d|$ implies $-2\langle c, c+d\rangle \leq -2|c+d|^2$.

as desired.

Set $a = \delta_{1/\|x\|}(x)$ and $b = \delta_{1/\|y\|}(y)$ so that $|a| = |b| = \eta$. Write

$$s = \frac{\|x\|}{\|x\| + \|y\| - A} \quad \text{and} \quad t = \frac{\|y\|}{\|x\| + \|y\| - A}.$$

Note that $s, t \in (0,1)$ since $4A \leq \min\{\|x\|, \|y\|\}$. These values have been chosen so that

$$\left|\delta_{\frac{1}{\|x\|+\|y\|-A}}(xy)\right| = \left|\delta_{\frac{\|x\|}{\|x\|+\|y\|-A}}\left(\delta_{\frac{1}{\|x\|}}(x)\right) \cdot \delta_{\frac{\|y\|}{\|x\|+\|y\|-A}}\left(\delta_{\frac{1}{\|y\|}}(y)\right)\right| = \left|\delta_s\,(a)\,\delta_t\,(b)\right|.$$

Write $a = (a_1, a_2)$ and $b = (b_1, b_2)$ as before. For ease of notation, we write $\delta_\lambda(a_1) = \delta_\lambda((a_1, 0))$ and $\delta_\lambda(a_2) = \delta_\lambda((0, a_2))$ for $\lambda > 0$, and we similarly use the shorthand $\delta_\lambda(b_1)$ and $\delta_\lambda(b_2)$. Write $v = \delta_s(a_2) + \delta_t(b_2) + (0, R_2(\delta_s(a), \delta_t(b)))$ where R_2 is as defined in Section 2. The bounds (2.4) and (2.6) give

$$\begin{aligned}
|v| &\overset{(2.4)}{\leq} s^2|a_2| + t^2|b_2| + C_2\left(|\delta_s(a_1)||\delta_t(b_2)| + |\delta_s(a_2)||\delta_t(b_1)| + |\delta_s(a_2)||\delta_t(b_2)|\right) \\
&\leq s^2|a_2| + t^2|b_2| + st(C_2+1)\left(|a_2|(|b_1| + |b_2|) + |b_2|(|a_1| + |a_2|)\right) \\
&\qquad\qquad - \left(|\delta_s(a_2)|(|\delta_t(b_1)| + |\delta_t(b_2)|) + |\delta_t(b_2)|(|\delta_s(a_1)| + |\delta_s(a_2)|)\right) \\
&\overset{(2.6)}{\leq} s^2|a_2| + t^2|b_2| + \tfrac{1}{4}st\left(|a_2| + |b_2|\right) \\
&\qquad\qquad - \left(|\delta_s(a_2)|(|\delta_t(b_1)| + |\delta_t(b_2)|) + |\delta_t(b_2)|(|\delta_s(a_1)| + |\delta_s(a_2)|)\right) \\
&\leq s|a_2| + t|b_2| - \tfrac{1}{2}st\left(|a_2| + |b_2|\right) \\
&\qquad\qquad - \left(|\delta_s(a_2)|(|\delta_t(b_1)| + |\delta_t(b_2)|) + |\delta_t(b_2)|(|\delta_s(a_1)| + |\delta_s(a_2)|)\right).
\end{aligned}$$

This last inequality follows from the following argument. Since $4A \leq \|y\|$, we have

$$\tfrac{3}{4}t - (1-s) = \frac{\frac{3}{4}\|y\|}{\|x\| + \|y\| - A} - \frac{\|y\| - A}{\|x\| + \|y\| - A} = \frac{A - \frac{1}{4}\|y\|}{\|x\| + \|y\| - A} \leq 0$$

so $(\frac{3}{4}t + s) \leq 1$ (and similarly $(\frac{3}{4}s + t) \leq 1$), and thus

$$s^2|a_2| + t^2|b_2| - s|a_2| - t|b_2| + \tfrac{3}{4}st\left(|a_2| + |b_2|\right) = s|a_2|\left((\tfrac{3}{4}t + s) - 1\right) + t|b_2|\left((\tfrac{3}{4}s + t) - 1\right) \leq 0.$$

Moreover, the inequalities $(\frac{3}{4}t + s) \leq 1$ and $(\frac{3}{4}s + t) \leq 1$ imply that

$$|v| \leq s^2|a_2| + t^2|b_2| + \tfrac{1}{4}st\left(|a_2| + |b_2|\right) \leq |a_2| + |b_2|.$$

Hence

$$\begin{aligned}
|v|^2(1 + st) &\leq |v|^2 + st|v|\left(|a_2| + |b_2|\right) \\
&\leq \left(|v| + \tfrac{1}{2}st\left(|a_2| + |b_2|\right)\right)^2 \\
&\leq \left(s|a_2| + t|b_2| - \left(|\delta_s(a_2)|(|\delta_t(b_1)| + |\delta_t(b_2)|) + |\delta_t(b_2)|(|\delta_s(a_1)| + |\delta_s(a_2)|)\right)\right)^2 \\
&\leq \left(s|a_2| + t|b_2|\right)^2 - \left(|\delta_s(a_2)|(|\delta_t(b_1)| + |\delta_t(b_2)|) + |\delta_t(b_2)|(|\delta_s(a_1)| + |\delta_s(a_2)|)\right)^2.
\end{aligned}$$

This last inequality follows from the fact that $(k - \ell)^2 \leq k^2 - \ell^2$ when $\ell \leq k$ and

$$\begin{aligned}
|\delta_s(a_2)|(|\delta_t(b_1)| + |\delta_t(b_2)|) + |\delta_t(b_2)|(|\delta_s(a_1)| + |\delta_s(a_2)| \\
&\leq st\left(|a_2|(|b_1| + |b_2|) + |b_2|(|a_1| + |a_2|)\right) \\
&\leq s|a_2| + t|b_2|
\end{aligned}$$

since $|a| = |b| = \eta < \frac{1}{2}$. Therefore (3.3) together with the fact that $2\langle u_1, u_2\rangle \leq st|u_1|^2 + \frac{|u_2|^2}{st}$ for any $u_1, u_2 \in \mathbb{R}^N$ gives

$$\begin{aligned}
|\delta_s(a)\delta_t(b)|^2 &= |\delta_s(a_1) + \delta_t(b_1)|^2 + |v + (0, R_1(\delta_s(a), \delta_t(b)))|^2 \\
&\leq (|\delta_s(a_1)| + |\delta_t(b_1)|)^2 - |\delta_s(a_1)||\delta_t(b_1)| \left| \frac{\delta_s(a_1)}{|\delta_s(a_1)|} - \frac{\delta_t(b_1)}{|\delta_t(b_1)|} \right|^2 \\
&\qquad + |v|^2 (1 + st) + |R_1(\delta_s(a), \delta_t(b))|^2 \left(1 + \frac{1}{st}\right) \\
&\overset{(2.5)}{\leq} (s|a_1| + t|b_1|)^2 + (s|a_2| + t|b_2|)^2 - |\delta_s(a_1)||\delta_t(b_1)| \left| \frac{\delta_s(a_1)}{|\delta_s(a_1)|} - \frac{\delta_t(b_1)}{|\delta_t(b_1)|} \right|^2 \\
&\qquad - (|\delta_s(a_2)|(|\delta_t(b_1)| + |\delta_t(b_2)|) + |\delta_t(b_2)|(|\delta_s(a_1)| + |\delta_s(a_2)|))^2 \\
&\qquad + \left(1 + \frac{1}{st}\right) C_1^2 |\delta_s(a_1)|^2 |\delta_t(b_1)|^2 \left| \frac{\delta_s(a_1)}{|\delta_s(a_1)|} - \frac{\delta_t(b_1)}{|\delta_t(b_1)|} \right|^2 .
\end{aligned}$$

Since

$$|\delta_s(a)| \geq s^r |a| > \eta \left(\frac{\|x\|}{\|x\| + \|y\|} \right)^r \geq \eta \left(\frac{m\rho}{\rho + \rho} \right)^r = \eta \left(\frac{m}{2} \right)^r$$

(and similarly $|\delta_t(b)| > \eta(\frac{m}{2})^r$), we may apply (3.1) to obtain

$$\begin{aligned}
&- (|\delta_s(a_2)|(|\delta_t(b_1)| + |\delta_t(b_2)|) + |\delta_t(b_2)|(|\delta_s(a_1)| + |\delta_s(a_2)|))^2 \\
&\qquad \leq \left(|\delta_s(a_1)||\delta_t(b_1)| \left| \frac{\delta_s(a_1)}{|\delta_s(a_1)|} - \frac{\delta_t(b_1)}{|\delta_t(b_1)|} \right| \right)^2 - \alpha \left\| NH(\delta_s(a)\delta_t(b)) \right\|^{2r}
\end{aligned}$$

and thus, by the Cauchy-Schwarz inequality,

$$\begin{aligned}
&|\delta_s(a)\delta_t(b)|^2 \\
&\quad \leq (s|a| + t|b|)^2 - |\delta_s(a_1)||\delta_t(b_1)| \left| \frac{\delta_s(a_1)}{|\delta_s(a_1)|} - \frac{\delta_t(b_1)}{|\delta_t(b_1)|} \right|^2 \\
&\qquad + \left(\left(1 + \frac{1}{st}\right) C_1^2 + 1 \right) |\delta_s(a_1)|^2 |\delta_t(b_1)|^2 \left| \frac{\delta_s(a_1)}{|\delta_s(a_1)|} - \frac{\delta_t(b_1)}{|\delta_t(b_1)|} \right|^2 - \alpha \left\| NH(\delta_s(a)\delta_t(b)) \right\|^{2r} \\
&\quad \leq (s|a| + t|b|)^2 - \alpha \left\| NH(\delta_s(a)\delta_t(b)) \right\|^{2r} \\
&\qquad + \left(\left((st + 1) C_1^2 + st \right) |a_1||b_1| - 1 \right) |\delta_s(a_1)||\delta_t(b_1)| \left| \frac{\delta_s(a_1)}{|\delta_s(a_1)|} - \frac{\delta_t(b_1)}{|\delta_t(b_1)|} \right|^2 \\
&\quad \overset{(2.7)}{\leq} (s|a| + t|b|)^2 - \alpha \left\| NH(\delta_s(a)\delta_t(b)) \right\|^{2r} - \tfrac{1}{2} |\delta_s(a_1)||\delta_t(b_1)| \left| \frac{\delta_s(a_1)}{|\delta_s(a_1)|} - \frac{\delta_t(b_1)}{|\delta_t(b_1)|} \right|^2 \\
&\quad \leq (s|a| + t|b|)^2 - \alpha \left\| NH(\delta_s(a)\delta_t(b)) \right\|^{2r} - d_{\mathbb{R}^n}(sa_1, \ell_{sa_1 + tb_1})^2
\end{aligned}$$

by Lemma 3.3. Since $sa_1 = \frac{x_1}{\|x\| + \|y\| - A}$, $tb_1 = \frac{y_1}{\|x\| + \|y\| - A}$, and $|\delta_s(a)\,\delta_t(b)| = \left| \delta_{\frac{1}{\|x\| + \|y\| - A}}(xy) \right|$, we have

$$\begin{aligned}
|\delta_s(a)\delta_t(b)| - (s|a| + t|b|) &\leq \frac{-\alpha \| NH(\delta_s(a)\delta_t(b)) \|^{2r} - d_{\mathbb{R}^n}(sa_1, \ell_{sa_1 + tb_1})^2}{|\delta_s(a)\delta_t(b)| + (s|a| + t|b|)} \\
&= \frac{-\alpha \left\| NH\left(\delta_{\frac{1}{\|x\| + \|y\| - A}}(xy) \right) \right\|^{2r} - \frac{d_{\mathbb{R}^n}(x_1, \ell_{x_1 + y_1})^2}{(\|x\| + \|y\| - A)^2}}{|\delta_s(a)\delta_t(b)| + (s|a| + t|b|)}
\end{aligned}$$

$$= \frac{-\alpha \frac{\|NH(xy)\|^{2r}}{(\|x\|+\|y\|-A)^{2r}} - \frac{d_{\mathbb{R}^n}(x_1,\ell_{x_1+y_1})^2}{(\|x\|+\|y\|-A)^2}}{|\delta_s(a)\delta_t(b)| + (s|a|+t|b|)} = (*).$$

We proved in particular above that $|\delta_s(a)\delta_t(b)| \leq s|a| + t|b|$. Since $|a| = |b| = \eta$, this gives $|\delta_s(a)\delta_t(b)| \leq 2\eta < 2^{2r+3}m^{-2r}/\eta$. Thus

$$\begin{aligned}
(*) &\leq -\frac{\eta m^{2r}}{2^{2r+4}} \left(\frac{\alpha \|NH(xy)\|^{2r}}{(\|x\|+\|y\|-A)^{2r}} + \frac{d_{\mathbb{R}^n}(x_1,\ell_{x_1+y_1})^2}{(\|x\|+\|y\|-A)^2} \right) \\
&\leq -\frac{\eta m^{2r}}{2^{2r+4}} \left(\frac{\alpha \|NH(xy)\|^{2r}}{(\|x\|+\|y\|)^{2r-1}} + \frac{d_{\mathbb{R}^n}(x_1,\ell_{x_1+y_1})^2}{\|x\|+\|y\|} \right) (\|x\|+\|y\|-A)^{-1}.
\end{aligned}$$

Since $|a| = |b| = \eta$, the definitions of s and t give

$$s|a| + t|b| = \frac{\|x\|}{\|x\|+\|y\|-A}|a| + \frac{\|y\|}{\|x\|+\|y\|-A}|b| = \frac{\eta(\|x\|+\|y\|)}{\|x\|+\|y\|-A}.$$

In other words,

$$\left|\delta_{\frac{1}{\|x\|+\|y\|-A}}(xy)\right| = |\delta_s(a)\delta_t(b)| \leq \eta \left(\frac{\|x\|+\|y\| - \frac{m^{2r}}{2^{2r+4}}\left(\frac{\alpha\|NH(xy)\|^{2r}}{(\|x\|+\|y\|)^{2r-1}} + \frac{d_{\mathbb{R}^n}(x_1,\ell_{x_1+y_1})^2}{\|x\|+\|y\|} \right)}{\|x\|+\|y\|-A} \right) = \eta$$

according to the definition of A. Therefore, the definition of the HS norm gives

$$\|xy\| \leq \|x\| + \|y\| - \frac{m^{2r}}{2^{2r+4}} \left(\frac{\alpha \|NH(xy)\|^{2r}}{(\|x\|+\|y\|)^{2r-1}} + \frac{d_{\mathbb{R}^n}(x_1,\ell_{x_1+y_1})^2}{\|x\|+\|y\|} \right). \quad \square$$

Corollary 3.5. *Fix $v, w \in \mathbb{G}$. Suppose $v, w \in \mathbb{G}$ satisfy $\|v\|, \|v^{-1}w\| \in [m\rho, \rho]$ for some $\rho > 0$. Then*

$$\frac{m^{2r}}{2^{2r+4}} \left(\frac{\alpha \|NH(w)\|^{2r}}{(d(0,v)+d(v,w))^{2r-1}} + \frac{d_{\mathbb{R}^n}(v_1,\ell_{w_1})^2}{d(0,v)+d(v,w)} \right) \leq d(0,v) + d(v,w) - d(0,w). \tag{3.5}$$

Proof. Write $x = v$ and $y = v^{-1}w$ in Lemma 3.4. $\square$

3.2. **Proof of Theorem 3.1.** We are now ready to prove Theorem 3.1. This theorem controls the deviation of horizontal segments by the excess in a four point triangle inequality. We restate it here for convenience. Again, we have fixed $m = 2^{-217}$.

Theorem 3.1. *Suppose $a, z, v, w \in \mathbb{G}$ satisfy*

$$m\rho \leq \min\{d(a,z), d(a,v), d(z,v), d(v,w)\}$$

and

$$\max\{d(a,z), d(a,v), d(z,v), d(v,w), d(a,w)\} \leq \rho$$

for some $\rho > 0$. Then there is a constant $C_0 = C_0(\mathbb{G}) > 0$ so that

$$\sup_{t\in[0,1]} d(L_{av}(t), L_{aw})^{2r^2} + \sup_{t\in[0,1]} d(L_{vw}(t), L_{aw})^{2r^2} \leq C_0 \rho^{2r^2-1}\Delta$$

where $\Delta := d(a,z) + d(z,v) + d(v,w) - d(a,w)$.

We may assume without loss of generality from here on out that $a = 0$. Indeed, the metric is left invariant, and horizontal segments commute with left multiplication in the following sense:

$$c^{-1}L_{ab}(t) = c^{-1}a\delta_t(\tilde{\pi}(a^{-1}b)) = c^{-1}a\delta_t(\tilde{\pi}((c^{-1}a)^{-1}(c^{-1}b))) = L_{(c^{-1}a)(c^{-1}b)}(t)$$

for any $b, c \in \mathbb{G}$ and any $t \in [0,1]$. We will first establish the important tools used in the proof of the theorem.

Lemma 3.6. *Under the assumptions of Theorem 3.1, we have*

$$\rho^{-(2r-1)}\|NH(w)\|^{2r} + \rho^{-1}d_{\mathbb{R}^n}(v_1, \ell_{w_1})^2 \lesssim d(0,v) + d(v,w) - d(0,w) \tag{3.6}$$

and

$$\rho^{-(2r-1)}\|NH(v)\|^{2r} + \rho^{-1}d_{\mathbb{R}^n}(z_1, \ell_{v_1})^2 \lesssim d(0,z) + d(z,v) - d(0,v). \tag{3.7}$$

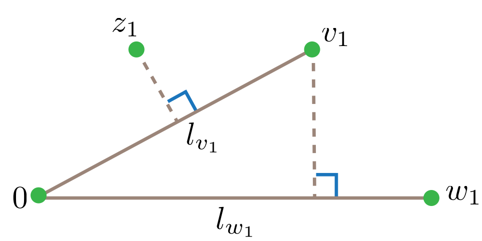

Figure 2.

Proof. Inequality (3.6) follows immediately from Corollary 3.5. Inequality (3.7) follows from Corollary 3.5 as well with z in place of v and v in place of w. $\square$

Consider the following version of [21, Lemma 3.6]:

Lemma 3.7. *Fix a constant $M \geq 1$. For any $0 < \omega < M$, if $g, h, p, q \in \mathbb{G}$ satisfy $d(g,h) \leq \omega$, $\|p\| \leq 1$ and $d(p,q) \leq \omega$, then*

$$\sup_{t\in[0,1]} d(g\delta_t(p), h\delta_t(q)) \leq C'\omega^{1/r}$$

for some constant $C' = C'(\mathbb{G}, M) > 0$.

In other words, given two horizontal segments whose starting points and "directions" are close, points along the segments are close as well. Note that the original lemma in [21] is stated for $0 < \omega < 1$, but the triangle inequality gives

$$\begin{aligned} d(g\delta_t(p), h\delta_t(q)) \leq d(g\delta_t(p), g) + d(g,h) + d(h, h\delta_t(q)) &\leq 1 + \omega + (1+\omega) \\ &\leq 4\omega \leq (4M^{1-\frac{1}{r}})\omega^{1/r} \end{aligned}$$

when $1 \leq \omega < M$. We will now use this lemma to prove Theorem 3.1.

Proof of Theorem 3.1. It suffices to prove the theorem in the case $\rho = 1$. Indeed, given arbitrary a, z, v, w satisfying the assumptions of the theorem, their dilations $\delta_{1/\rho}(a), \delta_{1/\rho}(z), \delta_{1/\rho}(v), \delta_{1/\rho}(w)$ satisfy the assumptions with $\rho = 1$, and horizontal lines commute with dilations in the following sense: $L_{\delta_s(p)\delta_s(q)}(t) = \delta_s(L_{pq}(t))$.

Note that $\tilde{\pi}(v) \in \mathbb{R}^n \times \{0\}$ and $L_w \subset \mathbb{R}^n \times \{0\}$. Say $p_v \in L_w$ is the nearest point in L_w to $\tilde{\pi}(v)$ (in the Euclidean norm).

We begin with the bound on $d(L_v(t), L_w)$. Since $\delta_s(p) = sp$ for any $p \in \mathbb{R}^n \times \{0\}$, we have

$$d(\tilde{\pi}(v), p_v)^{2r} \lesssim |\tilde{\pi}(v) - p_v|^2 = d_{\mathbb{R}^n}(v_1, \ell_{w_1})^2 \lesssim \Delta$$

according to (2.2) and (3.6). (Note that the constant from (2.2) here depends only on $\mathbb{G}$ since $\tilde{\pi}(v)$ and p_v lie in the unit ball of $\mathbb{G}$). In other words, $d(\tilde{\pi}(v), p_v) \leq C\Delta^{1/2r}$ for some constant $C = C(\mathbb{G}) > 0$. Therefore, we may apply Lemma 3.7 with $\omega = C\Delta^{1/2r}$ (and $M = C3^{1/2r}$) to get for any $t \in [0, 1]$

$$d(L_v(t), L_w) = d(\delta_t(\tilde{\pi}(v)), L_w) \leq d(\delta_t(\tilde{\pi}(v)), \delta_t(p_v)) \lesssim \Delta^{1/2r^2}.$$

We now bound $d(L_{vw}(t), L_w)$. Using (3.6), (3.7), and the above inequality, we have

$$\begin{aligned} d(v, p_v)^{2r} &\lesssim d(v, \tilde{\pi}(v))^{2r} + d(\tilde{\pi}(v), p_v)^{2r} \\ &= \|NH(v)\|^{2r} + d(\tilde{\pi}(v), p_v)^{2r} \\ &\lesssim \left[\|NH(v)\|^{2r} + d_{\mathbb{R}^n}(z_1, \ell_{v_1})^2\right] + \left[\|NH(w)\|^{2r} + d_{\mathbb{R}^n}(v_1, \ell_{w_1})^2\right] \\ &\lesssim [d(0, z) + d(z, v) - d(0, v)] + [d(0, v) + d(v, w) - d(0, w)] = \Delta. \end{aligned}$$

This is the bound on the distance between the starting points of our segments. We will now bound their "directions". That is, we will bound $d(\tilde{\pi}(v^{-1}w), \tilde{\pi}(p_v^{-1}w))$. Since the HS norm is invariant under rotations of $\mathbb{R}^n \times \{0\} \subset \mathbb{R}^N$ which fix the other $N - n$ coordinates, we may assume without loss of generality that the segment L_w lies along the x_1 axis in $\mathbb{R}^N$. Under this assumption, we have

$$\tilde{\pi}(w) = (w_1, 0) = (w_1^1, 0, 0), \quad \tilde{\pi}(v) = (v_1, 0) = (v_1^1, v_1^2, 0), \quad \text{and} \quad p_v = (p_1^1, 0, 0)$$

where $w_1^1, v_1^1, p_1^1 \in \mathbb{R}$, $w_1^1 > 0$, and $v_1^2 \in \mathbb{R}^{n-1}$. In particular, it follows that $d_{\mathbb{R}^n}(v_1, \ell_{w_1}) \geq |v_1^2|$. Thus

$$\|\tilde{\pi}(v^{-1}w)^{-1}\tilde{\pi}(p_v^{-1}w)\| = \|(-(w_1^1 - v_1^1), v_1^2, 0)(w_1^1 - p_1^1, 0, 0)\| = \|(v_1^1 - p_1^1, v_1^2, P)\|$$

where P is a BCH polynomial generated by w_1^1, v_1^1, p_1^1, and v_1^2. Arguing as in the proof of Lemma 3.2, we may see that the polynomial P is a finite sum of constant multiples of terms of the form

$$[Z_1, [Z_2, \cdots, [Z_{k-2}, [Z_{k-1}, Z_k]] \cdots]] \tag{3.8}$$

where each Z_i is either w_1^1, v_1^1, p_1^1, or v_1^2. (Note the abuse of notation here in which we identify $(w_1^1, 0, 0)$, $(v_1^1, 0, 0)$, $(p_1^1, 0, 0)$, and $(0, v_1^2, 0)$ with their associated vectors in the first layer of the Lie algebra.) Note that the points $(w_1^1, 0, 0)$, $(p_1^1, 0, 0)$, and $(v_1^1, 0, 0)$ are co-linear. Thus, $[w_1^1, v_1^1] = [w_1^1, p_1^1] = [p_1^1, v_1^1] = 0$. In particular, it follows that, in each non-vanishing term of the form (3.8), $[Z_{k-1}, Z_k]$ equals

$$[w_1^1, v_1^2], \quad [p_1^1, v_1^2], \quad \text{or} \quad [v_1^1, v_1^2].$$

Since $\eta < 1$, we have $|p_1^1| \leq |w_1^1| = \eta\|\tilde{\pi}(w)\| \leq d(0, w) \leq 1$, and similarly we get $|v_1^1| \leq 1$ and $|v_1^2| \leq 1$. Therefore,

$$|[Z_1, [Z_2, \cdots, [Z_{k-2}, [Z_{k-1}, Z_k]] \cdots]]| \lesssim \prod_{i=1}^{k} |Z_i| \leq |v_1^2| \leq d_{\mathbb{R}^n}(v_1, \ell_{w_1})$$

for each term of the form (3.8). Hence we may conclude

$$d(\tilde{\pi}(v^{-1}w), \tilde{\pi}(p_v^{-1}w))^{2r} \lesssim |(v_1^1 - p_1^1, v_1^2, P)|^2 \lesssim d_{\mathbb{R}^n}(v_1, \ell_{w_1})^2 \lesssim \Delta$$

since $|v_1^1 - p_1^1|$ vanishes if $v_1^1 \in [0, w_1^1]$ and is bounded by $d_{\mathbb{R}^n}(v_1, \ell_{w_1})$ otherwise. Once again, we apply Lemma 3.7 with $\omega = C\Delta^{1/2r}$ (for some (possibly different) constant $C = C(\mathbb{G}) > 0$) to get for any $t \in [0,1]$

$$d(L_{vw}(t), L_w) = d(v\delta_t(\tilde{\pi}(v^{-1}w)), L_w) \leq d(v\delta_t(\tilde{\pi}(v^{-1}w)), p_v\delta_t(\tilde{\pi}(p_v^{-1}w))) \lesssim \Delta^{1/2r^2}$$

since $p_v\delta_t(\tilde{\pi}(p_v^{-1}w)) = (tw_1^1 + (1-t)p_1^1, 0, 0)$ lies in L_w. This completes the proof of the theorem. □

4. Using Theorem 3.1 to prove Theorem 1.4

We will now apply the estimates from Section 3 to prove the Traveling Salesman Theorem (Theorem 1.4) in $\mathbb{G}$. In this section, we will follow the proof of the Traveling Salesman Theorem in the Heisenberg group [22, Theorem I] given in [22]. Many of the arguments therein hold in any metric space. As such, this section will provide a rough outline of the proof of Theorem 1.4. Full proofs will be provided for the results whose proofs differ significantly from those in [22].

4.1. **Preliminaries: arcs.** First, we will recall the notation from [22]. Fix a connected $\Gamma \subset \mathbb{G}$ with $\mathcal{H}^1(\Gamma) < \infty$. According to [27, Lemma 3.4], we may assume without loss of generality that Γ is compact. Fix also a 1-Lipschitz, arc-length parameterization $\gamma : \mathbb{T} \to \Gamma$ (where $\mathbb{T}$ is a circle in $\mathbb{R}^2$). Such a parameterization exists according to Lemma 2.10 in [22] and Corollary 3.8 in [27]. (The proof of this depends on the Banach space structure of the ambient space, but, since any metric space may be embedded isometrically into a Banach space via the Kuratowski embedding, the same arguments hold in $\mathbb{G}$.) Orient $\mathbb{T}$ so that we may discuss a particular direction of flow along Γ. Since β_Γ is scale invariant (i.e. $\beta_{\delta_\lambda(\Gamma)}(\delta_\lambda(B)) = \beta_\Gamma(B)$), we may assume without loss of generality that $\operatorname{diam}(\Gamma) = 1$.

An *arc* τ in γ is the restriction $\gamma|_{I_\tau}$ where $I_\tau = [a(\tau), b(\tau)]$ is some interval in $\mathbb{T}$ compatible with the orientation chosen above. Given two arcs τ and ζ, the notation $\zeta \subset \tau$ means $I_\zeta \subset I_\tau$, and we will write $\operatorname{diam}(\tau)$ to represent $\operatorname{diam}(\tau(I_\tau))$.

For any $L > 0$, we define a *prefiltration* $\mathcal{F}^0 = \bigcup_n \mathcal{F}_n^0$ to be a collection of arcs in γ satisfying the following three conditions for any $n \in \mathbb{N}$:

(1) For $\tau \in \mathcal{F}_n^0$, we have $L2^{-100n} \leq \operatorname{diam}(\tau) < L2^{-100n+3}$.
(2) The domains of any two distinct arcs in $\mathcal{F}_n^0$ are disjoint in $\mathbb{T}$.
(3) For any $k \in \mathbb{N}$, if the domains of the arcs $\zeta \in \mathcal{F}_{n+k}^0$ and $\tau \in \mathcal{F}_n^0$ intersect non-trivially, then $\zeta \subset \tau$.

According to [22, Lemma 2.13], given any prefiltration $\mathcal{F}^0$, one may construct a *filtration* $\mathcal{F} = \cup_n \mathcal{F}_n$ generated by $\mathcal{F}^0$ i.e. a collection of arcs in γ satisfying the following for any $n \in \mathbb{N}$:

(1) Given $\zeta \in \mathcal{F}_{n+1}$, there is a unique $\tau \in \mathcal{F}_n$ such that $\zeta \subset \tau$.
(2) For $\tau \in \mathcal{F}_n$, we have $L2^{-100n-10} \leq \operatorname{diam}(\tau) < L2^{-100n+4}$.
(3) The domains of any two distinct arcs in $\mathcal{F}_n$ are either disjoint or intersect in (one or both of) their endpoints.
(4) $\bigcup_{\tau \in \mathcal{F}_n} \tau = \mathbb{T}$.
(5) For each arc $\tau^0 \in \mathcal{F}_n^0$, there is a unique arc $\tau \in \mathcal{F}_n$ such that $\tau^0 \subset \tau$. Moreover, if I_0 and I are the domains of τ_0 and τ respectively, then the image of each of the connected components of $I \setminus I_0$ under γ has diameter less than $L2^{-100n-10}$.

4.2. **Preliminaries: balls.** For each $n \in \mathbb{Z}$, choose a 2^{-n} separated net $\mathbb{X}_n$ of Γ (i.e. a set $\mathbb{X}_n \subset \Gamma$ such that $d(x,y) \geq 2^{-n}$ for any $x, y \in \mathbb{X}_n$, and such that, for any $z \in \Gamma$, there is some $x \in \mathbb{X}$ with $d(z,x) < 2^{-n}$). Define a *multiresolution* of Γ as follows:

$$\hat{\mathcal{G}} = \{B(x, 10/2^n) \,:\, x \in \mathbb{X}_n \text{ and } n \in \mathbb{Z}\}.$$

We will use [22, Lemma 2.6] (which holds here with the same proof since $\mathbb{G}$ is Q-regular) to prove the Traveling Salesman Theorem (Theorem 1.4) by establishing the bound

$$\sum_{B \in \hat{\mathcal{G}}} \beta_\Gamma(B)^{2r^2} \operatorname{diam}(B) \leq C\mathcal{H}^1(\Gamma) \tag{4.1}$$

where C depends only on $\mathbb{G}$. As in [22, Lemma 2.9], it suffices to prove inequality (4.1) when the sum is taken over the family $\mathcal{G}$ of balls in $\hat{\mathcal{G}}$ with radius less than $1/100$.

For a ball $B = B(x,r)$, write $\alpha B = B(x, \alpha r)$. Fix an integer $\kappa > 3$. Define $\mathcal{B}$ to be the collection of balls $2B$ where $B \in \mathcal{G}$. According to [22, Lemma 2.11], since $\mathbb{G}$ is doubling, there is a constant $D = D(\kappa) > 0$ and a decomposition $\mathcal{B} = \cup_{i=1}^D \mathcal{B}_i$ into pairwise disjoint families of balls satisfying the following for each i:

(1) if $2B_1, 2B_2 \in \mathcal{B}_i$ have the same radius $2r$, then $d(2B_1, 2B_2) > 2\kappa r$.
(2) for any $2B_1, 2B_2 \in \mathcal{B}_i$, the ratio of their radii equals 2^{100j} for some $j \in \mathbb{Z}$.

Fix $i \in \{1, \dots, D\}$. From each ball $B \in \mathcal{G}$ with $2B \in \mathcal{B}_i$, we may construct a set $Q(B)$ (called a *cube*) so that the family $\Delta(\mathcal{B}, i)$ of cubes constructed from double-balls in $\mathcal{B}_i$ satisfies the following (see [22, Lemma 2.12]):

(1) $2B \subset Q(B) \subset (1 + 2^{-98})2B$.
(2) Fix $2B, 2B' \in \mathcal{B}_i$. If $Q(B) \cap Q(B') \neq \emptyset$ and the radius of B is larger than the radius of B', then $Q(B') \subset Q(B)$.
(3) Fix balls $2B, 2B' \in \mathcal{B}_i$ of equal radius $2r$. Then $d(Q(B), Q(B')) > (\kappa - 1)r$.

Given any cube $Q(B) \in \Delta(\mathcal{B}, i)$, define

$$\Lambda(Q(B)) = \{\tau = \gamma|_I \,:\, I \text{ is a connected component of } \gamma^{-1}(\Gamma \cap Q(B)) \text{ and } \gamma(I) \cap B \neq \emptyset\}.$$

These are the arcs of γ inside $Q(B)$ that meet B. According to [22, Lemma 2.17], the collection of arcs $\mathcal{F}^{0,i} = \bigcup_{Q(B) \in \Delta(\mathcal{B},i)} \Lambda(Q(B))$ is a prefiltration for some $L_i > 0$. As discussed above, this induces a filtration $\mathcal{F}^i$. In particular, for each $\tau^0 \in \mathcal{F}^{0,i}$, there is a unique $\tau \in \mathcal{F}^i$ with $\tau^0 \subset \tau$. We can therefore define the collection of extensions of arcs in $\Lambda(Q(B))$ as

$$\Lambda'(Q(B)) = \{\tau \in \mathcal{F}^i \,:\, \tau \supset \tau^0 \text{ and } \tau^0 \in \Lambda(Q(B))\}$$

for each cube $Q(B) \in \Delta(\mathcal{B}, i)$.

For any arc τ of γ with domain I_τ, write $L_\tau = L_{a(\tau)b(\tau)}$ and define

$$\beta(\tau) = \sup_{t \in I_\tau} \frac{d(\gamma(t), L_{a(\tau)b(\tau)})}{\operatorname{diam}(\tau)} = \sup_{p \in \tau} \frac{d(p, L_\tau)}{\operatorname{diam}(\tau)}.$$

As in [22], write $\mathcal{G} = \mathcal{G}_1 \cup \mathcal{G}_2$ where

$$\mathcal{G}_1 = \{B \in \mathcal{G} \,:\, \text{there is } \tau \in \Lambda'(Q(B)) \text{ such that } \beta(\tau) \geq 10^{-10}\beta_\Gamma(B)\}$$

and

$$\mathcal{G}_2 = \{B \in \mathcal{G} \,:\, \beta(\tau) < 10^{-10}\beta_\Gamma(B) \text{ for all } \tau \in \Lambda'(Q(B))\}.$$

4.3. **Non-flat balls.** In this section, we prove

Proposition 4.1.

$$\sum_{B \in \mathcal{G}_1} \beta_\Gamma(B)^{2r^2} \operatorname{diam}(B) \lesssim \mathcal{H}^1(\Gamma).$$

This is half of the estimate (4.1) (and thus half of the proof of Theorem 1.4). We first need to introduce some notation. There are D different filtrations of γ to consider, but we will treat them individually. Fix $i \in \{1, \dots, D\}$ and write $\mathcal{F} := \mathcal{F}^i$. Recall that $\mathcal{F} = \bigcup_k \mathcal{F}_k$ by the definition of a filtration. Given $\tau \in \mathcal{F}_k$ and $j \in \mathbb{N}$, write

$$\mathcal{F}_{\tau,j} = \{\tau' \in \mathcal{F}_{k+j} \,:\, \tau' \subset \tau\}.$$

This is the collection of arcs j layers lower in the filtration which are contained in τ. Define

$$d_\tau = \max_{\tau' \in \mathcal{F}_{\tau,1}} \sup_{z \in L_{\tau'}} d(z, L_\tau) \quad \text{for any } \tau \in \mathcal{F}.$$

According to [22, Lemma 3.4], $d_\tau \leq 2 \operatorname{diam}(\tau)$. We now prove the following version of Lemma 3.5 in [22]. This is the first place in this section where our proof differs significantly from the arguments in [22], so details are included. In particular, it is in the proof of this lemma that we use the curvature bound from Theorem 3.1.

Lemma 4.2. *For any $\tau \in \mathcal{F}$, we have*

$$\frac{d_\tau^{2r^2}}{\operatorname{diam}(\tau)^{2r^2-1}} \leq C'' \left(\left(\sum_{\tau' \in \mathcal{F}_{\tau,3}} d(\gamma(a(\tau')), \gamma(b(\tau'))) \right) - d(\gamma(a(\tau)), \gamma(b(\tau))) \right) \tag{4.2}$$

for some $C'' = C''(\mathbb{G}) > 0$.

Proof. Fix some $\tau \in \mathcal{F}_k \subset \mathcal{F}$. As in the proof of Lemma 3.5 in [22], we have

$$\frac{d_\tau^{2r^2}}{\operatorname{diam}(\tau)^{2r^2-1}} \leq 2^{2r^2} \operatorname{diam}(\tau) \leq 2^{2r^2} L 2^{-100k+4} = 2^{2r^2+217} L 2^{-100k-213}. \tag{4.3}$$

Thus if

$$\left(\left(\sum_{\tau' \in \mathcal{F}_{\tau,3}} d(\gamma(a(\tau')), \gamma(b(\tau'))) \right) - d(\gamma(a(\tau)), \gamma(b(\tau))) \right) \geq L 2^{-100k-213},$$

we are done. Hence we may assume that

$$\left(\left(\sum_{\tau' \in \mathcal{F}_{\tau,3}} d(\gamma(a(\tau')), \gamma(b(\tau'))) \right) - d(\gamma(a(\tau)), \gamma(b(\tau))) \right) < L 2^{-100k-213}. \tag{4.4}$$

Write $\mathcal{F}_{\tau,1} = \{\tau_i\}_{i=1}^M$ arranged in order of the orientation of $\mathbb{T}$. Set

$$P = \bigcup_{i=1}^{M-1} \{\gamma(b(\tau_i))\}.$$

We will prove

$$d(P, \{\gamma(a(\tau)), \gamma(b(\tau))\}) \geq L 2^{-100k-113}. \tag{4.5}$$

(The proof of this is nearly identical to the proof of (18) in [22].) Suppose (4.5) is not true. That is, there is some j so that (without loss of generality) $d(\gamma(b(\tau_j)), \gamma(a(\tau))) < L 2^{-100k-113}$. Say ξ

is the sub arc of τ defined on $[a(\tau), b(\tau_j)]$. The arc ξ must contain at least one arc in $\mathcal{F}_{\tau,1}$, so we have $\mathrm{diam}(\xi) \geq L2^{-100(k+1)-10}$. Thus there is some point $w \in \xi$ so that

$$\min\{d(\gamma(a(\tau)), w), d(w, \gamma(b(\tau_j)))\} \geq L2^{-100(k+1)-12}.$$

Note that $w \in \tilde{\tau}$ for some $\tilde{\tau} \in \mathcal{F}_{\tau,2}$. Since $\mathrm{diam}(\tilde{\tau}) \leq L2^{-100(k+2)+4}$, the triangle inequality gives

$$\min\{d(\gamma(a(\tilde{\tau})), d(\gamma(a(\tau))), d(\gamma(a(\tilde{\tau})), \gamma(b(\tau_j)))\} \geq L2^{-100k-113}.$$

Therefore, according to the triangle inequality and the negation of (4.5), we have

$$\begin{aligned}\sum_{\tau' \in \mathcal{F}_{\tau,3}} d(\gamma(a(\tau')), \gamma(b(\tau'))) &\geq d(\gamma(a(\tau)), d(\gamma(a(\tilde{\tau}))) + d(\gamma(a(\tilde{\tau})), \gamma(b(\tau_j))) + d(\gamma(b(\tau_j)), \gamma(b(\tau))) \\ &> L2^{-100k-113} + d(\gamma(a(\tau)), d(\gamma(b(\tau_j))) + d(\gamma(b(\tau_j)), \gamma(b(\tau))) \\ &\geq L2^{-100k-213} + d(\gamma(a(\tau)), \gamma(b(\tau))).\end{aligned}$$

This contradicts (4.4). A similar argument in the case of $\gamma(b(\tau))$ proves (4.5).

For any $i \in \{1, \ldots, M\}$, we can repeat the above proof of (4.5) replacing $\gamma(a(\tau))$ with $\gamma(a(\tau_i))$ and $\gamma(b(\tau_j))$ with $\gamma(b(\tau_i))$ to conclude

$$d(\gamma(a(\tau_i)), \gamma(b(\tau_i))) \geq L2^{-100k-113}. \tag{4.6}$$

Indeed, if $d(\gamma(a(\tau_i)), \gamma(b(\tau_i))\}) < L2^{-100k-113}$, we set $\xi = \tau_i$ and follow the previous arguments to obtain

$$\begin{aligned}&\sum_{\tau' \in \mathcal{F}_{\tau,3}} d(\gamma(a(\tau')), \gamma(b(\tau'))) \\ &\quad\geq d(\gamma(a(\tau)), \gamma(a(\tau_i))) + d(\gamma(a(\tau_i)), \gamma(a(\tilde{\tau}))) + d(\gamma(a(\tilde{\tau})), \gamma(b(\tau_i))) + d(\gamma(b(\tau_i)), \gamma(b(\tau))) \\ &\quad> d(\gamma(a(\tau)), \gamma(a(\tau_i))) + L2^{-100k-113} + d(\gamma(a(\tau_i)), \gamma(b(\tau_i))) + d(\gamma(b(\tau_i)), \gamma(b(\tau))) \\ &\qquad\geq L2^{-100k-213} + d(\gamma(a(\tau)), \gamma(b(\tau)))\end{aligned}$$

which again contradicts (4.4). This proves (4.6).

Fix $i \in \{2, \ldots, M-1\}$. We will first establish an estimate on the distance from $L_{\gamma(a(\tau))\gamma(b(\tau_i))}$ to L_τ. Combining (4.5) and (4.6) allows us to bound

$$\min\{d(\gamma(a(\tau)), \gamma(a(\tau_i))), d(\gamma(a(\tau)), \gamma(b(\tau_i))), d(\gamma(a(\tau_i)), \gamma(b(\tau_i))), d(\gamma(b(\tau_i)), \gamma(b(\tau)))\}$$

from below by $L2^{-100k+4}2^{-117} \geq 2^{-117}\,\mathrm{diam}(\tau)$. Therefore, the assumptions of Theorem 3.1 are satisfied with $m = 2^{-217}$ and $\rho = \mathrm{diam}(\tau)$ where

$$a = \gamma(a(\tau)), \quad z = \gamma(a(\tau_i)), \quad v = \gamma(b(\tau_i)), \quad \text{and} \quad w = \gamma(b(\tau)).$$

Theorem 3.1 then gives

$$\begin{aligned}\frac{d(L_{\gamma(a(\tau))\gamma(b(\tau_i))}(t), L_\tau)^{2r^2}}{\mathrm{diam}(\tau)^{2r^2-1}} &= \frac{d(L_{av}(t), L_{aw})^{2r^2}}{\rho^{2r^2-1}} \\ &\lesssim d(\gamma(a(\tau)), \gamma(a(\tau_i))) + d(\gamma(a(\tau_i)), \gamma(b(\tau_i))) \\ &\qquad + d(\gamma(b(\tau_i)), \gamma(b(\tau))) - d(\gamma(a(\tau)), \gamma(b(\tau)))\end{aligned} \tag{4.7}$$

for any $t \in [0, 1]$.

We now establish an estimate on the distance from L_{τ_i} to $L_{\gamma(a(\tau))\gamma(b(\tau_i))}$. Pairing this with (4.7) will give (4.2). Choose an arc $\hat{\tau} \in \mathcal{F}_{\tau,2}$ so that $\hat{\tau}$ is contained in the arc defined on $[a(\tau), a(\tau_i)]$ and $b(\hat{\tau}) \neq a(\tau_i)$. Such an arc always exists because, if it did not, then the only arc in $\mathcal{F}_{\tau,2}$ would be the arc defined on $[a(\tau), a(\tau_i)]$, and this violates the diameter bounds (2) in the definition of a filtration. We may follow the proof of (4.5) to conclude that

$$\min\{d(a(\tau), b(\hat{\tau})), d(b(\hat{\tau}), a(\tau_i))\} \geq L2^{-100k-213}. \tag{4.8}$$

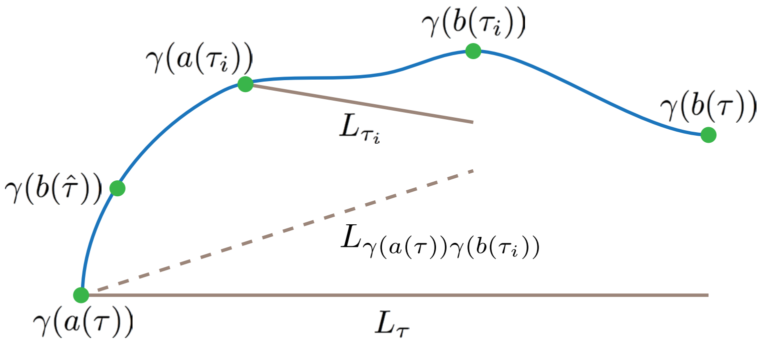

FIGURE 3.

Indeed, assume (without loss of generality) that $d(a(\tau), b(\hat{\tau})) < L2^{-100k-213}$, set ξ to be the arc defined on $[a(\tau), b(\hat{\tau})]$, and note that ξ must contain an arc in $\mathcal{F}_{\tau,2}$. Thus $\operatorname{diam}(\xi) \geq L2^{-100(k+2)-10}$, and we can choose $\tilde{\tau} \in \mathcal{F}_{\tau,3}$ with $\operatorname{diam}(\tilde{\tau}) \leq L2^{-100(k+3)+4}$ so that

$$\min\{d(\gamma(a(\tilde{\tau})), d(\gamma(a(\tau))), d(\gamma(a(\tilde{\tau})), \gamma(b(\hat{\tau})))\} \geq L2^{-100k-213}.$$

Applying the triangle inequality leads to a contradiction of (4.4) just as before. This proves (4.8). We have therefore bounded

$$\min\{d(\gamma(a(\tau)), \gamma(b(\hat{\tau}))), d(\gamma(a(\tau)), \gamma(a(\tau_i))), d(\gamma(b(\hat{\tau})), \gamma(a(\tau_i))), d(\gamma(a(\tau_i)), \gamma(b(\tau_i)))\}$$

from below by $2^{-217} \operatorname{diam}(\tau)$ as before. The assumptions of Theorem 3.1 are satisfied with $m = 2^{-217}$ and $\rho = \operatorname{diam}(\tau)$ where

$$a = \gamma(a(\tau)), \quad z = \gamma(b(\hat{\tau})), \quad v = \gamma(a(\tau_i)), \quad \text{and} \quad w = \gamma(b(\tau_i)).$$

This gives

$$\begin{aligned} \frac{d(L_{\tau_i}(t), L_{\gamma(a(\tau))\gamma(b(\tau_i))})^{2r^2}}{\operatorname{diam}(\tau)^{2r^2-1}} &= \frac{d(L_{vw}(t), L_{aw})^{2r^2}}{\rho^{2r^2-1}} \\ &\lesssim d(\gamma(a(\tau)), \gamma(b(\hat{\tau}))) + d(\gamma(b(\hat{\tau})), \gamma(a(\tau_i))) \\ &\qquad + d(\gamma(a(\tau_i)), \gamma(b(\tau_i))) - d(\gamma(a(\tau)), \gamma(b(\tau_i))) \end{aligned} \tag{4.9}$$

for any $t \in [0, 1]$.

Fix $t \in [0, 1]$. Choose $p \in L_{\gamma(a(\tau))\gamma(b(\tau_i))}$ so that

$$d(L_{\tau_i}(t), L_{\gamma(a(\tau))\gamma(b(\tau_i))}) = d(L_{\tau_i}(t), p)$$

and $q \in L_\tau$ so that

$$d(p, L_\tau) = d(p, q).$$

Combining (4.7) and (4.9) gives

$$\frac{d(L_{\tau_i}(t), L_\tau)^{2r^2}}{\rho^{2r^2-1}} \leq \frac{d(L_{\tau_i}(t), q)^{2r^2}}{\rho^{2r^2-1}} \lesssim \frac{d(L_{\tau_i}(t), p)^{2r^2} + d(p, q)^{2r^2}}{\rho^{2r^2-1}}$$

$$
\begin{aligned}
&= \frac{d(L_{\tau_i}(t), L_{\gamma(a(\tau))\gamma(b(\tau_i))})^{2r^2}}{\rho^{2r^2-1}} + \frac{d(p, L_\tau)^{2r^2}}{\rho^{2r^2-1}} \\
&\lesssim d(\gamma(a(\tau)), \gamma(a(\tau_i))) + d(\gamma(a(\tau_i)), \gamma(b(\tau_i))) \\
&\qquad\qquad + d(\gamma(b(\tau_i)), \gamma(b(\tau))) - d(\gamma(a(\tau)), \gamma(b(\tau))) \\
&\quad + d(\gamma(a(\tau)), \gamma(b(\hat\tau))) + d(\gamma(b(\hat\tau)), \gamma(a(\tau_i))) \\
&\qquad\qquad + d(\gamma(a(\tau_i)), \gamma(b(\tau_i))) - d(\gamma(a(\tau)), \gamma(b(\tau_i))) \\
&\leq 2[d(\gamma(a(\tau)), \gamma(b(\hat\tau))) + d(\gamma(b(\hat\tau)), \gamma(a(\tau_i))) + d(\gamma(a(\tau_i)), \gamma(b(\tau_i))) \\
&\qquad\qquad + d(\gamma(b(\tau_i)), \gamma(b(\tau))) - d(\gamma(a(\tau)), \gamma(b(\tau)))] \\
&\leq 2\left(\left(\sum_{\tau' \in \mathcal{F}_{\tau,3}} d(\gamma(a(\tau')), \gamma(b(\tau')))\right) - d(\gamma(a(\tau)), \gamma(b(\tau)))\right).
\end{aligned}
$$

In the case $i = 1$ (similarly, $i = M$) where $a(\tau_1) = a(\tau)$ (similarly, $b(\tau_M) = b(\tau)$), choose $\hat\tau \in \mathcal{F}_{\tau,2}$ so that $\hat\tau$ is contained in the arc defined on $[a(\tau), a(\tau_2)]$ (similarly, the arc defined on $[a(\tau), a(\tau_M)]$) and $b(\hat\tau) \neq a(\tau_2)$ (similarly, $b(\hat\tau) \neq a(\tau_M)$). We may then apply Theorem 3.1 with $m = 2^{-217}$ and $\rho = \operatorname{diam}(\tau)$ (and noting that $a(\tau_2) = b(\tau_1)$) to

$$
a = \gamma(a(\tau)), \quad z = \gamma(b(\hat\tau)), \quad v = \gamma(a(\tau_2)), \quad \text{and} \quad w = \gamma(b(\tau))
$$

(similarly, $v = \gamma(a(\tau_M))$) to get, as in the proof of (4.7) and (4.9),

$$
\begin{aligned}
\frac{d(L_{\gamma(a(\tau_1))\gamma(b(\tau_1))}(t), L_\tau)^{2r^2}}{\operatorname{diam}(\tau)^{2r^2-1}} &\lesssim d(\gamma(a(\tau)), \gamma(b(\hat\tau))) + d(\gamma(b(\hat\tau)), \gamma(b(\tau_1))) \\
&\qquad + d(\gamma(b(\tau_1)), \gamma(b(\tau))) - d(\gamma(a(\tau)), \gamma(b(\tau)))
\end{aligned}
$$

for any $t \in [0,1]$. Similarly, we have the following bound for $i = M$:

$$
\begin{aligned}
\frac{d(L_{\gamma(a(\tau_M))\gamma(b(\tau_M))}(t), L_\tau)^{2r^2}}{\operatorname{diam}(\tau)^{2r^2-1}} &\lesssim d(\gamma(a(\tau)), \gamma(b(\hat\tau))) + d(\gamma(b(\hat\tau)), \gamma(a(\tau_M))) \\
&\qquad + d(\gamma(a(\tau_M)), \gamma(b(\tau))) - d(\gamma(a(\tau)), \gamma(b(\tau))).
\end{aligned}
$$

This gives the result. □

We conclude this subsection with the following estimate:

Proposition 4.3.

$$
\sum_{\tau \in \mathcal{F}} \beta(\tau)^{2r^2} \operatorname{diam}(\tau) \lesssim \mathcal{H}^1(\Gamma).
$$

Once this has been proven, we may argue exactly as in [22, Corollary 3.3] to prove Proposition 4.1, and this completes the subsection. (It is in that argument that the definition of $\mathcal{G}_1$ is used.)

Proof of Proposition 4.3. Summing equation (4.2) over all $\tau \in \mathcal{F}$ and all k gives

$$
\begin{aligned}
\sum_{\tau \in \mathcal{F}} \frac{d_\tau^{2r^2}}{\operatorname{diam}(\tau)^{2r^2-1}} &\leq C'' \sum_{k=1}^{\infty} \left(\left(\sum_{\tau \in \mathcal{F}_{k+3}} d(\gamma(a(\tau)), \gamma(b(\tau))) \right) - \left(\sum_{\tau \in \mathcal{F}_k} d(\gamma(a(\tau)), \gamma(b(\tau))) \right) \right) \\
&\leq 3C'' \sup_{k \in \mathbb{N}} \sum_{\tau \in \mathcal{F}_k} d(\gamma(a(\tau)), \gamma(b(\tau))) \leq 3C'' \ell(\gamma).
\end{aligned}
$$

For any $\tau \in \mathcal{F}$, say $\{\tau_k\}_{k=0}^{\infty}$ is a sequence of subarcs with $\tau_j \in \mathcal{F}_{\tau,j}$ chosen so that d_{τ_j} is maximal among all subarcs in $\mathcal{F}_{\tau,j}$. Arguing as in the proof of [22, Proposition 3.1], applying [22, Lemma 3.6], using Minkowski's integral inequality in ℓ^{2r^2}, and applying property (2) from the definition of filtrations gives

$$\begin{aligned}\left(\sum_{\tau\in\mathcal{F}} \beta(\tau)^{2r^2} \operatorname{diam}(\tau)\right)^{1/(2r^2)} &\leq \sum_{k=0}^{\infty}\left(\sum_{\tau\in\mathcal{F}} \frac{d_{\tau_k}^{2r^2}}{\operatorname{diam}(\tau)^{2r^2-1}}\right)^{1/(2r^2)} \\ &\leq \sum_{k=0}^{\infty} 2^{(-100k+14)\frac{2r^2-1}{2r^2}}\left(\sum_{\tau\in\mathcal{F}} \frac{d_{\tau_k}^{2r^2}}{\operatorname{diam}(\tau_k)^{2r^2-1}}\right)^{1/(2r^2)} \\ &\leq \big(3C''\ell(\gamma)\big)^{1/(2r^2)} \sum_{k=0}^{\infty} 2^{(-100k+14)\frac{2r^2-1}{2r^2}} \\ &\lesssim \ell(\gamma)^{1/(2r^2)} \qquad \square\end{aligned}$$

4.4. **Flat balls.** In this section, we will prove the other half of (4.1):

Proposition 4.4.

$$\sum_{B\in\mathcal{G}_2} \beta_\Gamma(B)^2 \operatorname{diam}(B) \lesssim \mathcal{H}^1(\Gamma).$$

To do so, we will follow the proof in Section 4 of [22] of a similar bound in the Heisenberg group. As stated at the beginning of that section, most of the arguments therein may be applied in any general metric space. The only Heisenberg-specific ingredients of the proof are Lemma 4.1 and equations (23) and (24). Therefore, in order to prove Proposition 4.4, it suffices to verify these three facts in $\mathbb{G}$.

Equation (23) in [22] requires

$$\operatorname{diam}(B(p,\lambda r)) \leq \lambda \operatorname{diam}(B(p,r))$$

for any $r > 0$, $p \in \mathbb{G}$, and $\lambda > 1$. In $\mathbb{G}$, we have

$$\operatorname{diam}(B(p,\lambda r)) = \lambda \operatorname{diam}(B(p,r)) \tag{4.10}$$

for any $r > 0$, $p \in \mathbb{G}$, and $\lambda > 0$. This follows from the fact that $\operatorname{diam}(B(p,r)) = 2r$ for any left invariant, homogeneous metric in $\mathbb{G}$ [14, Proposition 2.4]. Moreover, equation (24) in [22] is a result of

$$d(L(t_1), L(t_2)) = |t_1 - t_2|\|\tilde{\pi}(p^{-1}q)\| \tag{4.11}$$

for any horizontal segment $L = L_{pq} : [0,1] \to \mathbb{G}$ between $p, q \in \mathbb{G}$ and any $t_1, t_2 \in [0,1]$.

It remains to prove Lemma 4.1 from [22] in the Carnot group setting. We first establish the following:

Lemma 4.5. *There is a radius $0 < r_0 \leq \frac{1}{2}$ such that, for any horizontal segment L which intersects $B(0, r_0)$ non-trivially, L is never tangent to the unit sphere $\partial B(0,1) = \partial B_{\mathbb{R}^N}(0,\eta)$.*

Proof. Note that we need only consider those horizontal segments in $B_{\mathbb{R}^n}(0,2\eta) \times \mathbb{R}^{N-n}$. Indeed, the projection of the segment to $\mathbb{R}^n \times \{0\}$ is a Euclidean segment traversed at constant speed, and the restriction of a horizontal segment to a subinterval is still a horizontal segment. Hence a horizontal segment will intersect both $B(0, r_0)$ and $\partial B(0,1)$ if and only if its restriction to $B_{\mathbb{R}^n}(0,2\eta) \times \mathbb{R}^{N-n}$ (which is also a connected, horizontal segment) does as well.

Suppose by way of contradiction that there is a sequence of horizontal segments $L_j : [0,1] \to \mathbb{G}$ in $B_{\mathbb{R}^n}(0, 2\eta) \times \mathbb{R}^{N-n}$ which intersect $B(0, 1/j)$ non-trivially and lie tangent to $\partial B(0,1)$. Say $L_j = L_{p_j q_j}$ for some $p_j, q_j \in \mathbb{G}$. These horizontal segments are all 4-Lipschitz since

$$d(L_j(s), L_j(t)) = d(\delta_s(\tilde{\pi}(p_j^{-1}q_j)), \delta_t(\tilde{\pi}(p_j^{-1}q_j))) = \tfrac{1}{\eta}|s-t||(q_j)_1 - (p_j)_1| \leq 4|s-t|.$$

In particular, since each segment L_j meets $B(0,1/j)$, there is some $M_0 > 0$ so that $|p_j| < M_0$ for every $j \in \mathbb{N}$. Write $p_j = (p_j^1, p_j^2)$ and $q_j = (q_j^1, q_j^2)$. By definition, we have

$$L_j(s) = \left(p_j^1 + s(q_j^1 - p_j^1), p_j^2 + P\left(p_j, \left(s(q_j^1 - p_j^1), 0\right)\right)\right) \quad \text{for any } s \in [0,1],\ j \in \mathbb{N}$$

for some polynomial P given by the BCH formula. Therefore, by the uniform boundedness of $|p_j|$ and $|q_j^1|$, there is some $M_1 > 0$ so that $\left|d^i/ds^i(L_j)\right| < M_1$ for every $i, j \in \mathbb{N}$. The Arzelà-Ascoli theorem then gives a subsequence of these horizontal segments (also called $\{L_j\}$) converging uniformly in $\mathbb{R}^N$ (and thus in $\mathbb{G}$) to some C^∞ curve $L : [0,1] \to \mathbb{G}$ passing through the origin so that all derivatives of L_j converge uniformly to the corresponding derivatives of L. Note that L itself must also be a horizontal segment. Indeed, $p_j = L_j(0) \to p$ for some $p \in \mathbb{G}$, and $\tilde{\pi}(p_j^{-1}q_j) = p_j^{-1}L_j(1) \to (z, 0)$ for some $z \in \mathbb{R}^n$. Thus, for any $q \in \mathbb{G}$ satisfying $q_1 - p_1 = z$, we have

$$L(s) = \lim_{j\to\infty} L_j(s) = \lim_{j\to\infty} p_j \delta_s(\tilde{\pi}(p_j^{-1}q_j)) = p\delta_s((z,0)) = p\delta_s(\tilde{\pi}(p^{-1}q))$$

for every $s \in [0,1]$. Since L is a horizontal segment passing through the origin, it must be the case that L is a Euclidean line segment in $\mathbb{R}^n \times \{0\}$.[4] In particular, L cannot be tangent to the Euclidean sphere $\partial B_{\mathbb{R}^N}(0, \eta)$. Since the derivatives of the segments L_j converge uniformly to the derivatives of L, it is impossible that L_j lies tangent to the sphere for every j. This is a contradiction and completes the proof. □

The following is the Carnot group version of Lemma 4.1 from [22]. Note that, here, we have the constant r_0 included in the inequality, while, in [22], the constant is 1. This, however, is not a problem since the constant r_0 depends only on $\mathbb{G}$.

Lemma 4.6. *Let τ be a connected subarc. Then*

$$\sup_{x \in L_\tau} d(x, \tau) \leq r_0^{-1} \beta(\tau) \operatorname{diam}(\tau). \tag{4.12}$$

In particular, if we write $I_\tau = [a,b]$, we have

$$d(L_\tau(1), \gamma(b)) \leq r_0^{-1} \beta(\tau) \operatorname{diam}(\tau). \tag{4.13}$$

Proof. Recall that $\beta(\tau)\operatorname{diam}(\tau) = \sup_{p\in\tau} d(p, L_\tau)$. By the invariance of the metric under left translation, we may assume that $0 = \gamma(a) = L_\tau(0)$. We begin by proving (4.13). Choose $t_0 \in [0,1]$ so that $d(\gamma(b), L_\tau(t_0)) = d(\gamma(b), L_\tau) \leq \beta(\tau)\operatorname{diam}(\tau)$. Since $L_\tau(1) = \tilde{\pi}(\gamma(b))$ and $L_\tau(t_0) = \delta_{t_0}(\tilde{\pi}(\gamma(b)))$ are co-linear in $\mathbb{R}^n \times \{0\}$, it follows that

$$d(L_\tau(1), L_\tau(t_0)) = \left\|\delta_{t_0}(\tilde{\pi}(\gamma(b)))^{-1}\tilde{\pi}(\gamma(b))\right\| = \left\|\tilde{\pi}[\delta_{t_0}(\tilde{\pi}(\gamma(b)))^{-1}\gamma(b)]\right\| \leq d(\gamma(b), L_\tau(t_0)).$$

Therefore, we have

$$d(L_\tau(1), \gamma(b)) \leq d(L_\tau(1), L_\tau(t_0)) + d(L_\tau(t_0), \gamma(b)) \leq 2\beta(\tau)\operatorname{diam}(\tau) \leq r_0^{-1}\beta(\tau)\operatorname{diam}(\tau).$$

In order to prove (4.12), we will first show that the mapping $f : \tau \to L_\tau$ defined as

$$f(p) = L_\tau(t_0) \quad \text{where } t_0 = \sup\{t \in [0,1] \,:\, d(L_\tau(t), p) \leq r_0^{-1}\beta(\tau)\operatorname{diam}(\tau)\}$$

---

[4] Divide L into two segments: the segment ending at 0 and the segment starting at 0. Since both of these must also be horizontal, they must be Euclidean segments.

is continuous. (Note that $d(p, L_\tau) \leq \beta(\tau)\operatorname{diam}(\tau)$ for every $p \in \tau$, so f is well defined.) In order to prove that f is continuous, it suffices to prove for every $p \in \tau$ that L_τ does not lie tangent to the sphere centered at p with radius $r_0^{-1}\beta(\tau)\operatorname{diam}(\tau)$.

Fix $p \in \tau$. We may translate by p^{-1} and dilate by $r_0(\beta(\tau)\operatorname{diam}(\tau))^{-1}$ to reduce to the following problem: show that the horizontal segment $L = \delta_{r_0(\beta(\tau)\operatorname{diam}(\tau))^{-1}}(p^{-1}L_\tau)$ is never tangent to the sphere $\partial B(0,1)$. This follows from Lemma 4.5 since

$$d(0, L) = d(0, \delta_{r_0(\beta(\tau)\operatorname{diam}(\tau))^{-1}}(p^{-1}L_\tau)) = r_0(\beta(\tau)\operatorname{diam}(\tau))^{-1}d(p, L_\tau) \leq r_0$$

implies that the segment L intersects the ball $B(0, r_0)$ non-trivially. Therefore, f is continuous.

Since τ is connected and $f(\gamma(b)) = L_\tau(1)$ by (4.13), the continuous map f sends τ onto an interval in L_τ containing $[f(0), L_\tau(1)] \subset L_\tau$. Say $t_1 \in [0,1]$ is such that $L_\tau(t_1) = f(0)$. Then, for any $t \in [t_1, 1]$, we have $d(L_\tau(t), \tau) \leq d(L_\tau(t), p) \leq r_0^{-1}\beta(\tau)\operatorname{diam}(\tau)$ for some $p \in \tau$ by the surjectivity of f, and, for any $t \in [0, t_1]$, we have

$$d(L_\tau(t), 0) = t\|\tilde{\pi}(\gamma(b))\| \leq t_1\|\tilde{\pi}(\gamma(b))\| = d(L_\tau(t_1), 0) \leq r_0^{-1}\beta(\tau)\operatorname{diam}(\tau).$$

This proves the lemma. □

The following lemma reproves Lemma 4.3 of [22] in a more Carnot way.

Lemma 4.7. *Let $B \in \mathcal{G}$ be a ball of radius r. Let $Q = Q(B)$ and in particular suppose $3B \supset Q \supset 2B$. Suppose $\tau' \in \Lambda'(Q)$ and $\tau' \ni \operatorname{Center}(B)$. Suppose further that*

$$r_0^{-1}\beta(\tau')\operatorname{diam}(\tau') < h < \tfrac{1}{10}r.$$

Then there is an arc $\tilde{\tau} \subset \tau'$ with image in $2B$ such that $\operatorname{diam}(\tilde{\tau}) \geq 4r - 10h$.

Proof. Without loss of generality, we may suppose $\operatorname{Center}(B) = 0$. Write $I_{\tau'} = [a(\tau'), b(\tau')]$ as before, and set $L = L_{\tau'}$. By definition, τ' is an extension of some $\tau \in \Lambda(Q)$. Note that $\gamma(a(\tau)), \gamma(b(\tau)) \in \partial Q$. This says that $d(\gamma(a(\tau)), 0), d(\gamma(b(\tau)), 0) \in [2r, 3r]$. We also have that $\tau \ni 0$. Indeed, if $0 \notin \tau$, then there is a subarc ξ of $\tau' \setminus \tau$ containing 0 and a point in ∂Q, so $\operatorname{diam}(\xi) \geq 2r$. However, property (5) of a filtration gives $\operatorname{diam}(\xi) < 2^{-10}\operatorname{diam}(\tau) \leq 2^{-10}6r$ since the image of τ is contained in $Q \subset 3B$. This is impossible, so we must have $\tau \ni 0$.

Recall the definition of the continuous map $f : \tau' \to L_{\tau'}$ from the proof of the previous lemma:

$$f(p) = L_{\tau'}(t_0) \quad \text{where } t_0 = \sup\{t \in [0,1] : d(L_{\tau'}(t), p) \leq r_0^{-1}\beta(\tau')\operatorname{diam}(\tau')\}.$$

Write $x = f(\gamma(a(\tau)))$, $y = f(0)$, and $z = f(\gamma(b(\tau)))$. In particular, $x, y, z \in L$ and the assumptions of the lemma give

$$d(\gamma(a(\tau)), x) < h, \quad d(0, y) < h, \quad d(\gamma(b(\tau)), z) < h. \tag{4.14}$$

We then get by the triangle inequality that

$$d(x, y), d(z, y) \in [2r - 2h, 3r + 2h]. \tag{4.15}$$

We first claim that x, z are on opposite sides of y along L. In particular, if $x = L(t_1)$, $y = L(t_2)$, and $z = L(t_3)$, then $t_1 < t_2 < t_3$. Suppose not (e.g. $t_1 \geq t_2$.). We remind the reader that $L(0) = \gamma(a(\tau'))$. Define the 1-Lipschitz abelianization map $\pi : \mathbb{G} \to \mathbb{R}^n$ as $\pi(p_1, p_2) = \frac{1}{\eta}p_1$, and recall that η is the constant chosen in the definition of the HS norm. Notice that $\pi(L)$ is a Euclidean segment from $\pi(\gamma(a(\tau')))$ to $\pi(\gamma(b(\tau')))$. In fact, (4.11) implies that π is an isometry on L. Indeed, (2.1) gives

$$\begin{aligned}|\pi(L(t)) - \pi(L(s))| = |t-s||\pi(\gamma(a(\tau'))) - \pi(\gamma(b(\tau')))| &= |t-s|\|\tilde{\pi}(\gamma(b(\tau'))^{-1}\gamma(a(\tau')))\| \\ &= d(L(t), L(s))\end{aligned}$$

for each $s,t \in [0,1]$. In particular, this means that (4.15) holds for the projections $\pi(x)$, $\pi(y)$, and $\pi(z)$ as well. According to our assumption that $t_1 \geq t_2$, the point $\pi(x)$ lies further along $\pi(L)$ than $\pi(y)$ in the following sense:

$$|\pi(\gamma(a(\tau'))) - \pi(x)| \geq |\pi(\gamma(a(\tau'))) - \pi(y)|. \tag{4.16}$$

Using Properties (1) and (5) of prefiltrations and filtrations, respectively, we get from the fact that π is 1-Lipschitz that

$$|\pi(\gamma(a(\tau'))) - \pi(\gamma(a(\tau)))| \leq d(\gamma(a(\tau')), \gamma(a(\tau))) < \tfrac{1}{10}r.$$

Also, (4.14) implies that $|\pi(\gamma(a(\tau))) - \pi(x)| < h$. Putting this all together gives

$$|\pi(\gamma(a(\tau'))) - \pi(x)| < |\pi(\gamma(a(\tau'))) - \pi(\gamma(a(\tau)))| + |\pi(\gamma(a(\tau))) - \pi(x)| \leq \tfrac{1}{10}r + h < \tfrac{2}{10}r$$

while

$$\begin{aligned}|\pi(\gamma(a(\tau'))) - \pi(y)| &\geq |\pi(y) - \pi(x)| - |\pi(x) - \pi(\gamma(a(\tau)))| - |\pi(\gamma(a(\tau))) - \pi(\gamma(a(\tau')))| \\ &> 2r - 2h - h - \tfrac{1}{10}r > \tfrac{8}{5}r.\end{aligned}$$

This contradicts (4.16), and thus $t_1 < t_2$. We may similarly show $t_3 > t_2$ by proving that $\pi(z)$ cannot lie before $\pi(y)$ along $\pi(L)$.

We now know that τ is a curve so that the images of the endpoints of τ under f lie on opposite sides of y. We claim that there is a subcurve $\tau_1 \subset \tau$ which is a connected component of $\tau \cap 2B$ so that the endpoints of τ_1 are mapped to opposite sides of y via f. Suppose by way of contradiction that both endpoints of every connected component of $\tau \cap 2B$ are mapped to one side of y or both are mapped to the other. (Note that the endpoints of such a curve can never map onto y itself since then we would have $d(p,0) \leq d(p,y) + d(y,0) \leq 2h < \frac{1}{5}r$ for such an endpoint p, but these endpoints must lie in the boundary of $2B$.) Then the above property of τ and the fact that the endpoints of τ lie outside of the interior of $2B$ imply that there is at least one sub curve $\xi \subset \tau$ in $\tau \setminus 2B$ whose endpoints are mapped to opposite sides of y. By the continuity of f, there must be some p in the image of ξ so that $f(p) = y$. However, $d(p,0) \leq d(p,y) + d(y,0) \leq 2h \leq 2r$, and so $p \in 2B$ which is a contradiction. Therefore, such a τ_1 exists as claimed. Note that the endpoints p,q of τ_1 must lie on $\partial(2B)$. We have $d(p,f(p)) < h$ and $d(q,f(q)) < h$ by assumption, and thus

$$d(f(p),y) \geq d(p,0) - d(p,f(p)) - d(y,0) \geq 2r - 2h,$$

and similarly $d(f(q),y) \geq 2r - 2h$. Since $f(p)$ and $f(q)$ lie on opposite sides of y, it follows that $d(f(p),f(q)) \geq 4r - 4h$. Thus

$$d(p,q) \geq d(f(p),f(q)) - d(p,f(p)) - d(q,f(q)) \geq 4r - 6h$$

which proves the lemma. □

With the above lemmas established, we may now argue exactly as in Section 4 of [22] (with the constants therein adjusted appropriately to account for r_0) to conclude Proposition 4.4.

We now give a brief sketch of the rest of the overall strategy of Section 4 of [22]. In Lemma 4.4, we prove that if there exists a curve $\tau' \in \Lambda'(Q)$ with small $\beta(\tau')$ and another point x in some other curve $\xi \in \Lambda(Q)$ where $d(x, L_{\tau'})$ is sufficiently large, then there is actually a subcurve $\check{\xi} \subset \xi$ with

$$\operatorname{diam}(\check{\xi}) > 20\epsilon_0 r_0^{-1} \beta_\Gamma(B) \operatorname{diam}(B)$$

so that

$$d(\check{\xi}, \tau') > 20\epsilon_0 r_0^{-1} \beta_\Gamma(B) \operatorname{diam}(B)$$

where $\epsilon_0 = 10^{-10} r_0^{-1}$ (note that $r_0 = 1$ in [22] which allowed ϵ_0 to be just 10^{-10}).

In (the r_0^{-1} modification of) the proof, we said it suffices to show that

$$d(x,\tau') > 40\epsilon_0 r_0^{-1}\beta_\Gamma(B)\,\mathrm{diam}(B).$$

The lemma then follows from an easy triangle inequality argument. Indeed, consider any maximal length connected subcurve of $\check{\xi} \subseteq \xi$ in $B(x, 20\epsilon_0 r_0^{-1}\beta_\Gamma(B)\,\mathrm{diam}(B))$ that contains the center x. This must have diameter at least $20\epsilon_0 r_0^{-1}\beta_\Gamma(B)\,\mathrm{diam}(B)$ as it goes from the center of the ball to the outside. Furthermore, it must be of distance greater than

$$40\epsilon_0 r_0^{-1}\beta_\Gamma(B)\,\mathrm{diam}(B) - 20\epsilon_0 r_0^{-1}\beta_\Gamma(B)\,\mathrm{diam}(B) = 20\epsilon_0 r_0^{-1}\beta_\Gamma(B)\,\mathrm{diam}(B)$$

from τ', as required.

Lemma 4.5 and Proposition 4.6 of [22] say the following. Suppose we have $\tau, \xi \in \Lambda(Q)$ satisfying the last two lemmas. That is, τ contains Center(B) and has an extension $\tau' \in \Lambda'(Q)$ so that $\beta(\tau')$ is small and ξ contains a point x so that $d(x, L_{\tau'})$ is large. Then covering $\xi \cup \tau$ with balls $\{B_i\}$ requires that the sum of the diameters of the balls must exceed $4r$ by some definite constant. This follows from the last two results as τ contains a subcurve of length almost $4r$ whereas ξ contains a subcurve $\hat{\xi}$ that is far away from τ and has a definite length.

Note that as we are working with balls in $\mathcal{G}_2$, we are always in this situation. Indeed, curves in Q have small β but $\beta_\Gamma(B)$ itself is large, which means there must exist a point x that is far away from the curve τ going through the center of B. Now consider all the cubes Q associated to balls in $\mathcal{G}_2$, which form a nested set of cubes. Proposition 4.7 builds on Proposition 4.6 and says that for one such cube Q, the sum of the diameter of all the maximal subcubes Q' of Q plus the $\mathcal{H}^1$ measure of the remainder $R = Q\backslash \bigcup_i Q'$ must be larger than the diameter of Q by a multiplicative factor that is quantitatively greater than 1.

Proposition 4.8 is the main technical result of this section. We are still working with $\mathcal{G}_2$, although now we further decompose these into subfamilies $\{\mathcal{B}^M\}_{M=0}^\infty$ where $\beta_\Gamma(B) \in [2^{-M-1}, 2^{-M}]$ for each $B \in \mathcal{B}^M$. Let Δ denote the cubes associated to balls of some $\mathcal{B}^M$ and $\tilde{\Delta} \subset \Delta$ be any finite subset. We now perform the following construction $\tilde{\Delta}$ (the proof in [22] applies the construction to all of Δ, which we will remark upon below).

For each cube $Q \in \tilde{\Delta}$, we decompose it into $R_Q \cup \bigcup_i Q_i$ where Q_i are maximal subcubes of Δ in Q. Note that R_Q are precisely the parts of Γ in Q that not in any other Q_i. This means that all the R_Q are pairwise disjoint and so any intersection $R_Q \cap R_{Q'}$ is empty if $Q \neq Q'$. For each such Q, we define a weight function $w_Q : \mathcal{H} \to [0,\infty)$ supported on Q so that

(1) its total mass $\int_Q w(x)\, d\mathcal{H}^1(x)$ is $\mathrm{diam}(Q)$,
(2) its mass is on R_Q and each maximal subcube Q_i (which sums up to $\mathrm{diam}(Q)$) is proportional to $\mathcal{H}^1(R_Q)$ and $\mathrm{diam}(Q_i)$,
(3) $w_Q|_{R_Q}$ is proportional to a constant function,
(4) $w_Q|_{Q_i}$ is proportional to a constant function if $Q_i \in \Delta\backslash\tilde{\Delta}$,
(5) $w_Q|_{Q_i}$ is proportional to w_{Q_i} if $Q_i \in \tilde{\Delta}$ (which is defined in the same manner).

The main point is that if $Q_1 \supset Q_2 \supset ... \supset Q_N$ is a chain of maximal subcubes and $x \in Q_N$, then one can use Proposition 4.7 to show that $w_{Q_i}(x) \le q^{N-i}$, for some $q \in (0,1)$ depending only on M (see the calculations between equations (30) and (31) of [22]). This will allow us to conclude that $\sum_{Q\in\tilde{\Delta}} w_Q(x)$ is a geometric series bounded by a constant multiple of 2^M, which will then easily lead to a proof that $\sum_{Q\in\tilde{\Delta}} \mathrm{diam}(Q) \lesssim 2^M \mathcal{H}^1(\Gamma)$. As the summands are positive and the bound holds for any finite partial sum, we get that $\sum_{Q\in\Delta} \mathrm{diam}(Q) \lesssim 2^M \mathcal{H}^1(\Gamma)$, which easily leads to $\sum_{B\in\mathcal{G}_2} \beta_\Gamma(B)^2\,\mathrm{diam}(B) \lesssim \mathcal{H}^1(\Gamma)$.

The reason we work with $\tilde{\Delta}$ is so we can guarantee Property (1) of the w_Q since all iterations coming from Property (5) will eventually terminate. Without this truncation, one would have to separately deal with the integral of w_Q on the set of points lying in infinite chains $Q_1 \supset Q_2 \supset \ldots$. While this may be possible, the truncation argument allows us to sidestep this issue. We are thankful to the anonymous referee for pointing out this missing step.

Proposition 4.4, together with Proposition 4.1, finishes the proof of (4.1), and thus the proof of Theorem 1.4 is complete.

5. Step 2 groups

In this section, we will prove Theorem 1.5. In Theorem 1.2 (proven in [22]), the Traveling Salesman Theorem is established in the Heisenberg group, and the exponent on the β-numbers is 4. However, this is not the same exponent provided by Theorem 1.4. Indeed, the Heisenberg group has step $r = 2$, and we have proven the TST in step 2 groups where the exponent on the β-numbers is $2r^2 = 8$.

The increase from $2r$ to $2r^2$ occured when we appied Lemma 3.7 in the proof of Theorem 3.1. To avoid this, we will prove a step 2 version of Theorem 3.1 directly without appealing to this lemma. This allows us to replace any instance of $2r^2$ with $2r = 4$ in all of the arguments that follow. This will prove Theorem 1.5 and provide a true generalization of Theorem 1.2.

Theorem 5.1. *Suppose $a, z, v, w \in \mathbb{G}$ satisfy*

$$m\rho \leq \min\{d(a,z), d(a,v), d(z,v), d(v,w)\}$$

and

$$\max\{d(a,z), d(a,v), d(z,v), d(v,w), d(a,w)\} \leq \rho$$

for some $\rho > 0$. Then there is a constant $C_0 = C_0(\mathbb{G}) > 0$ so that

$$\sup_{t\in[0,1]} d(L_{av}(t), L_{aw})^4 + \sup_{t\in[0,1]} d(L_{vw}(t), L_{aw})^4 \leq C_0\rho^3\Delta$$

where $\Delta := d(a,z) + d(z,v) + d(v,w) - d(a,w)$.

The hypothesis of this theorem is the same as that of Theorem 3.1. However, the exponents in the conclusion are 4 and 3 rather than $2r^2 = 8$ and $2r^2 - 1 = 7$. Once this theorem has been proven, the rest of the arguments in Section 4 follow in exactly the same manner with all instances of $2r^2$ replaced with 4.

In the proof, we will bound d_∞ distances (rather than d) which are defined on a step 2 group $\mathbb{G}$ as

$$d_\infty(x,y) = N_\infty(y^{-1}x) \quad \text{where } N_\infty(p) = \max\{|p_1|, |p_2|^{1/2}\}$$

for any $p = (p_1, p_2) \in \mathbb{G}$. Though d_∞ is not a true metric (since a scaling constant is present in the triangle inequality), it is homogeneous and hence bi-Lipschitz equivalent to the HS-distance d in the sense of (2.3). This will suffice.

Proof of Theorem 5.1. As before, we may assume without loss of generality that $a = 0$ and $\rho = 1$, and we set $p_v = (p_1, 0) \in L_w$ to be the closest point in L_w to $\tilde{\pi}(v)$ in the Euclidean norm.

Since the BCH formula reduces to $X + Y + \frac{1}{2}[X,Y]$ in a step 2 Carnot group, we have

$$\delta_t(p_v)^{-1}\delta_t(\tilde{\pi}(v)) = (-tp_1, 0)(tv_1, 0) = \left(t(v_1 - p_1), -\tfrac{t^2}{2}[p_1, v_1]\right).$$

By definition and (3.6), we have $|v_1 - p_1| = d_{\mathbb{R}^n}(v_1, \ell_{w_1}) \lesssim \Delta^{1/2} \lesssim \Delta^{1/4}$ (since $\Delta \le 3$). Also,

$$|[p_1, v_1]|^{1/2} = |[p_1, v_1] - [v_1, v_1]|^{1/2} = |[p_1 - v_1, v_1]|^{1/2} \lesssim |p_1 - v_1|^{1/2}|v_1|^{1/2} \lesssim \Delta^{1/4}$$

since $|v_1| = |\tilde{\pi}(v)| \lesssim \|v\| \le 1$. Therefore, for each $t \in [0.1]$, we have

$$d(L_v(t), L_w) \lesssim d_\infty(\delta_t(\tilde{\pi}(v)), \delta_t(p_v)) = \max\left\{t|v_1 - p_1|, \left|\tfrac{t^2}{2}[p_1, v_1]\right|^{1/2}\right\} \lesssim \Delta^{1/4}.$$

We will now bound $d(L_{vw}(t), L_w)$. To do so, we compute

$$v\delta_t(\tilde{\pi}(v^{-1}w)) = (v_1, v_2)(t(w_1 - v_1), 0) = (tw_1 + (1-t)v_1, v_2 + \tfrac{t}{2}[v_1, w_1])$$

and

$$p_v\delta_t(\tilde{\pi}(p_v^{-1}w)) = (p_1, 0)(t(w_1 - p_1), 0) = (tw_1 + (1-t)p_1, \tfrac{t}{2}[p_1, w_1])$$

so that

$$\begin{aligned}
&(p_v\delta_t(\tilde{\pi}(p_v^{-1}w)))^{-1}v\delta_t(\tilde{\pi}(v^{-1}w)) \\
&\quad = \left((1-t)(v_1 - p_1), v_2 + \tfrac{t}{2}[v_1, w_1] - \tfrac{t}{2}[p_1, w_1] - \tfrac{1}{2}[tw_1 + (1-t)p_1, tw_1 + (1-t)v_1]\right) \\
&\quad = \left((1-t)(v_1 - p_1), v_2 + \tfrac{t}{2}[v_1 - p_1, w_1] + \tfrac{t(1-t)}{2}[v_1 - p_1, w_1] - \tfrac{(1-t)^2}{2}[p_1, v_1]\right).
\end{aligned}$$

As above, we have $|[v_1 - p_1, w_1]| \lesssim d_{\mathbb{R}^n}(v_1, \ell_{w_1})$. Finally, note that $NH(v) = (0, v_2)$ in step 2 groups, and so $|v_2|^{1/2} = N_\infty(NH(v)) \lesssim \|NH(v)\|$. Hence (3.6) and (3.7) give

$$\begin{aligned}
&\left|v_2 + \tfrac{t}{2}[v_1 - p_1, w_1] + \tfrac{t(1-t)}{2}[v_1 - p_1, w_1] - \tfrac{(1-t)^2}{2}[p_1, v_1]\right|^2 \\
&\qquad \lesssim \|NH(v)\|^4 + d_{\mathbb{R}^n}(v_1, \ell_{w_1})^2 \\
&\qquad \le \left[\|NH(v)\|^4 + d_{\mathbb{R}^n}(z_1, \ell_{v_1})^2\right] + \left[\|NH(w)\|^4 + d_{\mathbb{R}^n}(v_1, \ell_{w_1})^2\right] \\
&\qquad \lesssim [d(0, z) + d(z, v) - d(0, v)] + [d(0, v) + d(v, w) - d(0, w)] = \Delta.
\end{aligned}$$

Therefore, for each $t \in [0.1]$, we have

$$d(L_{vw}(t), L_w) \lesssim d_\infty(v\delta_t(\tilde{\pi}(v^{-1}w)), p_v\delta_t(\tilde{\pi}(p_v^{-1}w))) \lesssim \Delta^{1/4}. \quad \square$$

6. Singular integrals on 1-regular curves

Recall that if (X, d) is a metric space, an $\mathcal{H}^1$-measurable set $E \subset X$ is 1-*(Ahlfors)-regular*, if there exists a constant $1 \le C < \infty$, such that

$$C^{-1}r \le \mathcal{H}^1(B(x, r) \cap E) \le Cr$$

for all $x \in E$, and $0 < r \le \operatorname{diam}(E)$. In this section we are going to prove Theorem 1.6, which we reformulate in a more precise manner below.

Theorem 6.1. *Let $(\mathbb{G}, d)$ be Carnot group of step $r \ge 2$ equipped with a homogeneous metric d. Let $K_d : \mathbb{G} \setminus \{0\} \to [0, \infty)$ be defined by*

$$K_d(p) = \frac{d(NH(p), 0)^{2r^3} + d(NH(p^{-1}), 0)^{2r^3}}{d(p, 0)^{2r^3+1}},$$

and let E be a 1-regular set which is contained in a 1-regular curve. Then the corresponding truncated singular integrals

$$T^\varepsilon f(p) = \int_{E \setminus B_d(p, \varepsilon)} K_d(q^{-1} \cdot p) f(q) \, d\mathcal{H}^1(q)$$

are uniformly bounded in $L^2(\mathcal{H}^1|_E)$.

Proof. The proof follows as in the proof of [3, Theorem 1.1] once we have at our disposal Theorem 1.4 and Lemma 6.2. Nevertheless we will provide an outline for the convenience of the reader. To simplify notation we let $\mu = \mathcal{H}^1|_E$ and $K = K_d$. Since E is 1-regular there exists some constant $c_\mu \in (0,1]$ such that

$$c_\mu r \leq \mu(B(x,r)) \leq c_\mu^{-1} r, \qquad \forall x \in E, r > 0.$$

We first observe that the kernel K is a symmetric 1-dimensional *Calderón-Zygmund* (CZ)-kernel, see [3, Definition 2.6 and Lemma 2.7]. We will use the $T1$-theorem (which we explain more in the following) to prove that the operators T^ε are uniformly bounded on $L^2(\mu)$. For this reason, we need a system of dyadic-like cubes associated to the set E. These systems were introduced by David in [6] (see also [7]) for regular Euclidian sets and later generalized by Christ [4] to any regular set of a geometrically doubling metric space. In particular for the set E, there is a constant $c_d \in (0,1]$ and a family of partitions Δ_j of E, $j \in \mathbb{Z}$, with the following properties;

(D1) If $k \leq j$, $Q \in \Delta_j$ and $Q' \in \Delta_k$, then either $Q \cap Q' = \emptyset$, or $Q \subset Q'$.
(D2) If $Q \in \Delta_j$, then $\operatorname{diam} Q \leq 2^{-j}$.
(D3) Every set $Q \in \Delta_j$ contains a set of the form $B(p_Q, c_d 2^{-j}) \cap E$ for some $p_Q \in Q$.

We will call the sets in $\Delta := \cup \Delta_j$ the *dyadic cubes* of E. For a cube $S \in \Delta$, we define

$$\Delta(S) := \{Q \in \Delta : Q \subseteq S\}.$$

Given a cube $Q \in \Delta$ and $\lambda \geq 1$, we define

$$\lambda Q := \{x \in E : d(x,Q) \leq (\lambda - 1)\operatorname{diam}(Q)\}.$$

It follows from (D2), (D3) and the 1-regularity of E that if $Q \in \Delta_j$,

$$c_d 2^{-j} \leq \operatorname{diam}(Q) \leq 2^{-j} \text{ and } c_d c_\mu 2^{-j} \leq \mu(Q) \leq c_\mu^{-1} 2^{-j}.$$

To prove the $L^2(\mu)$ boundedness of the operator T^ε it suffices to verify that there exists a uniform bound $C < \infty$ that can depend on c_μ, c_d so that

$$\|T^\varepsilon \chi_S\|^2_{L^2(S)} \leq C\mu(S), \qquad \forall S \in \Delta, \forall \varepsilon > 0 \tag{6.1}$$

where $L^2(S) := L^2(\mu|_S)$. These conditions suffice by the $T1$ theorem of David and Journé, applied in the homogeneous metric measure space (E, d, μ), see [31, Theorem 3.21]. Notice that since K is symmetric, $(T^\varepsilon)^* = T^\varepsilon$ where $(T^\varepsilon)^*$ is the formal adjoint of T^ε, see also [2, Remark 2.3]. The statement in Tolsa's book is formulated for Euclidean spaces, but the proof works with minor standard changes in homogeneous metric measure spaces; the details can be found in the honors thesis of Fernando [12]. Observe that we may suppose that E is a 1-regular rectifiable curve as taking a subset can only decrease the $L^2(\mu)$-bound of $T^\varepsilon \chi_S$.

We will now decompose our singular integral dyadically. This approach was used in [2] and [3] and is inspired by [30]. Let $\psi : \mathbb{R} \to \mathbb{R}^+$ be a Lipschitz function so that $\chi_{B(0,1/2)} \leq \psi \leq \chi_{B(0,2)}$. For any $j \in \mathbb{Z}$ we let $\psi_j : \mathbb{G} \to \mathbb{R}$ such that $\psi_j(z) = \psi(2^j d(z,0))$ and we set $\phi_j := \psi_j - \psi_{j+1}$. Note that ϕ_j is supported on the annulus $B(0, 2^{1-j}) \backslash B(0, 2^{-2-j})$ and for any $N \in \mathbb{Z}$, $\sum_{n \leq N} \phi_n = 1 - \psi_{N+1}$, hence

$$\chi_{\mathbb{G} \backslash B(0,2^{-N})} \leq \sum_{n \leq N} \phi_n \leq \chi_{\mathbb{G} \backslash B(0,2^{-N-2})}. \tag{6.2}$$

For each $j \in \mathbb{Z}$, we let $K_{(j)} := \phi_j \cdot K$ and we define

$$T_{(j)} f(x) = \int K_{(j)}(y^{-1}x) f(y) \, d\mu(y).$$

for nonnegative functions $f \in L^2(\mu)$. For $N \in \mathbb{Z}$ let $S_N = \sum_{n \leq N} T_{(n)}$. As the kernel K is positive, (6.2) implies the following *pointwise* estimates for any nonnegative function $f \in L^2(\mu)$

$$0 \leq T^\varepsilon f \leq S_n f, \qquad \forall \varepsilon \geq 2^{-n}.$$

Thus, to establish the uniform bound (6.1), it suffices to show that there exists some absolute constant $C < \infty$ such that

$$\|S_n \chi_S\|^2_{L^2(S)} \leq C\mu(S), \qquad \forall S \in \Delta, \forall n \in \mathbb{Z}. \tag{6.3}$$

We now fix $S \in \Delta_\ell$ for some $\ell \in \mathbb{Z}$. We will show that for any $j \in \mathbb{Z}$ and $x \in E$, we have

$$T_{(j)}1(x) \lesssim_{c_\mu} \beta_E(x, 2^{1-j})^{2r^2}. \tag{6.4}$$

In order to prove (6.4) we need the following lemma which was first proven in the case of the Heisenberg group in [23, Lemma 3.3].

Lemma 6.2. *Let $(\mathbb{G}, d)$ be Carnot group of step $r \geq 2$ equipped with a homogeneous metric d. Then*

$$\frac{d(NH(a^{-1}b), 0)^r}{d(a,b)^{r-1}} \lesssim \max\{d(a, L), d(b, L)\} \tag{6.5}$$

for any $a, b \in \mathbb{G}$ and any horizontal line $L \subset \mathbb{G}$.

Proof. For any $p \in \mathbb{G}$, we will write $p = (p_1, \ldots, p_r)$ where $p_k \in \mathbb{R}^{v_k}$ and $v_k = \dim V_k$. As in the previous section, we will utilize the homogeneous norm

$$\|p\|_\infty = \max\{|p_k|^{1/k}\}_{k=1}^r.$$

For $x, y \in \mathbb{G}$ we will denote $d_\infty(x, y) := \|y^{-1}x\|_\infty$. Note that d_∞ is not a true metric since it does not satisfy the triangle inequality. Rather, there is a sub-additive constant $C_\infty \geq 1$. Regardless, it follows that d_∞ is globally equivalent to d in the sense of (2.3). Fix $a, b \in \mathbb{G}$ and a horizontal line $L \subset \mathbb{G}$. Note that

$$\|NH(a^{-1}b)\|_\infty \leq C_\infty(\|\tilde{\pi}(a^{-1}b)\|_\infty + \|a^{-1}b\|_\infty) \leq 2C_\infty\|a^{-1}b\|_\infty = 2C_\infty d_\infty(a, b). \tag{6.6}$$

If $d_\infty(a, b) < \max\{d_\infty(a, L), d_\infty(b, L)\}$, then

$$\frac{d_\infty(NH(a^{-1}b), 0)^r}{d_\infty(a, b)^{r-1}} \leq (2C_\infty)^r d_\infty(a, b) < (2C_\infty)^r \max\{d_\infty(a, L), d_\infty(b, L)\}.$$

Thus we may assume $d_\infty(a, b) \geq \max\{d_\infty(a, L), d_\infty(b, L)\}$.

Write $d := d_\infty(a, b)$, and choose $\ell_a, \ell_b \in L$ so that $d_\infty(a, L) = d_\infty(a, \ell_a)$ and $d_\infty(b, L) = d_\infty(b, \ell_b)$. Without loss of generality, we may assume that $\ell_a = 0$ so that $\ell_b = (x, 0, \ldots, 0)$. We have

$$\frac{\|NH(a^{-1}b)\|^r_\infty}{d^{r-1}} = d\left(\frac{\|NH(a^{-1}b)\|_\infty}{d}\right)^r = d\|NH(\delta_{1/d}(a^{-1}b))\|^r_\infty \lesssim d|NH(\delta_{1/d}(a^{-1}b))|.$$

This last inequality follows from (2.2) with a constant depending only on $\mathbb{G}$ since (6.6) implies

$$NH(\delta_{1/d}(a^{-1}b)) \in B_\infty(0, 2C_\infty)$$

for any choice of a and b. We can write $c = \ell_b^{-1}b$ so that $a^{-1}b = a^{-1}\ell_b c$ and $\|c\|_\infty = d_\infty(b, L)$. This gives

$$NH(\delta_{1/d}(a^{-1}b)) = \tilde{\pi}(\delta_{1/d}(a^{-1}\ell_b c))^{-1}\delta_{1/d}(a^{-1}\ell_b c) = (0, Q)$$

where Q is a Lie bracket polynomial determined by the BCH formula. As in the proof of Lemma 3.2, Q is a finite sum of constant multiples of terms of the form

$$[Z_1, [Z_2, \cdots, [Z_{k-2}, [Z_{k-1}, Z_k]] \cdots]] \tag{6.7}$$

where each Z_j is either a_i/d^i, c_i/d^i, or x/d. (Again, we are abusing notation and identifying each a_i and c_i with the associated vector in $V_i \subset \mathfrak{g}$.) The definition of $\|\cdot\|_\infty$ gives

$$\frac{|a_i|}{d^i} \le \left(\frac{d_\infty(a,\ell_a)}{d_\infty(a,b)}\right)^i = \left(\frac{d_\infty(a,L)}{d_\infty(a,b)}\right)^i \le \frac{d_\infty(a,L)}{d_\infty(a,b)}$$

since, by assumption, $d_\infty(a,L) \le d_\infty(a,b)$. Similarly, $|c_i|/d^i \le d_\infty(b,L)/d_\infty(a,b)$. We also have

$$C_\infty^{-2}|x| = C_\infty^{-2}\|\ell_b\|_\infty \le \|a\|_\infty + \|a^{-1}b\|_\infty + \|b^{-1}\ell_b\|_\infty = d_\infty(a,L) + d_\infty(a,b) + d_\infty(b,L) \le 3d.$$

Therefore, $|x|/d \le 3C_\infty^2$. Now, each nested bracket of the form (6.7) with $k \ge 2$ must contain at least one term a_i/d^i or c_i/d^i (since, otherwise, we would have $Z_j = x/d$ for each j, so the brackets would all vanish). Since $\max\{d_\infty(a,L), d_\infty(b,L)\}/d_\infty(a,b) \le 1$, this gives

$$|[Z_1,\cdots,[Z_{k-1},Z_k]]\cdots]| \lesssim \prod_{j=1}^k |Z_j| \lesssim \frac{\max\{d_\infty(a,L), d_\infty(b,L)\}}{d_\infty(a,b)}.$$

Since the sum in the BCH formula is finite, we have

$$\frac{d_\infty(NH(a^{-1}b),0)^r}{d_\infty(a,b)^{r-1}} \lesssim d_\infty(a,b)|NH(\delta_{1/d}(a^{-1}b))| \lesssim \max\{d_\infty(a,L), d_\infty(b,L)\}.$$

This completes the proof of the lemma. □

Let $A = E \cap A(x, 2^{-2-j}, 2^{1-j})$. Since ψ is Lipschitz, we have $\phi_j(y^{-1}x) \lesssim 2^{j+2} d(y,x)$. Hence

$$\begin{aligned} T_{(j)}1(x) = \int_E \phi_j(y^{-1}x)K(y^{-1}x)\, d\mu(y) &\lesssim 2^{j+2}\int_A \frac{d(NH(y^{-1}x),0)^{2r^3} + d(NH(x^{-1}y),0)^{2r^3}}{d(y,x)^{2r^3}}\, d\mu(y) \\ &\lesssim \sup_{y\in A} \frac{d(NH(y^{-1}x),0)^{2r^3} + d(NH(x^{-1}y),0)^{2r^3}}{d(x,y)^{2r^3}}. \end{aligned}$$

Observe that, if $y \in A$, it holds that $d(x,y) \ge 2^{-j-2}$. Moreover, there exists a horizontal line L such that

$$\begin{aligned} \beta_{\{x,y\}}(x, 2^{1-j}) = \frac{\max\{d(x,L), d(y,L)\}}{2^{1-j}} &\gtrsim \frac{\max\{d(x,L), d(y,L)\}}{d(x,y)} \\ &\overset{(6.5)}{\gtrsim} \frac{d(NH(y^{-1}x),0)^{2r^3} + d(NH(x^{-1}y),0)^r}{d(x,y)^r}. \end{aligned}$$

Hence (6.4) follows as $\beta_E(B(x,2^{1-j})) \ge \beta_{\{x,y\}}(x,2^{1-j})$.

If $Q \in \Delta_j$ for some $j \in \mathbb{Z}$ we define

$$\beta_E(Q) := \beta_E(p_Q, 2^{2-j}).$$

Note that if $R \in \Delta_j$ for some $j \in \mathbb{Z}$ then (6.4) implies that for any $\alpha > 0$

$$\int_R T_{(j)}1(x)^\alpha\, d\mu(x) \lesssim_{c_\mu} \beta_E(R)^{2r^2\alpha}\mu(R). \tag{6.8}$$

Using (6.4) and (6.8) and arguing exactly as in [3, pp 1416-1417] we deduce that if $S \in \Delta_j$ for some $j \in \mathbb{Z}$, then

$$\|S_n\chi_S\|_{L^2(S)}^2 \lesssim_{c_\mu} \sum_{Q\in\Delta(S^*)} \beta(Q)^{2r^2}\mu(Q) \tag{6.9}$$

where S^* is the unique cube in Δ_{j-2} such that $S \subset S^*$.

Now let $P \in \Delta_j, j \in \mathbb{Z}$, and denote by Σ_P the set of connected components of $B(p_P, 2^{1-j}) \cap E$ which intersect P. Note that $\sharp\Sigma_P \lesssim_{c_\mu} 1$, because each member of Σ_P has length comparable to $\mathcal{H}^1(E \cap B(p_P, 2^{1-j}))$. Therefore

$$\sum_{Q\in\Delta(P)} \beta_E(Q)^{2r^2}\mu(Q) \lesssim_{c_\mu} \sum_{\Gamma\in\Sigma_P} \int_{\mathbb{G}} \int_0^\infty \beta_\Gamma(x,t)^{2r^2} \frac{dt}{t^Q} d\mathcal{H}^Q(x) \lesssim_{C,c_\mu} \mu(B(p_P, 2^{1-j})),$$

where we applied Theorem 1.4 for every $\Gamma \in \Sigma_P$ in order to obtain the second inequality. Therefore for any $P \in \Delta$, we have

$$\sum_{Q\in\Delta(P)} \beta_E(Q)^{2r^2}\mu(Q) \lesssim_{C,c_\mu} \mu(P). \tag{6.10}$$

Now (6.3) follows by (6.9), (6.10) and the 1-regularity of μ. The proof is complete. □

Remark 6.3. If $(\mathbb{G}, d)$ is a Carnot group of step 2, Theorem 6.1 is valid for the simpler kernel

$$K_d(p) = \frac{d(NH(p),0)^8}{d(p,0)^9}.$$

This follows because of Theorem 1.5 and the fact that $NH(p)^{-1} = NH(p^{-1})$ in Carnot groups of step 2.

References

[1] Calderón, A.-P. Cauchy integrals on Lipschitz curves and related operators. *Proc. Nat. Acad. Sci. U.S.A. 74*, 4 (1977), 1324–1327.

[2] Chousionis, V., Fässler, K., and Orponen, T. Boundedness of singular integrals on $C^{1,\alpha}$ intrinsic graphs in the Heisenberg group. *Submitted* (2017).

[3] Chousionis, V., and Li, S. Nonnegative kernels and 1-rectifiability in the Heisenberg group. *Anal. PDE 10*, 6 (2017), 1407–1428.

[4] Christ, M. A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral. *Colloq. Math. 60/61*, 2 (1990), 601–628.

[5] Coifman, R. R., McIntosh, A., and Meyer, Y. L'intégrale de Cauchy définit un opérateur borné sur L^2 pour les courbes lipschitziennes. *Ann. of Math. (2) 116*, 2 (1982), 361–387.

[6] David, G. Opérateurs d'intégrale singulière sur les surfaces régulières. *Ann. Sci. École Norm. Sup. (4) 21*, 2 (1988), 225–258.

[7] David, G. *Wavelets and singular integrals on curves and surfaces*, vol. 1465 of *Lecture Notes in Mathematics*. Springer-Verlag, 1991.

[8] David, G., and Semmes, S. Singular integrals and rectifiable sets in $\mathbb{R}^n$: Beyond Lipschitz graphs. *Astérisque*, 193 (1991), 152.

[9] David, G., and Semmes, S. *Analysis of and on uniformly rectifiable sets*, vol. 38 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1993.

[10] David, G., and Semmes, S. Quantitative rectifiability and Lipschitz mappings. *Trans. Amer. Math. Soc. 337*, 2 (1993), 855–889.

[11] David, G. C., and Schul, R. The analyst's traveling salesman theorem in graph inverse limits. *Ann. Acad. Sci. Fenn. Math. 42* (2017), 649–692.

[12] Fernando, S. The T1 Theorem in Metric Spaces. Undergraduate honors thesis, University of Connecticut, 2017.

[13] Ferrari, F., Franchi, B., and Pajot, H. The geometric traveling salesman problem in the Heisenberg group. *Rev. Mat. Iberoam. 23*, 2 (2007), 437–480.

[14] Franchi, B., Serapioni, R., and Serra Cassano, F. On the structure of finite perimeter sets in step 2 Carnot groups. *J. Geom. Anal. 13*, 3 (2003), 421–466.

[15] Hahlomaa, I. Menger curvature and Lipschitz parametrizations in metric spaces. *Fund. Math. 185*, 2 (2005), 143–169.

[16] Hahlomaa, I. Curvature integral and Lipschitz parametrization in 1-regular metric spaces. *Ann. Acad. Sci. Fenn. Math. 32*, 1 (2007), 99–123.

[17] Hebisch, W., and Sikora, A. A smooth subadditive homogeneous norm on a homogeneous group. *Studia Math. 96*, 3 (1990), 231–236.

[18] Jones, P. W. Square functions, Cauchy integrals, analytic capacity, and harmonic measure. In *Harmonic analysis and partial differential equations (El Escorial, 1987)*, vol. 1384 of *Lecture Notes in Math.* Springer, Berlin, 1989, pp. 24–68.
[19] Jones, P. W. Rectifiable sets and the traveling salesman problem. *Invent. Math. 102*, 1 (1990), 1–15.
[20] Le Donne, E. A metric characterization of Carnot groups. *Proc. Amer. Math. Soc. 143*, 2 (2015), 845–849.
[21] Li, S. Coarse differentiation and quantitative nonembeddability for Carnot groups. *J. Funct. Anal. 266*, 7 (2014), 4616–4704.
[22] Li, S., and Schul, R. The traveling salesman problem in the Heisenberg group: upper bounding curvature. *Trans. Amer. Math. Soc. 368*, 7 (2016), 4585–4620.
[23] Li, S., and Schul, R. An upper bound for the length of a traveling salesman path in the heisenberg group. *Rev. Mat. Iberoam. 32*, 2 (2016), 391–417.
[24] Okikiolu, K. Characterization of subsets of rectifiable curves in $\mathbb{R}^n$. *J. London Math. Soc. (2) 46*, 2 (1992), 336–348.
[25] Pajot, H. *Analytic capacity, rectifiability, Menger curvature and the Cauchy integral*, vol. 1799 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002.
[26] Schul, R. Ahlfors-regular curves in metric spaces. *Ann. Acad. Sci. Fenn. Math. 32*, 2 (2007), 437–460.
[27] Schul, R. Subsets of rectifiable curves in Hilbert space—the analyst's TSP. *J. Anal. Math. 103* (2007), 331–375.
[28] Serra Cassano, F. Some topics of geometric measure theory in Carnot groups. In *Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. 1*, EMS Ser. Lect. Math. Eur. Math. Soc., Zürich, 2016, pp. 1–121.
[29] Stein, E., and Murphy, T. *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals*. Monographs in harmonic analysis. Princeton University Press, 1993.
[30] Tolsa, X. Uniform rectifiability, Calderón-Zygmund operators with odd kernel, and quasiorthogonality. *Proc. Lond. Math. Soc. (3) 98*, 2 (2009), 393–426.
[31] Tolsa, X. *Analytic capacity, the Cauchy transform, and non-homogeneous Calderón-Zygmund theory*, vol. 307 of *Progress in Mathematics*. Birkhäuser/Springer, Cham, 2014.

V. Chousionis: Department of Mathematics, University of Connecticut, 341 Mansfield Road U1009, Storrs, Connecticut 06269, USA, vasileios.chousionis@uconn.edu

S. Li: Department of Mathematics, University of Connecticut, 341 Mansfield Road U1009, Storrs, Connecticut 06269, USA, sean.li@uconn.edu

S. Zimmerman: Department of Mathematics, University of Connecticut, 341 Mansfield Road U1009, Storrs, Connecticut 06269, USA, scott.zimmerman@uconn.edu