The Large Scale Geometry of Nilpotent Lie Groups

SCOTT D. PAULS

In this paper, we prove results concerning the large scale geometry of connected, simply connected nonabelian nilpotent Lie groups equipped with left invariant Riemannian metrics. Precisely, we prove that there do not exist quasi-isometric embeddings of such a nilpotent Lie group into either a CAT_0 metric space or an Alexandrov metric space with curvature bounded below. The main technical aspect of this work is the proof of a limited metric differentiability of Lipschitz maps between connected graded nilpotent Lie groups equipped with left invariant Carnot-Carathéodory metrics and complete metric spaces.

1. Introduction.

In this paper we investigate the large scale geometry of connected nilpotent Lie groups equipped with left invariant Riemannian metrics by studying their quasi-isometric embeddings into various metric spaces. Let (N, g) be a connected nilpotent Lie group with a left invariant Riemannian metric and d be the induced distance function on N. If (X, d_X) is a complete metric space, then $f : N \to X$ is an (L, C)-quasi-isometric embedding if, for all $x, y \in N$,

$$\frac{1}{L} d(x, y) - C \leq d_X(f(x), f(y)) \leq L d(x, y) + C.$$

After studying certain invariants of these maps, we prove two main applications:

Theorem A. *There do not exist quasi-isometric embeddings of a connected nonabelian nilpotent Lie group equipped with a left invariant Riemannian metric into a CAT_0 metric space.*

Theorem B. *There do not exist quasi-isometric embeddings of a connected nonabelian nilpotent Lie group equipped with a left invariant Riemannian metric into a CBB_0 metric space.*

CAT_κ (resp. CBB_κ) metric spaces are spaces of curvature bounded above (resp. below) by κ is the sense of Topanogov triangle comparison. Thus, a CAT_0 metric space is a generalized space of nonpositive curvature while a CBB_0 metric space is a generalized space of nonnegative curvature. These include simply connected Riemannian manifolds of nonpositive and nonnegative curvature respectively. In the literature, CAT_0 spaces are also called Hadamard spaces and CBB_κ spaces are called Alexandrov spaces (with curvature bounded below). We retain the earlier notion for simplicity and consistency. We will discuss these in detail later in the paper. In [Wol64], Wolf showed that a connected nonabelian nilpotent Lie group equipped with a left invariant Riemannian metric must contain 2-planes of both positive and negative curvatures. Theorems A and B can be viewed as a large scale analogue of this theorem of J. Wolf.

While the large scale characteristics of such nilpotent Lie groups may be interesting in their own right, we shall see that such investigations reduce to problems which are motivated by the now standard arguments used in the proof of Mostow's rigidity theorem and its many extensions (see [Mos73], [Bal85], [BS87], [Pan89], [Ham91] as well as many others). In the proofs of these results, one attempts to show that two candidate spaces (e.g. two compact constant negative curvature spaces with isomorphic fundamental groups) are isometric by exhibiting an equivariant quasi-isometry between their universal covers and showing that the existence of such a quasi-isometry must imply the existence of a true equivariant isometry. A common element in many such proofs is to reduce to an examination of the ideal boundaries of the two spaces in question. The quasi-isometries between the two spaces induce quasiconformal maps on the boundaries which then become the focus of the study. In the cases of the papers mentioned above, the tangent cones at points on the ideal boundaries are isometric to connected simply connected graded nilpotent Lie group equipped with left invariant Carnot-Carathéodory metrics. Thus, a local analysis of the ideal boundaries involves an examination of the geometry of graded nilpotent Lie groups with left invariant Carnot-Carathéodory metrics. This provides a natural motivation for considering the geometry of such spaces and, in particular, of studying the quasi-conformal maps between them. It is a natural generalization to instead consider the quasi-conformal embeddings of such spaces. Unfortunately, the study of such embeddings seems quite intractable but does provide motivation for various lines of study - for example the study of biLipschitz embeddings of such spaces. As we shall see, the proofs of theorems A and B rest on a local analysis of biLipschitz embeddings of graded nilpotent Lie groups with left invariant Carnot-Carathéodory metrics into complete

metric spaces which are *locally CAT_κ* or CBB_κ. For metric spaces which are locally CAT_κ, i.e. for each point there is a closed ball about that point which is itself CAT_κ, we use the notation CBA_κ (for "curvature bounded above"). Note that, for a lower curvature bound, these distinctions are not necessary. In particular, we prove an intermediate theorem from which theorems A and B follow:

Theorem C. *Let G be a connected simply connected graded nonabelian nilpotent Lie group equipped with a left invariant Carnot-Carathéodory metric and $U \subset G$ be an open set. Then U does not admit a biLipschitz embedding into any CBA_κ metric space or into any CBB_κ metric space.*

Remark. Consider the special case of theorem C when the metric space target is a Riemannian manifold, M, of either nonpositive or nonnegative curvature. Note first that the Hausdorff dimension of M coincides with the topological dimension of M, while the Hausdorff dimension of G is strictly larger than its topological dimension. Since both dimensions are biLipschitz invariants, this proves theorem C in this case. This points out that the strength of the theorem lies in the case when the metric space is not smooth. In particular, there are examples of metric spaces with the specified curvature conditions which have the same type of disparity between their Hausdorff and topological dimension, for which the above argument does not apply.

To prove the main theorems from this intermediate theorem, we study some biLipschitz embedded invariants of graded nilpotent Lie groups with left invariant Carnot-Carathéodory metrics and use them to build local obstructions to quasi-isometric embeddings of nonabelian nilpotent Lie groups with left invariant Riemannian metrics into various metric spaces. Next we give a brief outline of the argument and the paper and show how Theorems A and B follow from Theorem C.

Let N be a connected simply connected nilpotent Lie group with a left invariant Riemannian metric g and let d be the induced distance function on N. If $f : N \to X$ is some (L, C)-quasi-isometric embedding of N into a complete metric space X, we first consider asymptotic cones of N and X and a map between them, denoted F, derived from f and the coning procedure. The asymptotic cone of a metric space (X, d_X) is a limit metric space of the pointed dilated spaces $(X, x_i, \lambda_i d_X)$ for some sequence $\lambda_i \to 0$. Since such spaces do not necessarily converge in the Gromov-Hausdorff topology, we use Gromov's ultrafilter construction, choosing a nonprincipal ultrafilter ω

(see section 5 for a definition) to form the asymptotic cone $(C_\infty^\omega X, d_X^\omega)$. For precise details of the construction, we refer the reader to either [Gro93] or [KL96]. For our purposes, there are a few keys pieces of information. First, if (N, g) is a nilpotent Lie group with a left invariant Riemannian metric, then Pansu ([Pan89]) proved that the asymptotic cone is unique and isometric to $(\widetilde{N}, d_{cc})$, a graded nilpotent Lie group equipped with a left invariant Carnot-Carathéodory metric (see section 3 for definitions). Second, an asymptotic cone of a CAT_0 space is also CAT_0 and an asymptotic cone to a CBB_0 space is also CBB_0. These facts (and references to their proofs) are reviewed in section 6. Third, the quasi-isometric embedding f gives rise to an L-biLipschitz map F between the cones $(\widetilde{N}, d_{cc})$ and $(C_\infty^\omega X, d_X^\omega)$. This fact is a simple consequence of the definition of asymptotic cones and can be found, for example, in section 2 of [KL96]. The construction of this map depends on the choice of ultrafilter. Thus, to prove theorems A and B, we must prove instead theorem C. The paper is devoted to proving this intermediate theorem.

The main technical result in the paper (theorem 4.3) is that, in an appropriate sense, such an L-biLipschitz map is differentiable in certain directions almost everywhere. This is a generalization of a classical theorem of Rademacher. Our proof follows the same lines as a covering argument used in Kleiner's proof of the differentiability of Lipschitz maps from $\mathbb{R}^n$ into metric spaces (in [Kle97]), which is a special case of differentiability theory of section 1.9 in [KS93]. In [Kir94], Kirchheim shows the metric differentiability of Lipschitz maps $f : \mathbb{R}^n \to X$ where the targets are complete metric spaces. Our differentiability result is an extension of Kirchheim's work. The interested reader should also consult [Pan89] and [MM95] for results concerning the differentiability of quasiconformal maps between Carnot-Carathéodory spaces.

To prove Theorem C, we construct a tangent cone of F at a point of differentiability. Using this tangent cone, we can compare the local geometry of the two asymptotic cones. The local geometry of connected graded nilpotent Lie groups with left invariant Carnot-Carathéodory metrics is well understood and the differentiability allows us to "push forward" this structure to a tangent cone to the asymptotic cone of X. When this asymptotic cone has additional structure, this allows us to measure the compatibility of these two objects. In the case of Theorem C, this structure provides estimates on the rate of growth of the spread between two tangent cone geodesics which show that they spread apart sublinearly but do not remain a bounded distance from one another. Comparing this to the spread of geodesics in the tangent cone to either a CAT_0 or CBB_0 space, we derive a contradiction

which proves Theorem C.

Sections 2 and 3 are reviews of the constructions mentioned above and discuss, respectively, nilpotent Lie groups and Carnot-Carathéodory metrics. Section 4 contains the proof of the limited differentiability of biLipschitz embeddings of connected graded nilpotent Lie groups with left invariant Carnot-Carathéodory metrics into complete metrics spaces. The main goal of this section is to prove theorem 4.3. Section 5 reviews the tangent cone construction and interprets theorem 4.3 as a statement about maps between tangent cones. Section 6 reviews the definition and some properties of CAT_0 and CBB_0 spaces. In section 7 we prove Theorem C, that there do not exist biLipschitz embeddings of a nonabelian connected graded nilpotent Lie group equipped with a Carnot-Carathéodory metric into either a CBA_κ or CBB_κ space.

The author wishes to thank Bruce Kleiner for suggesting this line of work as well as for many helpful discussions. Thanks are also due to Chris Croke for helpful discussions of this work and for many hours of help with the preparation of this document. The author also wishes to thank the referee for a careful reading of this manuscript and many useful comments which led to several simplifications in proof and many improvements to the exposition.

2. Nilpotent Lie Groups.

In this section, we recall some basic definitions of nilpotent Lie groups and some associated structures. Let G be a connected Lie group and $\mathfrak{g}$ be its Lie algebra equipped with $[\cdot,\cdot]$ a bracket. For the rest of this exposition, we assume that all Lie groups are connected and simply connected.

Definition 1. Let $\mathfrak{g}$ be a Lie algebra over a field K, we define the descending central sequence, $\{C^k(\mathfrak{g})\}$ as follows:

$$C^0(\mathfrak{g}) = \mathfrak{g}$$
$$C^{k+1}(\mathfrak{g}) = [C^k(\mathfrak{g}), \mathfrak{g}].$$

We use the descending central sequence to define nilpotency:

Definition 2. A Lie group G is called *nilpotent* if there exists an integer n such that $C^n(\mathfrak{g}) = 0$. The smallest such n is called the degree of nilpotency of G.

Recall that the exponential map, $v \mapsto e^v$, is a diffeomorphism for simply connected nilpotent Lie groups. For this paper, we will use the following notation: if e^v and e^w are exponential images of the Lie algebra elements v and w, then th e product, $e^v e^w$ is also the exponential image of a Lie algebra element which we denote $v \odot w$. In other words, $e^v e^w = e^{v \odot w}$.

For this work, Pansu's theorem tells us that after taking asymptotic cones, graded nilpotent Lie groups are our main objects of study.

Definition 3. A nilpotent Lie group is graded if its Lie algebra comes with a grading. A grading for a Lie algebra $\mathfrak{g}$ is a decomposition

$$\mathfrak{g} = \oplus_i \mathcal{V}^i$$

where the subspaces $\mathcal{V}^i$ satisfy the condition

$$[\mathcal{V}^i, \mathcal{V}^j] \subset \mathcal{V}^{i+j}$$

for all $i, j \in \mathbb{Z}$.

Example. The three dimensional Heisenberg group is defined by giving a three dimensional Lie algebra spanned by vectors X, Y, Z with the only nontrivial bracket relation given by $[X, Y] = Z$. Thus, the Lie algebra has a grading $\mathfrak{h} = \mathcal{V}^1 \oplus \mathcal{V}^2$ where $\mathcal{V}^1 = span\{X, Y\}$ and $\mathcal{V}^2 = span\{Z\}$.

Every graded nilpotent Lie group possesses a self-similarity.

Definition 4. Let $g \in G$ be an element in a graded nilpotent Lie group of nilpotency degree $k+1$ such that $g = e^{g_1 + g_2 + \cdots + g_k}$ where $g_i \in \mathcal{V}^i$. For $t \geq 0$, consider the 1-parameter group of automorphisms, $\{h_t\}$, of G given by

$$h_t(g) = e^{t g_1 + t^2 g_2 + \cdots + t^k g_k}.$$

We may sometimes analyze this action on Lie algebra elements. Thus we define $dh_t(v_1 + v_2 + \cdots + v_k) = t v_1 + t^2 v_2 + \cdots + t^k v_k$. If we wish to consider $h_{-|t|} g$ we use the convention that $h_{-|t|} g = h_{|t|} g^{-1}$. (Please note that this convention is not standard in the literature.)

For general Lie groups, the Campbell-Baker-Hausdorff formula provides a method for computing the product of two nearby group elements. We state it in the special case when N is a simply connected graded nilpotent Lie group. For a more general version and, for that matter, a proof, see [Var84] Section 2.15.

Theorem 2.1 (Campbell-Baker-Hausdorff Formula). *Let N be a simply connected graded nilpotent Lie group of nilpotency degree $k + 1$, let $\mathfrak{n}$ be its Lie algebra and let $e^{\cdot} : \mathfrak{n} \to N$ denote the exponential map at the identity. Then, given $X, Y \in \mathfrak{n}$*

$$X \odot Y = \sum_{i=1}^{k-1} \frac{(-1)^{i+1}}{i} \sum_{\substack{1 \leq j \leq i, \\ p_j + q_j > 0}} \frac{\sum_{j=1}^{i}(p_j + q_j)^{-1}}{p_1! q_1! \ldots p_n! q_n!}$$
$$\cdot (adX)^{p_1}(adY)^{q_1} \cdots (adX)^{p_n}(adY)^{q_n - 1} Y$$

where $(adX)Y = [X, Y]$ and if $q_n = 0$ then the last term in the sum is $(adX)^{p_1}(adY)^{q_1} \cdots (adX)^{p_n - 1} X$. In other words, we have the formula:

$$e^X e^Y = e^{X \odot Y} = e^{X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + \cdots}.$$

In the case that N is a graded nilpotent Lie group, we use the following convention. Let $C_i(X, Y)$ be the pieces of $X \odot Y$ which lie in $\mathcal{V}^i$. Thus, if $X, Y \in \mathcal{V}^1$, $C_1(X, Y) = X + Y$, $C_2(X, Y) = \frac{1}{2}[X, Y]$ and so on. Thus, the Campbell-Baker-Hausdorff formula yields $e^{X \odot Y} = e^{C_1(X,Y) + \cdots + C_k(X,Y)}$. Using the Campbell-Baker-Hausdorff formula, we easily confirm that h_t is a group automorphism.

In subsequent sections, we will use repeatedly several properties of h_t. We group them here as a lemma for easy reference:

Lemma 2.2. *Let N be a graded nilpotent Lie group and let h_t be the automorphism described above.*

- *For $n \in N$, $h_t h_s n = h_{ts} n$.*

- *For $v \in \mathcal{V}^1$, $h_t e^v h_s e^v = h_{t+s} e^v = e^{(t+s)v}$.*

- *If dL_g is the derivative of the left translation map $n \mapsto gn$, then $dh_t \circ dL_n = dL_{h_t n} \circ dh_t$.*

- *$dh_t(v \odot w) = dh_t(v) \odot dh_t(w)$.*

Proof. The first two are simple exercises using the Campbell-Baker-Hausdorff formula. The last two are restatements of the fact that h_t is a group automorphism in terms of dh_t. $\square$

3. Carnot-Carathéodory metrics.

Carnot-Carathéodory metrics arise in a variety of settings and have been the object of a great deal of study. In this work, we examine graded nilpotent Lie groups equipped with Carnot-Carathéodory metrics and hence the definitions and propositions will be tailored for this small set of spaces. For a more general and comprehensive introduction to the theory, the interested reader is encouraged to consult [Str86], [Gro96], and [BR96]. The rough idea of a Carnot-Carathéodory metric is to measure the distance between two points by taking the infimum of the length of piecewise C^1 paths connecting the points over a restricted set of paths, those tangent to some subbundle of the tangent bundle. More precisely (in our setting), if N is a connected graded nilpotent Lie group with Lie algebra grading $\mathcal{V}^1 \oplus ... \oplus \mathcal{V}^k$, we define a left invariant subbundle of the tangent bundle, $\mathcal{V}$, by letting the fiber at a point be the left translate of the subspace $\mathcal{V}^1$. In the rest of the paper, we will, by abuse of notation, refer to the subbundle and the fiber by $\mathcal{V}$. If we place a norm $|\cdot|$ on each fiber of $\mathcal{V}$, we define the Carnot-Carathéodory distance between $x, y \in N$ by

$$d_{cc}(x,y) = \inf_{\gamma}\left\{ \int |\gamma'(t)|dt \middle| \gamma \text{ connects } x \text{ to } y \text{ and } \gamma' \subset \mathcal{V} \right\}.$$

If no such paths exist, we set $d_{cc}(x,y) = \infty$. We call a path whose tangent is in the subbundle $\mathcal{V}$ at each point a horizontal path. Note that d_{cc} defines a left invariant metric on N. In general, if $|\cdot|$ is merely a norm, we call the quadruple $(N, \mathcal{V}, |\cdot|, d_{cc})$ a Carnot-Carathéodory group or a sub-Finsler group (if we wish to emphasize that $|\cdot|$ is only a norm) and d_{cc} a Carnot-Carathéodory or sub-Finsler metric. If $|\cdot|$ arises from an inner product, $\langle\cdot,\cdot\rangle$, on $\mathcal{V}$, we denote the resulting space by $(N, \mathcal{V}, \langle\cdot,\cdot\rangle, d_{cc})$ and we call it a sub-Riemannian group. This nomenclature and notation is somewhat different from the literature which often interchanges the terms Carnot-Carathéodory metric, sub-Riemannian metric, and singular Riemannian metric freely. In this exposition, any object adorned with a "cc" will be constructed to be with respect to d_{cc}. For example, $B_{cc}(x,r)$ is the d_{cc}-ball of radius r with center x.

It is a consequence of Chow's theorem ([Cho39] or see [Gro96] for a exhaustive discussion of this fact) that any two points in $(N, \mathcal{V}, |\cdot|, d_{cc})$ can be connected with a smooth horizontal path. This leaves open the question of geodesics in Carnot-Carathéodory spaces. We call a segment a geodesic if it is a local length minimizer. Again, in the literature, there are many different terms/meanings for geodesic paths and/or length minimizing paths

depending on whether the geodesics satisfy certain differential equations (see, for example, [LS94] and [Mon96]). For this exposition, these issues are not relevant and hence we will use the simplified language. In a sub-Riemannian group $(N, \mathcal{V}, \langle \cdot, \cdot \rangle, d_{cc})$, there is a nice class of geodesics which will be useful in later sections.

Definition 5. Let $(N, \mathcal{V}, \langle \cdot, \cdot \rangle, d_{cc})$ be a left invariant sub-Riemannian metric where N is a connected graded nilpotent Lie group and fix $n \in N$, $v \in \mathcal{V}$. Let V be a smooth left invariant vector field on N with $V(e^0) = v$. Let γ be an integral curve of V such that $\gamma(0) = n$. We call γ a radial geodesic eminating from n. If a basepoint is understood, we shall simply call γ a radial geodesic.

Note that radial geodesics are actually geodesics of d_{cc}. Indeed, considering the quotient $N/[N, N]$ endowed with the left invariant metric induced by the cc inner product on $\mathcal{V}$, the quotient map $N \to N/[N, N]$ is 1-Lipschitz and takes radial "geodesics" to geodesics in the quotient and preserves their length. Thus, these are geodesics of d_{cc} as well. In addition, we see that h_t is a homothety of d_{cc}. Since h_t is an automorphism of N and dh_t preserves $\mathcal{V}$ and acts as a homothety on $\mathcal{V}$, we see that h_t is a homothety on (N, d_{cc}).

In [Mit85], Mitchell describes the local structure of more general Carnot-Carathéodory spaces. If $(N, \mathcal{V}, \langle \cdot, \cdot \rangle, d_{cc})$ is a sub-Riemannian group, the local structure is much easier to determine given the existence of a left invariant metric, a homothety of that metric, and the uniqueness of the Haar measure. The first two following results can either be thought of as Mitchell's results in this special case or as consequences the extra structure in this special case.

Proposition 3.1. *Let $(N, \mathcal{V}, \langle \cdot, \cdot \rangle, d_{cc})$ be a sub-Riemannian group. Then,*

1. *For any $n \in N$, the tangent cone at n, $(C_n N, \overline{d}_{cc})$ exists, is unique, and is isometric to (N, d_{cc}).*

2. *Let $Q = \sum_{i=1}^{k} i \dim \mathcal{V}^i$. Then Q is the Hausdorff dimension of (N, d_{cc}).*

3. *Let $\langle \cdot, \cdot \rangle_R$ be a Riemannian completion of the inner product $\langle \cdot, \cdot \rangle$ on all of $\mathfrak{n}$ such that the grading is orthogonal. Let "dvol" denote the Riemannian volume measure associated to this left invariant Riemannian metric. Then, if $\mathcal{H}^Q$ denotes the Q dimensional Hausdorff measure, $\mathcal{H}^Q$ and dvol are constant multiples of one another.*

The last part of the previous theorem follows from the fact that Haar measure is unique up to scale - both the Hausdorff measure generated from d_{cc} and $dvol$ are left invariant measure on N. In addition, letting $\mathcal{L}$ be the pull back of Lebesgue measure on $\mathfrak{n}$ by the exponential map, $\mathcal{L}$ is also a constant multiple of $\mathcal{H}^Q$ and $dvol$. One should note that in Mitchell's work on more general Carnot-Carathéodory spaces, the Hausdorff and Riemannian measures are only commensurable on compact sets. However, we now see that in our setting, the left invariance of the metric and the presence of a homothety allows us have a global constant for the volume and Hausdorff measure comparisons. First, we start with an illustrative example:

Example. Let H^3 denote the three dimensional Heisenberg group with a left-invariant sub-Riemannian metric. It is well known that the Haar measure on H^3 is, in fact, the four dimensional Hausdorff measure with respect to d_{cc} on H^3 (see [Pan83]). This confirms the formula from Mitchell's theorem. Indeed,

$$\mathcal{V}^1 = span\left(\{X, Y\}\right), \ \mathcal{V}^2 = span\left(\{Z\}\right), \ \mathcal{V}^3 = \emptyset$$
$$\text{therefore } Q = 1 \cdot 2 + 2 \cdot 1 = 4.$$

Consider what h_r, with $r > 0$, does to the Riemannian volume. Denote by g the Riemannian metric given by the inner product $\langle \cdot, \cdot \rangle_R$ described above. At the identity, the metric $h_r^* g$, written in terms of the left invariant vector fields X, Y and Z looks like

$$\begin{pmatrix} r^2 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^4 \end{pmatrix}.$$

Thus, the Riemannian volume element determined by the metric g is multiplied by r^4 under h_r^*. Thus, picking C to work for a particular value of r_0, we can then use the dilation h_r to scale the picture so that the same constant works for any r value larger than zero.

This argument actually works for more general graded nilpotent Lie groups yielding the same comparative scaling between the Carnot-Carathéodory metric and the Riemannian metric.

In the proof of the limited differentiability of biLipschitz maps, we will need to be able to manipulate covers of sets in graded nilpotent Lie groups while using the measure $\mathcal{H}^Q$. In particular, we want to be able to take the analogue of a Vitali cover of a measurable set S and find a countable disjoint subcover which still covers almost all of S with respect to $\mathcal{H}^Q$. For

this section, we will use the language of [Fed69] (section 2.8). Assume that N is a connected graded nilpotent Lie group with a sub-Riemannian metric, d_{cc}. Let $\mathcal{H}^Q$ be the Q dimensional Hausdorff measure associated to d_{cc} as in last section. Let $\mathcal{C}$ be the collection of closed cc-balls in N. Then, mimicing the standard proof in $\mathbb{R}^n$ we prove the following lemma.

Lemma 3.2. $\mathcal{C}$ *is a Vitali cover. In other words, given a measurable set A covered by a fine subcollection of $\mathcal{C}$, we may find a countable disjoint subcover, $\mathcal{C}_A$ such that* $\mathcal{H}_{cc}^Q(A \setminus \cup_{C \in \mathcal{C}_A} C) = 0.$

4. Differentiability of biLipschitz maps.

Let $(N, \mathcal{V}, \langle \cdot, \cdot \rangle, d_{cc})$ be a graded nilpotent Lie group of nilpotency degree $k + 1$ with a sub-Riemannian metric. To set up the notation, we let $F : (N, d_{cc}) \to (X, d_X)$ be an L-biLipschitz map of N into a complete metric space. Let d be $F^* d_X$, the metric d_X pulled back through F. We want to show that, appropriately defined, d is differentiable in $\mathcal{V}$ directions at most points. As discussed in the introduction, to prove this, we follow the style of the covering argument in the appendix of [Kle97] which Kleiner used to prove a version of a result of Korevaar and Schoen's [KS93]. One can also consider this a version of Kirchheim's metric differentiability in [Kir94] although the argument is different.

4.1. Definitions and Preliminary lemmas.

In examining differentiability, we consider an analog of the directional derivative at x in the direction of $v \in \mathcal{V}$ by looking at difference quotients along the paths $y h_t e^v$ for $t > 0$:

$$\frac{d(y, y h_t e^v)}{t} \text{ where } y \in B_{cc}(x, t)$$

We consider whether such a "directional derivative" exists (i.e., the limit as $t \to 0$ exists) by examining the "lower" and "upper" limits. If U is some open set in N, we define these limits as follows. Examination of this type of quotient is inspired by Pansu's work (see [Pan89], [Pan83]) where he considered similar types of quotients to investigate differentiability properties of maps between subRiemannian groups.

Definition 6. The *lower derivate* of d is the function $\underline{\rho} : U \times \mathfrak{n} \to \mathbb{R}$ defined by:

$$\underline{\rho}(x, v) = \underline{\lim}_{t \to 0^+} \left\{ \frac{d(y, yh_t e^v)}{t} \bigg| y \in B_{cc}(x, t) \right\}.$$

Definition 7. The *upper derivate* of d is the function $\overline{\rho} : U \times \mathfrak{n} \to \mathbb{R}$ defined by:

$$\overline{\rho}(x, v) = \overline{\lim}_{t \to 0^+} \left\{ \frac{d(y, yh_t e^v)}{t} \bigg| y \in B_{cc}(x, t) \right\}.$$

Note that both $\underline{\rho}$ and $\overline{\rho}$ are measurable functions. Properly viewed, they are just the lim inf and lim sup of measurable functions of more than one variable (see [Fed69] page 152).

To understand and estimate the lower and upper derivates, we want to estimate the quotients for small values of t. To do this, we construct special boxes that aid in the estimates. For each $v \in \mathcal{V}$, denote by V the left invariant vector field on N determined by v. Choose $\langle \cdot, \cdot \rangle_R$ to be a Riemannian completion of the sub-Riemannian metric on $\mathcal{V}$ such that the grading is orthogonal.

Definition 8. Fix $v \in \mathcal{V}$ and let $V^\perp$ be the orthogonal complement of v in $\mathcal{V}$ with respect to $\langle \cdot, \cdot \rangle$. Let end $(\varepsilon, v) = \{w_1 + \cdots + w_k \in V^\perp \oplus \mathcal{V}^2 \oplus \cdots \oplus \mathcal{V}^k | w_1 \in V^\perp, w_i \in \mathcal{V}^i$ for $i > 1, \langle w_j, w_j \rangle_R^{\frac{1}{2}} < \varepsilon^j, \forall j\}$. Now, let

$$\text{End}(x, v, \varepsilon) = xe^{\text{end}(\varepsilon, v)}.$$

End (x, v, ε) forms the end of our box.

Note that $e^{\text{end}(\varepsilon, v)}$ is the exponential image of a carefully constructed rectangle which should be thought of as a square of side length 2ε. Recalling Theorem 3.6 in [KM96] or one of the various formulations of the "ball-box" estimates (see, for example [Gro96]), if $v = v_1 + \cdots + v_k \in \mathfrak{n}$ then $\sum_{i=1}^k \langle v_i, v_i \rangle_R^{\frac{1}{2i}} \leq A d_{cc}(e^0, e^v)$ where A is a constant depending only on the choice of $\langle \cdot, \cdot \rangle_R$ and the structure constants of the Lie group N. Clearly, the constant A is invariant under scaling by h_t: if $v' = dh_t v$ and v_i' are the components of v' in the various pieces of the grading, then $\sum_{i=1}^k \langle v_i', v_i' \rangle^{\frac{1}{2i}} = \sum_{i=1}^k t \langle v_i, v_i \rangle^{\frac{1}{2i}} \leq A t d_{cc}(e^0, e^v) = A d_{cc}(e^0, e^{dh_t v})$.

Definition 9. Define the estimate boxes by

$$\text{Box}(x, v, \varepsilon) = \bigcup_{z \in \text{End}(x, v, \varepsilon)} \left(\bigcup_{s \in [0,1]} zh_s e^v \right).$$

As in Kleiner's argument, one first thickens $xh_t e^v|_{t \in [0,1]}$ into the estimate box, $\mathrm{Box}\,(x, v, \varepsilon)$, and then uses smaller boxes, $\mathrm{Box}\,(y, tv, t\varepsilon)$ with y close to x, to cover it and help estimate the distance between the ends of the estimate box. To use this argument, we must ensure that the boxes we have constructed have metric properties (with respect to d_{cc}) similar to the boxes in $\mathbb{R}^n$, at least on the small scale. One important technical detail is to control the distances between points in the "far" end of the box with respect to corresponding distances in the "near" end. Let $z = yh_{t\varepsilon}e^{z_1 + z_2 + \cdots + z_k} \in End(y, tv, t\varepsilon) \subset Box(y, tv, t\varepsilon)$ with $d_{cc}(e^0, e^{z_1 + z_2 + \cdots + z_k}) \leq 1$ and $v \in \mathcal{V}$. Then,

$$
\begin{aligned}
(*) \quad & d_{cc}(yh_t e^v, zh_t e^v) \\
&= d_{cc}(e^0, e^{-tv}e^{t\varepsilon z_1 + (t\varepsilon)^2 z_2 + \cdots + (t\varepsilon)^k z_k}e^{tv}) \\
&= d_{cc}(e^0, e^{t\varepsilon z_1 + ((t\varepsilon)^2 z_2 + 2t^2\varepsilon c_2[z_1,v]) + \cdots + ((t\varepsilon)^k z_k + 2t^k\varepsilon c_{k-1}[z_{k-1},v] + \cdots)}) \\
&= td_{cc}(e^0, e^{\varepsilon z_1 + ((\varepsilon)^2 z_2 + 2\varepsilon c_2[z_1,v]) + \cdots + ((\varepsilon)^k z_k + 2\varepsilon c_{k-1}[z_{k-1},v] + \cdots)}) \\
&\leq tC(\varepsilon).
\end{aligned}
$$

Here, $C(\varepsilon)$ is a constant that depends on ε and goes to zero with ε. This constant exists by the compactness of the closed unit CC-ball. The first equality is due to the left invariance of d_{cc}. The second is a use of the Campbell-Baker-Hausdorff formula. The third follows from the fact the h_t is a homothety of d_{cc}.

As in Kleiner's argument, we will prove a relation between the upper and lower derivates by covering boxes which thicken the radial geodesic segment $xh_t e^v|_{t \in [0,1]}$ with well controlled smaller boxes. To do this, we make the following estimates.

Let

$$
\mathcal{B}(x, v, \beta, \varepsilon) = \left\{ \mathrm{Box}\,(y, tv, t\beta) \,\middle|\, \left| \frac{d(y, yh_t e^v)}{t} - \underline{\rho}(x, v) \right| < \varepsilon, y \in B_{cc}(x, t) \right\}.
$$

In other words, members of $\mathcal{B}$ are boxes of some variable height $t\beta$ such that the difference quotients of the endpoints of the central segment are infimal up to an error of ε. Notice that $\mathcal{B}(x, v, \beta, \varepsilon)$ is a collection of boxes indexed by t. Next, we make an estimate on the difference quotient between points on the ends.

Lemma 4.1. *Let* u_1, u_2 *be points on opposite ends of* $\mathrm{Box}\,(y, tv, t\beta) \in \mathcal{B}(x, v, \beta, \varepsilon)$. *Then,*

$$
\frac{d(u_1, u_2)}{t} \leq C_1(\beta) + \underline{\rho}(x, v) + \varepsilon
$$

where $C_1(\beta)$ is a constant depending on β which goes to zero with β.

Proof. This is just a computation:

$$
\begin{aligned}
&\frac{d(u_1, u_2)}{t} \\
&\leq \frac{d(u_1, y) + d(y, yh_t e^v) + d(yh_t e^v, u_2)}{t} \quad \text{(by the triangle inequality)} \\
&\leq \frac{Ld_{cc}(u_1, y) + d(y, yh_t e^v) + Ld_{cc}(yh_t e^v, u_2)}{t} \quad \text{(by L-Lipschitz)} \\
&\leq LA\beta + \underline{\rho}(x, v) + \varepsilon + L(C(\beta)) \quad \text{(by choice of } \mathcal{B} \text{ and equation } (*)).
\end{aligned}
$$

Let $C_1(\beta) = L(A\beta + C(\beta))$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

This end estimate provides the needed precision to make the later estimates. The "well controlled smaller boxes" mentioned above will be members of $\mathcal{B}$. Other than the end estimates, the other technical point we need is a measure comparison. We wish to use the estimate boxes to prove a relation between ρ and $\overline{\rho}$ almost everywhere with respect to $\mathcal{H}^Q$. Thus, we need to know the relative measure of the boxes in the CC balls. In light of the fact that the Hausdorff and Riemannian measures are constant multiples of one another (see Section 3), this is fairly straightforward. Recalling the general ball-box theorems for Carnot-Carathéodory spaces, it is useful to view the following lemma as a directional refinement of such ball-box theorems.

Lemma 4.2. *Fix $\beta > 0$ and $v \in \mathcal{V}$. There is a number, $R(\beta, v)$, such that $\mathrm{Box}\,(y, tv, t\beta) \subset B_{cc}(y, tR(\beta, v))$. Then, there exists a constant C so that*

$$
\frac{\mathcal{H}^Q(\mathrm{Box}\,(y, tv, t\beta))}{\mathcal{H}^Q(B_{cc}(y, t(R(\beta, v))))} = C
$$

where C depends only on d_{cc}, β, v and the constants relating the Hausdorff, Lebesgue and Riemannian measures.

Proof. The constant $R(\beta, v)$ exists and can be estimated by equation $(*)$. It is independent of y because d_{cc} is left invariant. Since Q is the Hausdorff dimension of (N, d_{cc}) and $\mathcal{H}^Q(h_t A) = t^Q \mathcal{H}^Q(A)$, $\mathcal{H}^Q(B_{cc}(y, tR(\beta, v))) = \alpha_Q (tR(\beta, v))^Q$ where $\alpha_Q = \mathcal{H}^Q(B_{cc}(e^0, 1))$. Thus we only need estimate $\mathcal{H}^Q(Box(y, tv, t\beta))$.

Let $\langle \cdot, \cdot \rangle_R$ be a left invariant Riemannian metric on N making the grading orthogonal and matching the sub-Riemannian metric's inner product on $\mathcal{V}$. Let $dvol$ be the volume form on N associated to $\langle \cdot, \cdot \rangle_R$. Let $\mathcal{L}^n$ be the pull back of the Lebesgue measure on the Lie algebra of N using the exponential mapping. By the discussion in section 3, there exist constants C_0 and C_1 so that,

$$\mathcal{H}^Q = C_0 dvol, \qquad dvol = C_1 \mathcal{L}^n$$

Now, the construction of the boxes shows that the preimage of Box $(y, tv, t\beta)$ in the Lie algebra (considered at the point y) is end $(t\beta, tv) \times [0, td_{cc}(e^0, e^v)]$. Thus, since

$$\mathcal{L}^n(\text{end } (t\beta, tv) \times [0, td_{cc}(e^0, e^v)]) = (2t\beta)^{Q-1} td_{cc}(e^0, e^v)$$

we have

$$\mathcal{H}^Q(\text{Box } (y, tv, t\beta)) = C_0 C_1 t^Q (2\beta)^{Q-1} d_{cc}(e^0, e^v).$$

Letting

$$C = \frac{C_0 C_1 (2\beta)^{Q-1} d_{cc}(e^0, e^v)}{\alpha_Q R(\beta, v)^Q}$$

we have the desired result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 1. Lemma 4.2 and 3.2 show that we may refine covers of the boxes defined above in the same way that we refine covers of closed balls. In other words, if A is a measurable set in N and $\mathcal{B}$ is a cover of A by boxes which is fine at every point of A, then we may find a countable disjoint subcover $\mathcal{B}'$ such that $\mathcal{H}^Q(A \setminus \cup_{B \in \mathcal{B}'} B) = 0$. This follows from standard covering arguments (i.e. a countable number of iterations of the "greedy" algorithm used to cover at least a fixed portion of A in one iteration). Following Federer, we say that such a collection of boxes is a *Vitali covering relation*.

4.2. Differentiability of distances Lipschitz to d_{cc}.

The following theorem is the main technical piece of the paper. As it deals with the differentiability of Lipschitz maps, it can be viewed as a measure-theoretic version of the Rademacher theorem.

Theorem 4.3. *Let $(N, \mathcal{V}, \langle \cdot, \cdot \rangle, d_{cc})$ be a sub-Riemannian group of nilpotency degree $k + 1$. Suppose d is a distance function on $U \times U$, U an open subset of N, which is L-Lipschitz to d_{cc}. Then, for $e^v, e^w \in N$,*

$$\Delta_x(e^v, e^w) = \overline{\lim}_{t \to 0^+} \left\{ \frac{d(yh_t e^v, yh_t e^w)}{t} \middle| y \in B_{cc}(x, t) \right\}$$

defines a family of real-valued measurable functions on $N \times N$, indexed by $x \in U$, enjoying the following properties. There is a subset $\mathcal{U}$ of U of full measure such that for every $x \in \mathcal{U}$,

- *For $v \in \mathcal{V}$,*

$$\lim_{t \to 0^+} \left\{ \left| \frac{d(y, y h_t e^v)}{t} - \Delta_x(e^0, e^v) \right| \middle| y \in B_{cc}(x, t) \right\} = 0$$

- *For $v \in \mathcal{V}$, $\Delta_x(e^0, h_t e^v) = t \Delta_x(e^0, e^v)$.*

- *For $v \in \mathcal{V}$, $\Delta_x(e^0, e^v) = \Delta_x(e^v, e^0) = \Delta_x(e^0, e^{-v})$.*

- *For $v, w, \in \mathfrak{n}$ such that $-v \odot w \in \mathcal{V}$, $\Delta_x(e^v, e^w) = \Delta_x(e^0, e^{-v} e^w)$.*

Remark 2. Note that since d is L-Lipschitz to d_{cc} and h_t are homotheties of d_{cc}, it is immediate that Δ_x is L-Lipschitz to d_{cc}.

We will also see that if, for some $x \in \mathcal{U}$, $(C_x N, \overline{d})$ exists, we can interpret this proposition as a statement about $\overline{d}$. One should view the function Δ_x as being something like an infinitesimal sub-Finsler metric and think of the last three claims of the theorem as giving a limited homothety, symmetry and left invariance for the function Δ_x. It would be interesting to find conditions under which this construction actually converged to a sub-Finsler metric.

Remark 3. This theorem leaves open the question of complete metric differentiabilty in this setting and other settings.

We prove Theorem 4.3 through several lemmas.

Lemma 4.4. *If $v \in \mathcal{V}$ and x an approximate continuity point (see proof for the definition) of $\underline{\rho}(\cdot, v)$ and $\tau \in \mathbb{R}$, then $\overline{\rho}(x, \tau v) \leq |\tau| \underline{\rho}(x, v)$. In particular, for such points x, $\underline{\rho}(x, v) = \overline{\rho}(x, v)$.*

Proof. Fix $\varepsilon_0 > 0$. We aim to show that for small enough $t > 0$ and $y \in B_{cc}(x, t)$,

$$(+) \qquad \frac{d(y, y h_t h_\tau e^v)}{t} \leq |\tau| \underline{\rho}(x, v) + \varepsilon_0.$$

For simplicity, we will first consider only $\tau > 0$. The nonpositive case is an easy consequence of the positive case and will be discussed at the end

of the proof. To prove the positive case, we will consider special boxes which will easily estimate the central segment of $\mathrm{Box}\,(y, t\tau v, t\beta)$ (for some appropriate β) and still have their endpoint quotients close to the infimum $\underline{\rho}$. First, pick $\varepsilon, \beta \in (0, \infty)$. Since x is an approximate continuity point of $\underline{\rho}(\cdot, v)$, the set $Z = \{z \mid |\underline{\rho}(z, v) - \underline{\rho}(x, v)| < \varepsilon\}$ has density in $B_{cc}(x, r)$ approaching 1 as r approaches 0. The same is true if we replace $B_{cc}(x, r)$ with boxes $\mathrm{Box}\,(y, v, \tau\beta) \subset B_{cc}(x, r)$. This follows from lemma 4.2, the fact that the boxes yield an alternate base for the topology of N, and the scaling properties of both balls and boxes with respect to the homothety.

For the rest of the proof, fix t_0 such that for $t \le t_0$, $y \in B_{cc}(x, t)$,

$$\frac{\mathcal{H}^Q(Z \cap \mathrm{Box}\,(y, t\tau v, t\beta))}{\mathcal{H}^Q(\mathrm{Box}\,(y, t\tau v, t\beta))} \ge 1 - \varepsilon.$$

To get estimate $(+)$, we fix a box $B_0 = Box(y, t\tau v, t\beta)$ with $t \le t_0$ and $y \in B_{cc}(x, t)$. Let $Z_0 = Z \cap B_0$. Next, let $\mathcal{B}_Z = \{B \in \mathcal{B}(z, v, \beta, \varepsilon) \mid B \subset B_0, z \in Z_0\}$. Recalling the definition of ρ, we see that for $z_0 \in Z_0$ and if t is sufficiently small (i.e. t such that $\frac{d(y, y h_t e^v)}{t}$ must be close to infimal and $y \in B_{cc}(z_0, t)$), then there are infinitely many boxes in $\mathcal{B}(z_0, v, \beta, \varepsilon)$ with height $t\beta$. Using Lemma 4.2, we see that $\mathcal{B}_Z$ is a Vitali covering relation (see the remark after 4.2) of Z_0. Thus, we may select from the cover a disjoint finite subcover such that $\mathcal{H}^Q(\cup_{B \in \mathcal{B}_Z} B) > \mathcal{H}^Q(B_0) - 2\varepsilon\mathcal{H}^Q(B_0)$ which, by abuse of notation, we denote this by $\mathcal{B}_Z$ as well. Consider the picture at this point, we have fixed a particular box and covered it with smaller boxes, all contained in the fixed box, such that the end estimates on the small boxes are approximately infimal. Next, we will extract a particular curve from the fixed box and approximate the difference quotient of the endpoints of its central segment using the cover.

Using the exponential map as used in lemma 4.2, we realize B_0 as the diffeomorphic image of $R \times [0, t\tau] \subset \mathbb{R}^{n-1} \times \mathbb{R}$. We can calculate $\mathcal{H}^Q(B_0)$, up to constant multiple, by calculating the Lebesgue measure of $R \times [0, t\tau]$. Using this diffeomorphic identification, we denote by $\mathcal{R}_Z$ the image of the collection $\mathcal{B}_Z$ in $R \times [0, t\tau]$. The fact that $\frac{\mathcal{H}^Q(\cup_{B \in \mathcal{B}_Z} B \cap B_0)}{\mathcal{H}^Q(B_0)} \ge 1 - 2\varepsilon$ implies that $\frac{\mathcal{L}^n(\cup_{S \in \mathcal{R}_Z} S \cap (R \times [0, t\tau]))}{\mathcal{L}^n(R \times [0, t\tau])} \ge 1 - 2\widetilde{C}_0\varepsilon$ where $\widetilde{C}_0$ is a constant depending on the structure constants of the Lie group and the constant relating $\mathcal{H}^Q$ to $\mathcal{L}^n$. Thus, using Fubini's theorem on $\mathbb{R}^{n-1} \times \mathbb{R}$, we can conclude that there exists a point $r_0 \in R$ such that $\frac{\mathcal{L}^1(\cup_{S \in \mathcal{R}_Z} S \cap r_0 \times [0, t\tau])}{\mathcal{L}^1(r_0 \times [0, t\tau])} \ge 1 - 2\widetilde{C}_0\varepsilon$. Let $\gamma(t) = y' h_t h_\tau e^v$

be the fiber which is diffeomorphically identified with $r_0 \times [0, t\tau]$. Then,

$$\frac{\mathcal{H}^1(\cup_{B \in \mathcal{B}_Z} B \cap \gamma)}{\mathcal{H}^1(\gamma)} \geq 1 - 2C_0\varepsilon$$

where C_0 is a (potentially) different constant depending on the same data. Now, we will do end estimates on the smaller boxes and sum them to get an estimate for $\frac{d(y', y'h_t h_\tau e^v)}{t}$. Let e_i, $i = 1, 2, 3, \ldots, n$ be the finite list of points along γ, listed in order following the parameter, such that each e_{2i-1} lies in the end of one of the small boxes and e_{2i} lies in the other. For e_i, e_{i+1} a pair which are in the ends of a single box, Box $(z_i, t_i v, t_i \beta)$, the construction yields the following estimate. By Lemma 4.1, we have:

$$(\dagger) \qquad\qquad \frac{d(e_i, e_{i+1})}{t_i} \leq C_1(\beta) + \underline{\rho}(x, v) + \varepsilon.$$

Now using the triangle inequality for d, the fact that the smaller boxes cover most of B_0 shows that $d(y', y'h_t h_\tau e^v) \leq \sum_{i=1}^{\frac{n}{2}} d(e_{2i-1}, e_{2i}) + error$. Let $A = \gamma \setminus \cup_{B \in \mathcal{B}_Z} B$. The error is the sum of the distances between the endpoints of the adjacent boxes. This *error* is less that the one dimensional d-Hausdorff measure of A, $\mathcal{H}_d^1(A)$ because for any connected set C, $\mathcal{H}_d^1(C) \geq \operatorname{diam}_d(C)$ (see [Fed69] 2.10.12). Using the Lipschitz property of d, $\mathcal{H}_d^1(A) \leq L\mathcal{H}^1(A)$. Next, we estimate $\mathcal{H}^1(A)$. Since $\frac{\mathcal{H}^1(\cup_{B \in \mathcal{B}_Z} B \cap \gamma)}{\mathcal{H}^1(\gamma)} \geq 1 - 2C_0\varepsilon$, we have that $\frac{\mathcal{H}^1(A)}{\mathcal{H}^1(\gamma)} = 1 - \frac{\mathcal{H}^1(\cup_{B \in \mathcal{B}_Z} B \cap \gamma)}{\mathcal{H}^1(\gamma)} \leq 1 - (1 - 2C_1\varepsilon)$ and using the fact that $\mathcal{H}_d^1(A) \leq L\mathcal{H}^1(A)$ and the Lipschitz property of d, $error \leq \mathcal{H}_d^1(A) \leq L\mathcal{H}^1(A) \leq 2C_1\varepsilon L\mathcal{H}^1(\gamma) \leq 2C_1\varepsilon\alpha_1 t\tau L d_{cc}(e^0, e^v)$. Letting $C_2(\varepsilon) = 2C_1\varepsilon\tau L d_{cc}(e^0, e^v)$ we have:

$$d(y', y'h_t h_\tau e^v)$$

$$\leq \left(\sum_{i=1}^{\frac{n}{2}} d(e_{2i-1}, e_{2i})\right) + tC_2(\varepsilon)$$

$$\leq \left(\sum_{i=1}^{\frac{n}{2}} t_i\right)(C_1(\beta) + \underline{\rho}(x, v) + \varepsilon) + tC_2(\varepsilon) \text{ (by inequality ($\dagger$))}.$$

Next, we note that $\frac{\sum_{i=1}^{\frac{n}{2}} t_i}{t} \leq \tau$. To see this, note that since the curve $\gamma(t)$ is rectifiable with respect to $d_{cc}{}^1$, we can pick time indices s_i and s_{i+1} such

---

[1]This use of rectifiability is what prevents the same argument from proving differentiability in nondistributional directions.
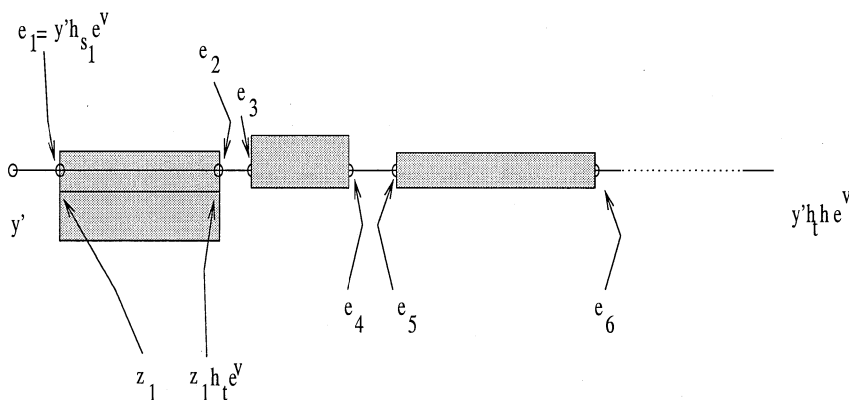
Figure 1: A schematic of the covering of the line segment by small boxes.

that

$$\sum_{i=1}^{\frac{n}{2}} t_i = \sum_{i=1}^{\frac{n}{2}} \frac{d_{cc}(z_i, z_i h_{t_i} e^v)}{d_{cc}(e^0, e^v)} = \sum_{i=1}^{\frac{n}{2}} \frac{d_{cc}(y' h_{s_i} e^v, y' h_{s_{i+1}} e^v)}{d_{cc}(e^0, e^v)}$$

By left invariance, this equals $\sum_{i=1}^{\frac{n}{2}} |s_{i+1} - s_i| \leq t\tau$. (See Figure 1.)

We use the estimate and this fact to investigate the desired quotient.

$$\frac{d(y', y' h_t h_\tau e^v)}{t} \leq |\tau|(C_1(\beta) + \underline{\rho}(x, v) + \varepsilon) + C_2(\varepsilon).$$

Using the Lipschitz property as in the proof of lemma 4.1, we see:

$$\frac{d(y, y h_t h_\tau e^v)}{t} \leq |\tau|(C_1(\beta) + \underline{\rho}(x, v) + \varepsilon) + C_2(\varepsilon) + C_1(\beta).$$

Picking β small enough this yields,

$$\frac{d(y, y h_t h_\tau e^v)}{t} \leq |\tau|\underline{\rho}(x, v) + \frac{\varepsilon_0}{2} + C_3(\varepsilon)$$

where $C_3(\varepsilon)$ depends only on ε. Thus, picking ε so that $C_3(\varepsilon) < \frac{\varepsilon_0}{2}$ allows us to conclude that $\frac{d(y, y h_t h_\tau e^v)}{t} \leq |\tau|\underline{\rho}(x, v) + \varepsilon_0$ Taking the lim sup and letting ε_0 tend to zero, we have the desired result for $\tau > 0$.

To see that the same result holds true for negative τ, we make the following observations. First, $h_{-\tau}e^v = h_\tau e^{-v}$ since $v \in \mathcal{V}$. Second, for small enough t, the density of Z in $\text{Box}(y, -t\tau v, t\alpha)$ is still greater than $1 - \varepsilon$. Therefore, by the symmetry of the distance d, the same argument goes through for e^{-v}. Since the case of $\tau = 0$ is trivial, the lemma is valid for all $\tau \in \mathbb{R}$. $\qquad \square$

Lemma 4.5. *Retaining the assumptions of Theorem 4.3, define $\rho(x, \cdot) = \overline{\rho}(x, \cdot)$. Then, there exists a $\mathcal{U} \subset U$ of full measure such that for $x \in \mathcal{U}$, $\rho(x, \tau v) = |\tau|\rho(x, v)$ for all $v \in \mathcal{V}$. We will call $\mathcal{U}$ the set of **metric regular** points of d.*

Proof. Let $n_0 = \dim \mathcal{V}$. Let Q be a countable subset of $\mathcal{V}$ so that e^Q is dense in $e^{\mathcal{V}}$. Since $\underline{\rho}$ is measurable function, it follows that for fixed $v \in \mathcal{V}$, the set

$$U_v = \{x \mid x \text{ is an approximate continuity point of } \underline{\rho}(\cdot, v)\}$$

has full measure in U. Therefore, $\mathcal{U} = \{x \mid x \in U_v \text{ for all } v \in Q\} = \cap_{v \in Q} U_v$ is a Borel set of full measure in U. For a proof of this fact see, for example, Theorem 2.9.13 in [Fed69]. On this set, we have the desired properties. Using Lemma 4.4, we see that for $v \in Q$ and $\tau \in \mathbb{Q}$,

$$\underline{\rho}(x, \tau v) \le \overline{\rho}(x, \tau v) \le |\tau|\underline{\rho}(x, v)$$
$$\underline{\rho}\left(x, \frac{1}{\tau}v\right) \le \overline{\rho}(x, \frac{1}{\tau}v) \le \left|\frac{1}{\tau}\right|\underline{\rho}(x, v)$$

therefore,

$$\underline{\rho}(x, \tau v) = |\tau|\underline{\rho}(x, v).$$

Given the monotonicity in τ of $\underline{\rho}$ and $\overline{\rho}$, this holds for all τ in $\mathbb{R}$ and $v \in Q$. Note in particular that $\underline{\rho}(x, -v) = \underline{\rho}(x, v)$ and therefore we have symmetry of this type. To extend these properties to all of $\mathcal{V}$ we check that $\rho(x, \cdot)$ is continuous. Fix $v \in \mathcal{V}$ and consider a sequence $v_i \in Q$ which

converges to v. Then,

$$\rho(x, v_i) - \rho(x, v)$$

$$= \overline{\lim}_{t \to 0^+} \left\{ \frac{d(x', x' h_t e^{v_i})}{t} \middle| x' \in B_{cc}(x, t) \right\}$$

$$- \overline{\lim}_{t \to 0^+} \left\{ \frac{d(x'', x'' h_t e^v)}{t} \middle| x'' \in B_{cc}(x, t) \right\}$$

$$\leq \overline{\lim}_{t \to 0^+} \left\{ \frac{d(x', x' h_t e^{v_i}) - d(x', x' h_t e^v)}{t} \middle| x' \in B_{cc}(x, t) \right\}$$

$$\leq \overline{\lim}_{t \to 0^+} \left\{ \frac{d(x' h_t e^{v_i}, x' h_t e^v)}{t} \middle| x' \in B_{cc}(x, t) \right\}$$

$$\leq L \overline{\lim}_{t \to 0^+} d_{cc}(e^{v_i}, e^v)$$

which tends to zero as $i \to \infty$. A similar argument shows that $\lim_{i \to \infty} \rho(x, v_i) - \rho(x, v) \geq 0$ as well. □

Lemma 4.6. *Suppose $x_0 \in N$ is a metric regular point (see Lemma 4.5). Then, if $v, w \in \mathfrak{n}$ such that $-v \otimes w \in \mathcal{V}$ then*

$$\overline{\lim}_{t \to 0^+} \left\{ \frac{d(y h_t e^v, y h_t e^w)}{t} \middle| y \in B_{cc}(x_0, t) \right\} = \rho(x_0, -v \otimes w).$$

Proof. This fact follows from the covering argument in Lemma 4.4. Fix $\varepsilon_0 > 0$. Using the same notation, picking β sufficiently small, the density of Z_0 in $Box(x_0 h_t e^v, t(-v \otimes w), \beta)$ is bigger than $1 - \varepsilon$. Now, we can run the same covering argument to produce the estimate $\frac{d(x_0 h_t e^v, x_0 h_t e^w)}{t} \leq \underline{\rho}(x_0, -v \otimes w) + \varepsilon_0.$ □

Proof of Theorem 4.3. Let

$$\Delta_x(e^v, e^w) = \overline{\lim}_{t \to 0^+} \left\{ \frac{d(y h_t e^v, y h_t e^w)}{t} \middle| y \in B_{cc}(x, t) \right\}.$$

This clearly satisfies the first part of the claim with $\mathcal{U}$ being the set of metric regular points of d. We now prove the other items in reverse order. By lemma 4.6, we see that $\Delta_x(e^v, e^w) = \Delta_x(e^0, e^{-v \otimes w})$ for $-v \otimes w \in \mathcal{V}$ proving the fourth part of the theorem. Lemma 4.5 implies that $\Delta_x(e^0, e^v) = \Delta_x(e^0, e^{-v})$ and the symmetry of the distance function d immediately shows that $\Delta_x(e^0, e^v) = \Delta_x(e^v, e^0)$, proving the third part of the claim. Lemma 4.5 also shows that for $v \in \mathcal{V}$, $\Delta_x(e^0, h_t e^v) = t \Delta(e^0, e^v)$ proving the second part of the theorem. □

5. Tangent cones.

The differentiability result of the last section allows us to compare the local structures of a sub-Riemannian group and some metric space to which it is biLipschitz equivalent. Presuming the existence of tangent cones to image points of F in X, we would like construct a "derivative" mapping between the tangent cone at a metric regular point in N to the tangent cone at its F image in X. To do this, we will first review the tangent cone construction and apply it to this setting, interpreting the results of theorem 4.3. The goal of this section is to prove the following:

Proposition 5.1. *Let $(N, \mathcal{V}, \langle \cdot, \cdot \rangle, d_{cc})$ be a sub-Riemannian group, (X, d_X) be a metric space and $F : N \to X$ be an L-biLipschitz map between them. Letting $d = F^* d_X$ we let Δ_X be the function constructed in Theorem 4.3. Suppose there exists an $x \in X$ in the image of the metric regular points of F such that a tangent cone, $(C_x X, \overline{d}_X)$, exists. Then,*

- *There exists an L-biLipschitz map $\overline{F} : (C_{F^{-1}(x)} N, \overline{d}_{cc}) \to (C_x X, \overline{d}_X)$ such that for $n^{-1} m \in C_{F^{-1}(x)} N$ that is in the exponential image of $\mathcal{V}$, $\overline{d}_X(\overline{F}(n), \overline{F}(m)) = \Delta_{F^{-1}(x)}(n, m)$.*

- *$\overline{F}$ takes radial geodesics to geodesic lines.*

We begin by considering the structure of the tangent cone to (N, d_{cc}) at some point. By proposition 3.1, we know that the tangent cone at any point is unique and isometric to (N, d_{cc}). We will consider the tangent cone at e^0 for simplicity. Given a complete metric space (X, d), one usually constructs the tangent cone at x as the Gromov-Hausdorff limit of the sequence of dilated spaces $(X, x, \lambda_i d)$ for some sequence $\{\lambda_i\}$ tending towards infinity as $i \to \infty$. For (N, d_{cc}), one may construct the dilated spaces using the homothety. Thus one considers the convergence of $(N, h^*_{\lambda_i} d_{cc})$. Now, since (N, d_{cc}) is isometric to $(N, h^*_{\lambda_i} d_{cc})$ via the isometry h_{λ_i}, the sequence is, in fact, the constant sequence and thus converges in the Gromov-Hausdorff topology for any choice of sequence $\{\lambda_i\}$. (See [Gro81] for an explanation for the convergence of pointed spaces. The reader may also want to consult [Gro91] which provides a somewhat shorter explanation for compact spaces).

Before continuing, we reiterate the assumptions about (X, d_X): first, there exists $x \in Im(F) \subset X$ and the tangent cone to X at x exists, i.e. there exists a sequence $\{\lambda_i\}$ such that $(C_x X, \overline{d}_X) = \text{GH-lim}_{i \to \infty}(X, x, \lambda_i d_X)$ (the Gromov-Hausdorff limit of $(X, x, \lambda_i d_X)$). To construct the map $\overline{F}$, we first need to review the concept and usage of ultrafilters and ultralimits.

Definition 10. A nonprincipal ultrafilter is a finitely additive probability measure ω on subsets of $\mathbb{N}$ such that

- ω is zero on any finite subset of $\mathbb{N}$.

- For every subset $S \subset \mathbb{N}$, $\omega(S)$ is either zero or one.

It is a consequence of the axiom of choice that nonprincipal ultrafilters exist (see [Dug66] Chapter 10, Theorem 7.3). We can use an ultrafilter to choose a distinguished "convergent" point for a sequence. As an illustrative example, let $\mathbb{R}$ be the line with its usual metric and let $\{x_i\}$ be a bounded sequence in $\mathbb{R}$. Think of $\{x_i\}$ as the image of a map $s : \mathbb{N} \to \mathbb{R}$. Then there is a unique element $x_\omega \in X$ ($x_\omega = \omega\text{-lim } s$) such that for any neighborhood U of x_ω, $\omega(s^{-1}(U)) = 1$. In this way, the nonprincipal ultrafilter picks a limit point of the sequence. Next we point out a useful fact about ultralimits of sequences in $\mathbb{R}$. If $\{a_i\}$ and $\{b_i\}$ are bounded sequences and $a_i \leq b_i$ then $a_\omega = \omega\text{-lim } a_i \leq b_\omega = \omega\text{-lim } b_i$. Indeed, if $b_\omega < a_\omega$ then we can find open neighborhoods U_a and U_b of a_ω and b_ω which separate the ultralimits. Since $\omega(s^{-1}(U_a)) = 1 = \omega(s^{-1}(U_b))$, if the inverse images are disjoint, this violates the additivity of the ultrafilter while if they intersect they must intersect in an infinite set by the additivity of the ultrafilter. Thus, for indices in this intersection, $\lim b_i < \lim a_i$, contradicting the assumption that $a_i \leq b_i$ for all i.

Let ω be a nonprincipal ultrafilter and define a limit metric space $(C_x^\omega X, x, d_X^\omega)$ as follows. Let $\overline{X} = \{\{x_i\} | x_i \in (X, x, \lambda_i d_X), \lambda_i d_X(x_i, x) < \infty\}$. Then $C_X^\omega X = \overline{X}/\sim$ where $\{x_i\} \sim \{y_i\}$ if $\omega\text{-lim}_{i \to \infty} \lambda_i d_X(x_i, y_i) = 0$ (note that here $d_X(x_i, y_i)$ is a bounded map from $\mathbb{N}$ to $\mathbb{R}$) and $d_X^\omega(\{x_i\}, \{y_i\}) = \omega\text{-lim}_{i \to \infty} \lambda_i d_X(x_i, y_i)$. Note that the choice of ultrafilter forces convergence of the dilated spaces and allows us to group the many choices we otherwise have to make into the choice of the ultrafilter. Also note that our assumption that the Gromov-Hausdorff limit of $(X, x, \lambda_i d_X)$ exists allows us to find an isometry between $C_x X$ and $C_x^\omega X$. For a proof of this fact, as well as a more general discussion of ultrafilters and ultralimits, see [KL96] Section 2.4.2 and, in particular, Lemma 2.4.3.

Using left invariance, we may assume without loss of generality that $F^{-1}(x) = e^0$. We define $\overline{F} : (C_{e^0} N, e^0, \overline{d}_{cc}) \to (C_x^\omega X, x, d_X^\omega)$ by $\overline{F}(n) = \{F(h_{\lambda_i^{-1}} n)\}$. Note that $\overline{F}$ is L-biLipschitz if F is:

By definition,

$$d_X^\omega(\overline{F}(n), \overline{F}(m)) = \omega\text{-lim}_{i \to \infty} \lambda_i d_X(F(h_{\lambda_i^{-1}} n), F(h_{\lambda_i^{-1}} m)).$$

Using the biLipschitz property,

$$\omega\text{-}\lim_{i\to\infty}\frac{1}{L}d_{cc}(n,m) \leq \omega\text{-}\lim_{i\to\infty}\lambda_i d_X(F(h_{\lambda_i^{-1}}n), F(h_{\lambda_i^{-1}}m))$$

$$\leq \omega\text{-}\lim_{i\to\infty}Ld_{cc}(n,m).$$

Upon taking ultralimits, we have:

$$\frac{1}{L}\overline{d}_{cc}(n,m) \leq d_X^\omega(\overline{F}(n),\overline{F}(m)) \leq L\overline{d}_{cc}(n,m).$$

As mentioned above, Lemma 2.4.3 in [KL96] shows that $(C_x X, \overline{d}_X)$ is isometric to (C_x^ω, d_X^ω) and therefore we may view $\overline{F}$ as an L-biLipschitz map between $(C_{e^0}N, \overline{d}_{cc})$ and $(C_x X, \overline{d}_X)$. In addition, by construction, $\overline{d}_X(\overline{F}(n), \overline{F}(m)) = \omega\text{-}\lim_{i\to\infty}\lambda_i d_X(F(h_{\lambda_i^{-1}}n), F(h_{\lambda_i^{-1}}m)) = \lim_{i\to\infty}\frac{d(h_{\lambda_i^{-1}}n, h_{\lambda_i^{-1}}m)}{\lambda_i}$ which by Theorem 4.3 applied to F^*d_X converges to $\Delta_x(n,m)$ when $n^{-1}m$ is in the exponential image of $\mathcal{V}$. Thus we have proved the first part of Proposition 5.1. For the second part, consider a radial geodesic, $h_t e^v$ in N where $v \in \mathcal{V}$. Using the definition of $\overline{F}$, let $\gamma(t) = \overline{F}(h_t e^v)$. We claim $\gamma(t)$ minimizes between any two time indices $t_1 < t_2$. Assume for now that $t_1 > 0$. Indeed,

$$\begin{aligned}
\overline{d}_X(\gamma(t_1), \gamma(t_2)) &= \overline{d}_X^\omega(\overline{F}(h_{t_1}e^v), \overline{F}(h_{t_2}e^v))\\
&= \Delta_x(h_{t_1}e^v, h_{t_2}e^v) = \Delta_x(e^0, (h_{t_1}e^v)^{-1}(h_{t_2}e^v))\\
&= \Delta_x(e^0, h_{t_1}e^{-v}h_{t_2}e^v) = \Delta_x(e^0, h_{t_2-t_1}e^v)\\
&= |t_2 - t_1|\Delta_x(e^0, e^v).
\end{aligned}$$

Here we used the left invariance of Δ_x and Lemma 2.2. For $t_1 < 0$, we simply recall our convention regarding $h_t n$ for negative values of t. When $t_1, t_2 \geq 0$ we observe that

$$\begin{aligned}
\overline{d}_X(\gamma(-t_1), \gamma(t_2)) &= \overline{d}_X^\omega(\overline{F}(h_{-t_1}e^v), \overline{F}(h_{t_2}e^v))\\
&= \Delta_x(h_{t_1}e^{-v}, h_{t_2}e^v) = \Delta_x(e^0, (h_{t_1}e^{-v})^{-1}(h_{t_2}e^v))\\
&= \Delta_x(e^0, h_{t_1}e^v h_{t_2}e^v) = \Delta_x(e^0, h_{t_2+t_1}e^v)\\
&= |t_2 + t_1|\Delta_x(e^0, e^v).
\end{aligned}$$

Thus, $\gamma(t)$ minimizes between any two time indices. In addition, by left invariance, all radial geodesics are taken to geodesics under the map $\overline{F}$.

The reader may wonder why such trouble is taken to use an ultrafilter construction to produce convergence. In fact, if X is a separable and locally

compact metric space, then one can use more standard (e.g., Arzela-Ascoli type) arguments to produce uniform convergence on compact sets. However, since asymptotic cones are not always locally compact, we use the ultrafilter construction to resolve convergence issues.

6. Generalized notions of curvature.

6.1. Curvature bounded above.

Again, we follow the notation and definitions of [KL96]. To determine curvature bounds of metric spaces, we will need to compare to model spaces. For $\kappa \in \mathbb{R}$, let (M_κ^2, d_κ) be the two dimensional constant curvature κ model space (i.e., $\mathbb{R}^2, S_\kappa^2$, or $\mathbb{H}_\kappa^2$). Denote by diam (κ) the diameter of the model space M_κ^2. Let (X, d) be a complete metric space. Given a triangle, $\triangle$, in X with minimizing geodesic sides s_1, s_2, s_3, we call the triangle $\triangle_\kappa \subset M_\kappa^2$ with the same side lengths as $\triangle$ a comparison triangle for $\triangle$.

Definition 11. (X, d) is a CAT_κ space if

- For every pair of points $x, y \in X$ with $d(x, y) < $ diam (κ), there is a geodesic segment joining x to y.

- Let $\triangle$ be geodesic triangle in X with sides s_1, s_2, s_3 of lengths l_1, l_2, l_3 such that $l_1 + l_2 + l_3 < 2$diam (κ). For any two points $x, y \in \triangle$, $d(x, y) \leq d_\kappa(x_\kappa, y_\kappa)$ where x_κ and y_κ are the corresponding points on $\triangle_\kappa$.

Another way of saying the second condition in the definition is that the triangles in X are thinner than their counterparts in M_κ^2. As mentioned in the introduction, we sometimes will consider metric spaces which have this triangle property locally. Hence, we say a metric space, X, is a CBA_κ metric space if for each point $x \in X$, there exists an r such that the closure of the ball of radius r about x is a CAT_κ metric space in its own right. In the next several lemmas, we outline the properties of CAT_κ spaces that will be of use below. A nicely self-contained discussion of these facts can be found in [Bal95].

Theorem 6.1 (Hadamard-Cartan Theorem for CAT_0 spaces). *Let* X *be a simply connected complete CAT_0 metric space. Then, for any two points $x, y \in X$, there exists a unique geodesic connecting x to y.*

Lemma 6.2. *Let (X, d) be a CAT_0 space and suppose $\gamma_1(t)$ and $\gamma_2(t)$ are unit speed geodesic rays in X. Then the function $d(\gamma_1(t), \gamma_2(t))$ either grows linearly in t or is bounded.*

Proof. The key fact is that in a CAT_0 space, the distance function is convex (for a proof, see [Bal95] Proposition 5.4). If the function $f(t) = d(\gamma_1(t), \gamma_2(t))$ is linear or bounded, we are done. Assuming that the function is not linear but is unbounded, we can find an increasing sequence $\{t_i\}$ such that $f''(t_i) > 0$ for all i. One can now easily show that $f(t)$ must grow at least linearly in this case. But, $f(t)$ is bounded above by $2t + d(\gamma_1(0), \gamma_2(0))$ simply by measuring along the two geodesics. Thus, $f(t)$ grows linearly or is bounded. $\square$

For the sake of completeness, we quote some results concerning the two coning operations we use in this argument.

Lemma 6.3. *Let (X, d) be a CAT_κ space.*

1. *For all $x \in X$, $C_x X$ exists and is a CAT_0 space.*

2. *If $\kappa = 0$ and ω is a nonprincipal ultrafilter, then $C_\infty^\omega X$ is a CAT_0 space.*

Proof. For 1) This is proved in [Nik95], 2) see [KL96] or [Bal95].

Also note that the lemma implies that the tangent cone to a CBA_κ space is also a CAT_0 space.

6.2. Curvature bounded below.

Curvature bounded below for metric spaces is defined in much the same way as for CAT_κ spaces. A standard resource for this material is [BGP92]. The definition and first uses of these spaces were due to A. D. Alexandrov and are often referred to as "Alexandrov spaces". Using the notation of the previous section, we make the following definitions.

Definition 12. Let (X, d) be a complete metric space. Then, (X, d) is a CBB_κ space if

- (X, d) is a locally compact geodesic space.

- Let $\triangle$ be geodesic triangle in X with sides s_1, s_2, s_3 of lengths l_1, l_2, l_3 such that $l_1 + l_2 + l_3 < 2\text{diam}\,(\kappa)$. For any two points $x, y \in \triangle$, $d(x, y) \geq d_\kappa(x_\kappa, y_\kappa)$ where x_κ and y_κ are the corresponding points on $\triangle_\kappa$.

As with CAT_0 spaces, there is a natural picture, CBB_0 triangles are "fatter" than Euclidean triangles. As mentioned in the introduction, by the usual arguments using Toponogov's theorem the localized version of this definition curvature is equivalent to the global and, for the purposes of this exposition, do not distinguish between the two. We end this section with some useful theorems concerning the geometry of CBB_κ spaces.

Lemma 6.4. *Let (X, d) be a CBB_κ space.*

1. *Any tangent cone to X at x is a CBB_0 space.*

2. *If $\kappa = 0$ and ω is a nonprincipal ultrafilter then $C_\infty^\omega X$ is a CBB_0 space.*

Proof. The first claim is discussed in [BGP92]. The discussion in Section 2.4.2 in [KL96] reduces the second claim to the claim that the tangent cone to a CBB_0 space is CBB_0 as well. The latter fact is also discussed in [BGP92].

Theorem 6.5 (Grove-Petersen [GP]). *Let X be a locally compact, complete, noncompact CBB_0 metric space. If X contains a line (i.e., an unbounded geodesic which minimizes for all time) then X splits isometrically as $E \times \mathbb{R}$ where E is some totally convex subset of X and $\mathbb{R}$ is the real line with its usual metric.*

For our purposes, we use this theorem to prove an analogous result to lemma 6.2.

Lemma 6.6. *Let (X, d) be a CBB_0 space and let $\gamma_1(t)$ and $\gamma_2(t)$ be two lines in X. Then, $\gamma_1(t)$ and $\gamma_2(t)$ are either parallel or diverge linearly.*

Proof. Since these are lines, theorem 6.5 says that X splits isometrically as $X_0 \times \mathbb{R}$ where $\mathbb{R}$ is the geodesic γ_1. Now, $\gamma_2(t)$ is a geodesic line in the product and hence its projection to each factor must be a geodesic line as well. Therefore, the projection of $\gamma_2(t)$ in X_0 is either a point or a line. If it is a point, then γ_1 and γ_2 are parallel. If its projection in X_0 is not a point, then since it is a line, this implies that $d(\gamma_1(t), \gamma_2(t))$ must be linear in t. $\square$

7. Applications of Proposition 5.1.

The purpose of this section is to use the differentiability result (albeit weakly) to examine the possible quasi-isometric embeddings of N_0, a nonabelian simply connected connected nilpotent Lie group equipped with a left invariant Riemannian metric. As discussed in the introduction, we wish to prove that quasi-isometric embeddings of such a Lie group into CAT_0 or CBB_0 cannot exist. After taking asymptotic cones, denoting by N the resulting nonabelian simply connected connected graded nilpotent Lie group equipped with a left-invariant sub-Riemannian metric, we reduce to proving the following theorem.

Theorem C. *Let G be a connected simply connected graded nonabelian nilpotent Lie group equipped with a left invariant Carnot-Carathéodory metric and $U \subset G$ be an open set. Then U does not admit a biLipschitz embedding into any CAT_0 or CBB_0 metric space.*

This follows quickly from the next lemma which estimates the growth rate between radial geodesics.

Lemma 7.1. *Suppose $f : (N, \mathcal{V}, \langle \cdot, \cdot \rangle, d_{cc}) \to (X, d)$ is an L-biLipschitz map such that radial geodesics in N are mapped to d-geodesic lines in X. Then, there exists a pair of geodesics $\gamma_1(t)$ and $\gamma_2(t)$ in X which are images of the radial geodesics e^{tv} and $e^w e^{tv}$, and constants $C_1, C_2 > 0$ and $0 < \alpha < \beta < 1$ such that,*

$$C_1 |t|^\alpha \leq d(\gamma_1(t), \gamma_2(t)) \leq C_2 |t|^\beta$$

for $|t| > 1$.

Proof. This is an exercise in using the Campbell-Baker-Hausdorff formula. Recall that since N is nilpotent, the Campbell-Baker-Hausdorff expansion has only finitely many terms. Suppose $\gamma_1(t) = f(h_t e^v)$ and $\gamma_2(t) = f(e^w h_t e^v)$ for $v, w \in \mathcal{V}$ such that $[v, w] \neq 0$. Using the biLipschitz property, we reduce immediately to consider the rate of growth of the function $d_{cc}(h_t e^v, e^w h_t e^v)$ for $|t| > 1$.

$$d_{cc}(h_t e^v, e^w h_t e^v) = d_{cc}(e^0, h_t e^{-v} e^w h_t e^v)$$
$$= d_{cc}(e^0, e^{w + t[w,v] + H(w,v,t)})$$

where $H(v, w, t)$ are the higher order bracket terms in the Campbell-Baker-Hausdorff expansion. Note that the order n bracket terms in the expansion

each have a coefficient which of the form Bt^m where B is a constant from the Campbell-Baker-Hausdorff expansion and $m < n$. Therefore, for $0 < b < 1$ this is equal to,

$$d_{cc}\left(e^0, h_{|t|^b} e^{\frac{1}{|t|^b} w + \frac{t}{|t|^{2b}}[w,v] + H'(w,v,t)}\right)$$
$$= |t|^b d_{cc}\left(e^0, e^{\frac{1}{|t|^b} w + \frac{t}{|t|^{2b}}[w,v] + H'(w,v,t)}\right).$$

Again, $H'(w,v,t)$ are higher order bracket terms but now, the terms or order n each have a coefficient of the form $\pm Bt^{m-nb}$, where B a constant. Now, picking $0 < \alpha < \beta < 1$ such that $m - n\alpha > 0$ for at least one value of n and $m - n\beta \leq 0$ for all values of n, we have that:

$$C_1 |t|^\alpha \leq d_{cc}(h_t e^v, e^w h_t e^v) \leq C_2 |t|^\beta$$

where C_1 is the minimum of the function $t \mapsto d_{cc}\left(e^0, e^{\frac{1}{|t|^\alpha} w + \frac{t}{|t|^{2\alpha}}[w,v] + H'(w,v,t)}\right)$ on $(-\infty, -1] \cup [1, \infty)$ and C_2 is the maximum of the function $t \mapsto d_{cc}\left(e^0, e^{\frac{1}{|t|^\beta} w + \frac{t}{|t|^{2\beta}}[w,v] + H'(w,v,t)}\right)$ on the same domain. C_1 exists because the function in question is nonnegative and goes to infinity as $t \to \pm\infty$ while C_2 exists because the function in question is nonegative and goes to zero as $t \to \pm\infty$. Thus, using the L-biLipschitz property of the map, we have the desired result with the constants C_i adjusted by either L or $\frac{1}{L}$. $\qquad \square$

Proof of Theorem C. Suppose $f : N \to X$ is the L-biLipschitz map we wish to investigate and X is either CBA_κ or CBB_κ. Then, Proposition 5.1 lets us reduce to an examination of a biLipschitz map $\overline{F} : (N, d_{cc}) \to (C_x X, \overline{d})$ where (N, d_{cc}) and $(C_x X, \overline{d})$ are the appropriate tangent cones and $\overline{F}$ takes radial geodesics to geodesic lines. Lemma 6.3 or Lemma 6.4 implies that the tangent cone $C_x X$ is either CAT_0 or CBB_0. Now we satisfy the assumptions of Lemma 7.1 and hence have a pair of geodesics with the specified divergence rate. However, this contradicts Lemma 6.2 or Lemma 6.6. $\qquad \square$

References.

[Bal85] W. Ballmann, *Nonpositively curved manifolds of higher rank*, Annals of Mathematics, **122(2)** (1985), 597–609.

[Bal95] W. Ballmann, *Lectures on Spaces of Nonpositive Curvature*, DMV Seminar, Birkhäuser, Boston, Basel, Berlin, 1995, Band 25.

[BGP92] Yu. Burago, M. Gromov, and G. Perel'man. *A.D. Aleksan-drov spaces with curvatures bounded below*, Uspekhi Mat. Nauk, **47**(2(284)) (1992), 3–51, 222; Translated in Russian Math. Surveys, **47(2)** (1992), 1–58.

[BR96]   A. Bellaïche and Risler, editors, *Sub-Riemannian Geometry*, **144** of Progress in Mathematics, Birkhäuser, Basel, 1996.

[BS87]   K. Burns and R. Spatzier, *Manifolds of nonpositive curvature and their buildings*, Institute Hautes Études Scientifique Publications Mathématiques, **65** (1987), 35–59.

[Cho39]   W. L. Chow. *Systeme von linearen partiellen differential gleichugen erster ordnug*, Annals of Mathematics, **117** (1939), 98–105.

[Dug66]   J. Dugundji, *Topology*, Allyn and Bacon, Boston, 1966.

[Fed69]   H. Federer, *Geometric measure theory*, Springer-Verlag New York Inc., New York, 1969; Die Grundlehren der mathematischen Wissenschaften, Band 153.

[GP]   K. Grove and P. Petersen, *On the excess of metric spaces and manifolds*, preprint.

[Gro81]   M. Gromov, *Structures métriques pour les variétés riemanniennes*, Textes Mathématiques [Mathematical Texts], **1** CEDIC, Paris, 1981, Edited by J. Lafontaine and P. Pansu.

[Gro93]   M. Gromov. *Asymptotic invariants of infinite groups*, in 'Geometric group theory', **2** (Sussex, 1991); London Math. Soc. Lecture Note Ser., **182**, 1–295, Cambridge Univ. Press, Cambridge, 1993.

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

[Gro91]   K. Grove. *Metric and topological measurements of manifolds*, **II**, 511–519, Mathematical Society of Japan, Tokyo, 1991.

[Ham91]   U. Hamenstädt, *A geometric characterization of locally symmetric spaces*, Journal of Differnetial Geometry, **34(1)** (1991), 193–221.

[Kir94]   B. Kirchheim, *Rectifiable metric spaces: Local structure and the regularity of the Hausdorff measure*, Proc. Amer. Math. Soc., **121** (1994), 113–123.

[KL96]   B. Kleiner and B. Leeb, *Rigidity of quasi-isometries for symmetric spaces and Euclidean buildings*, preprint, April 1996.

[Kle97]   B. Kleiner, *The local structure of length spaces with curvature bounded above*, preprint, April 1997.

[KM96]   R. Karidi and R. Montgomery, *Convexity of nilpotent balls*, March 1996, preprint available at http://math.stanford.edu/~karidi/papers.html.

[KS93]   N.J. Korevaar and R.M. Schoen, *Sobolev spaces and harmonic maps for metric space targets*, Communications in Analytic Geometry, **1(3-4)** (1993), 561–659.

[LS94]   W.S. Liu and H.J. Sussmann, *Abnormal sub-Riemannian minimizers*, Lecture Notes in Pure and Applied Mathematics, **152**, 705–716, Dekker, New York, 1994.

[Mit85]   J. Mitchell, *On Carnot-Carathéodory metrics*, Journal of Differential Geometry, **21** (1985), 35–45.

[MM95]   G.A. Margulis and G.D. Mostow, *The differential of a quasiconformal mapping of a Carnot-Carathéodory space*, Geometric and Functional Analysis, **5(2)** (1995), 402–433.

[Mon96]   R. Montgomery, *Survey of singular geodesics*, Bellaïche and Risler [BR96], **144** (1996), 325–339.

[Mos73]   G.D. Mostow, *Strong rigidity of locally symmetric spaces*, Annals of Mathematical Studies, **78**, Princeton University Press, Princeton, N.J., 1973.

[Nik95]   I. Nikolaev, *The tangent cone of an Aleksandrov space of curvature $\leq k$*, Manuscripta Mathematica, **86(2)** (1995), 137–147.

[Pan83]   P. Pansu, *Croissance des boules et des géodésiques fermé dans les nilvariétés*, Ergodic Theory Dynamical Systems, **3(3)** (1983), 415–445.

[Pan89]   P. Pansu, *Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un*, Annals of Mathematics (2), **129(1)** (1989), 1–60.

[Str86]   R.S. Strichartz, *Sub-Riemannian geometry*, Journal of Differential Geometry, **24** (1986), 221–263.

[Var84]   V.S. Varadarajan, *Lie groups, Lie algebras and their representations*, Springer-Verlag, New York, 1984.

[Wol64]   J.A. Wolf, *Curvature in nilpotent Lie groups*, Proc. Amer. Math. Soc., **15** (1964), 271–274.

RICE UNIVERSITY
HOUSTON, TX, 77005
E-mail address: pauls@math.rice.edu