BILIPSCHITZ EMBEDDINGS AND NONEMBEDDINGS OF METRIC SPACES AND RELATED PROBLEMS

BY

CHRISTOPHER GARTLAND

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2020

Urbana, Illinois

Doctoral Committee:

Professor Jang-Mei Wu, Chair
Professor Jeremy Tyson, Director of Research
Professor Marius Junge
Professor Denka Kutzarova

Abstract

This thesis is concerned with problems relating to the Lipschitz category of metric spaces. We are chiefly interested in building machinery that can be used to deduce the existence or nonexistence of biLipschitz embeddings from one class of metric spaces into another. We will discuss two families of results along these lines.

The first family deals with the problem of biLipschitz embeddability of metric spaces into Banach spaces with Radon-Nikodým property (henceforth, RNP spaces). A major role in this story is played by differentiation theories of Lipschitz functions on metric measure spaces. Nonabelian Carnot groups are prime examples of spaces which support a good differentiation theory, and as a consequence they do not biLipschitz embed into any RNP space, as observed independently by Cheeger-Kleiner and Lee-Naor as a corollary of Pansu's theorem. In search of a nonlinear, metric characterization of the RNP, Ostrovskii found another class of metric spaces that do not biLipschitz embed into RNP spaces, namely spaces containing thick families of geodesics. His proof used an elementary martingale argument and involved no differentiation theory. Our first result is that any metric space containing a thick family of geodesics also contains a subset and a probability measure on that subset that supports a weak differentiation theory for RNP-valued Lipschitz functions. A corollary is a new nonembeddability result: the product of a Carnot group and an RNP space does not contain a biLipschitz copy of a thick family of geodesics. A second result from this project is that, if the metric space is a nonRNP Banach space, a subset consisting of a thick family of geodesics can be constructed to support a true differentiation theory of RNP-valued Lipschitz functions, like the one supported by Carnot groups. An intriguing question is whether the only obstructions to biLipschitz embeddability of complete metric spaces into RNP spaces, like the ones arising from differentiation theory, are local. If this question has a positive answer, it would imply that every complete, topologically discrete metric space biLipschitz embeds into an RNP space. Our third result is a proof of this statement in the special case where the metric space (X, d) is *essentially uniformly discrete*, meaning there is a $\theta > 0$ such that $|B_\theta(p)| < \infty$ for every $p \in X$. This generalizes a result of Kalton who proved that every uniformly discrete metric space biLipschitz embeds into an RNP space. Like Kalton, we prove our result by showing that the Lipschitz free space of X

has the RNP.

The second family of results contained in this thesis is on the calculation of Markov convexity exponents of Carnot groups and applications. Markov convexity, developed by Lee-Naor-Peres and Mendel-Naor, is a biLipschitz and Lipschitz quotient invariant of metric spaces arising as a nonlinear generalization of the property of p-convexity of Banach spaces. It depends only on the finite subsets of the metric space and is thus of a different nature than theories of differentiation, which necessitate the existence of cluster points. Our first main result from this family is that every Carnot group G of step r is Markov p-convex for all $p \in [2r, \infty)$. Our second result is that this is sharp whenever G is a Carnot group with $r \leq 3$ or a model filiform group; such groups are not Markov p-convex for any $p \in (0, 2r)$. This continues a line of research started by Li who proved this sharp result when G is the Heisenberg group. Finally, we obtain the following corollaries of these theorems, which are not attainable by differentiation methods: let G be a Carnot group of step r such that $r \leq 3$, G is a free Carnot group, or G is a jet space group. Let G' be any Carnot group of step $r' < r$.

1. For any lattice $\Gamma \leq G$, the biLipschitz distortion of the Γ-ball of radius R (with respect to a fixed finite generating set) into G' is $\gtrsim \frac{\ln(R)^{\frac{1}{2r'} - \frac{1}{2r}}}{\ln(\ln(R))^{\frac{1}{2r'} + \frac{1}{2r}}}$.
2. G is not a Lipschitz quotient of any subset of G'.
3. G is not a Lipschitz quotient of any subset of L^p (or any p-convex space) for any $p \in (1, 2r)$.
4. The model filiform group of infinite step is not a Lipschitz quotient of any subset of a superreflexive Banach space.

The main question left open by this work is whether there is some Carnot group of step r that is Markov p-convex for some $p < 2r$.

Acknowledgments

First, I thank my family and girlfriend for the support they've given me in graduate school. I am also very appreciative to have had Jeremy Tyson as my thesis advisor and thank him for the career and mathematical guidance he's given me. Thanks goes to my doctoral committee members Jang-Mei Wu, Marius Junge, and Denka Kutzarova for conversations that helped shape this thesis. Finally, I owe much gratitude to the people who kindly wrote letters of recommendation supporting my job applications: Philipp Hieronymi, Sean Li, Mikhail Ostrovskii, and Jeremy Tyson.

Table of Contents

Chapter 1

Introduction

1.1 Background

1.1.1 BiLipschitz Embeddings and Assouad's Theorem

A metric space is a pair (X, d), where X is a set and $d : X \times X \to [0, \infty)$ is a *metric*, or *distance*, satisfying for all $x, y, z \in X$,

- $d(x, y) = 0 \Leftrightarrow x = y$ (Positive definiteness)
- $d(x, y) = d(y, x)$ (Symmetry)
- $d(x, z) \le d(x, y) + d(y, z)$ (Triangle inequality)

We will often suppress notation and just write X instead of (X, d). Although the most important examples of metric spaces are the (finite dimensional) *Euclidean spaces* $(\mathbb{R}^n, (x, y) \mapsto \|x - y\|_2)$, the population is much more varied - metric spaces range from graphs to (sub)Riemannian manifolds to infinite dimensional Banach spaces. Metric spaces also support an extremely rich mapping theory; one may consider, for example, categories whose mappings are continuous, uniformly continuous, quasisymmetric, quasi-isometric, coarse, Lipschitz, or isometric, to name a few. The first two categories are fundamental to basic calculus. The next two are central to geometric group theory and the proof of Mostow's rigidity theorem ([Mos73]), and the fifth gained popularity for its application to the Novikov conjecture ([Yu00]). It is, however, the sixth category, and in particular biLipschitz embeddings, that are the concern of this thesis. Let us state the relevant definitions here.

Definition 1.1. A map $f : X \to Y$ between metric spaces (X, d_X) and (Y, d_Y) is called *Lipschitz* or a *Lipschitz map* if there exists an $L < \infty$ such that $d_Y(f(x), f(y)) \le L d_X(x, y)$ for all $x, y \in X$. The smallest such L is called the *Lipschitz constant* of f. We also say that f is *L-Lipschitz* an *L-Lipschitz map*. f is called *biLipschitz* or a *biLipschitz embedding* if there are $0 < D < \infty$ and $L < \infty$ such that $d_X(x, y) \le D d_Y(f(x), f(y)) \le L d_X(x, y)$ for all $x, y \in X$. The least such L is called the *biLipschitz* *distortion*, or just *distortion* of f. We also say that f is *L-biLipschitz* or an *L-biLipschitz embedding.* If f is also surjective, f is called an *L-biLipschitz equivalence*, or just *biLipschitz equivalence* and X and Y are said to be *biLipschitz equivalent.* A 1-biLipschitz embedding is called an *isometric embedding*, and a 1-biLipschitz equivalence is called an isometry. In this case X and Y are said to be isometric. If a set X is equipped with two metrics such that the identity map between them is biLipschitz, then the metrics are *biLipschitz equivalent*, or just *equivalent.*

Given $x \in X$ and $r \geq 0$, we let $B_r(x)$ denote the *ball of radius r centered at x*; $B_r(x) := \{y \in X : d_X(x,y) \leq r\}$. A surjective map $f : X \to Y$ is a *Lipschitz quotient* if there exist $0 < D < \infty$ and $L < \infty$ such that $B_r(f(x)) \subseteq f(B_{Dr}(x)) \subseteq B_{Lr}(f(x))$ for all $r > 0$, $x \in X$. We also say that f is an *L-Lipschitz quotient map.*

The Lipschitz category also has its applications, especially to approximations in computer science. See [Vem04] for applications of the Johnson-Lindenstrauss lemma ([JL84]) and [LN06] and [NY18] for a discussion of the Sparsest Cut problem and the Goemans-Linial conjecture. Even so, our reasons for studying biLipschitz embeddings are pure mathematical, and we hold that the theory has intrinsic interest for the wealth of its results and diversity of its tools. Of the embeddings in the metric categories we named, biLipschitz embeddings are the second most rigid behind isometric ones. Why study biLipschitz maps instead of isometric ones then? We'll give two reasons. The first is that it often happens that a particular mathematical structure may be naturally equipped with an equivalence class of metrics, but not with any one particular metric. For example, consider a finitely generated group Γ and a finite generating set S with $e \notin S$ and $S^{-1} = S$. Define the *Cayley* graph of (Γ, S) by letting the vertex set be Γ and the edge set being all pairs (g,h) such that $g^{-1}h \in S$. Then we may equip this graph with the shortest path metric and obtain a metric on Γ. This metric generally depends on the choice of generating set S, but any two finite generating sets yield biLipschitz equivalent metrics. Thus, knowing only the algebraic structure of Γ, the statement "Γ biLipschitz embeds into X" is well-defined, but the statement "Γ isometrically embeds into X" is not. Another reason is that isometric embeddings are in many cases simply too rigid and lead to a void theory rather than a rich one. Doubling Hölder spaces furnish such an example.

Definition 1.2. A metric space X is called *doubling* if there is $C < \infty$ such that for every $r > 0$ and $x \in X$, there is a finite set $Y \subseteq X$ with $|Y| \leq C$ and $\cup_{y \in Y} B_{r/2}(y) \supseteq B_r(x)$. A metric space (X, d) is called *Hölder* if there is $q > 1$ such that d^q satisfies the triangle inequality.

The isometric category does not allow for a rich embedding theory for these spaces into Euclidean spaces.

Theorem 1.1 ([LDRW18]). *An infinite Hölder space does not isometrically embed into any Euclidean space.*

This is in sharp contrast to the biLipschitz category, which allows the for the beautiful *Assouad embedding theorem*.

Theorem 1.2 ([Ass83])**.** *Every doubling Hölder space biLipschitz embeds into some Euclidean space.*

The doubling assumption in Assouad's theorem is easily seen to be necessary, so the obvious question is whether the Hölder assumption is necessary.

Question 1.1. Does every doubling metric space admit a biLipschitz embedding into a Euclidean space?

The answer to this question is trivial if "biLipschitz" is replaced with "isometric". Indeed, let $X = \{w, x, y, z\}$ and

$$d(a,b) := \begin{cases} 0 & a = b \\ 1 & \{a,b\} \in \{\{w,x\},\{x,y\},\{y,z\},\{z,x\}\} \\ 2 & \{a,b\} \in \{\{w,y\},\{x,z\}\} \end{cases}$$

X can be pictured as the vertices of a square, where the distance between adjacent vertices is 1 and opposite vertices is 2. X is doubling since $|X| < \infty$, but admits no isometric embedding into a Euclidean space. This is because Euclidean spaces have the property that if $w, x, y, z \in \mathbb{R}^n$ and $\|w - x\| = \|x - y\| = \frac{1}{2}\|w - y\| = \|z - y\| = \|w - z\|$, then $x = \frac{w+y}{2} = z$. Of course, this argument does not apply equally well to biLipschitz embeddings, since $|X| < \infty$ implies that it does biLipschitz embed into $\mathbb{R}$. Nevertheless, the answer to Question 1.1 is a resounding NO, and the examples and tools used to provide this answer are the starting points for the research in this thesis.

1.1.2 The Heisenberg Group

Definition 1.3. Let $\mathbb{H}$ denote $\mathbb{R}^3$ equipped with the binary operation $(x, y, t) * (x', y', t') = (x + x', y + y', t + t' - 2xy' + 2x'y)$. This binary operation is a group product, and $\mathbb{H}$ is called the *Heisenberg group*. We denote the abelianization map $\pi^{\mathrm{ab}} : \mathbb{H} \to \mathbb{R}^2$, $\pi^{\mathrm{ab}}(x, y, t) = (x, y)$. Let $\|(x, y, t)\|_K := ((x^2 + y^2)^2 + t^2)^{\frac{1}{4}}$ and $d_K((x, y, z), (x', y', z')) := \|(x, y, z)^{-1} * (x', y', t')\|_K$. d_K is a metric on $\mathbb{H}$, called the *Korányi metric*. When we refer to the Heisenberg group, we are typically referring to the metric space $(\mathbb{H}, d_K)$.

The Heisenberg group is doubling, and it was observed by Semmes ([Sem96]) that Pansu's differentiation theorem implies that no Euclidean space admits a biLipschitz embedding of the Heisenberg group, thus negatively answering Question 1.1 (see Section 1.3.3 for background on Pansu's theorem). Later, Cheeger-Kleiner and Lee-Naor independently extended this observation to Banach spaces with the Radon-Nikodým property (henceforth, RNP, see Section 1.3.1 for further background).

Definition 1.4. A Banach space V has the *RNP*, or is an *RNP space* if every Lipschitz map $\mathbb{R} \to V$ is differentiable Lebesgue-almost everywhere.

Theorem 1.3 ([CK06], Theorem 6.1; [LN06], Section 1.2)**.** *$\mathbb{H}$ does not biLipschitz embed into any RNP space.*

Li found a fundamentally different proof of the non-biLipschitz embeddability of the Heisenberg group into Hilbert space using Markov convexity (see Section 1.1.6 for background on Markov convexity). The proof is fundamentally different because Markov convexity is a finitary notion, as opposed to differentiation which requires cluster points.

Theorem 1.4 ([Li16], Theorem 1.1, Corollary 1.3)**.** *$\mathbb{H}$ is Markov p-convex if and only if $p \geq 4$. Consequently, $\mathbb{H}$ does not biLipschitz embed into Hilbert space.*

A question left open by the work of Li is whether a similar statement holds true for Carnot groups of higher step ($\mathbb{H}$ has step 2, see Section 1.3.3 for background on Carnot groups).

Question 1.2. If G is a Carnot group of step r, is G Markov p-convex if and only if $p \geq 2r$?

The Banach space $L^1([0,1])$ is not an RNP space, but despite this fact it also admits no biLipschitz embedding of the Heisenberg group.

Theorem 1.5 ([CK10])**.** *$\mathbb{H}$ does not biLipschitz embed into $L^1([0,1])$*

The proof method used here by Cheeger-Kleiner is still differentiation-based. A stronger, quantitative version of this theorem was found by Naor-Young in [NY18] using method of quantitative rectifiability. Importantly, it also solved a strong version of the Goemans-Linial conjecture (see [NY18] for a discussion).

The differentiation theorem of Pansu is a generalization of Rademacher's theorem from Euclidean spaces to Carnot groups. In [Che99], Cheeger found an even vaster generalization of Rademacher's theorem for a class of metric measure spaces called PI spaces (see Section 1.3.4). In addition to the Heisenberg group, one of the first and most important examples of a PI space is Laakso space, our next example of a doubling space non-biLipschitz embeddable into Euclidean spaces.

1.1.3 Laakso Space

Definition 1.5. We define a sequence of metric graphs $G_0, G_1, \ldots$ recursively, as follows:

- G_0 consists of two vertices connected by a single edge, whose length is 1. The metric on G_0 is denoted d_0.

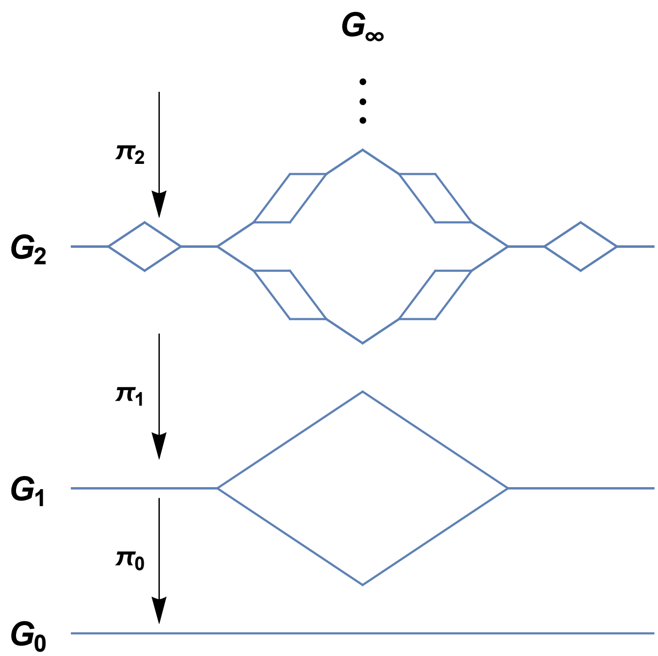

Figure 1.1: The Laakso diamond graphs. Each new graph G_{i+1} is obtained from G_i by replacing each edge with a copy of G_1, scaled down so that the diameter of G_{i+1} remains 1. There are 1-Lipschitz surjections $\pi_{i+1} : G_{i+1} \to G_i$, and the Laakso space, G_∞, is defined to be the inverse limit of this system.

- Each new graph G_{i+1} is obtained from G_i by replacing each edge of G_i with a copy of G_1, shown in Figure 1.1, scaled down so that the diameter of G_{i+1} remains 1. That is, the length of each edge in G_{i+1} is one fourth of the length of each edge in G_i. G_{i+1} is equipped with the shortest path metric d_{i+1}.

There are canonical 1-Lipschitz surjections $\pi_{i+1} : G_{i+1} \to G_i$ defined by collapsing each scaled down copy of G_1 back onto the edge which it replaced. *Laakso space*, G_∞, is defined to be the inverse limit metric space of the system. Specifically, G_∞ is the set $\{x \in \Pi_{i=0}^{\infty} G_i : \forall i, \pi_i(x_{i+1}) = x_i\}$ equipped with the metric d_∞ defined by $d_\infty(x, y) := \lim_{i\to\infty} d_i(x_i, y_i)$.

In the form we present, Laakso space was actually first introduced by Lang-Plaut in [LP01], inspired by a construction of Laakso in [Laa00]. In any case, it has become conventional in the field to refer to G_∞ as Laakso space.

In [LP01, Theorem 2.3], Lang-Plaut proved that Hilbert space does not admit a biLipschitz embedding of Laakso space. In [CK09], Cheeger-Kleiner extended this to RNP spaces using the theory of differentiation on metric measure spaces.

Theorem 1.6 ([CK09], Corollary 1.7)**.** *G_∞ satisfies the differentiability nonembeddability criterion into RNP spaces (see Definition 1.26). Consequently, G_∞ does not biLipschitz embed into any RNP space.*

In [Ost14b], Ostrovskii proved a nonlinear characterization of RNP spaces as those that do not admit biLipschitz embeddings of a class of metric spaces called thick families of geodesics (see Theorem 1.8 and Definition 1.6). As a consequence, he provided a different proof of the non-biLipschitz embeddability of Laakso space into RNP spaces.

Theorem 1.7 ([Ost16], Example 3.3)**.** *G_∞ contains a thick family of geodesics. Consequently, G_∞ does not biLipschitz embed into any RNP space.*

Ostrovskii's method of proof via martingales and is quite elementary compared to that of Cheeger-Kleiner.

1.1.4 The RNP and Thick Families of Geodesics

It's been known since 1973 that Lipschitz maps from separable Banach spaces to RNP spaces are, in a suitable sense, differentiable almost everywhere. This is due independently to [Aro76], [Chr73], and [Man73] (see [BL00, section 6.6]). It follows that the RNP is inherited under biLipschitz embeddability of Banach spaces, since it is inherited under isomorphic embeddability. It is then natural to ask for a purely metric characterization of the RNP - one that does not rely on the linear structure. This question was asked by Bill Johnson in 2009 and answered in 2014 by Ostrovskii (see [Ost14b, Section 1]).

Theorem 1.8 ([Ost14b], Corollary 1.5)**.** *A Banach space does not have the RNP if and only if it admits a biLipschitz embedding of a thick family of geodesics.*

Definition 1.6. Let (X, d) be a metric space and $p, q \in M$. A p-q *geodesic* is an isometric embedding from some closed bounded interval into X mapping the left endpoint of the interval to p and the right endpoint to q. The distance function d is said to be *geodesic* if there exists a p-q geodesic for every $p, q \in X$. A family of p-q geodesics Γ with common domain $[a, b]$ is said to be *concatenation closed* if for every $c \in [a, b]$ and $\gamma_1, \gamma_2 \in \Gamma$ with $\gamma_1(c) = \gamma_2(c)$, the concatenated curve γ defined by $\gamma(t) = \gamma_1(t)$ if $t \in [a, c]$, $\gamma(t) = \gamma_2(t)$ if $t \in [c, b]$, also belongs to Γ.

Given $\alpha > 0$, a concatenation closed family of p-q geodesics Γ sharing a common domain $[a, b]$ is said to be *α-thick* or an *α-thick family of geodesics* if for every $\gamma \in \Gamma$ and $a = t_0 < t_1 < \dots t_k = b$, there exist $a = q_0 < s_1 < q_1 < s_2 < \dots s_j < q_j = b$ and $\tilde{\gamma} \in \Gamma$ such that

- $\{t_i\} \subseteq \{q_i\}$
- $\gamma(q_i) = \tilde{\gamma}(q_i)$
- $\sum_{i=1}^{j} d(\gamma(s_i), \tilde{\gamma}(s_i)) \geq \alpha$

A concatenation closed family of p-q geodesics Γ sharing a common domain $[a, b]$ is said to be *thick* or a *thick family of geodesics* if it is α-thick for some $\alpha > 0$.

Remark 1.1. Informally, a family of geodesics is concatenation closed if for any γ_1, γ_2 in the family, the geodesic obtained by concatenating an initial segment of γ_1 and a terminal segment of γ_2 also belongs to the family. Informally, a concatenation closed family of p-q geodesics is α-thick if for any geodesic γ in the family and any finite set of points F in the image of γ, there is another geodesic $\tilde{\gamma}$ in the family that intersects γ at each point of F (but possibly more points), and so that the deviation of $\tilde{\gamma}$ from γ between their points of intersection adds up to at least α.

On the other hand, according to another intriguing result of Ostrovskii, the Heisenberg group does not admit a biLipschitz embedding of a thick family of geodesics. This is due to Theorem 1.4 the fact that Markov convexity is inherited under biLipschitz embeddings, and the following result of Ostrovskii.

Theorem 1.9 ([Ost14a],Theorem 1.5)**.** *Metric spaces admitting a biLipschitz embedding of a thick family of geodesics are not Markov p-convex for any p. Consequently, $\mathbb{H}$ does not admit a biLipschitz embedding of a thick family of geodesics.*

So although containing a thick family of geodesics is a necessary condition for the non-biLipschitz embeddability of Banach spaces into RNP Banach spaces, the same is not true of general metric spaces, even for quasi-convex ones such as $\mathbb{H}$.

Question 1.3. When does a metric space fail to biLipschitz embed into RNP spaces?

1.1.5 Uniformly Discrete Metric Spaces

The proofs of Theorems 1.3, 1.6, and 1.8 actually imply something stronger than non-biLipschitz embeddability, namely non-*local* biLipschitz embeddability.

Definition 1.7. A metric space (X, d_X) is said to *locally biLipschitz embed* into a class of metric spaces $\mathcal{Y}$ if for every $x \in X$, there are an open set $U_x \subseteq X$ and a metric space $Y_x \in \mathcal{Y}$ such that $x \in U_x$ and U_x biLipschitz embeds into Y_x.

Differentiation methods (and also the closely related martingale methods) are inherently local, so in fact we know that the Heisenberg group and Laakso space do not locally biLipschitz embed into RNP spaces. As far as we are aware, these are the only known techniques used to prove non-biLipschitz embeddability. Thus, a specific form of Question 1.3 is:

Question 1.4. Are the only obstructions to the biLipschitz embeddability of complete metric space into RNP spaces local? That is, if a complete metric space X locally biLipschitz embeds into RNP spaces, must X biLipschitz embed into some RNP space?

An example where the hypothesis is trivially satisfied is when the metric space is discrete. In this case, Question 1.4 takes the form:

Question 1.5. Does every complete, discrete metric space biLipschitz embed into an RNP space?

The strongest partial result towards a positive answer to Question 1.5 is due to Kalton.

Definition 1.8. A metric space (X, d) is called *uniformly discrete* if there is $\theta > 0$ such that $d(x, y) \geq \theta$ for all $x \neq y \in X$.

Theorem 1.10 ([Kal04], Proposition 4.4)**.** *If X is uniformly discrete, then* $\mathrm{LF}(X)$ *has the RNP.*

This theorem implies uniformly discrete metric spaces isometrically embed into RNP spaces, since every metric space X isometrically embeds into its Lipschitz free space $\mathrm{LF}(X)$ (see Definition 1.17).

1.1.6 The Ribe Program and Markov Convexity

See [Nao12] and [Nao18] for good surveys on the Ribe program.

Definition 1.9. A Banach space V is *finitely representable* in another W if there exists $\lambda < \infty$ such that for any finite dimensional $F \subseteq V$, there is an injective linear map $T : F \to W$ with $\|T\|\|T^{-1}\| \leq \lambda$. Properties of Banach spaces that are preserved under mutual finite representability are called *local*.

In [Rib76], Ribe showed that if two Banach spaces E, F are uniformly homeomorphic, then they are mutually finitely representable. This theorem implies that local properties are really metric properties, suggesting that each should have a reformulation that involves only the metric structure of the Banach space and not the linear structure. The research program concerned with finding these reformulations is known as the *Ribe program*. The program was initiated by Bourgain in [Bou86] in which he made the first substantial contribution by characterizing superreflexive Banach spaces as those which do not admit biLipschitz embeddings of the binary trees of depth k with uniform control on the biLipschitz distortion. Another major contribution to the Ribe program is a purely metric reformulation of p-convexity (see Section 1.3.1 for background on superreflexivity and p-convexity). The metric property *Markov p-convexity* was originally defined by Lee-Naor-Peres in [LNP09] and proved by Mendel-Naor in [MN13] to be a reformulation of p-convexity. Here are the specifics:

Definition 1.10 ([MN13], Definition 1.2)**.** Let $\{X_t\}_{t\in\mathbb{Z}}$ be a Markov chain on a state space Ω. Given an integer $k \geq 0$, we denote by $\{\tilde{X}_t(k)\}_{t\in\mathbb{Z}}$ the process which equals X_t for time $t \leq k$ and evolves independently (with respect to the same transition probabilities) for time $t > k$. Fix $p > 0$. A metric space (M, d) is called *Markov p-convex* if there is $\Pi < \infty$ so that for every Markov chain $\{X_t\}_{t\in\mathbb{Z}}$ on a state space Ω, and for every $f : \Omega \to M$,

$$\sum_{k=0}^{\infty}\sum_{t\in\mathbb{Z}} \frac{\mathbb{E}[d(f(X_t), f(\tilde{X}_t(t-2^k)))^p]}{2^{kp}} \leq \Pi^p \sum_{t\in\mathbb{Z}} \mathbb{E}[d(f(X_{t+1}), f(X_t))^p]$$

Set $\Pi_p(M)$ equal to the least value of Π so that the above inequality holds (whenever it exists). $\Pi_p(M)$ is called the *Markov p-convexity constant* of M.

Definition 1.11. Fix $p > 0$. A metric space (X, d) is *4-point p-convex* if there exist a symmetric function $\rho : X \times X \to [0, \infty)$ and constants $C, K < \infty$ such that for all $x, y, w, z \in X$,

$$\frac{1}{C}d(x, y) \leq \rho(x, y) \leq Cd(x, y)$$

and

$$\rho(y, x)^p + \frac{\rho(y, z)^p}{2} + \frac{\rho(y, w)^p}{2} - \frac{\rho(x, z)^p}{2^p} - \frac{\rho(x, w)^p}{2^p} \geq \frac{\rho(z, w)^p}{K}$$

Theorem 1.11 (Theorem 1.3, [MN13])**.** *A metric space that is 4-point p-convex is Markov p-convex, and a Banach space is p-convex if and only if it is 4-point p-convex if and only if it is Markov p-convex.*

We have already seen an application of Markov convexity in Theorem 1.9. Here is another very interesting application.

Theorem 1.12 ([LNP09], Lemma 3.8)**.** *If a finitely generated group Γ admits a nonconstant, bounded harmonic function, then Γ is not Markov p-convex for any p. Consequently, Γ does not biLipschitz embed into any superreflexive Banach space.*

In addition to Theorem 1.4, the following is known about the Markov convexities of Carnot groups.

Theorem 1.13 ([Li14], Proposition 7.2 and Theorem 7.4)**.** *Every graded nilpotent Lie group of step r is Markov $2(r!)^2$-convex.*

1.2 Summary of Results

1.2.1 Thick Families of Geodesics and Differentiation

See Chapter 2 for further discussion and proofs of the statements in this subsection.

We sought to study nonembeddability into RNP spaces (Question 1.3) in more detail. The Heisenberg group (Theorem 1.9) shows that it can not be determined using thick families of geodesics.

Question 1.6. Is there a more general RNP non-biLipschitz embeddability criterion that works equally well for thick families of geodesics and the Heisenberg group? The Heisenberg group satisfies the differentiability nonembeddability criterion into RNP spaces (see Definition 1.26 and the proceeding examples) - do thick families of geodesics also satisfy this criterion?

We prove that the answer to this question is yes if the notion of differentiability is weakened. The type of RNP LDS (see Definition 1.24) we construct is weaker than a true RNP LDS because the almost everywhere approximation of RNP-valued Lipschitz functions by their derivative only holds on some sequence of scales tending to 0 instead of all scales. More specifically, we prove Theorems 2.3 and 2.9, which can be summarized as:

Theorem 1.14 (Summary of Theorems 2.3 and 2.9)**.** *For any complete metric space M containing a thick family of geodesics, there exist a compact subset X_∞, Borel probability measure μ_∞ on X_∞, Lipschitz map $\pi : X_\infty \to [0,1]$, Borel subset $S_\infty \subseteq X_\infty$, a sequence of scales $r_i(x) \searrow 0$ for almost every $x \in X_\infty$, and a nonprincipal ultrafilter $\mathcal{U}(x)$ on $\mathbb{N}$ for each $x \in S_\infty$ such that:*

2.3 $\mu_\infty(S_\infty) > 0$, and for every $x \in S_\infty$ the tangent cone $T_x^{r_i(x),\mathcal{U}(x)} X_\infty$ admits no continuous injection into $\mathbb{R}$.

2.9 For every RNP space B and Lipschitz map $f : X_\infty \to B$, for μ_∞-almost every $x \in X_\infty$, f is differentiable at x with respect to π along the sequence of scales $(r_i(x))_{i=0}^\infty$.

As a corollary, we obtain a new proof of nonembeddability into RNP spaces:

Corollary 1.1. *A metric space M containing a thick family of geodesics does not biLipschitz embed into any RNP space.*

The proof is the same as for the true differentiation nonembeddability criterion into RNP spaces (see Theorem 1.21).

Proof. Let B be an RNP space and assume there is a biLipschitz map $f : M \to B$. We may assume M is complete. Let $X_\infty \subseteq M$, μ_∞, S_∞, $r_i(x)$, and $\mathcal{U}(x)$ be as in the statement of Theorem 1.14. Since $\mu_\infty(S_\infty) > 0$, there exist a point $x \in S_\infty$ and a nonprincipal ultrafilter $\mathcal{U}(x)$ such that f is differentiable at x along $(r_i(x))_{i=0}^\infty$ with respect to π and $T_x^{r_i(x),\mathcal{U}(x)} X_\infty$ admits no continuous injection into $\mathbb{R}$. The function f being differentiable with respect to π at x along $(r_i(x))_{i=0}^\infty$ implies that there exists a unique linear map

$f'(x) : \mathbb{R} \to B$ such that, for every nonprincipal ultrafilter $\mathcal{U}$, the blowup of f at x, $f_x : T_x^{r_i(x),\mathcal{U}} X_\infty \to B$, exists and factors though the blowup of π at x, $\pi_x : T_x^{r_i(x),\mathcal{U}} X_\infty \to \mathbb{R}$, and $f'(x) : \mathbb{R} \to B$. That is, $f_x = f'(x) \circ \pi_x$. Since $T_x^{r_i(x),\mathcal{U}(x)} X_\infty$ admits no continuous injection into $\mathbb{R}$, π_x cannot be injective, which by the factorization implies f_x cannot be injective, in turn implying f cannot be biLipschitz. $\square$

Theorem 1.14 actually proves a stronger statement, Corollary 2.1. We postpone the proof until Section 2.7. We chose to give a separate proof Corollary 1.1 because it is easier and requires no knowledge of Carnot groups.

Corollary 2.1. *A complete metric space M containing a thick family of geodesics does not biLipschitz embed into the product metric space $G \times B$, where G is a Carnot group and B is an RNP space.*

At the time of this writing, Theorems 1.8 and 1.9 were the only known nontrivial means by which one could prove nonembeddability of thick families of geodesics into metric spaces. Suppose G is a nonabelian Carnot group, such as the Heisenberg group, and B is an RNP which is not superreflexive, such as ℓ^1. Then G embeds into no RNP space by the differentiation nonembeddability criterion, so Theorem 1.8 does not apply to $G \times B$, and B is not Markov p-convex for any p, so Theorem 1.9 does not apply to $G \times B$. That non-superreflexive spaces are not Markov p-convexity for any p follows from the fundamental theorem of Mendel-Naor on Markov convexity ([MN13, Theorem 1.3]), and Pisier's renorming theorem (Theorem 1.17). Thus, Corollary 2.1 is a genuinely new nonembeddability result.

In our second result, Theorem 2.10, we restrict our attention from a general metric containing a thick family of geodesics to a nonRNP Banach space B. This is indeed a "restriction" since every such B contains a thick family of geodesics by Theorem 1.8. In this setting, we prove that the subset X_∞ and measure μ_∞ can be constructed to satisfy the true RNP differentiation nonembeddability criterion (not just the weakened form described in Theorem 1.14). That it satisfies the true RNP differentiation criterion is a consequence of the fact that it is an inverse limit of an admissible system of graphs, defined in [CK15]. In that article, Cheeger and Kleiner proved that such spaces are PI spaces. They also gave a necessary and sufficient condition for these spaces to satisfy the differentiation nonembeddability criterion into RNP spaces, stated in [CK15, Theorem 10.2]. We verify this condition for our subset $X_\infty \subseteq B$, and thus our result can be viewed as a converse to [CK15, Theorem 10.2].

Theorem 2.10. *Every nonRNP Banach space contains a biLipschitz copy of a metric measure space satisfying the differentiation nonembeddability criterion. The metric measure space is an inverse limit of admissible graphs, as in [CK15], with nonEuclidean tangent cones at almost every point.*

1.2.2 Essentially Uniformly Discrete Spaces

See Chapter 3 for further discussion and proofs of the statements in this subsection.

Our main result is a generalization of Theorem 1.10 and a step closer to a positive answer to Question 1.5.

Theorem 1.15. *If X is essentially uniformly discrete, then* $\mathrm{LF}(X)$ *has the RNP.*

Let us also mention other metric spaces whose Lipschitz free space is known to have the RNP: (1) Proper, countable spaces ([Dal15, Theorem 2.1]). (2) Proper biLipschitz-Hölder spaces ([Jen68, Theorem 4.1]). Recall that a metric space (X, ρ) is *proper* if its closed and bounded subsets are compact and *biLipschitz-Hölder* if it is biLipschitz equivalent to a Hölder space. Actually, in both of [Dal15, Theorem 2.1] and [Jen68, Theorem 4.1], the Lipschitz free spaces are shown to be isomorphic to separable dual spaces, which is strictly stronger than RNP ([MO80], [Pis16, Corollary 2.15]). In light of this and Theorems 1.10 and 1.15, one may ask if uniformly discrete and essentially uniformly discrete countable spaces biLipschitz embed into separable dual spaces. For uniformly discrete spaces, this is an open question (equivalent to [dLPP19, Problem 1.3]), and indeed we do not have an essentially uniformly discrete counterexample either.

We will give the proof of Theorem 1.15 at the conclusion of this subsubsection, after stating the relevant definitions and collecting the main ingredients that are proven in Chapter 3.

Definition 1.12. Let (X, d) be a topologically discrete metric space, i.e., every set is open in the metric topology. For each $p \in X$ and $r \geq 0$, we let $B_r(p) := \{q \in X : \rho(p, q) \leq r\}$. For each $p \in X$, let $rad(p) := \sup\{r \leq \mathrm{diam}(X) : B_r(p) = \{p\}\}$, $essrad(p) := \sup\{r \leq \mathrm{diam}(X) : |B_r(p)| < \infty\}$. We say that X is *θ-uniformly discrete* if $0 < \theta = \inf_{p \in X} \mathrm{rad}(p)$ and *θ-essentially uniformly discrete* if $0 < \theta = \inf_{p \in X} \mathrm{essrad}(p)$. We say that X is uniformly discrete if there exists a $\theta > 0$ such that X is θ-uniformly discrete, and similarly for essentially uniform discreteness.

Note the following implications:

uniformly discrete $\Rightarrow$ *essentially uniformly discrete* $\Rightarrow$ *complete*

The first implication is obvious, and we explain how to prove the second. We'll show that every Cauchy sequence is eventually constant, which is equivalent to discrete and completeness. Suppose we have a Cauchy sequence taking values in an essentially uniformly discrete metric space. Since it is Cauchy, it eventually belongs to a ball of arbitrarily small radius B. By definition of essentially uniformly discreteness, the radius of B can be chosen small enough so that $|B| < \infty$. Thus, our Cauchy sequence eventually belongs to a finite set. This implies it must be eventually constant.

Since X is discrete, finitely supported functions on X are Lipschitz. Let $\mathrm{Lip}_{\mathrm{fin}}(X)$ denote the subspace of $\mathrm{Lip}_0(X)$ consisting of the finitely supported functions. The inclusion $\mathrm{Lip}_{\mathrm{fin}}(X) \hookrightarrow \mathrm{Lip}_0(X)$ dualizes to a quotient $\mathrm{Lip}_0(X)^* \twoheadrightarrow \mathrm{Lip}_{\mathrm{fin}}(X)^*$. Let **res** denote the restriction of this map to $\mathrm{LF}(X)$, $\mathrm{res} : \mathrm{LF}(X) \to \mathrm{Lip}_{\mathrm{fin}}(X)^*$.

Theorem 3.2. *If X is bounded and countable, then $\mathrm{Lip}_{\mathrm{fin}}(X)^*$ is separable.*

Theorem 3.1. *If X is bounded and countable, then $\mathrm{res} : \mathrm{LF}(X) \to \mathrm{Lip}_{\mathrm{fin}}(X)^*$ is an isomorphic embedding if and only if X is essentially uniformly discrete.*

Proof of Theorem 1.15. We'll cite [Pis16, Chapter 2] for standard results we need on RNP (our definition of RNP is different but equivalent to that in [Pis16], see [Pis16, Remark 2.17]). Assume X is essentially uniformly discrete. The RNP is separably determined; that is, if every separable closed subspace of a Banach space has the RNP, then so does the entire space ([Pis16, Corollary 2.12]). Clearly, any separable subspace of $\mathrm{LF}(X)$ is contained in $\mathrm{LF}(Y)$ for some countable Y, so it suffices to prove $\mathrm{LF}(Y)$ has the RNP for any countable $Y \subseteq X$. Let $Y \subseteq X$ be countable. By [Kal04, Proposition 4.3], $\mathrm{LF}(Y)$ isomorphically embeds into the ℓ^1-direct sum $\oplus_{i=1}^{\infty} \mathrm{LF}(B_i(0))$, where $B_i(0)$ denotes the ball of radius i in Y centered at the basepoint $0 \in Y$. Since an ℓ^1 sum of RNP spaces has the RNP, it suffices to assume that Y is bounded. But now Theorem 3.1 kicks in (essentially uniform discreteness passes to subsets), and we get that $\mathrm{LF}(Y)$ isomorphically embeds into $\mathrm{Lip}_{\mathrm{fin}}(Y)^*$. Separable dual spaces have the RNP, so Theorem 3.2 implies $\mathrm{Lip}_{\mathrm{fin}}(Y)^*$ has the RNP. Since $\mathrm{LF}(Y)$ isomorphically embeds into the RNP space $\mathrm{Lip}_{\mathrm{fin}}(Y)^*$, $\mathrm{LF}(Y)$ has the RNP. □

1.2.3 Markov Convexity of Carnot Groups

See Chapter 4 for further discussion and proofs of the statements in this subsection.

We present in this subsection our results on the calculation of the Markov convexities of Carnot groups (Question 1.2). We now state our main theorems, which sharpen Theorems 1.4 and 1.13.

Theorem 4.1. *Every graded nilpotent Lie group of step r, equipped with a left invariant metric homogeneous with respect to the dilations induced by the grading, is 4-point p-convex - and consequently Markov p-convex - for every $p \in [2r, \infty)$.*

Theorem 4.2. *For every $p > 0$, $r \geq 1$, coarsely dense set $N \subseteq J^{r-1}(\mathbb{R})$, and $R \geq 3$, let $B_N(R) := \{x \in N : d_{CC}(0,x) \leq R\}$. Then*

$$\Pi_p(B_N(R)) \gtrsim \frac{\ln(R)^{\frac{1}{p}-\frac{1}{2r}}}{\ln(\ln(R))^{\frac{1}{p}+\frac{1}{2r}}}$$

where the implicit constant can depend on r, p but not on N, R.

Recall that a subset N of a metric space (X, d_X) is *coarsely dense* if there exists $C < \infty$ such that $X = \cup_{x' \in N} \{x \in X : d_X(x, x') \leq C\}$. See Section 1.3.3 for the definition of $J^{r-1}(\mathbb{R})$. Theorem 4.1 is restated and proved at the end of Section 4.3.2, and similarly for Theorem 4.2 at the end of Section 4.4.2.

We can extend this result to other groups using the notion of subquotients. X is a *Lipschitz subquotient* of Y with constant C if there is a metric space Z such that Z embeds isometrically into Y and X is a Lipschitz quotient of Z with constant C, or, equivalently, there is a a metric space Z such that Z is a Lipschitz quotient of Y with constant C and X isometrically embeds into Z. It follows from [MN13, Proposition 4.1] that if X is a Lipschitz subquotient of Y with constant C then $\Pi_p(X) \leq C\Pi_p(Y)$.

Every free Carnot group of step $r \geq 2$ has $J^{r-1}(\mathbb{R})$ (in fact every graded nilpotent Lie group of step r with 2-dimensional horizontal layer) as a graded quotient group, and the projection map $\mathbb{R}^k \twoheadrightarrow \mathbb{R}$ dualizes to a graded embedding $J^{r-1}(\mathbb{R}) \hookrightarrow J^{r-1}(\mathbb{R}^k)$. See [BLU07, Chapter 14] for background on free Carnot groups and [War05] for background on the jet spaces groups $J^{r-1}(\mathbb{R}^k)$.

Corollary 1.2. *Let G be a Carnot group of step r that has $J^{r-1}(\mathbb{R})$ as a graded subquotient group, for example G may be a free Carnot group, $J^{r-1}(\mathbb{R}^k)$, or any Carnot group if $r \leq 3$. The set of $p > 0$ for which G is Markov p-convex is exactly $[2r, \infty)$.*

Proof. This follows from Theorems 4.1 and 4.2 and the preceding discussion. □

Recall that a subgroup $\Gamma \leq G$ of a Lie group G is a *lattice* if the subspace topology on Γ is discrete and G/Γ carries a G-invariant, Borel probability measure.

Corollary 1.3. *Let G be a Carnot group of step r that has $J^{r-1}(\mathbb{R})$ as a graded subquotient group, for example G may be a free Carnot group, $J^{r-1}(\mathbb{R}^k)$, or any Carnot group if $r \leq 3$ (by Lemma 1.3). Let $\Gamma \leq G$ be a lattice equipped with the word metric with respect to a finite generating set (which exists by [Rag72, Theorem 2.21]), and let $B_\Gamma(R)$ denote the ball of radius R in Γ centered at the identity. Then for any $p > 0$,*

$$\Pi_p(B_\Gamma(R)) \gtrsim \frac{\ln(R)^{\frac{1}{p} - \frac{1}{2r}}}{\ln(\ln(R))^{\frac{1}{p} + \frac{1}{2r}}}$$

Proof. Let G, Γ, p be as above. The inclusion $\Gamma \hookrightarrow G$ is a biLipschitz embedding onto a coarsely dense subset when Γ is equipped with the word metric with respect to a finite generating set (this can be proven using Mostow's theorem that lattices in nilpotent Lie groups are cocompact ([Mos62]) and applying the fundamental theorem of geometric group theory). Thus it suffices to prove the conclusion for any coarsely dense $N'' \subseteq G$. Let N'' be such a subset. By assumption, there is a Carnot group G' and a graded quotient homomorphism $q : G \to G'$ such that $J^{r-1}(\mathbb{R})$ is a graded subgroup of G'. Then q is a Lipschitz quotient map, so there is a constant $C < \infty$ such that for any $R \geq 3$,

$$\Pi_p(B_{N''}(R)) \gtrsim \Pi_p(B_{q(N'')}(R/C))$$

Thus it suffices to prove the conclusion for any coarsely dense subset $N' \subseteq G'$. Let N' be such a subset. Fix $B >> 1$ and let $N \subseteq J^{r-1}(\mathbb{R})$ be a coarsely dense, B-separated subset (each pair of distinct points in N is separated by a distance at least B - such sets always exist by Zorn's Lemma). Then since $J^{r-1}(\mathbb{R})$ is a graded subgroup of G', there is a biLipschitz embedding $N \to G'$. If B is chosen large enough, we map postcompose with a nearest neighbor map $G' \to N'$ to obtain another biLipschitz embedding $N \to N'$. Then the conclusion follows from Theorem 4.2. □

The following quantitative nonembeddability estimate follows from the previous corollary and Theorem 4.1.

Corollary 1.4. *Let G be a Carnot group of step r that has $J^{r-1}(\mathbb{R})$ as a graded subquotient group, for example G may be a free Carnot group, $J^{r-1}(\mathbb{R}^k)$, or any Carnot group if $r \leq 3$. Let $\Gamma \leq G$ be a lattice equipped with the word metric with respect to a finite generating set, and let $B_\Gamma(R)$ denote the ball of radius R in Γ centered at the identity. Let G' be any graded nilpotent Lie group of step $r' < r$. Then we have the following estimate for $c_{G'}(B_\Gamma(R))$, the biLipschitz distortion of $B_\Gamma(R)$ in G':*

$$c_{G'}(B_\Gamma(R)) \gtrsim \frac{\ln(R)^{\frac{1}{2r'}-\frac{1}{2r}}}{\ln(\ln(R))^{\frac{1}{2r'}+\frac{1}{2r}}}$$

where the implicit constant depends on G and G' but not on R.

Such quantitative nonembeddability estimates have been the subject of much attention for embeddings of Heisenberg groups into certain Banach spaces, see [ANT13] and [LN14] for uniformly convex Banach space targets and [NY18] for L^1 targets. In particular, it can be deduced from [ANT13] and [Ass83] that the biLipschitz distortion of the ball of radius R in a lattice in the Heisenberg group into Hilbert space equals, up to universal factors, $\sqrt{\ln(R)}$. Thus, our estimates in the previous corollary cannot be sharp when $r = 2$ and $r' = 1$. However, these estimates seem to be the first of their type when the target is allowed to be a nilpotent group of step larger than 1. Other quantitative nonembeddability estimates of between Carnot groups were obtained in [Li14], but they are of a different flavor. Since our estimates are not sharp for $r = 2, r' = 1$, we speculate that they are not sharp for larger values of r, r' either. Next, we obtain new results on the nonexistence Lipschitz subquotient maps.

Corollary 1.5. *Let G be a Carnot group of step r that has $J^{r-1}(\mathbb{R})$ as a graded subquotient group, for* *example G may be a free Carnot group, $J^{r-1}(\mathbb{R}^k)$, or any Carnot group if $r \leq 3$. Let G' be any graded nilpotent Lie group of step r'.*

1. *G is not a Lipschitz subquotient of L^p (or any p-convex space) for any $p \in (1, 2r)$.*
2. *If $r > r'$, G is not a Lipschitz subquotient of G'.*

Proof. These follow from the previous corollary, the fact that Markov p-convexity is preserved under Lipschitz subquotients, Theorem 1.11, and the classical fact that L^p is $\max(2, p)$-convex for $p > 1$. □

Essentially all of the previously know results of this flavor are proved as a corollary of Pansu differentiation (Theorem 1.19), which applies when the domain is a (finite dimensional) Carnot group and the target is an RNP Banach space or (finite dimensional) Carnot group. There is also a more recent differentiation theorem of Le Donne-Li-Moisala ([LDLM18]) which applies when the domain is a "scalable" group filtrated by (finite dimensional) Carnot groups and the target is an RNP space. However, there does not seem to be a clear way to deduce Corollary 1.5 in full generality from any of these methods.

We may use Markov convexity again to prove nonexistence of subquotient maps onto some "infinite step" graded Lie groups. See Section 1.3.3 for the definitions of inverse limits, $J^\infty(\mathbb{R}^k)$, and the free Carnot group on k generators, F_k^∞.

Corollary 1.6. *Let $G_0 \leftarrow G_1 \leftarrow \dots$ be an inverse system of graded nilpotent Lie groups such that for every r, there is an i with $J^{r-1}(\mathbb{R})$ a graded subquotient of G_i, and let G_∞ be the inverse limit group. For example, G_∞ may be $J^\infty(\mathbb{R}^k)$ or F_k^∞. Then G_∞ is not a Lipschitz subquotient of any superreflexive space.*

Proof. Pisier's renorming theorem (Theorem 1.17), states that any superreflexive Banach space is p-convex for some $p \in [2, \infty)$. Thus it suffices to show that G_∞ is not Markov p-convex for any $p \in (0, \infty)$. For every $r \geq 1$, $J^{r-1}(\mathbb{R})$ is a Lipschitz subquotient of G_∞, so since Markov p-convexity is preserved under Lipschitz quotients, the conclusion follows from Corollary 1.2. □

Finally, we provide a positive result on the existence of embeddings using one of the main results of [LNP09]. A *metric tree* is the vertex set of a weighted graph-theoretical tree equipped with the shortest path metric.

Theorem 1.16 (Theorem 4.1, [LNP09])**.** *If T is a metric tree and T is Markov p-convex, then T biLipschitz embeds into L^p.*

Corollary 1.7. *If a metric tree T is a Lipschitz subquotient of a graded nilpotent Lie group G of step r, then T biLipschitz embeds into L^p for every $p \geq 2r$.*

Proof. This follows from Theorem 4.1, the fact that Markov convexity is inherited by Lipschitz subquotients, and Theorem 1.16. □

We conclude this introduction with a conjecture and a question.

Conjecture 1.1. For each graded nilpotent Lie group G, the set of p for which G is Markov p-convex is the same as that of the largest Carnot subgroup of G.

Question 1.7. Let Γ be a lattice in a Carnot group that does not biLipschitz embed into some other Carnot group G. What is the infimal α so that $\limsup_{R\to\infty} \dfrac{c_G(B_\Gamma(R))}{\ln(R)^\alpha} < \infty$?

1.3 Preliminaries

1.3.1 Banach Spaces

Definition 1.13. A *normed space* is a pair $(V, \|\cdot\|)$ where V is a vector space over $\mathbb{R}$ and $\|\cdot\| : V \to [0,\infty)$ is a *norm*, satisfying for all $x, y \in V$ and $c \in \mathbb{R}$,

- $\|x\| = 0 \Rightarrow x = 0$ (Positive definiteness)
- $\|cx\| = |c|\|x\|$ (Absolute homogeneity)
- $\|x+y\| \leq \|x\| + \|y\|$ (Triangle inequality)

We will often suppress notation and just write V instead of $(V, \|\cdot\|)$. These axioms imply that $(x,y) \mapsto \|x-y\|$ is a metric on V, and we will always treat normed spaces as metric spaces equipped with this norm. A normed space for which the associated metric is complete is called a *Banach space*.

A linear map between normed spaces is *bounded* if it is Lipschitz, an *isomorphic embedding* if it is a biLipschitz embedding, and an *isomorphism* if it a biLipschitz equivalence. Two norms on a vector space are *equivalent* if the identity map is an isomorphism.

We let B_V denote the closed unit ball of V centered at the origin.

Every finite dimensional normed space is a Banach space, and any two norms on a finite dimensional space are equivalent.

Example 1.1. Given any Banach space $(V, \|\cdot\|)$, there is a new Banach space $(V^*, \|\cdot\|_{V^*})$ called the *dual space* of V, consisting of the linear functionals $\lambda : V \to \mathbb{R}$ for which $\sup_{v\in B_1(0)} |\lambda(v)| < \infty$, equipped with the norm $\|\lambda\|_{V^*} := \sup_{v\in B_V} |\lambda(v)|$.

Example 1.2. See [Pis16, Chapter 1] for the following discussion. Given a Banach space $(V, \|\cdot\|)$, a measure space $(\Omega, \mathcal{A}, \mu)$, and $p \in [1, \infty]$, we get a new Banach space $L^p(\mu; V)$ of (equivalence classes of) Bochner measurable functions equipped with the norm

$$\|f\|_{L^p(\mu;V)} := \left(\int \|f\|^p d\mu\right)^{\frac{1}{p}}$$

A function $\Omega \to V$ is called *Bochner measurable* if it is a pointwise μ-almost everywhere limit of a sequence of simple functions.

When $V = \mathbb{R}$, we get the classical *Lebesgue space* $L^p(\mu)$.

Bochner measurable functions are Borel measurable, but generally not conversely.

Definition 1.14. A Banach space V is *p-convex* for some $p \in [2, \infty)$ if there exists an equivalent norm $\|\cdot\|$ and $K < \infty$ such that for every $\epsilon \in [0, 2]$,

$$\sup\{\|(x+y)/2\| : \|x\|, \|y\| \leq 1, \|x-y\| \geq \epsilon\} \leq 1 - \epsilon^p/K$$

Example 1.3. For any $q \in (1, \infty)$ and measure μ such that $L^q(\mu)$ is infinite dimensional, $L^q(\mu)$ is p-convex if and only if $p \geq \max(2, q)$. Every finite dimensional normed space is p-convex for all $p \geq 2$.

Definition 1.15. Given a Banach space V, there is a canonical linear isometric embedding $J : V \to V^{**}$ defined by $J(v)(\lambda) = \lambda(v)$. V is *reflexive* if J is surjective.

Definition 1.16. A Banach space V is *superreflexive* if every Banach space that is finitely representable in V is reflexive.

An deep and important fact is *Pisier's renorming theorem.*

Theorem 1.17 ([Pis16], Theorem 11.37)**.** *A Banach space is superreflexive if and only if it is p-convex for some $p \in [2, \infty)$.*

We recall again the definition of RNP spaces.

Definition 1.4. *A Banach space V has the* RNP*, or is an* RNP space *if every Lipschitz map $\mathbb{R} \to V$ is differentiable Lebesgue-almost everywhere.*

Example 1.4. Separable dual spaces and reflexive spaces have the RNP. In particular, $\ell^1 = c_0^*$ has the RNP. $L^1([0,1])$ and c_0 do not have the RNP. For $L^1([0,1])$, an example of a nowhere differentiable Lipschitz map is furnished by $t \mapsto \mathbf{1}_{[0,t]}$, and for c_0, $t \mapsto (\sin(nt)/n)_{n=1}^{\infty}$.

It will be helpful to keep in mind the following chain of implications regarding Banach spaces.

$$p\text{-}convex \Rightarrow superreflexive \Rightarrow reflexive \Rightarrow RNP$$

See [Pis16, Chapter 2] for further background on the RNP and its many characterizations. We will state a few important definitions and theorems as they concern this thesis, and the proofs can be found in [Pis16, Chapter 2].

Theorem 1.18. *A Banach space V has the RNP if and only if for every probability space $(\Omega, \mathcal{A}, \mathbb{P})$, filtration $(\mathcal{A}_n)_{n=0}^{\infty}$, and $(\mathcal{A}_n)_{n=0}^{\infty}$-adapted martingale $(M_n)_{n=0}^{\infty}$ taking values in a bounded subset of $L^{\infty}(\mathbb{P}; V)$, there exists an $M \in L^{\infty}(\mathbb{P}; V)$ such that $\mathbb{E}^{\mathcal{A}_n}(M) = M_n$ for all $n \in \mathbb{N}$. Moreover, $\mathbb{E}^{\mathcal{A}_n}(M) = M_n$ for all $n \in \mathbb{N}$ if and only if $M_n \overset{n\to\infty}{\to} M$ $\mathbb{P}$-almost surely.*

1.3.2 Lipschitz Free Spaces

Definition 1.17. Let (X, ρ) be a metric space with distinguished basepoint $0 \in X$. Let $\mathrm{Lip}_0(X)$ denote the Banach space of Lipschitz functions $f : X \to \mathbb{R}$ satisfying $f(0) = 0$ equipped with the norm $\|f\|_{\mathrm{Lip}_0(X)} := \sup_{p\neq q} \frac{|f(p)-f(q)|}{\rho(p,q)}$. Then X isometrically embeds into $\mathrm{Lip}_0(X)^*$ via $\delta = p \mapsto \delta_p$, where $\delta_p(f) = f(p)$. The linear span of $\{\delta_p\}_{p\in X}$ in $\mathrm{Lip}_0(X)^*$ is denoted by $\mathrm{LF}_{\mathrm{fin}}(X)$, and its closure by $\mathrm{LF}(X)$. $\mathrm{LF}(X)$ is a Banach space called the *Lipschitz free space over X*.

Lipschitz free spaces are a very well-studied class of Banach spaces. See [Ost13, Chapter 10] and [Wea99] (note that Lipschitz free space are called Arens-Eells spaces in that text) for textbook introductions to Lipschitz free spaces and [God15] for a survey on more recent research.

We'll recall four fundamental facts about Lipschitz free space. The first is that $\mathrm{LF}(X)^* = \mathrm{Lip}_0(X)$ ([Wea99, Theorem 2.2.2]). Let $\Delta \subseteq X \times X$ denote the diagonal and set $\tilde{X} := X \times X \setminus \Delta$. Then ρ is nonvanishing on $\tilde{X}$. Let $\ell^1(\tilde{X})/\rho$ denote the Banach space of countably supported measures μ on $\tilde{X}$ equipped with the norm $\|\mu\| = \int \rho d|\mu|$. The second fact is that there is a linear quotient map $\pi : \ell^1(\tilde{X})/\rho \to \mathrm{LF}(X)$ defined on the canonical basis by $\pi(\delta_{(p,q)}) = \delta_p - \delta_q$. The third fundamental fact is that if $0 \in Y \subseteq X$, the natural inclusion $\mathrm{LF}(Y) \hookrightarrow \mathrm{LF}(X)$ is an isometric embedding. This is due to the *McShane extension theorem*: every Lipschitz function from Y to $\mathbb{R}$ can be extended to a Lipschitz function on all of X without increasing the Lipschitz norm ([Wea99, Theorem 1.5.6(a)]). The fourth and final fact is the *universal linear extension property*: Given any Lipschitz map $f : X \to V$ into a Banach space V with $f(0) = 0$, there exists a unique bounded linear map $T_f : \mathrm{LF}(X) \to V$ with $f = T_f \circ \delta$. Moreover, $\|T_f\| = \|f\|_{\mathrm{Lip}_0(X)}$.

1.3.3 Carnot Groups

The next several subsubsections don't follow any particular reference, but ones we recommend are [BLU07] for Carnot groups and [LD17] for graded nilpotent groups. We mostly follow [War05] for the subsection on jet spaces.

Graded Nilpotent and Stratified Lie Algebras and their Lie Groups

Definition 1.18. A *graded nilpotent Lie algebra* $(\mathfrak{g}, [\cdot,\cdot])$ *of step* r is a Lie algebra equipped with a grading $\mathfrak{g} = \oplus_{i=1}^r \mathfrak{g}_i$, meaning $\mathfrak{g}_r \neq 0$, $[\mathfrak{g}_i, \mathfrak{g}_j] \subseteq \mathfrak{g}_{i+j}$ if $i+j \leq r$, and $[\mathfrak{g}_i, \mathfrak{g}_j] = 0$ if $i + j > r$. A *stratified* Lie algebra $(\mathfrak{g}, [\cdot,\cdot])$ of *step* r is a graded nilpotent Lie algebra of step r such that the Lie subalgebra generated by $\mathfrak{g}_1$ is all of $\mathfrak{g}$. The grading is called a *stratification*, $\mathfrak{g}_1$ is often called the *horizontal layer* (or stratum), and $\mathfrak{g}$ is said to be *horizontally generated*. Whenever a Lie algebra $\mathfrak{g}$ (not presumed to be equipped with a grading) admits a stratification, it is unique (Lemma 2.16, [LD17]). A *graded nilpotent Lie group of step* r is a simply connected Lie group whose Lie algebra is graded nilpotent of step r. A graded nilpotent Lie group whose Lie algebra is stratified is a *Carnot group*. A *graded homomorphism* or *map* is a Lie group homomorphism between graded nilpotent Lie groups whose derivative is a graded Lie algebra homomorphism. One graded nilpotent Lie group G' is a *graded subgroup* of another graded nilpotent Lie group G if there is an injective graded homomorphism from G' into G. One graded nilpotent Lie group G' is a *graded quotient group* of another graded nilpotent Lie group G if there is a surjective graded homomorphism from G onto G'. One graded nilpotent Lie group G' is a *graded subquotient group* of another graded nilpotent Lie group G if there is another graded nilpotent Lie group G'' such that G'' is a graded subgroup of G and G' is a graded quotient group of G'', or, equivalently, there is another graded nilpotent Lie group G'' such that G'' is a graded quotient group of G and G' is a graded subgroup of G''.

Given a graded nilpotent Lie group G and its Lie algebra $\mathfrak{g}$, since $\mathfrak{g}$ is nilpotent and G is simply connected, the exponential map is a diffeomorphism, and thus we can use it to equip $\mathfrak{g}$ with a graded nilpotent Lie group structure such that it becomes graded isomorphic to G. The Baker-Campbell-Hausdorff formula provides a formula for the group product on $\mathfrak{g}$ in terms of the Lie algebra structure (Section 2, [War05]):

$$xy = \sum_{n>0} \frac{(-1)^{n+1}}{n} \sum_{\substack{0<p_i+q_i \\ i\leq i\leq n}} C_{p,q}^{-1} (\mathrm{ad}x)^{p_1} (\mathrm{ad}y)^{q_1} \dots (\mathrm{ad}x)^{p_n} (\mathrm{ad}y)^{q_n-1} y \tag{1.1}$$

where $(\mathrm{ad}x)y = [x,y]$ and $C_{p,q} = p_1!q_1!\dots p_n!q_n!\,(\sum_{i=1}^n p_i + q_i)$. In this formula and what follows, whenever $\mathfrak{g}$ is a graded nilpotent Lie algebra, we equip it with the product defined by (1.1) and simultaneously think of $\mathfrak{g}$ as a graded nilpotent Lie group and Lie algebra. We will always use juxtaposition to denote the group product.

Definition 1.19. Every graded nilpotent Lie group G has a canonical family of *dilations* $\delta_t : G \to G$ parametrized by $t \in (0, \infty)$ whose derivative $\delta_t' : \mathfrak{g} \to \mathfrak{g}$ is defined by

$$\delta_t'(x) := tx_1 + t^2 x_2 + \ldots t^r x_r$$

where $\mathfrak{g}$ is the Lie algebra, and $x_i \in \mathfrak{g}_i$ is the $\mathfrak{g}_i$-component of $x \in \mathfrak{g}$.

$t \mapsto \delta_t$ is an automorphic $\mathbb{R}_{>0}$-action on G. It can be deduced that a Lie group homomorphism θ between graded nilpotent Lie groups is a graded homomorphism if and only if it is δ_t-equivariant, that is, $\theta(\delta_t(x)) = \delta_t(\theta(x))$, where we've abused (and will continue to do so) notation and written δ_t for the dilation on both the domain and codomain.

Norms and Metrics

Definition 1.20. Let G be a graded nilpotent Lie group. A *homogeneous quasi-norm* on G is a continuous function $N : G \to \mathbb{R}$ such that for all $x \in G$ and $t \in \mathbb{R}_{>0}$,

- $N(x) \geq 0$ (Positive semi-definite)
- $N(x^{-1}) = N(x)$ (Symmetry)
- $N(\delta_t(x)) = tN(x)$ (Homogeneity)

If additionally $N(x) = 0$ implies $x = 0$, then N is a *positive definite* homogeneous quasi-norm, and if $N(xy) \leq N(x) + N(y)$ for all $x, y \in G$ (triangle inequality), N is a *homogeneous norm.*

For any two positive definite homogeneous quasi-norms N, N' on G, the continuity, homogeneity, and positive definiteness of N, N', together with the compactness of the unit sphere in $\oplus_{i=1}^r \mathbb{R}^{\dim(\mathfrak{g}_i)}$, imply that N and N' are biLipschitz equivalent, that is, there is a constant $0 < C < \infty$ such that

$$C^{-1}N(x) \leq N'(x) \leq CN(x)$$

for all $x \in G$.

Positive definite homogeneous norms always exist, most famously those considered in [HS90]. Thus any positive definite homogeneous quasi-norm N satisfies the *quasi-triangle inequality*: there is a $0 < C < \infty$ such that for all $x, y \in G$,

$$N(xy) \leq C(N(x) + N(y))$$

Typically one requires that every homogeneous quasi-norm N satisfies the quasi-triangle inequality. Although it turns out that the quasi-norms we consider in this article do satisfy the quasi-triangle inequality, we only need to know this for positive-definite quasi-norms and thus do not explicitly make this requirement.

There is a bijective correspondence between homogeneous, positive definite quasi-norms N on G and left-invariant, homogeneous quasi-metrics d_N on G via $N \mapsto d_N$ defined by

$$d_N(x,y) := N(y^{-1}x)$$

Positive definiteness of N implies positive definiteness of d_N, symmetry of N implies symmetry of d_N, homogeneity of N implies the homogeneity of d_N (meaning $d_N(\delta_t(x), \delta_t(y)) = td_N(x,y)$), and the quasi-triangle inequality of N implies the quasi-triangle inequality of d_N. The left-invariance of d_N is automatic from the definition. N satisfies the triangle inequality if and only if d_N does. The inverse of $N \mapsto d_N$ is $d \mapsto N_d$, where $N_d(x) := d(0,x)$. In addition to those determined by the homogeneous, positive definite norms from [HS90], there are canonical left-invariant, homogeneous metrics on Carnots groups called *Carnot-Caratheodory metrics*, denoted d_{CC}. These metrics are also geodesic. See [BLU07] or [LD17] for further information.

Whenever dealing with a graded nilpotent Lie group, we will automatically assume it is equipped with a left-invariant, homogeneous quasi-metric. By the preceding discussion, this quasi-metric is well-defined up to biLipschitz equivalence, so any biLipschitz-invariant property of metric spaces we may well attribute to a graded nilpotent Lie group G knowing only the algebraic structure of its graded Lie algebra. The δ_t-equivariance of graded group maps implies that any graded map between graded nilpotent Lie groups is Lipschitz, and thus graded group embeddings are biLipschitz embeddings, graded quotient maps are Lipschitz quotient maps, and graded group isomorphisms are biLipschitz equivalences.

Pansu's Differentiation Theorem

A Carnot group is in particular, a locally compact group, and thus supports a Haar measure. It turns our that nilpotent Lie groups are unimodular, so we make no distinction between the left and right Haar measures. It also turns out that Haar measure is homogenous with respect to the dilations, that is, there is some $s > 0$ such that $\lambda(\delta_t(E)) = t^s\lambda(E)$, where $E \subseteq G$ is Borel and λ is Haar measure. Given a Carnot groups G, we will always consider it equipped with a left-invariant homogeneous metric d and Haar measure λ so that (G, d, λ) is a metric measure space.

As far as the biLipschitz theory is concerned, the following theorem is the most fundamental.

Definition 1.21. Let (G, d_G) be a Carnot group and (X, d_X) a Carnot group or a Banach space, $f : G \to X$

a map, and $p \in G$. We say that f is *Pansu differentiable* at p if there exists a Carnot group homomorphism $\theta : G \to X$ such that, for each $K \subseteq G$ compact,

$$\sup_{q \in K} d_X \left(\delta_{\frac{1}{t}}(f(p)^{-1} * f(p * \delta_t(q))), \theta(q) \right) \overset{t \searrow 0}{\to} 0$$

Theorem 1.19 ([Pan89])**.** *Every Lipschitz map from a Carnot group to another Carnot group or RNP Banach space is Pansu differentiable almost everywhere. Consequently, if G biLipschitz embeds into H, then G is a Carnot subgroup of H.*

In the preceding definition and theorem, we need to interpret the notion of "Carnot group homomorphism" when the target is a Banach space V to mean a map that factors $G \overset{\pi^{\mathrm{ab}}}{\to} G^{\mathrm{ab}} \cong \mathfrak{g}_1 \overset{T}{\to} V$ where T is a linear map.

Pansu's theorem has another important consequence, the *unique lifting theorem.*

Theorem 1.20. *For any Carnot group G and Lipschitz maps $f, g : \mathbb{R} \to G$, if $f(p) = g(p)$ for some $p \in G$ and $\pi_{\mathrm{ab}} \circ f = \pi_{\mathrm{ab}} \circ g$ ($\pi_{\mathrm{ab}} : G \to G^{\mathrm{ab}}$ denotes the abelianization), then $f = g$.*

Model Filiform Groups and Jet Spaces over $\mathbb{R}$

We follow [War05] (especially Example 4.3) throughout this subsection. The *model filiform group* of step $r \geq 1$ is the Carnot group with stratified Lie algebra $\mathfrak{g} = (\mathbb{R}X \oplus \mathbb{R}Y_1) \oplus_{i=2}^{r} \mathbb{R}Y_i$, where X, Y_1 is a basis for $\mathfrak{g}_1$ and Y_i is a basis for $\mathfrak{g}_i$ for $2 \leq i \leq r$, and the nontrivial bracket relations are given by $[X, Y_i] = Y_{i+1}$ for $1 \leq i \leq r-1$. Clearly, for $s \geq r$, there is a canonical Carnot group quotient map from the model filiform group of step s to that of step r. The model filiform group of step 2 is frequently called the *Heisenberg group*, and the one of step 3 the *Engel group*. The corresponding Lie algebras are the *Heisenberg algebra* and *Engel algebra*.

The *jet space over* $\mathbb{R}$ of step $r \geq 0$, denoted $J^{r-1}(\mathbb{R})$, is a certain Carnot group of step r graded isomorphic to the model filiform group of step r. There are also jet space groups $J^{r-1}(\mathbb{R}^k)$ over higher dimensional Euclidean space, but we will focus on $k = 1$ in this discussion. As a set, $J^{r-1}(\mathbb{R})$ consists of equivalence classes of pairs (x, f) where $x \in \mathbb{R}$ and $f \in C^{r-1}(\mathbb{R})$. Two pairs $(x, f), (y, g)$ are equivalent if $x = y$ and $f^{(k)}(x) = g^{(k)}(y)$ for all $0 \leq k \leq r-1$. We define maps $\pi_x, \pi_i : J^{r-1}(\mathbb{R}) \to \mathbb{R}$, $0 \leq i \leq r-1$, by $\pi_x([(y, g)]) = y$ and $\pi_i([(y, g)]) = g^{(i)}(y)$. These maps are obviously well-defined and the direct sum map $\pi_x \oplus_{i=0}^{r-1} \pi_{r-1-i} : J^{r-1}(\mathbb{R}) \to \mathbb{R} \times \mathbb{R}^r$ is a bijection. For $v \in J^{r-1}(\mathbb{R})$, the quantity $\pi_x(v)$ is referred to as the *x-coordinate* and $\pi_i(v)$ as the *u_i-coordinate*. We equip $J^{r-1}(\mathbb{R})$ with a topological vector space structure so that this map is a linear homeomorphism, and from this point on will represent elements of $J^{r-1}(\mathbb{R})$ using these coordinates. We will especially represent elements as pairs $(y, v) \in J^{r-1}(\mathbb{R}) = \mathbb{R} \times \mathbb{R}^r$ so that $y \in \mathbb{R}$, $v \in \mathbb{R}^r$, and $\pi_x((y, v)) = y$. Although we won't explicitly use it, the group operation on $J^{r-1}(\mathbb{R})$ is given by

$$\pi_x((x, u_{r-1}, \dots u_0) * (y, v_{r-1}, \dots v_0)) = x + y$$

$$\pi_i((x, u_{r-1}, \dots u_0) * (y, v_{r-1}, \dots v_0)) = u_i + v_i + \sum_{j=i+1}^{r-1} u_j \frac{y^{j-i}}{(j-i)!}$$

Given $y \in \mathbb{R}$ and $g \in C^{r-1}(\mathbb{R})$, we get an element $[j^{r-1}(y)](g) \in J^{r-1}(\mathbb{R})$ defined by

$$\pi_x([j^{r-1}(y)](g)) = y$$

$$\pi_i([j^{r-1}(y)](g)) = g^{(i)}(y)$$

called the *jet* of g at y. The following two Lemmas are essentially all we need to know about jet spaces. The first is a special case of [RW10]. Although their lemma is stated for C^r functions, the proof works the same in the case of $C^{r-1,1}$ functions.

Lemma 1.1 (pages 4-5, [RW10])**.** *For any $[a, b] \subseteq \mathbb{R}$ and $\phi \in C^{r-1,1}([a, b])$,*

$$d_{CC}([j^{r-1}(b)](\phi), [j^{r-1}(a)](\phi)) \leq \left(1 + \left\|\phi^{(r)}\right\|_{L^\infty([a,b])}\right) |b - a|$$

Lemma 1.2. *There is a constant $c > 0$ such that for all $(x, u), (x, v) \in J^{r-1}(\mathbb{R})$,*

$$d_{CC}((x, u), (x, v)) \geq c|\pi_0(u - v)|^{\frac{1}{r}}$$

Proof. By left invariance of d_{CC} and the ball-box theorem (see Corollary 2.2 of [Jun19], there is a constant $c > 0$ such that for all $(x, u), (x, v) \in J^{r-1}(\mathbb{R})$,

$$d_{CC}((x, u), (x, v)) \geq c|\pi_0((x, v)^{-1}(x, u))|^{\frac{1}{r}}$$

and by Lemma 3.1 from [Jun17],

$$\pi_0((x, v)^{-1}(x, u)) = \pi_0(u - v)$$

□

The following lemma will be used to obtain lower bounds on the Markov convexity of Carnot groups of step 2 or 3.

Lemma 1.3. *Every Carnot group of step 2 or 3 contains the model filiform group of the corresponding step (the Heisenberg or Engel group) as a graded subquotient group.*

Proof. Let G be a Carnot group of step 2 with stratified Lie algebra $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$. Since $\mathfrak{g}$ has step 2, there is a nonzero $V_2 \in \mathfrak{g}_2$. Since $\mathfrak{g}$ is horizontally generated, there exist $U, V_1 \in \mathfrak{g}_1$ such that $[U, V_1] = V_2$. Recall that the Heisenberg algebra has first layer generated by linearly independent vectors X, Y_1, second layer generated by $Y_2 \neq 0$, and nontrivial bracket relation $[X, Y_1] = Y_2$. Then it easily follows that $X \mapsto U$, $Y_1 \mapsto V_1$, $Y_2 \mapsto V_2$ is a graded algebra embedding into $\mathfrak{g}$. This proves that the Heisenberg group is a graded subgroup of G.

Now assume G is of step 3 with stratified Lie algebra $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \mathfrak{g}_3$. By the grading property, any subspace of $\mathfrak{g}_3$ is an ideal, and thus there is a graded algebra quotient map onto another step 3 stratified Lie algebra whose third layer is one dimensional. Thus we may assume $\mathfrak{g}_3 = \mathbb{R}W$, $W \neq 0$, and prove that the Engel algebra embeds into $\mathfrak{g}$. Since $\mathfrak{g}$ is horizontally generated, $W = [U_1, [U_2, U_3]]$ for some $U_1, U_2, U_3 \in \mathfrak{g}_1$. First we claim that there is a 2-dimensional subspace of the span of U_1, U_2, U_3 that generates a Lie subalgebra of step 3. After proving the claim, we'll show that this subalgebra must be graded algebra-isomorphic to the Engel algebra. To prove the claim, we'll show that at least one of the following is nonzero:

1. $[U_1, [U_1, U_2]]$
2. $[U_1, [U_1, U_3]]$
3. $[U_2, [U_2, U_3]]$
4. $[U_3, [U_3, U_2]]$
5. $[U_1 + U_2, [U_1 + U_2, U_3]]$
6. $[U_1 + U_3, [U_1 + U_3, U_2]]$

Assume that all terms are 0. First let's see that $[U_2, [U_3, U_1]] = W$.

$$0 \overset{(5)}{=} [U_1 + U_2, [U_1 + U_2, U_3]] = [U_1, [U_1, U_3]] + [U_1, [U_2, U_3]] + [U_2, [U_1, U_3]] + [U_2, [U_2, U_3]]$$

$$\overset{(2),(3)}{=} W + [U_2, [U_1, U_3]] = W - [U_2, [U_3, U_1]]$$

Using $(6), (1), (4)$ in place of $(5), (2), (3)$ shows $[U_3, [U_1, U_2]] = W$. Putting these together yields:

$$[U_1, [U_2, U_3]] + [U_2, [U_3, U_1]] + [U_3, [U_1, U_2]] = 3W \neq 0$$

in violation of the Jacobi identity. This proves the claim.

So now the situation is that there are $Z_1, Z_2 \in \mathfrak{g}_1$ with $[Z_1, [Z_1, Z_2]] = zW$ for some $z \neq 0$. Recall that the Engel algebra has first layer spanned by X, Y_1, second layer by Y_2, and third layer by Y_3 with nontrivial bracket relations $[X, Y_1] = Y_2$ and $[X, Y_2] = Y_3$. Let $z' \in \mathbb{R}$ such that $[Z_2, [Z_1, Z_2]] = z'W$. Then since $[Z_1, [Z_1, Z_2]] = zW \neq 0$, the map from the Engel algebra into $\mathfrak{g}$ defined by

$$X \mapsto Z_1, \qquad Y_1 \mapsto Z_2 - \frac{z'}{z} Z_1, \qquad Y_2 \mapsto [Z_1, Z_2], \qquad Y_3 \mapsto zW$$

is a graded algebra embedding. □

Remark 1.2. The analogue of Lemma 1.3 is false for groups of step larger than 3. Let $\mathfrak{g}$ be the stratified Lie algebra $\mathfrak{g} = \oplus_{i=1}^4 \mathfrak{g}_i$ with $\mathfrak{g}_1 = \mathbb{R}X_{11} \oplus \mathbb{R}X_{12}$, $\mathfrak{g}_2 = \mathbb{R}X_2$, $\mathfrak{g}_3 = \mathbb{R}X_{31} \oplus \mathbb{R}X_{32}$, $\mathfrak{g}_4 = \mathbb{R}X_4$ and nontrivial brackets $[X_{11}, X_{12}] = X_2$, $[X_{11}, X_2] = X_{31}$, $[X_{12}, X_2] = X_{32}$, $[X_{11}, X_{31}] = X_4$, $[X_{12}, X_{32}] = X_4$. The only graded quotient maps from $\mathfrak{g}$ onto another step 4 stratified Lie algebra or graded embeddings into $\mathfrak{g}$ from another step 4 stratified Lie algebra are isomorphisms.

Infinite Step Carnot groups

Given an inverse system of graded nilpotent Lie groups $G_1 \overset{\rho_1}{\leftarrow} G_2 \overset{\rho_2}{\leftarrow} \dots$, where each ρ_i is a graded quotient map, we define the *inverse limit metric group*, G_∞, to be the subgroup of $(\oplus_{i=1}^\infty G_i)_\infty$ consisting of those sequences $(x_i)_{i=1}^\infty$ for which $\rho(x_{i+1}) = x_i$ for all $i \geq 1$, where $(\oplus_{i=1}^\infty G_i)_\infty$ is the ℓ_∞-sum of the pointed metric spaces $(G_i, d_{CC}, 0)$. G_∞ inherits a left-invariant homogeneous metric from $(\oplus_{i=1}^\infty G_i)_\infty$ (where the dilations δ_t are defined on G_∞ in the obvious way), and each G_i is a Lipschitz quotient of G_∞.

Definition 1.22. $J^\infty(\mathbb{R}^k)$ is the inverse limit metric group, equipped with the induced δ_t-action, associated to the natural inverse system formed by the jet space groups, $J^0(\mathbb{R}^k) \overset{\rho_1}{\leftarrow} J^1(\mathbb{R}^k) \overset{\rho_2}{\leftarrow} \dots$. See [War05] for background on jet space groups. Similarly, F_k^∞ is the inverse limit metric group, equipped with the induced δ_t-action, associated to the natural inverse system formed by the free Carnot groups on k generators, $F_k^1 \overset{\rho_1}{\leftarrow} F_k^2 \overset{\rho_2}{\leftarrow} \dots$. See Chapter 14 of [BLU07] for background on free Carnot groups.

1.3.4 Lipschitz Differentiability Spaces

For additional information on Lipschitz differentiability spaces, see [KM16], and note that we only consider single chart spaces in this thesis.

Definition 1.23. Let (X, d) be a metric space, $\psi : X \to \mathbb{R}^k$ a Lipschitz map, and $p \in X$. Let $(V, \|\cdot\|)$ be a Banach space and $f : X \to V$ a map. We say that f is *differentiable with respect to ψ at p* if there exists a unique linear map $D_\psi f_p : \mathbb{R}^k \to V$ such that

$$\lim_{x \to p} \frac{\|f(x) - f(p) - Df_p(\psi(x) - \psi(p))\|}{d(x,p)} = 0$$

In this case, $D_\psi f_p$ is called the *derivative* of f with respect to ψ at p.

Definition 1.24. Let (X, d, μ) be a *metric measure space*, meaning (X, d) is a metric space and μ is a complete Borel measure on X. (X, d, μ) is an (single chart) *RNP Lipschitz differentiability space* (henceforth *RNP LDS*) if there exists a Lipschitz map $\psi : X \to \mathbb{R}^k$ such that, for every RNP space V, every Lipschitz map $X \to V$ is differentiable with respect to ψ μ-almost everywhere.

Example 1.5. *Rademacher's theorem* states that $\mathbb{R}^n$ is an RNP LDS when equipped with the Euclidean metric and Lebesgue measure. The chart ψ is the identity map.

Example 1.6. *Pansu's differentiation theorem* implies (with slight modification) that every Carnot group G is an RNP LDS. The chart ψ is the abelianization map $G \to G^{\mathrm{ab}} \cong \mathfrak{g}_1$.

Example 1.7 ([CK15], Theorem 9.1)**.** Laakso space is an RNP LDS when equipped with a certain probability measure. The chart ψ is the projection $\pi_0 : G_\infty \to G_0 \cong [0, 1]$.

The concept of a LDS was first conceived by Cheeger in [Che99] in the context of doubling metric measure spaces admitting a Poincaré inequality, *PI spaces*. A metric measure space (X, d, μ) is doubling if there exists $C < \infty$ such that for all $p \in X$ and $r > 0$, $\mu(B_{2r}(x)) \leq C\mu(B_r(x))$. Since we never work directly with Poincaré inequalities in this thesis, we omit their definitions. The systematic study of PI spaces was initiated by Heinonen-Koskela in [HK98]. [CK09, Theorem 1.5] states that PI spaces can de decomposed (up to a null set) into a countable union of RNP LDS's.

Definition 1.25. Let (X, d) be a metric space. Given a point $p \in X$, a sequence $(r_i)_{i=0}^\infty$ decreasing to 0 and a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$, we define the *tangent cone* of X at p, $T_p^{r_i,\mathcal{U}} X$, to be the $\mathcal{U}$-ultralimit of the sequence of pointed spaces $(X, p, \frac{1}{r_i} d)$. Given a Lipschitz map $X \to V$ into a Banach space, the blowup of f at p, $f_p : T_p^{r_i,\mathcal{U}} X \to V$, is the $\mathcal{U}$-ultralimit of the sequence of maps $\frac{1}{r_i}(f - f(p)) : (X, p, \frac{1}{r_i} d) \to V$, if it exists (this is slightly abusive since the notation f_p does not reflect the dependence on r_i and $\mathcal{U}$). The $\mathcal{U}$-ultralimit exists if the limit exists in the usual sense or if V is finite dimensional. Also observe that if f_p exists and f is a biLipschitz embedding, then so is f_p (with distortion bounded by that of f).

Proposition 1.1. *Let (X, d) be a metric space, $\psi : X \to \mathbb{R}^k$ a Lipschitz map, and $p \in X$. Let $(V, \|\cdot\|)$ be a Banach space and $f : X \to V$ a Lipschitz map. If f is differentiable with respect to ψ at p with derivative*

$D_\psi f_p$, *then for every sequence* $(r_i)_{i=0}^\infty$ *decreasing to 0 and nonprincipal ultrafilter* $\mathcal{U}$ *on* $\mathbb{N}$, $f_p : T_p^{r_i,\mathcal{U}} X \to V$ *exists and* $f_p = D_\psi f_p \circ \psi_p$.

Proof. Assume f is differentiable with respect to ψ at p with derivative $D_\psi f_p$. Let $(r_i)_{i=0}^\infty$ and $\mathcal{U}$ be as above. Let $x \in T_p^{r_i,\mathcal{U}} X$. This means x is the equivalence class of a sequence $(x_i)_{i=0}^\infty \in X^\mathbb{N}$ with $\sup_i \frac{d(p,x_i)}{r_i} < \infty$, $[(x_i)_{i=0}^\infty]$. First assume f_p exists. Then

$$\|f_p(x) - D_\psi f_p \circ \psi_p(x)\| = \left\| \mathcal{U}\text{-}\lim_{i\to\infty} \frac{f(x_i) - f(p)}{r_i} - D_\psi f_p \left(\mathcal{U}\text{-}\lim_{i\to\infty} \frac{\psi(x_i) - \psi(p)}{r_i} \right) \right\|$$

$$= \mathcal{U}\text{-}\lim_{i\to\infty} \frac{\|f(x_i) - f(p) - D_\psi f_p\left(\psi(x_i) - \psi(p)\right)\|}{r_i}$$

$$\leq \left(\sup_i \frac{d(p,x_i)}{r_i} \right) \mathcal{U}\text{-}\lim_{i\to\infty} \frac{\|f(x_i) - f(p) - D_\psi f_p\left(\psi(x_i) - \psi(p)\right)\|}{d(p,x_i)} = 0$$

where the last equality holds since the usual topological limit exists by definition of derivative, and usual topological convergence implies $\mathcal{U}$-ultraconvergence. Since $x \in T_p^{r_i,\mathcal{U}} X$ was arbitrary, we get $f_p = D_\psi f_p \circ \psi_p$. This argument can also be turned around to prove that f_p exists. □

As far as the biLipschitz theory is concerned, the following theorem is the most fundamental.

Theorem 1.21 ([CK09], Theorem 1.6)**.** *Suppose* (X, d, μ) *is an RNP LDS with chart* $\psi : X \to \mathbb{R}^k$*. If there exists a Borel set* $E \subseteq X$ *such that* $\mu(E) > 0$ *and, for every* $p \in E$*, there exist a sequence* $(r_i)_{i=0}^\infty$ *decreasing to 0 and a nonprincipal ultrafilter* $\mathcal{U}$ *on* $\mathbb{N}$ *such that* $\psi_p : T_p^{r_i,\mathcal{U}} X \to \mathbb{R}^k$ *is not injective, then* X *does not biLipschitz embed into any RNP space.*

Proof. We proceed by contradiction. Assume there exists a Borel set $E \subseteq X$ such that $\mu(E) > 0$ and, for every $p \in E$, there exist a sequence $(r_i)_{i=0}^\infty$ decreasing to 0 and a nonprincipal ultrafilter $\mathcal{U}$ on $\mathbb{N}$ such that $\psi_p : T_p^{r_i,\mathcal{U}} X \to \mathbb{R}^k$ is not injective and that there is an RNP space V and a biLipschitz embedding $f : X \to V$. Then by definition of RNP LDS with chart ψ, f is differentiable with respect to ψ μ-almost everywhere. Then since $\mu(E) > 0$, there exists a point $p \in E$ such that f is differentiable at p. By the preceding proposition, f_p exists and $f_p = D_\psi f_p \circ \psi_p$. Since f is a biLipschitz embedding, so is $f_p = D_\psi f_p \circ \psi_p$. In particular, $D_\psi f_p \circ \psi_p$ is injective, contradicting ψ_p is not injective. □

Definition 1.26. A metric space (X, d) satisfies the *differentiability nonembeddability criterion* into RNP spaces if there exists a Borel measure μ so that (X, d, μ) satisfies the hypotheses of Theorem 1.21.

Example 1.8. The Heisenberg group (or any nonabelian Carnot group) satisfies the differentiability nonembeddability criterion. The existence of the dilations δ_t and that fact that closed balls in $\mathbb{H}$ are compact implies every tangent cone $T_p^{r_i,\mathcal{U}}\mathbb{H} = \mathbb{H}$ canonically and $\pi_p^{\text{ab}} : T_p^{r_i,\mathcal{U}}\mathbb{H} \to \mathbb{R}^2 = \pi_p^{\text{ab}} : \mathbb{H} \to \mathbb{R}^2$ canonically.

Example 1.9. Laakso space satisfies the differentiability nonembeddability criterion; we can take $E = G_\infty$.

Chapter 2

Thick Families of Geodesics and Differentiation

2.1 Introduction

This chapter is devoted to proving the results stated in Section 1.2.1.

2.1.1 Outline

We discuss our methods of proof in Section 2.1.2. Section 2.1.3 sets notation and terminology not already covered in Chapter 1.

Sections 2.2-2.6 are concerned with the proof of Theorem 1.14, Section 2.7 contains the proof of Corollary 2.1, and Section 2.8 contains the construction of the inverse limit of graphs in nonRNP Banach spaces from Theorem 2.10.

For an efficient reading of Sections 2.2-2.6, we advise the reader to start with Section 2.2, skip ahead to Section 2.6, and then refer back to the between sections as they are needed to understand the proof of Theorem 2.9.

In Section 2.2, we give the axioms for thick inverse systems of graphs whose inverse limit we are able to prove the weak form of differentiation of. Also included in this section are frequently used consequences of the axioms and a proof of one of the main theorems of the article, Theorem 2.1. This theorem asserts the existence of the thick inverse system of graphs in any metric space containing a thick family of geodesics. In Section 2.3, we define the set S_∞ and prove $\mu_\infty(S_\infty) > 0$. We also include results on asymptotic local geometry of the graphs. Section 2.4 covers the use of conditional expectation in approximating functions on X_∞ via functions on X_i. Also in this section is the definition of the derivative of RNP space-valued Lipschitz functions on X_∞. A relevant maximal operator and corresponding maximal inequality are defined and proved in Section 2.5. Section 2.6 contains the proof of the main theorem, Theorem 2.9, the weak form of differentiability.

2.1.2 Discussion of Proof Methods

The subset X_∞ of a metric space M containing a thick family of geodesics from Theorem 1.14 is constructed as an inverse limit of graphs. Cheeger and Kleiner proved in [CK15] that inverse limits of certain "admissible" inverse systems of graphs, such as Laakso spaces, are PI spaces and hence RNP Lipschitz differentiability spaces. It is this result which lead us to believe that X_∞ could be constructed to satisfy some kind of RNP Lipschitz differentiability. However, our space X_∞ cannot be constructed to be a PI space in any obvious way, and thus the theory of [CK09] does not apply; we are required to construct derivatives of RNP-valued Lipschitz functions and prove their defining approximation property by hand. To do so, we use only the almost sure differentiability of Lipschitz maps $\mathbb{R} \to B$ and the almost sure convergence of B-valued martingales for RNP spaces, which are quite classical compared to the asymptotic norming property of RNP spaces used in [CK09]. We also make heavy use of the uniform topology on Banach spaces of Lipschitz functions, in contrast to the Sobolev space techniques employed in [CK09] and [CK15].

Apart from these differences in proof techniques, the inverse systems of graphs we consider are fundamentally different from the admissible systems in [CK15] for two reasons. Firstly, in [CK15], the graphs are equipped with geodesic metrics, and the metrics on our graphs are only geodesic along directed edge paths. In fact, the inverse limit space need not even be quasiconvex, while PI spaces are always quasiconvex. Secondly, in [CK15], the lengths of edges in the sequence of graphs decrease by a constant factor $m \geq 2$ in each stage of the sequence, independent of the stage or edge. In our graphs, the edge lengths decrease by factors going to ∞. We make frequent use of this rapid decay in a number of independent results, such as (2.10), (2.11), and Lemma 2.3. Loosely, the rapid decay in edge length allows us to well-control the local geometry near a point along scales proportional to the lengths of edges containing the projections of the point, at the cost of control over the geometry along other scales, which would be necessary to prove true RNP differentiability.

The uniform topology on Lipschitz algebras has been studied before within the context of Lipschitz differentiability spaces. For, example, in [Sch14], Schioppa showed how to associate a Weaver derivation (which involves continuity with respect to uniform topology) to an Alberti representation, and Alberti representations were demonstrated by Bate in [Bat15] to be intimately connected to Lipschitz differentiability. Schioppa constructs the partial derivative of a function by taking its derivative along curve fragments and averaging them together with respect to the Alberti representation. Our procedure for constructing the derivative of a function (see Theorem 2.6), is very similar in nature; indeed, Lemma 2.6 gives Alberti representations of μ_i, which (after taking a suitable limit) give rise to an Alberti representation of μ_∞. We also note that in [Bat15], Bate gives necessary and sufficient conditions for a collection of Alberti representations to induce a Lipschitz differentiable structure on a metric measure space using what he called *universality* (see Definition 7.1 from [Bat15]). Our representation from Lemma 2.6 will generally fail this property (or at least doesn't obviously satisfy it - we don't actually provide an example), which is consistent with our discussion that the space $(X_\infty, d, \mu_\infty)$ is not a true Lipschitz differentiability space (again, we don't actually provide an example of this). We believe it is possible to find a weakened form of universality corresponding to the weakened form of differentiation from Theorem 2.9.

The construction of the inverse limit of admissible graphs, X_∞, of Theorem 2.10 is achieved by fine tuning two of the aspects of Ostrovskii's construction of a thick family of geodesics in nonRNP spaces. His construction is also essentially an inverse limit of a system of graphs, but the system is not "admissible" in the sense of [CK15] for two reasons. Firstly, the metrics on his system are not uniformly quasiconvex, which is a necessary condition for the inverse limit metric space to be a PI space. Secondly, the lengths of edges in a graph in an admissible system must be constant, but in the system of [Ost14b], the ratio of lengths of two edges in a graph may become unbounded.

The second obstacle is easily overcome in the following way: the length of an edge in a graph in the system from [Ost14b] corresponds to the coefficient α_i of some convex combination $z = \alpha_1 z_1 + \dots \alpha_n z_n$ with $\|z - z_i\| > \delta$ and $\|z\|, \|z_i\| < 1$. By density of the dyadic rationals in $(0,1)$, we may make small adjustments $z_i \to z_i'$ to obtain $z = q_1 z_1' + \dots q_n z_n'$ with each q_i a dyadic rational, all while maintaining $\|z - z_i'\| > \delta$ and $\|z\|, \|z_i'\| < 1$. We then 'split up' the convex combination into terms whose coefficients have numerator equal to 1. For example, $\frac{1}{2}z_1' + \frac{1}{4}z_2' + \frac{1}{4}z_3' \to \frac{1}{4}z_1' + \frac{1}{4}z_1' + \frac{1}{4}z_2' + \frac{1}{4}z_3'$. The edges corresponding to this convex combination now all have length $\frac{1}{4}$. The first obstacle can be overcome by constructed X_i with rapidly decreasing edge length, similar to construction in the proof of Theorem 2.1. Using the rapid decrease in edge length to control the quasiconvexity of the graphs is similar to the proof of Lemma 2.2.

2.1.3 Notation and Terminology

Given two metric spaces (X, d_X) and (Y, d_Y) and a Lipschitz map $f : X \to Y$, we define $\mathrm{Lip}\,(f) := \sup_{x \neq y} \frac{d_Y(f(x), f(y))}{d_X(x,y)}$. For a metric space (X, d) with basepoint x_0, define $\mathrm{Lip}_0(X; B)$ to be the Banach space of Lipschitz functions $f : X \to B$ satisfying $f(x_0) = 0$, equipped with the norm $\|f\|_{\mathrm{Lip}_0(X;B)} := \mathrm{Lip}(f)$. When $B = \mathbb{R}$, we simply recover the Lipschitz free space from Definition 1.17. Note that, when $\mathrm{diam}(X) \leq 1$, $\|f\|_{\mathrm{Lip}_0(X;B)} \leq \|f\|_{L^\infty(X)}$ (we shall generally find ourselves in this situation).

A finite, metric *graph* (or just graph) is a metric space X equipped with a finite set of vertices, $V(X)$, and a finite set of edges, $E(X)$, satisfying some properties.

- $V(X) \subseteq X$, and $E(X) \subseteq \mathcal{P}(X)$, the power set of X.

- Each $e \in E(X)$ is isometric to a compact interval $[a, b]$, and under any isometry $[a, b] \to e$, a and b get mapped to vertices, called the vertices of e, and no other point $c \in (a, b)$ gets mapped to a vertex.

- If $e_1, e_2 \in E(X)$ with $e_1 \neq e_2$, then $e_1 \cap e_2$ is empty, or $e_1 \cap e_2$ consists of one or two vertices.

The graph is *directed* if each edge is equipped with a direction, which is simply an ordering of its two vertices. The first vertex is called the *source*, and the second is called the *sink*. We say that the edge is directed from the source to the sink.

If A is a Borel subset of a finite graph, $|A|$ denotes its length measure. If x, y are points in a finite graph, $|x - y|$ denotes the distance between x and y with respect to the length metric, the metric given by the infimal length of paths between x and y. A length minimizing path from x to y will be denoted $[x, y]$ (so that $|x - y| = |[x, y]|$), and is frequently referred to as a *shortest path*. Since shortest paths need not be unique, the notation "$[x, y]$" does not unambiguously define one set, but it should be clear from context what is being referred to. In any case, as far as this article is concerned, the nonuniqueness of shortest paths don't pose any problems.

2.2 Inverse Systems of Nested Graphs

We begin this section by listing some axioms for a "thick inverse system" of nested metric graphs, see Definition 2.1. We introduce thick inverse systems for two reasons: one - we are able to prove our differentiation theorem, Theorem 1.14, for the inverse limit of these systems, and two - we are able to prove that a thick inverse system can be found in any metric space containing a thick family of geodesics, see Theorem 2.1.

2.2.1 Axioms and Terminology

Definition 2.1. An inverse system of nested metric measure directed graphs satisfying the following Axioms (A1) - (A6) and equipped with the measure from Definition 2.2 will be called a **thick inverse system**.

We use the notation $(X_0, d, \mu_0) \overset{\leftarrow}{\subseteq} (X_1, d, \mu_1) \overset{\leftarrow}{\subseteq} \dots$ for a system of nested metric directed graphs. The maps $X_{i+1} \to X_i$ are denoted π_i^{i+1}. Let $i \geq 0$ and $j \geq i$, and define $\boldsymbol{\pi_i^j} := \pi_i^{i+1} \circ \pi_{i+1}^{i+2} \circ \dots \pi_{j-1}^j : X_j \to X_i$.

Graph and Length Axioms:

(A1) X_0 has two vertices, denoted 0 and 1, and one edge directed from 0 to 1, with length 1. We identify X_0 with $I := [0, 1]$.

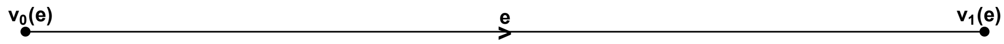

Figure 2.1: A directed edge e of X_i, shown in black.

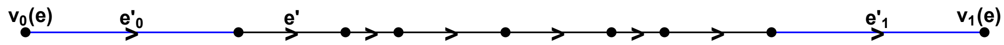

Figure 2.2: The directed subdivision of e in X_i'. Terminal subedges e_0' and e_1' are shown in blue, and nonterminal subedges are shown in black.

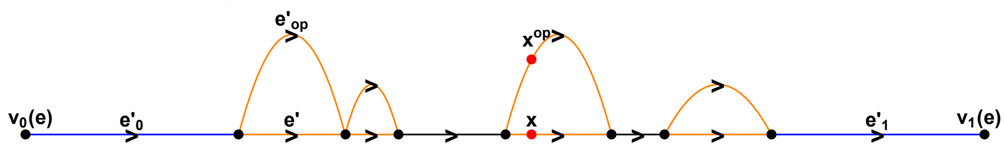

Figure 2.3: The set $(\pi_i^{i+1})^{-1}(e)$ in X_{i+1}. Terminal intervals are shown in blue, nonterminal intervals are shown in black, and circles are shown in orange. Examples of subedges e' and opposite subedge e'_{op} are labeled, as are example point and opposite point x and x^{op}, shown in red.

(A2) There is a directed subdivision of X_i, denoted X_i', satisfying the properties below. It will be helpful to refer to Figures 2.1, 2.2, and 2.3 while reading (A2).

(i) For each edge $e' \in E(X_i')$, $(\pi_i^{i+1})^{-1}(e') = e' \cup e'_{\mathrm{op}}$, where either $e'_{\mathrm{op}} = e'$, or e'_{op} is an edge having the same source and sink vertices as e', but whose interior is disjoint from the rest of X_{i+1}. The edge e'_{op} is called the **opposite edge** of e' in X_{i+1} (we may write e_{op} or e^{op} depending on the presence of other super or subscripts). We also define $(e'_{\mathrm{op}})_{\mathrm{op}} := e'$.

For future use, we note that, with respect to the length metric, the diameter of $(\pi_i^{i+1})^{-1}(e')$ equals $|e'|$. Thus, with respect to d,

$$\mathrm{diam}((\pi_i^{i+1})^{-1}(e')) \leq |e'| \tag{2.1}$$

Given a point $x \in e'$, we similarly define x^{op} to be the unique point of e'_{op} for which $\pi_i^{i+1}(x^{\mathrm{op}}) = x$, and call x^{op} the **opposite point** of x (in X_{i+1}). (If $e'_{\mathrm{op}} = e'$, then $x^{\mathrm{op}} = x$. We may also write x_{op} depending on the presence of other super or subscripts.) Again, we also define $(x^{\mathrm{op}})^{\mathrm{op}} := x$. If $e'_{\mathrm{op}} = e'$, we call $(\pi_i^{i+1})^{-1}(e')$ an **interval** (it is an interval topologically). If $e'_{\mathrm{op}} \neq e'$, we call $(\pi_i^{i+1})^{-1}(e')$ a **circle** (it is a circle topologically).

(ii) If e_0' and e_1' are terminal edges in the subdivision of some edge $e \in E(X_i)$ (meaning they share a vertex with e), then $(\pi_i^{i+1})^{-1}(e_0')$ and $(\pi_i^{i+1})^{-1}(e_1')$ are intervals (so not circles). We refer to these edges of X_i' and X_{i+1} as **terminal intervals** (sometimes terminal edges) of X_{i+1} and also as **terminal subintervals** (sometimes terminal subedges) of e. We note that a subedge $e' \in E(X_i')$ of e is not a terminal subinterval if and only if it is contained in the interior of e.

Metric Axioms:

(A3) For any $i \geq 0$, d is geodesic when restricted to any directed edge path of X_i, meaning there is an isometry from a compact interval to this edge path.

(A4) $\pi_i^{i+1} : X_{i+1} \to X_i'$ acts identically on any $e' \in E(X_i') \subseteq E(X_{i+1})$, and it collapses any $e'_{\text{op}} \in E(X_{i+1}) \setminus E(X_i')$ isometrically onto e'.

Thickness Axiom:

Suppose $e' \in E(X_i')$ is an edge such that $(\pi_i^{i+1})^{-1}(e')$ is a circle. For any $t \in e'$, let $t^{\text{op}} \in e'_{\text{op}}$ denote the opposite point. Define the **height** of e' by $\mathbf{ht(e')} := \max_{p \in e'} d(p, p^{\text{op}})$ (the height is between 0 and $|e'|$ and is a measure of how close the circle is to being to a standard circle; it equals $|e'|$ if and only if the circle is isometric to a standard circle of diameter $|e'|$).

(A5) There is a constant $\alpha > 0$ (independent of i) such that $\operatorname{ht}(e') \geq \alpha |e'|$ for every $e' \in E(X_i')$.

(A6) Let P be a directed edge path from 0 to 1 in X_i', and let $E_{\text{circ}}(P)$ denote the set of edges $e' \subseteq P$ along the path for which $(\pi_i^{i+1})^{-1}(e')$ is a circle. Then there is a constant $\beta > 0$ (independent of i and P) such that $|\cup E_{\text{circ}}(P)| \geq \beta$.

Measure Definition:

Definition 2.2. Define $\boldsymbol{(\mu_i)_{i=0}^{\infty}}$ to be the unique sequence of probability measures satisfying the following recursion: μ_0 is length (Lebesgue) measure on $X_0 = [0, 1]$. Restricted to any edge of X_{i+1}, μ_{i+1} is a constant multiple of length measure and for any $e' \in E(X_i')$, $\mu_{i+1}(e') = \mu_i(e')$ if $e'_{\text{op}} = e'$, and $\mu_{i+1}(e') = \mu_{i+1}(e'_{\text{op}}) = \frac{1}{2}\mu_i(e')$ if $e'_{\text{op}} \neq e'$.

2.2.2 Elementary Consequences of Axioms

Throughout this subsection, fix a thick inverse system, using the same notation as in the previous subsection. We begin with a proposition that lists, without proof, some elementary consequences of the axioms. We use these facts often and without mention. Then we prove some less immediate facts about the metric structure that will be needed for subsequent results.

Proposition 2.1. *The following are true:*

- *The map $\pi_i^{i+1} : X_{i+1} \to X_i$ is a projection onto $X_i \subseteq X_{i+1}$; $\pi_i^{i+1}\big|_{X_i} = id_{X_i}$.*

- *π_i^{i+1} is direction preserving, and by induction the same is true for π_i^j, $j \geq i$. Thus, π_i^j restricted to any directed edge path is an isometry.*

- *The restriction of μ_i to any interval or circle of X_i is a constant multiple (with constant ≤ 1) of length measure.*

- *$(\pi_i^{i+1})_\#(\mu_{i+1}) = \mu_i$.*

Definition 2.3. For any $i \geq 0$, define $\Delta_i^E := \min_{e \in E(X_i)} |e|$ and
$\boldsymbol{\delta_i^E} := \max_{e' \in E(X_i')} \frac{|e'|}{\Delta_i^E}$. The maximum is well-defined because each graph has finitely many edges.

Definition 2.4. For any $j \geq i \geq 0$ and $x \in X_j$, define $\boldsymbol{e_i(x)}$ and $\boldsymbol{e_i'(x)}$ to be edges of X_i and X_i', respectively, containing $\pi_i^j(x)$. These edges are unique except when x is a vertex, and the set of vertices form a measure 0 set.

Lemma 2.1. *If $(\delta_i')_{i=0}^\infty$ is a positive decreasing sequence with $\delta_0' \leq \frac{1}{2}$, and if $\delta_i^E \leq \delta_i'$, then for any $j \geq i \geq 0$ and $x_i \in X_i$, $diam((\pi_i^j)^{-1}(x_i)) \leq 2\delta_i'|e_i(x_i)|$.*

Proof. Assume δ_i', δ_i^E, and x_i are as above. Let $x_j \in (\pi_i^j)^{-1}(x_i)$, and for $k = i \ldots j$, set $x_k := \pi_k^j(x_j)$. By (2.1), $d(x_k, x_{k+1}) \leq |e_k'(x_k)|$. By a repeated application of the definition of δ_i^E, we have $|e_k'(x_k)| \leq \delta_i^E \cdot \delta_{i+1}^E \cdot \ldots \delta_k^E |e_i(x_i)| \leq (\delta_i')^{k+1-i}|e_i(x_i)|$, where the least inequality holds since $\delta_k^E \leq \delta_k'$ and δ_k' is decreasing. Then we have $d(x_i, x_j) \leq \sum_{k=i}^{j-1} d(x_k, x_{k+1}) \leq \left(\sum_{k=i}^{j-1} (\delta_i')^{k+1-i}\right) |e_i(x_i)| \leq 2\delta_i'|e_i(x_i)|$, where the last inequality holds since $\delta_i' \leq \delta_0' \leq \frac{1}{2}$. $\square$

Definition 2.5. For any $i \geq 0$, $e \in E(X_i)$ and $e' \in E(X_i')$ with e' a nonterminal subedge of e, define $\Delta_i^d(e, e') := d(e', X_i \setminus e)$. This is positive by compactness and since e' belongs to the interior of e (since it is nonterminal). Define $\Delta_i^d(e)$ to be the minimum of $\Delta_i^d(e, e')$ over all nonterminal $e' \subseteq e$, and define Δ_i^d to be the minimum of $\Delta_i^d(e)$ over all $e \in E(X_i)$. Define $\boldsymbol{\delta_i^d} := \max_{e'} \frac{|e'|}{\Delta_i^d}$, where the max is over all nonterminal edges $e' \in E(X_i')$.

Lemma 2.2. *If $\delta_i^d < \frac{1}{2}$ and $\Pi_{i=0}^\infty \frac{1}{1-2\delta_i^d} \leq L$, then $\mathrm{Lip}\left(\pi_i^j\right) \leq \Pi_{k=i}^{j-1} \mathrm{Lip}\left(\pi_k^{k+1}\right) \leq \Pi_{i=0}^\infty \frac{1}{1-2\delta_i^d} \leq L$ for any $j \geq i \geq 0$.*

Proof. It suffices to prove $\mathrm{Lip}\left(\pi_k^{k+1}\right) \leq \frac{1}{1-2\delta_k^d}$. Let $x_{k+1}, y_{k+1} \in X_{k+1}$, and set $x_k := \pi_k^{k+1}(x_{k+1})$, $y_k := \pi_k^{k+1}(y_{k+1})$. We need to show that $d(x_{k+1}, y_{k+1}) \geq (1 - 2\delta_k^d) d(x_k, y_k)$. We consider two cases; either x_k and y_k belong to the same edge of X_k, or they belong to different edges. Assume they belong to the same edge. Then there are again two cases; either x_{k+1} and y_{k+1} belong to opposite edges of a circle, or they belong to a directed edge path. The conclusion holds in this second case since the map π_k^{k+1} is an isometry on directed edges paths (so we get an ever better bound of 1). Now suppose they belong to opposite edges of a circle. Without loss of generality, assume $y_{k+1} \in e'$ and $x_{k+1} \in e'_{\text{op}}$ for some $e' \in E(X'_k)$. Then $y_{k+1} = y_k$, and a shortest path between them, $[x_{k+1}, y_{k+1}]$, passes through one of the vertices of the circle, say v. Then since x_k and v belong to an edge, $d(x_k, v) = |v - x_k|$ (recall that $|p - q|$ denotes the distance with respect to the length metric), and since v and y_k belong to an edge, so $d(v, y_k) = |y_k - v|$. Without loss of generality, assume $|v - x_k| \leq |y_k - v|$. This implies $x_k \in [v, y_k]$, in turn implying $|x_k - v| + |y_k - x_k| = |y_k - v|$. Then we have

$$d(x_{k+1}, y_{k+1}) \geq d(v, y_{k+1}) - d(x_{k+1}, v) = d(v, y_k) - d(v, x_k)$$

$$= |y_k - v| - |x_k - v| = |y_k - x_k| = d(y_k, x_k)$$

Our conclusion holds in this case (again with an ever better bound of 1).

Finally, assume that x_k and y_k do not belong to the same edge of X_k. We consider three cases now: both points belong to a terminal interval of X_k, neither point does, or one does and the other does not. Our conclusion holds in the first case, since π_k^{k+1} acts identically on X_k (so $y_{k+1} = y_k$ and $x_{k+1} = x_k$), and terminal intervals belong to X_k by definition. Assume the second case holds. The edges $e'_k(x_k)$ and $e'_k(y_k)$ are nonterminal by assumption. Then by definition of δ_k^d, since y_k and x_k do not belong to the same edge of X_k, $|e'_k(y_k)|, |e'_k(x_k)| \leq \delta_k^d d(x_k, y_k)$. Then we have

$$d(x_{k+1}, y_{k+1}) \geq d(x_k, y_k) - d(x_{k+1}, x_k) - d(y_{k+1}, y_k) \geq d(x_k, y_k) - |e'_k(x_k)| - |e'_k(y_k)|$$

$$\geq d(x_k, y_k) - 2\delta_k^d d(x_k, y_k) = (1 - 2\delta_k^d) d(x_k, y_k)$$

And our desired conclusion holds in this case. For the third and final case, assume without loss of generality that y_k belongs to a terminal interval and x_k does not. Then we get $y_{k+1} = y_k$ and $|e'_k(x_k)| \leq \delta_k^d d(x_k, y_k)$. Making the obvious adjustments to the argument above yields

$$d(x_{k+1}, y_{k+1}) = d(x_{k+1}, y_k) \geq d(x_k, y_k) - d(x_{k+1}, x_k) \geq d(x_k, y_k) - |e'_k(x_k)|$$

$$\geq d(x_k, y_k) - \delta_k^d d(x_k, y_k) = (1 - \delta_k^d) d(x_k, y_k)$$

□

Lemma 2.3. *If $x_{k+1}, y_{k+1} \in X_{k+1}$ do not belong to opposite open edges of a circle, then* $d(\pi_k^{k+1}(x_{k+1})), \pi_k^{k+1}(y_{k+1})) \geq \frac{1}{1+2\delta_k^d} d(x_{k+1}, y_{k+1})$ *(loosely, π_k^{k+1} collapses circles, but is close to an isometry* *away from them).*

Proof. Let $x_{k+1}, y_{k+1} \in X_{k+1}$, and set $x_k := \pi_k^{k+1}(x_{k+1})$, $y_k := \pi_k^{k+1}(y_{k+1})$. As before, there are two cases; either x_k and y_k belong to the same edge of X_k, or they belong to different edges. Assume they belong to the same edge. Again, as before, there are two cases; either x_{k+1} and y_{k+1} belong to opposite edges of a circle, or they belong to a directed edge path. The first case doesn't hold by assumption, and the conclusion holds in this second case since the map π_k^{k+1} is an isometry on directed edges paths (so we get an ever better bound of 1).

Finally, assume that x_k and y_k do not belong to the same edge of X_k. As before, three cases: both points belong to a terminal interval of X_k, neither point does, or one does and the other does not. Our conclusion holds the first case, since π_k^{k+1} acts identically on X_k (so $y_{k+1} = y_k$ and $x_{k+1} = x_k$, and intervals belong to X_k by definition. Assume the second case holds. The edges $e'_k(x_k)$ and $e'_k(y_k)$ are nonterminal by assumption. Then by definition of δ_k^d, since y_k and x_k do not belong to the same edge of X_k, $|e'_k(y_k)|, |e'_k(x_k)| \leq \delta_k^d d(x_k, y_k)$. Then we have

$$d(x_{k+1}, y_{k+1}) \leq d(x_k, x_{k+1}) + d(x_k, y_k) + d(y_k, y_{k+1}) \leq |e'_k(x_k)| + d(x_k, y_k) + |e'_k(y_k)| \leq (1 + 2\delta_k^d) d(x_k, y_k)$$

And our desired conclusion holds in this case. For the third and final case, assume without loss of generality that y_k belongs to a terminal interval and x_k does not. Then we get $y_{k+1} = y_k$ and $|e'_k(x_k)| \leq \delta_k^d d(x_k, y_k)$. Making the obvious adjustments to the argument above yields

$$d(x_{k+1}, y_{k+1}) = d(x_{k+1}, y_k) \leq d(x_k, x_{k+1}) + d(x_k, y_k) \leq |e'_k(x_k)| + d(x_k, y_k) \leq (1 + \delta_k^d) d(x_k, y_k)$$

□

Remark 2.1. Note that since $\frac{1}{1-2\delta} > 1 + \delta$, if the hypotheses of Lemma 2.3 are satisfied, then

$$\Pi_{k=0}^{\infty}(1 + \delta_k^d) \leq L \tag{2.2}$$

2.2.3 Existence of Inverse System

Let M be a metric space.

Theorem 2.1. *If M contains a thick family of geodesics, then for any positive sequence $(\delta'_i)_{i=0}^{\infty}$, M contains a thick inverse system with $\delta_i^E, \delta_i^d \leq \delta'_i$ for every i (see Definitions 2.1, 2.3, and 2.5).*

Proof. Assume M contains an α'-thick family of geodesics Γ for some $\alpha' > 0$. Let $(\delta'_i)_{i=0}^{\infty}$ be a positive sequence. We'll construct the inverse sequence $X_0 \overleftarrow{\subseteq} X_1 \overleftarrow{\subseteq} \ldots$ inductively. Let γ be any element of Γ, and set X_0 equal to the image of γ in M. Equip X_0 with the necessary graph structure. Assume X_i, X_{i-1}, and π^i_{i-1} have been constructed for some $i \geq 0$, satisfy the Graph, Metric, and Thickness Axioms, and also satisfy the additional hypothesis that the geodesic parametrization of each directed 0-1 edge path belongs to Γ. For each edge $e \in E(X_i)$, let $v_0(e)$ and $v_1(e)$ denote the source and sink vertices of e, respectively. The edge e is mapped isometrically onto $I_e := [d(0, v_0(e)), d(v_1(e), 1)]$ via π^i_0. Denote the inverse of this map $\gamma_e : I_e \to e \subseteq X_i$. Note that, for any geodesic parametrization γ of a 0-1 edge path whose image contains e, we must have $\gamma\big|_{I_e} = \gamma_e$, so γ_e extends to a geodesic parametrization of a directed 0-1 edge path.

Now we provide a more quantitative reformulation of Definition 1.6. By a *partition* T of an interval $[a, b]$, we mean a finite subset of $[a, b]$ equipped with the order induced from $[a, b]$, such that the least element is a and the greatest element is b. For any $t \in T$ other than b, we define t^+ to be the immediate successor of t, and we simply define $b^+ := b$, and for any $t \in T$ other than a, we define t^- to be the immediate predecessor of t, and we simply define $a^- := a$. For each partition T^e of I_e, and $\tilde{\gamma}^e \in \Gamma$ with $\tilde{\gamma}^e\big|_{T^e} \equiv \gamma_e\big|_{T^e}$ define the *deviations* of $(T^e, \tilde{\gamma}^e)$ and T^e, respectively:

$$\mathrm{dev}\,(T^e, \tilde{\gamma}^e) := \sum_{t \in T^e} \max_{s \in [t, t^+]} d(\gamma_e(s), \tilde{\gamma}^e(s))$$

$$\mathrm{dev}\,(T^e) := \sup_{\substack{\tilde{\gamma}^e \in \Gamma \\ \tilde{\gamma}^e|_{T^e} \equiv \gamma_e|_{T^e}}} \mathrm{dev}\,(T^e, \tilde{\gamma}^e)$$

Note that $\mathrm{dev}\,(T^e) \leq |e|$.

For any fixed partition T^e of I_e, let $T^e_{sup/2}$ and $\tilde{\gamma}^e_{sup/2}$ denote a partition of I_e and a geodesic in Γ, respectively, with

$$T^e_{sup/2} \supseteq T^e \tag{2.3}$$

$$\tilde{\gamma}^e_{sup/2}\big|_{T^e_{sup/2}} \equiv \gamma_e\big|_{T^e_{sup/2}} \tag{2.4}$$

$$\mathrm{dev}\left(T^e_{sup/2}, \tilde{\gamma}^e_{sup/2}\right) \geq \frac{1}{2} \sup_{T' \supseteq T^e} \mathrm{dev}\,(T') \tag{2.5}$$

Now, we can always choose $T^e_{sup/2}$ and $\tilde{\gamma}^e_{sup/2}$ such that the above properties remain true, and also such that for every $t \in T^e_{sup/2}$,

$$\gamma_e\big|_{[t,t^+]} \equiv \tilde{\gamma}^e_{sup/2}\big|_{[t,t^+]} \quad \text{or} \quad \gamma_e((t, t^+)) \cap \tilde{\gamma}^e_{sup/2}((t, t^+)) = \emptyset \tag{2.6}$$

To see this, take any $T^e_{sup/2}$ and $\tilde{\gamma}^e_{sup/2}$ as above, and let $t \in T^e_{sup/2}$. If

$\max_{s\in[t,t^+]} d(\gamma_e(s),\tilde{\gamma}^e(s)) = 0$, then γ_e and $\tilde{\gamma}^e_{sup/2}$ agree on all of $[t,t^+]$ and we are done. Otherwise, let $s_{max} = \text{argmax}_{s\in[t,t^+]} d(\gamma_e(s),\tilde{\gamma}^e(s))$. Then by continuity, there exists a largest, nonempty open subinterval (a,b) of $[t,t^+]$ containing s_{max} such that $\gamma_e((a,b))\cap\tilde{\gamma}^e_{sup/2}((a,b)) = \emptyset$. Since it is the largest, $\gamma_e(a) = \tilde{\gamma}^e_{sup/2}(a)$ and $\gamma_e(b) = \tilde{\gamma}^e_{sup/2}(b)$. We add these new points a and b to the partition $T^e_{sup/2}$, and modify $\tilde{\gamma}^e_{sup/2}$ so that it agrees with γ_e on $[t,a]\cup[b,t^+]$, and remains unchanged on $[a,b]$. This new curve still belongs to Γ because Γ is concatenation closed. It is clear that (2.3), (2.4), and (2.5) remain valid, and that we gain (2.6).

We use the partition $T^e_{sup/2}$ of I_e to subdivide e into smaller edges by taking the image of $T^e_{sup/2}$ under γ_e to be new vertices. Each new subedge equals $\gamma_e([t,t^+])$ for a unique $t \in T^e_{sup/2}$. Denote this edge e^t, and recall the *height* of e^t, defined in the Thickness Axioms,

$$\text{ht}\left(e^t\right) = \max_{s\in[t,t^+]} d(\gamma_e(s),\tilde{\gamma}^e_{sup/2}(s))$$

Set $\alpha := \frac{\alpha'}{4}$, and split the new subedges of e up into two groups, $E^e_{<\alpha}$ and $E^e_{\geq\alpha}$, where $e^t \subseteq e$ belongs to $E^e_{<\alpha}$ if $\text{ht}\,(e^t) < \alpha|e^t|$ and e^t belongs to $E^e_{\geq\alpha}$ if $\text{ht}\,(e^t) \geq \alpha|e^t|$. Name the collection of corresponding time intervals $(T^e_{sup/2})_{<\alpha}$ and $(T^e_{sup/2})_{\geq\alpha}$.

It follows from Definition 1.6 and the observation that γ^e extends to a geodesic in Γ, that for any 0-1 directed edge path P, and any choice of partition T^e for each $e \subseteq P$,

$$\sum_{e\subseteq P} \text{dev}\left(T^e_{sup/2},\tilde{\gamma}^e_{sup/2}\right) \geq \frac{\alpha'}{2}$$

It follows from this that

$$\frac{\alpha'}{2} \leq \sum_{e\subseteq P}\left(\sum_{e^t\in E^e_{<\alpha}} \text{ht}\left(e^t\right) + \sum_{e^t\in E^e_{\geq\alpha}} \text{ht}\left(e^t\right)\right) \leq \sum_{e\subseteq P}\left(\sum_{e^t\in E^e_{<\alpha}} \alpha|e^t| + \sum_{e^t\in E^e_{\geq\alpha}} |e^t|\right)$$

$$\leq \sum_{e\subseteq P}\left(\alpha|e| + \sum_{e^t\in E^e_{\geq\alpha}} |e^t|\right) = \alpha + \sum_{e\subseteq P}\sum_{e^t\in E^e_{\geq\alpha}} |e^t|$$

implying

$$|\cup E^P_{\geq\alpha}| \geq 2\beta \tag{2.7}$$

where $\beta := \frac{\alpha'}{8}$ and $E^P_{\geq\alpha} = \cup_{e\subseteq P} E^e_{\geq\alpha}$. Now that the preliminaries have been established, we are ready to choose a specific partition of e and apply the above results.

Set $\Delta^E_i := \min_{e\in E(X_i)} |e|$, and for each $e \in E(X_i)$, subdivide e into three edges $e'_0 < e_{mid} < e'_1$ such that $|e'_0| = \min(\frac{\beta}{2}|e|, \delta'_i\Delta^E_i) = |e'_1|$. Set $\Delta^d_i(e) := d(e_{mid}, X_i \setminus e)$. Since e_{mid} belongs to the interior of e, compactness gives us $\Delta_i^d(e) > 0$. Then set $\Delta_i^d := \min_{e \in E(X_i)} \Delta_i^d(e)$ and $\epsilon_i := \min(\delta_i' \Delta_i^E, \delta_i' \Delta_i^d)$. Now, for each $e \in E(X_i)$, choose a partition T^e of $I_e = [a,b] = [d(0, v_0(e)), d(v_1(e), 1)]$ such that

$$a^+ - a = \min\left(\frac{\beta}{2}|e|, \delta_i' \Delta_i^E\right) = b - b^- \tag{2.8}$$

(this implies $\gamma_e([a, a^+]) = e_0'$, $\gamma_e([b^-, b]) = e_1'$, and $\gamma_e([t, t^+]) \subseteq e_{mid}$ for $t \in T^e \setminus \{a, b^-, b\}$) and for any $t \in T^e \setminus \{a, b^-, b\}$

$$t^+ - t \leq \epsilon_i \tag{2.9}$$

For each $e \in E(X_i)$, fix $T^e_{sup/2} \supseteq T^e$ and $\tilde{\gamma}^e_{sup/2}$ as before. As explained in the previous paragraph, $T^e_{sup/2}$ induces a subdivision of e. Doing this for each e gives us the total subdivided graph X_i'. By (2.8) and (2.9), any subedge $e^t \subseteq e$ satisfies $|e^t| \leq \delta_i' \Delta_i^E$, so $\delta_i^E \leq \delta_i'$, as required. Furthermore, any nonterminal subedge e^t of e is contained in e_{mid}, by definition, and so by (2.9) we get $|e^t| \leq \epsilon_i \leq \delta_i' \Delta_i^d$, implying $\delta_i^d \leq \delta_i'$, as required.

It remains to construct X_{i+1} and π_i^{i+1}. We explain how to use segments of the curve $\tilde{\gamma}^e_{sup/2}$ as new edges to add to our graph X_i' to obtain X_{i+1}. Let $e \in E(X_i)$. There are three options for a subedge $e' \in E(X_i')$ of e: e' is a terminal subedge, (meaning $e' = e_0' = e^t$ or $e_1'^t$ for $t \in \{d(0, v_0(e)), d(v_1(e), 1)^-\}$), $e' = e^t$ for some $t \in (T^e_{sup/2})_{\geq \alpha} \setminus \{d(0, v_0(e)), d(v_1(e), 1)^-\}$ (meaning $\operatorname{ht}(e^t) \geq \alpha |e^t|$), or $e' = e^t$ for some $t \in (T^e_{sup/2})_{<\alpha} \setminus \{d(0, v_0(e)), d(v_1(e), 1)^-\}$ (meaning $\operatorname{ht}(e^t) < \alpha |e^t|$). In the first two cases, we set $e'_{\text{op}} = e'$, so that $(\pi_i^{i+1})^{-1}(e')$ is a circle, and in the third case, set $e'_{\text{op}} = e^t_{\text{op}} := \tilde{\gamma}^e_{sup/2}([t, t^+])$, so that the intersection of the interiors of e^t and e^t_{op} is empty, $(\pi_i^{i+1})^{-1}(e')$ is a circle, and $\operatorname{ht}(e^t) \geq \alpha |e^t|$. We define π_i^{i+1} in the unique way so that (A4) holds. It is clear that the Graph Axioms, Metric Axioms, and (A5) hold. Our additional hypothesis that the geodesic parametrization of every 0-1 directed edge path belongs to Γ also holds (again using concatenation closed). It remains to verify Axiom (A6).

To verify (A6), we fix a path P and compute $|\cup E_{\text{circ}}(P)|$. For each $e \subseteq P$, set $E_{\text{circ}}(e) = \{e' \in E_{\text{circ}}(P) : e' \subseteq e\}$. Then by (2.7) and (2.8),

$$|\cup E_{\text{circ}}(P)| = \sum_{e \subseteq P} |\cup E_{\text{circ}}(e)| = \sum_{e \subseteq P} \left(|\cup E^e_{\geq\alpha}| - |(\cup E^e_{\geq\alpha}) \cap (e_0' \cup e_1')|\right) \geq \sum_{e \subseteq P} \left(|\cup E^e_{\geq\alpha}| - |e_0' \cup e_1'|\right)$$

$$\overset{(2.8)}{\geq} \sum_{e \subseteq P} \left(|\cup E^e_{\geq\alpha}| - \left(\frac{\beta}{2}|e| + \frac{\beta}{2}|e|\right)\right) = |\cup E^P_{\geq\alpha}| - \beta \overset{(2.7)}{\geq} 2\beta - \beta = \beta$$

□

From here till the end of Section 2.6, fix a complete metric space (M, d) containing a thick family of geodesics, a positive sequence $(\delta_i')_{i=0}^\infty$ decreasing to 0 quickly enough so that $\delta_0' < \frac{1}{2}$ and $\Pi_{i=0}^\infty \frac{1}{1-2\delta_i'} \leq L$ for some $L < \infty$ (this also implies $\sum_i \delta'_i < \infty$), and a thick inverse system afforded to us by the theorem.

Definition 2.6. Denote the closure of $X_{<\infty} := \cup_{i=0}^{\infty} X_i$ inside M as $\boldsymbol{X_\infty}$. We fix $0 \in I = X_0 \subseteq X_\infty$ to be the basepoint. By Lemma 2.2, the maps π_i^j are uniformly L-Lipschitz, so we get L-Lipschitz extensions $\boldsymbol{\pi_i^\infty} : X_\infty \to X_i$. Summarizing:

$$\forall j \in \{i, i+1, \dots \infty\}, \quad \text{Lip}\left(\pi_i^j\right) \leq L \tag{2.10}$$

We also extend the definitions of $e_i(x)$ and $e'_i(x)$ (see Definition 2.4) in the obvious way when $x \in X_\infty$.

Remark 2.2. By Lemma 2.1, we get

$$\forall j \in \{i, i+1, \dots \infty\},\ x_i \in X_i, \quad \text{diam}((\pi_i^j)^{-1}(x_i)) \leq 2\delta'_i |e_i(x_i)| \tag{2.11}$$

Since each (X_i, d) is a finite graph, each (X_i, d) is compact and thus totally bounded. Then (2.11), together with our choice that $\delta'_i \to 0$, imply $X_{<\infty}$ is totally bounded. Then since M is complete, X_∞ is compact.

The maps $\pi_i^\infty : X_\infty \to X_i$ are each L-Lipschitz and act identically on $X_i \subseteq X_\infty$. These two facts imply, for any $p, q \in X_\infty$,

$$d(p, q) = \lim_{i \to \infty} d(\pi_i^\infty(p), \pi_i^\infty(q)) \tag{2.12}$$

This implies that the maps π_i^∞ generate the topology on X_∞, i.e., the topology on X_∞ is the weakest one such that each map π_i^∞ is continuous. Equivalently, the subalgebra of $C(X_\infty)$ consisting of those continuous functions that factor through some π_i^∞ is dense. We denote this subalgebra by $C_{\text{unif}}(X_{<\infty})$. The compatibility condition of the probability measures $((\pi_i^{i+1})_\#(\mu_{i+1}) = \mu_i)$ gives us a well-defined, bounded, positive linear functional $\lambda_{<\infty}$ on $C_{\text{unif}}(X_{<\infty})$. By density this extends to a unique positive linear functional λ_∞ on all of $C(X_\infty)$.

Definition 2.7. Define $\boldsymbol{\mu_\infty}$ to be the Radon measure representing the linear functional λ_∞ on $C(X_\infty)$. The measure μ_∞ is a probability measure uniquely characterized by:

$$\forall i \geq 0, \quad (\pi_i^\infty)_\#(\mu_\infty) = \mu_i \tag{2.13}$$

Remark 2.3. Although we won't make explicit use it, we believe it is worth mentioning the following fact: the metric space X_∞ and maps $(\pi_i^\infty)_{i=0}^\infty$ satisfy the universal property of an inverse limit space. This means that for any metric space Y and uniformly Lipschitz sequence of maps $(f_i)_{i=0}^\infty$, $f_i : Y \to X_i$, there exists a unique Lipschitz map $f_\infty : Y \to X_\infty$ such that $\pi_i^\infty \circ f_\infty = f_i$ for any i.

2.3 Asymptotic Local Properties of $(X_i)_{i=0}^\infty$ and Special Subsets of X_∞

2.3.1 Deep Points and their Natural Scales

Recall the definition of *terminal intervals* of X_{i+1} from Axiom (A2)(ii).

Definition 2.8. We define the set of **deep points**, $\boldsymbol{D}$, to be all those $x \in X_\infty$ such that $\pi_{i+1}^\infty(x)$ eventually (in i) does not belong to a terminal interval of X_{i+1}. The set D is a $G_{\delta\sigma}$ (and hence Borel) set.

Theorem 2.2. $\mu_\infty(D) = 1$.

Proof. Let e be an edge of X_i and e'_0 and e'_1 its terminal subintervals. By Definition 2.2 and Definition 2.3, $\mu_{i+1}(e'_0 \cup e'_1) = \mu_{i+1}(e'_0) + \mu_{i+1}(e'_1) \le 2\delta_i^E \mu_i(e)$. Summing over all $e \in E(X_i)$, we get that the total measure of the union of terminal intervals in $E(X_{i+1})$ is bounded by $2\delta_i^E$. Since $\sum_i \delta_i^E \le \sum_i \delta'_i < \infty$, Borel-Cantelli implies that the set of $x \in X_\infty$ such that $\pi_{i+1}^\infty(x)$ eventually (in i) does not belong to a terminal interval in X_{i+1} has measure 1. $\square$

Structure of $(\pi_{i-1}^i)^{-1}(e)$

We now discuss some geometric properties of $(\pi_{i-1}^i)^{-1}(e)$. While reading this section, it will be helpful to refer to Figure 2.3 for a picture of what $(\pi_{i-1}^i)^{-1}(e)$ typically looks like.

Definition 2.9. Given a deep point or, more generally, a nonvertex x and $i \ge 0$, define $\boldsymbol{r_i(x)} := |e_i(x)|$. We call $r_i(x)$ the sequence of *natural scales* of X_∞ at x.

Lemma 2.4. *For any deep point x and $R \ge 1$, $B^i_{Rr_i(x)}(\pi_i^\infty(x))$ is eventually (in i, depending on x and R) contained in $(\pi_{i-1}^i)^{-1}(e_{i-1}(x))$, where B^i indicates a ball in the space (X_i, d).*

Proof. Let $x \in D$ and $R \ge 1$. Set $x_i := \pi_i^\infty(x)$ and assume i is large enough so that $e'_{i-1}(x)$ is not a terminal interval. Then by Definition 2.5, $d(x_{i-1}, X_{i-1} \setminus e_{i-1}(x)) \ge \frac{|e'_{i-1}(x)|}{\delta_{i-1}^d} = \frac{r_i(x)}{\delta_{i-1}^d} \ge \frac{r_i(x)}{\delta'_{i-1}}$. Combining this with (2.10) yields

$$d(x_i, X_i \setminus (\pi_{i-1}^i)^{-1}(e_{i-1}(x))) \ge \frac{1}{L} d(x_{i-1}, X_{i-1} \setminus e_{i-1}(x)) \ge \frac{r_i(x)}{L\delta'_{i-1}}$$

Thus, as soon as i is large enough so that $\delta'_{i-1} < \frac{1}{LR}$, we get $B^i_{Rr_i(x_i)} \subseteq (\pi_{i-1}^i)^{-1}(e_{i-1}(x))$. $\square$

Lemma 2.5. *1. There exists $C \geq 1$ such that for any $i \geq 0$ and $e \in E(X_{i-1})$, μ_i restricted to $(\pi^i_{i-1})^{-1}(e)$ is C-doubling with respect to the length metric.*

2. For any shortest path $[x, y] \subseteq (\pi^i_{i-1})^{-1}(e)$, $\mu_i(B_r(x)) \leq 4\mu_i([x, y])$, where $r = |x - y|$.

Proof. Let $i \geq 0$ and $e \in E(X_{i-1})$. Recall the definition of *circles* and *intervals* from Axiom (A2)(i). By the discussion there, $(\pi^i_{i-1})^{-1}(e) = \cup_{e' \subseteq e}(\pi^i_{i-1})^{-1}(e')$ consists of a sequence of intervals and circles, glued together in a directed way along alternating sink and source vertices. This sequence begins and ends with *terminal intervals*, defined in Axiom (A2)(ii). With respect to the length metric and length measure, $(\pi^i_{i-1})^{-1}(e)$ is doubling. This follows by analyzing the worst case scenario for a ball. This scenario occurs near points where two circles are glued together. It is possible to have a geodesic ball of radius r such that the geodesic ball of radius $2r$ has 4 times the length. This implies length measure is doubling with doubling constant 4. Let $c \in (0, 1]$ such that μ_{i-1} restricted to e equals c times length measure, and for any $e \supseteq e' \in E(X'_{i-1})$, μ_i restricted to $(\pi^i_{i-1})^{-1}(e') \subseteq (\pi^i_{i-1})^{-1}(e)$ equals c or $\frac{c}{2}$ times length measure (c if it's an interval, $\frac{c}{2}$ if it's a circle). It follows that μ_i restricted to $(\pi^i_{i-1})^{-1}(e)$ is bounded above by c times length measure and below by $\frac{c}{2}$ times length measure. Since length measure it doubling with doubling constant 4, this implies μ_i is doubling with doubling constant bounded by 8 (this isn't sharp).

The second statement can also be observed by examining the worst case scenario where x is a vertex shared by two adjacent circles and y belongs to one of these circles. Then $B_r(x)$ will consist of four copies of an interval of length $r = |x - y|$, and the μ_i measure of any of these new intervals is the same as that of $[x, y]$. This implies the second statement. $\square$

Remark 2.4. It's also clear from the description of $(\pi^i_{i-1})^{-1}(e_{i-1}(x))$ given in the preceding section that if $x, y \in (\pi^{i+1}_i)^{-1}(e)$ and x and y do not belong to opposite edges of a circle, then x and y belong to a directed (and thus geodesic) edge path, and so $d(x, y) = |y - x|$. On the other hand, if $y \in B^i_{Rr_i(x)}(x_i)$ and x_i and y belong to opposite edges of a circle, then $|y - x_i| \leq |e_i(x)|$. In either case, we have, for $R \geq 1$ and i sufficiently large,

$$\forall y \in B^i_{Rr_i(x)}(x_i), \quad |y - x_i| \leq Rr_i(x) \tag{2.14}$$

2.3.2 Points having a NonEuclidean Tangent

Theorem 2.3. *There exists a Borel $S_\infty \subseteq X_\infty$ such that $\mu_\infty(S_\infty) > 0$, and for all $x \in S_\infty$, there exists a nonprincipal ultrafilter $\mathcal{U}(x)$ (depending on x) on $\mathbb{N}$ such that the tangent cone $T^{r_i(x), \mathcal{U}(x)}_x X_\infty$ does not embed (even topologically) into $\mathbb{R}$.*

Before beginning the proof of the theorem, we require a lemma:

Lemma 2.6. *For each $i \geq 0$, there is a finite set of directed 0-1 edge paths of X_i, $\mathcal{P}_i$, and a probability measure $\mathbb{P}_i$ on $\mathcal{P}_i$ such that for every edge $e \in E(X_i)$,*

$$\frac{\mu_i(e)}{|e|} = \sum_{\substack{P \in \mathcal{P}_i \\ e \subseteq P}} \mathbb{P}_i(P)$$

and it follows that, for any $A \subseteq e \in E(X_i)$ Borel,

$$\mu_i(A) = \sum_{P \in \mathcal{P}_i} \mathbb{P}_i(P)|A \cap P| \tag{2.15}$$

Proof. The proof is by induction on i. The base case $i = 0$ holds trivially with $\mathcal{P}_0 = \{X_0\}$, $\mathbb{P}_0 = \delta_{X_0}$. Assume the statement holds for some $i \geq 0$. Let $P \in \mathcal{P}_i$. Let P_{op} be the unique 0-1 directed edge path in X_{i+1} such that $e' \subseteq P$ if and only if $e'_{\text{op}} \subseteq P_{\text{op}}$ for every $e' \in E(X'_i)$. Let $\mathcal{P}_{i+1} = \{P, P_{\text{op}}\}_{P \in \mathcal{P}_i}$. For each $P \in \mathcal{P}_i$, define $\mathbb{P}_{i+1}(P_{\text{op}}) := \mathbb{P}_{i+1}(P) := \frac{1}{2}\mathbb{P}_i(P)$ if $P_{\text{op}} \neq P$, and $\mathbb{P}_{i+1}(P_{\text{op}}) = \mathbb{P}_{i+1}(P) := \mathbb{P}_i(P)$ if $P_{\text{op}} = P$. By Definition 2.2, $(\mathcal{P}_{i+1}, \mathbb{P}_{i+1})$ satisfies the desired property. ☐

Remark 2.5. This lemma gives an Alberti representation of the measure μ_i. In [Bat15], Bate used a property he called *universality* of Alberti representations to characterize Lipschitz differentiability spaces. Our representation of the measure μ_∞ (which can be constructed by taking limits of the representations of μ_i) will generally fail this universality condition, which is consistent with our discussion in Section 2.1.2 that $(X_\infty, d, \mu_\infty)$ is not a true Lipschitz differentiability space.

Proof of Theorem 2.3. Let $i \geq 0$ and $E_{\text{circ}}(X'_i)$ the set of edges $e' \in E(X'_i)$ such that $(\pi_i^{i+1})^{-1}(e')$ is a circle. Set $S_i := (\pi_i^\infty)^{-1}(E_{\text{circ}}(X'_i))$, so S_i is closed. By (2.13), (2.15), and Axiom (A6),

$$\mu_\infty(S_i) \overset{(2.13)}{=} \mu_i(E_{\text{circ}}(X'_i)) \overset{(2.15)}{=} \sum_{P \in \mathcal{P}_i} \mathbb{P}_i(P)|E_{\text{circ}}(X'_i) \cap P| \overset{(A6)}{\geq} \sum_{P \in \mathcal{P}_i} \mathbb{P}_i(P)\beta = \beta$$

Because of this, we set $S_\infty := \limsup_{i \to \infty} S_i$ (an $F_{\sigma\delta}$, and hence Borel, set) and get

$$\mu_\infty(S_\infty) \geq \beta > 0$$

By definition, S_∞ has the following property: for any $x \in S_\infty$, there is a subsequence $i_j(x)$ of i for which $\pi^\infty_{i_j(x)}(x) \in E_{\text{circ}}(X'_{i_j(x)})$. Thus, each pointed metric space $(X_\infty, \frac{1}{r_{i_j(x)}}d, x)$ contains a circle whose height (see Axiom (A5) for definition of height) is bounded below by α, and the point x belongs to this circle. Let

$\mathcal{U}(x)$ be any nonprincipal ultrafilter on $\mathbb{N}$ containing $\{i_j(x)\}_{j=0}^{\infty}$, which exists by Zorn's lemma. Then the $\mathcal{U}(x)$-ultralimit of this sequence of pointed metric spaces must also contain such a circle (and the point x will again belong to this circle), which obviously doesn't topologically embed into $\mathbb{R}$. $\square$

Remark 2.6. As described in the proof, each of the pointed spaces $(X_\infty, \frac{1}{r_{i_j}(x)}d, x)$ contain a circle of height α which contains x. Let e and e_{op} be the opposite edges of this circle. We can extend e in both directions to a 0-1 edge path. Since e_{op} has the same vertices as e, this also extends e_{op} to a 0-1 edge path. Unioning the circle $e \cup e_{\text{op}}$ with the extension to a 0-1 edge path results in a space consisting of two 0-1 geodesics whose union contains a circle of height α, and that coincide with each other outside that circle. Passing to the ultralimit, we see that the tangent cone $T_x^{r_i(x),\mathcal{U}(x)}X_\infty$ contains two bi-infinite geodesics whose union contains a circle of height α, and that coincide with each other outside that circle. Both geodesics get mapped down isometrically onto $\mathbb{R}$ under the blowup $(\pi_0^\infty)_x : T_x^{r_i(x),\mathcal{U}(x)}X_\infty \to \mathbb{R}$.

2.4 Approximation of Functions on X_∞ via X_i

We begin this section by introducing our fundamental tool for approximating functions on X_∞ by functions on X_i, the conditional expectation. The main results are Theorems 2.4 and 2.5. We then use this tool to define the derivative of Lipschitz functions on X_∞. The main result on the derivative is Theorem 2.6.

2.4.1 Conditional Expectation

Let $i \geq 0$ and $j \in \{i, i+1, \dots \infty\}$.

Definition 2.10. The **conditional expectation** is a bounded linear map $\mathbb{E}_i^j : L^1(\mu_j; B) \to L^1(\mu_i; B)$ uniquely characterized by the identity

$$\int_{X_i} \phi \cdot \mathbb{E}_i^j(h) d\mu_i = \int_{X_j} (\phi \circ \pi_i^j) \cdot h d\mu_j \tag{2.16}$$

for all $h \in L^1(\mu_j; B)$ and $\phi \in L^\infty(\mu_i)$. It is a standard tool in probability theory whose existence can be proven by elementary theorems of measure theory. See Chapter 1 of [Pis16] for background.

It follows from L^p-L^q duality that the conditional expectation is also contractive from $L^p(\mu_j; B) \to L^p(\mu_i; B)$ for any $p \in [1, \infty]$. The majority of this section is dedicated to proving the following theorem:

Theorem 2.4. *For every $i \geq 0$, $\mathbb{E}_i^\infty$ maps $\mathrm{Lip}_0(X_\infty; B)$ into $\mathrm{Lip}_0(X_i; B)$ with operator norm bounded by L^2.*

Such a result does not hold for general metric measure spaces (easy examples on $[0,1]$ show that conditional expectation need not preserve Lipschitz or even continuous functions), but will in our specific instance.

The proof will come at the end of this subsection and is preceded by several lemmas. We give an outline of the proof structure here:

- Show that for every $j < \infty$, $\mathbb{E}_i^j : \mathrm{Lip}_0\,(X_j; B) \to \mathrm{Lip}_0\,(X_i; B)$ has operator norm uniformly bounded by L.

- Noting that $\mathbb{E}_i^j := \mathbb{E}_i^{i+1} \circ \mathbb{E}_{i+1}^{i+2} \circ \ldots \mathbb{E}_{j-1}^j$, to prove the previous item, it suffices to consider the case $j = i+1$ and prove that $\|\mathbb{E}_i^{i+1}\|_{\mathrm{Lip}_0(X_{i+1;B}) \to \mathrm{Lip}_0(X_i;B)} \leq 1 + \delta_i'$, because by (2.2) we obtain

$$\|\mathbb{E}_i^j\|_{\mathrm{Lip}_0(X_j) \to \mathrm{Lip}_0(X_i)} \leq \Pi_{k=i}^{j-1}(1+\delta_k') \leq L \tag{2.17}$$

  for every $\infty > j \geq i \geq 0$. This is accomplished with Lemma 2.7.

- Extend the domain to X_∞ by approximating with maps factoring through some X_i, Lemma 2.8 (we gain another factor of L here).

Explicit Formula for and Boundedness of $\mathbb{E}_i^{i+1}$

Lemma 2.7. *For each $i \geq 0$ and $h \in \mathrm{Lip}_0\,(X_{i+1}; B)$,*

$$[\mathbb{E}_i^{i+1}(h)](p) = \frac{h(p) + h(p^{\mathrm{op}})}{2} \tag{2.18}$$

(recall the definition of p^{op} from Axiom (A2)(i)). Furthermore,

$$\|\mathbb{E}_i^{i+1}\|_{\mathrm{Lip}_0(X_{i+1}) \to \mathrm{Lip}_0(X_i)} \leq 1 + \delta_i'$$

Proof. Let $i \geq 0$ and $h \in \mathrm{Lip}_0\,(X_{i+1})$. It is a relatively simple exercise to check that (2.18) satisfies (2.16) using Definition 2.2. We now bound the operator norm. Let $x, y \in X_i$. No two points of $X_i \subseteq X_{i+1}$ can belongs to opposite edges of a circle in X_{i+1}, so also x^{op} and y^{op} do not belong to opposite edges of a circle. Thus the hypotheses for Lemma 2.3 are met. Then

$$\|\mathbb{E}_i^{i+1}(h)(x) - \mathbb{E}_i^{i+1}(h)(y)\| = \frac{\|h(x) + h(x^{\mathrm{op}}) - h(y) - h(y^{\mathrm{op}})\|}{2}$$

$$\leq \frac{\|h(x) - h(y)\|}{2} + \frac{\|h(x^{\mathrm{op}}) - h(y^{\mathrm{op}})\|}{2} \leq \frac{\|h\|_{\mathrm{Lip}_0(X_{i+1})}}{2}(d(x,y) + d(x^{\mathrm{op}}, y^{\mathrm{op}}))$$

$$\overset{\text{Lemma 2.3}}{\leq} \frac{\|h\|_{\mathrm{Lip}_0(X_{i+1})}}{2}(d(x,y) + (1+2\delta_i')d(x,y)) = (1+\delta_i')\|h\|_{\mathrm{Lip}_0(X_{i+1})}d(x,y)$$

□

Extending Domain to $\mathrm{Lip}_0\,(X_\infty; B)$

For Y a metric space and $K \geq 1$, we say a subspace $V \subseteq \mathrm{Lip}_0\,(Y; B)$ is K-*uniformly dense* in $\mathrm{Lip}_0\,(Y; B)$ if the closure with respect to the topology of uniform convergence of compacta (equivalently, pointwise convergence on any dense subset) of the ball of radius K in V contains the unit ball of $\mathrm{Lip}_0\,(Y; B)$.

Each Banach space $\mathrm{Lip}_0\,(X_i; B)$ can be identified as a closed subspace of $\mathrm{Lip}_0\,(X_\infty; B)$ by pulling back under the map π_i^∞. Denote the image of this identification by $\mathrm{Lip}_0\,(X_i; B)_\pi$. We then obtain the (nonclosed) subspace $\cup_{i<\infty}\mathrm{Lip}_0\,(X_i; B)_\pi \subseteq \mathrm{Lip}_0\,(X_\infty; B)$. We note that, for any $f \in \mathrm{Lip}_0\,(X_i; B)$,

$$\|f\|_{\mathrm{Lip}_0(X_i;B)} \leq \|f \circ \pi_i^\infty\|_{\mathrm{Lip}_0(X_\infty;B)} \leq \|f\|_{\mathrm{Lip}_0(X_i;B)}\|\pi_i^\infty\|_{\mathrm{Lip}} \leq L\|f\|_{\mathrm{Lip}_0(X_i;B)}$$

so that the embeddings $\mathrm{Lip}_0\,(X_i; B)_\pi \hookrightarrow \mathrm{Lip}_0\,(X_\infty; B)$ are uniformly bounded but not isometric.

Lemma 2.8. *For any Banach space* B*,* $\cup_{i<\infty}\mathrm{Lip}_0\,(X_i; B)_\pi \subseteq \mathrm{Lip}_0\,(X_\infty; B)$ *is* L*-uniformly dense.*

Proof. Let f be in the unit ball of $\mathrm{Lip}_0\,(X_\infty; B)$. Let g_i be the restriction to X_i of f. Then g_i belongs to the unit ball of $\mathrm{Lip}_0\,(X_i; B)$. Then $g_i \circ \pi_i^\infty$ belongs to the ball of radius L of $\cup_{i<\infty}\mathrm{Lip}_0\,(X_i; B)_\pi$. Clearly $g_i \circ \pi_i^\infty$ converges pointwise to f on the dense subset $X_{<\infty}$. □

Proof of Theorem 2.4. Let $i \geq 0$. Let f be in the unit ball of $\mathrm{Lip}_0\,(X_\infty; B)$. Let f_j be a sequence in the ball of radius L of $\cup_{i<\infty}\mathrm{Lip}_0\,(X_i; B)_\pi$ converging uniformly to f, which exists by Lemma 2.8. Then since $\mathbb{E}_i^\infty$ is bounded on L^∞, $\mathbb{E}_i^\infty(f_j)$ converges uniformly to $\mathbb{E}_i^\infty(f)$. Furthermore, for every j, by Lemma 2.7 and (2.17), $\|\mathbb{E}_i^\infty(f_j)\|_{\mathrm{Lip}_0(X_i;B)} \leq L\|f_j\|_{\mathrm{Lip}_0(X_\infty;B)} \leq L^2$. This implies $\|\mathbb{E}_i^\infty(f)\|_{\mathrm{Lip}_0(X_i;B)} \leq L^2$. □

Measure Representation of Conditional Expectation

We conclude our discussion of conditional expectation with a small theorem we will use once in the proof of Theorem 2.9. We begin with a standard but useful martingale convergence lemma.

Lemma 2.9. *For any Lipschitz map* $h : X_\infty \to \mathbb{R}$ *(not necessarily vanishing at 0) and* $i \geq 0$*,* $\mathbb{E}_i^\infty(h)$ *is Lipschitz and* $\mathbb{E}_i^\infty(h) \overset{i\to\infty}{\to} h$ *uniformly.*

Proof. Let $h : X_\infty \to \mathbb{R}$ be Lipschitz so that $h - h(0) \in \mathrm{Lip}_0\,(X_\infty)$. Then Theorem 2.4 implies $h = \mathbb{E}_i^\infty(h - h(0)) + h(0)$ is Lipschitz.

The Stone-Weierstrass theorem for algebras of continuous functions implies $\cup_{j<\infty} C(X_j)_\pi$ is uniformly dense in $C(X_\infty)$, where $C(X_j)_\pi$ is defined to be the continuous real-valued functions on X_∞ factoring through X_j. Then since $\mathbb{E}_i^\infty(h) \overset{i\to\infty}{\to} h$ (since it is eventually constant) for all $h \in \cup_{j<\infty} C(X_j)$, since $\sup_i \|E_i^\infty\|_{L^\infty(\mu_\infty)\to L^\infty(\mu_\infty)} = 1 < \infty$, and since μ_∞ and μ_i are fully supported on X_∞ and X_i, the claim follows. $\square$

Theorem 2.5. *For each $i \geq 0$, and $p \in X_i$, there exists a unique Borel probability measure μ_∞^p supported on $(\pi_i^\infty)^{-1}(p)$ such that for any $h \in C(X_\infty; B)$,*

$$[\mathbb{E}_i^\infty(h)](p) = \int_{(\pi_i^\infty)^{-1}(p)} h d\mu_\infty^p \tag{2.19}$$

Proof. Let $p \in X_i$. First we assume $B = \mathbb{R}$. Since, by Lemma 2.9 and the usual Stone-Weierstrass theorem, $\mathbb{E}_i^\infty$ preserves continuous functions and has uniform-uniform operator norm 1 (since μ_∞ and μ_i are fully supported), the map $h \mapsto [\mathbb{E}_i^\infty(h)](p)$ is a norm 1 linear functional on $C(X_\infty)$. Further, if $h \geq 0$, $[\mathbb{E}_i^\infty(h)](p) \geq 0$. Thus, our linear functional is represented by a probability measure μ_∞^p on X_∞. It remains to show μ_∞^p is supported on $(\pi_i^\infty)^{-1}(p)$. Consider the Lipschitz function $h_p : X_\infty \to R$ defined by $h_p(x) = d(x, (\pi_i^\infty)^{-1}(p))$. This function vanishes on $(\pi_i^\infty)^{-1}(p)$ and is strictly positive on $X_\infty \setminus (\pi_i^\infty)^{-1}(p)$. Thus, it suffices to show $[\mathbb{E}_i^\infty(h_p)](p) = 0$. Let $\epsilon > 0$. By Lemma 2.9, $\mathbb{E}_i^\infty(h_p) \overset{i\to\infty}{\to} h_p$ uniformly, so there exists $j \geq i$ such that $|[\mathbb{E}_j^\infty(h_p)](x)| < \epsilon$ for all $x \in (\pi_i^j)^{-1}(p)$ (since h_p vanishes on $(\pi_i^\infty)^{-1}(p))$). Since $\mathbb{E}_j^\infty(h_p)$ is a Lipschitz function on X_j, we may apply (2.18) (this was originally stated for functions vanishing at 0 but easily extends to the general case) and induction to conclude $|[\mathbb{E}_i^j(\mathbb{E}_j^\infty(h_p))](p)]| < \epsilon$. Since $\mathbb{E}_j^\infty \circ \mathbb{E}_i^j = \mathbb{E}_i^\infty$, we take $\epsilon \to 0$ and obtain the desired conclusion for $B = \mathbb{R}$.

Now we extend to general B. Define a map $E : C(X_\infty; B) \to C(X_i; B_{\text{weak}})$ by

$$[E(h)](p) := \int_{(\pi_i^\infty)^{-1}(p)} h d\mu_\infty^p$$

where B_{weak} indicated the space B equipped with the weak topology. We need to show $E = \mathbb{E}_i^\infty$, which we already know holds for $B = \mathbb{R}$. First, let us quickly verify that E indeed maps into the desired space. Let $h \in C(X_\infty; B)$ and $b^* \in B^*$. By an elementary property of the Bochner integral (see Chapter 1 of [Pis16], especially (1.7)) and the fact that $E = \mathbb{E}_i^\infty$ on real-valued continuous functions, $b^* \circ E(h) = E(b^* \circ h) = \mathbb{E}_i^\infty(b^* \circ h)$. We already know $\mathbb{E}_i^\infty$ maps real-valued continuous functions to real-valued continuous functions, so this shows $b^* \circ E(h)$ is continuous, completing our verification. By another elementary fact on B-valued conditional expectation (again see see Chapter 1 of [Pis16], (1.7)), $\mathbb{E}_i^\infty(b^* \circ h) = b^* \circ \mathbb{E}_i^\infty(h)$ μ_i-almost everywhere, for every $b^* \in B^*$. Thus, μ_i-almost everywhere, $b^* \circ E(h) = b^* \circ \mathbb{E}_i^\infty(h)$ for every $b^* \in B$, implying $E(h) = \mathbb{E}_i^\infty(h)$ μ_i-almost everywhere. But since both $E(h)$ and $\mathbb{E}_i^\infty(h)$ are continuous functions from X_i into the Hausdorff space B_{weak}, and since μ_i is fully-supported, $E(h) = \mathbb{E}_i^\infty(h)$ everywhere. $\square$

2.4.2 The Derivative and Fundamental Theorem of Calculus

We define the derivative of Lipschitz functions on X_∞ in this section. To do so, we must (and do) assume that B has the RNP. We also prove an inequality in Theorem 2.7 that should be thought of as an adapted version of the fundamental theorem of calculus.

Definition 2.11. For any $h_i \in \cup_{j<\infty}\mathrm{Lip}_0(X_j; B)$, since X_i is a finite graph equipped with a measure mutually absolutely continuous with length measure and with a distance geodesic on edges, the fact that B has the RNP allows us to take the derivative of h_i μ_i-almost everywhere defined by the usual formula $\boldsymbol{h_i'(x)} = \lim_{t\to 0} \frac{h_i(x+t)-h_i(x)}{t}$. We make sense of $x+t$ for t small by identifying the directed edge contained x with an interval, and the limit is an almost everywhere, norm limit. Equivalently, h_i' is characterized by

$$\lim_{r\to 0} \sup_{y\in B_r^i(x)} \frac{\|h_i(y) - h_i(x) - h_i'(x)(\pi(y) - \pi(x))\|}{r} = 0 \tag{2.20}$$

for μ_i-almost every $x \in X_i$, where $\pi := \pi_0^\infty$. The map $h_i \mapsto h_i'$ is a linear contraction $\mathrm{Lip}_0(X_\infty; B) \to L^\infty(\mu_\infty; B)$

Theorem 2.6. *There exists a unique bounded linear map $h \mapsto h' : \mathrm{Lip}_0(X_\infty; B) \to L^\infty(\mu_\infty; B)$, called the **derivative**, that*

1. *satisfies $\mathbb{E}_i^\infty(h)' \overset{i\to\infty}{\to} h'$ μ_∞-almost everywhere*
2. *restricts to the usual derivative on $\cup_{j<\infty}\mathrm{Lip}_0(X_j; B)$*
3. *has operator norm bounded by L^2.*

Proof. Note that uniqueness and the second statement already follow from the first statement. Let $h \in \mathrm{Lip}_0(X_\infty; B)$ with $\|h\|_{\mathrm{Lip}_0(X_\infty;B)} \le 1$, and for any $i \ge 0$, let $h_i := \mathbb{E}_i^\infty(h)$, so that $\|h_i\|_{\mathrm{Lip}_0(X_i;B)} \le L^2$ (by Theorem 2.4). Then the intermediate averages $x \mapsto \frac{h_i(x+t)-h_i(x)}{t}$ are uniformly (in t) $L^\infty(\mu_i; B)$-bounded by L^2. The DCT then implies that $\frac{h_i(\cdot+t)-h_i(\cdot)}{t} \overset{t\to 0}{\to} h_i'(\cdot)$ in $L^1(\mu_i; B)$. Then since the conditional expectation $\mathbb{E}_i^{i+1} : L^1(\mu_{i+1}; B) \to L^1(\mu_i; B)$ is continuous,

$$\mathbb{E}_i^{i+1}(h_{i+1}') = \mathbb{E}_i^{i+1}\left(\lim_{t\to 0} \frac{h_{i+1}(\cdot+t) - h_{i+1}}{t}\right) = \lim_{t\to 0} \mathbb{E}_i^{i+1}\left(\frac{h_{i+1}(\cdot+t) - h_{i+1}}{t}\right)$$

$$= \lim_{t\to 0} \frac{\left[\mathbb{E}_i^{i+i}(h_{i+1})\right](\cdot+t) - \left[\mathbb{E}_i^{i+i}(h_{i+1})\right]}{t} = \lim_{t\to 0} \frac{h_i(\cdot+t) - h_i}{t} = h_i'$$

The second to last equality says that conditional expectation commutes with precomposition with a translation, which can be directly verified by (2.18). Thus, the sequence $(h_i')_{i=0}^{\infty}$ forms a martingale uniformly bounded in $L^\infty(\mu_\infty; B)$ by L^2. Since B has the RNP property, the martingale converges μ_∞-almost everywhere to some function in $L^\infty(\mu_\infty; B)$ with norm bounded by L^2. We define h' to be this limit. $\square$

Theorem 2.7 (Fundamental Theorem of Calculus)**.** *For all $g \in \mathrm{Lip}_0(X_\infty; B)$, $i \geq 1$, $e \in E(X_{i-1})$, and $x, y \in (\pi^i_{i-1})^{-1}(e)$,*

$$\|[\mathbb{E}_i^\infty(g)](y) - [\mathbb{E}_i^\infty(g)](x)\| \leq 2|y-x| \fint_{[x,y]} \mathbb{E}_i^\infty(\|g'\|) d\mu_i$$

Proof. Let g, i, e and x, y be as above. Set $g_i := \mathbb{E}_i^\infty(g)$. First assume that x and y belong to a directed edge path. Then the usual Lebesgue fundamental theorem of calculus implies $\int_x^y g_i' ds = g_i(y) - g_i(x)$, where, for any positive Radon ν on X_i and $f \in L^1(X_i, \nu; B)$, $\int_x^y f d\nu$ is interpreted as $\int_{[x,y]} f d\nu$ if $x \leq y$ along the path, and $-\int_{[y,x]} f d\nu$ if $y \leq x$. If x and y don't belong to a directed edge path, there exists an intermediate point z on the shortest path from x to y such that the path is directed from x to z, and then anti-directed from z to y, or vice versa. We then still have $\int_x^y g_i' ds = g_i(y) - g_i(x)$ if we interpret $\int_x^y f d\nu$ as $\int_{[x,z]} f d\nu - \int_{[y,z]} f d\nu$ if $x \leq z$ and $y \leq z$ or $-\int_{[z,x]} f d\nu + \int_{[z,y]} f d\nu$ if $z \leq x$ and $z \leq y$. For future use, we also note that $\left\|\int_x^y f d\nu\right\| \leq \int_{[x,y]} \|f\| d\nu$.

As explained in the proof of Lemma 2.5, μ_i restricted to $[x,y] \subseteq (\pi^i_{i-1})^{-1}(e)$ is bounded below by $\frac{c}{2}$ times length measure and above by c times length measure. This implies that for any $f \in L^1(\mu_i)$ with $f \geq 0$, we have

$$\frac{1}{2} \fint_{[x,y]} f ds \leq |y-x| \fint_{[x,y]} f d\mu_i \leq 2 \fint_{[x,y]} f ds$$

Combining the last two paragraphs yields:

$$\|g_i(y) - g_i(x)\| = \left\|\int_x^y g_i' ds\right\| \leq \int_{[x,y]} \|g_i'\| ds \leq 2|y-x| \fint_{[x,y]} \|g_i'\| d\mu_i$$

$$= 2|y-x| \fint_{[x,y]} \|\mathbb{E}_i^\infty(g)'\| d\mu_i = 2|y-x| \fint_{[x,y]} \|\mathbb{E}_i^\infty(g')\| d\mu_i$$

$\square$

2.5 Maximal Operator and $L^1 \to L^{1,w}$ Inequality

Definition 2.12. Let $i \geq 0$ and $h_i \in L^1(\mu_i)$. For any nonvertex $x_i \in X_i$ and $i \geq 0$, define

$$[M_i(h_i)](x_i) := \left(\sup_{y_i \in (\pi^i_{i-1})^{-1}(e_{i-1}(x_i))} \fint_{[x_i, y_i]} |h_i| d\mu_i \right)$$

Now let $h \in L^1(\mu_\infty)$, and set $h_i := \mathbb{E}_i^\infty(h)$. For any nonvertex $x \in X_\infty$, set $x_i := \pi_i^\infty(x)$ and define the **maximal function**

$$[\boldsymbol{M(h)}](x) := \sup_{i\geq 0}[M_i(h_i)](x_i) \tag{2.21}$$

Theorem 2.8 (Maximal Inequality). *There exists a constant $C \geq 1$ such that for any $h \in L^1(\mu_\infty; B)$ and $p \in (1, \infty]$,*

$$\|M(\|h\|)\|_{L^{1,w}(\mu_\infty)} \leq \frac{Cp}{p-1}\|h\|_{L^p(\mu_\infty;B)} \tag{2.22}$$

Proof. As is typical, the proof is an application of a relevant covering lemma, Lemma 2.10, which we state and prove following this proof. This lemma is a combination of the Vitali covering lemma for doubling metric measure spaces and the covering lemma for atoms in a filtration of finite σ-algebras. Let $h \in L^1(\mu_\infty; B)$, $h_i := \mathbb{E}_i^\infty(h)$, and $p \in (1, \infty]$. After making the usual "covering lemma-to-maximal inequality" argument, we will have a $C \geq 1$ (independent of h or p, given to us by Lemma 2.10) such that

$$\|M(\|h\|)\|_{L^{1,w}(\mu_\infty)} \leq C\|h^*\|_{L^1(\mu_\infty)}$$

where h^* is Doob's maximal function; $h^*(x) := \sup_{i\geq 0}\|h_i\|(x)$. By Doob's maximal inequality ([Pis16, Theorem 1.25]),

$$\|h^*\|_{L^p(\mu_\infty)} \leq \frac{p}{p-1}\|h\|_{L^p(\mu_\infty;B)}$$

Combining these two inequalities with the simple inequality $\|h^*\|_{L^1(\mu_\infty)} \leq \|h^*\|_{L^p(\mu_\infty)}$ yields the desired conclusion. $\square$

Lemma 2.10 (Covering Lemma). *Let Γ be a collection of closed subsets of X_∞, such that for each $\gamma \in \Gamma$, there is an $i \geq 1$, a (not necessarily directed) shortest path $[p_\gamma, q_\gamma] \subseteq X_i$, and an edge $e_\gamma \in X_{i-1}$ such that:*

- $\gamma = (\pi_i^\infty)^{-1}([p_\gamma, q_\gamma])$
- *$[p_\gamma, q_\gamma]$ is completely contained in $(\pi_{i-1}^i)^{-1}(e_\gamma)$.*

Then there exists a subfamily $\Gamma' \subseteq \Gamma$, such that

- *The sets in Γ' are essentially pairwise disjoint*
- *For each $\gamma' \in \Gamma'$, there exists a closed set containing γ', denoted γ'_C, such that $\bigcup_{\gamma'\in\Gamma'}\gamma'_C \supseteq \bigcup\Gamma$ and $\mu_\infty(\gamma'_C) \leq C\mu_\infty(\gamma')$.*

Proof. First, consider the collection of sets $E_\Gamma := \{(\pi_{i-1}^\infty)^{-1}(e_\gamma)\}_{\gamma\in\Gamma}$. This set covers $\bigcup\Gamma$ by assumption. It is a collection of atoms in the filtration $(\mathcal{A}_i)_{i=0}^\infty$, where $\mathcal{A}_i$ is the σ algebra on X_∞ generated by preimages of edges in $E(X_i)$ under the map π_i^∞. Thus we may find an essentially disjoint subcollection that still covers $\cup\Gamma$. We consider a single one these sets, $(\pi_{i-1}^\infty)^{-1}(e)$. Let Γ_e be the collection of those $\gamma\in\Gamma$ with $[p_\gamma, q_\gamma] \subseteq (\pi_{i-1}^i)^{-1}(e)$. Since preimages under π_i^∞ preserve unions and essential disjointness, it suffices to work directly with the paths $[p_\gamma, q_\gamma]$. The path $[p_\gamma, q_\gamma]$ is contained in a geodesic ball $B_r(p_\gamma)$, where $r = |p_\gamma - q_\gamma|$. By Lemma 2.5, $\mu_i(B_r(p_\gamma)) \le 4\mu_i([p_\gamma, q_\gamma])$. By the 5r covering lemma, we can then find a pairwise disjoint subcollection of $\{B_r(p_\gamma)\}_{\gamma\in\Gamma_e}$, say $\{B_r(p_{\gamma'})\}_{\gamma'\in\Gamma'_e}$, such that $\{B_{5r}(p'_\gamma)\}_{\gamma'\in\Gamma'_e}$ covers $\bigcup\{B_r(p_\gamma)\}_{\gamma\in\Gamma_e}$ (and thus covers $\bigcup\Gamma_e$). We set $\Gamma' := \bigcup_{e\in E_\Gamma}\{(\pi_i^\infty)^{-1}([p', q'])\}_{[p',q']\in\Gamma_e}$. By Lemma 2.5, $\mu_i(B_{5r}(p)) \le \mu_i(B_{8r}(p)) \le 4^3\mu_i(B_r(p))$. We set $\Gamma' = \bigcup_{e\in E}\Gamma'_e$, $C = 4\cdot 4^3$, and $\gamma'_C = B_{5r}(p_{\gamma'})$. □

2.6 Proof of Weak Form of RNP Differentiability, Theorem 2.9

For each deep point $x \in D \subseteq X_\infty$ (a full measure set), recall the natural scale $r_i(x) = |e_i(x)|$, where $e_i(x)$ is the unique edge of X_i containing $\pi_i^\infty(x)$. Let $\pi := \pi_0^\infty$.

Theorem 2.9. *For every RNP space B and Lipschitz map $f : X_\infty \to B$, for μ_∞-almost every $x \in X_\infty$, f is differentiable at x with respect to π along the sequence of scales $(r_i(x))_{i=0}^\infty$. More specifically, for almost every $x \in D$ and any $R \ge 1$,*

$$\limsup_{i\to\infty} \sup_{y\in B_{Rr_i(x)}(x)} \frac{\|f(y) - f(x) - f'(x)(\pi(y) - \pi(x))\|}{r_i(x)} = 0$$

where f' is the derivative of f from Theorem 2.6.

Proof. Let B be an RNP space, $f : X_\infty \to B$ Lipschitz, and $R \ge 1$. The conclusion of the theorem is clearly invariant under postcomposition of f with a translation, so we may assume $f \in \mathrm{Lip}_0(X_\infty; B)$. For each $n \ge 0$, let $f_n := \mathbb{E}_n^\infty(f) \circ \pi_n^\infty \in \mathrm{Lip}_0(X_\infty; B)$ (see Section 2.4.1 for relevant definitions). Let

$$(*) := \limsup_{i\to\infty} \sup_{y\in B_{Rr_i(x)}(x)} \frac{\|f(y) - f(x) - f'(x)(\pi(y) - \pi(x))\|}{r_i(x)}$$

(so $(*)$ is a function of x). For every x, the triangle inequality implies

$$(*) \le \limsup_{n\to\infty}\limsup_{i\to\infty} \sup_{y\in B_{Rr_i(x)}(x)} \frac{\|f(y) - f(x) - f_n(y) + f_n(x)\|}{r_i(x)}$$

$$+\frac{\|f_n(y) - f_n(x) - f'_n(x)(\pi(y) - \pi(x))\|}{r_i(x)} + \frac{\|(f'_n(x) - f'(x))(\pi(y) - \pi(x))\|}{r_i(x)}$$

For almost every x and every fixed n, the second term equals 0 by (2.20), and so

$$(*) \leq \limsup_{n\to\infty} \limsup_{i\to\infty} \sup_{y\in B_{Rr_i(x)}(x)} \frac{\|f(y)-f(x)-f_n(y)+f_n(x)\|}{r_i(x)} + \frac{\|(f_n'(x)-f'(x))(\pi(y)-\pi(x))\|}{r_i(x)}$$

$$\leq \limsup_{n\to\infty} \limsup_{i\to\infty} \sup_{y\in B_{Rr_i(x)}(x)} \frac{\|f(y)-f(x)-f_n(y)+f_n(x)\|}{r_i(x)} + LR\|f_n'(x)-f'(x)\|$$

By Theorem 2.6, the second term here also equals 0 for almost every x, and so

$$(*) \leq \limsup_{n\to\infty} \limsup_{i\to\infty} \sup_{y\in B_{Rr_i(x)}(x)} \frac{\|f(y)-f(x)-f_n(y)+f_n(x)\|}{r_i(x)}$$

Let $k_n := f - f_n$, so that $\sup_n \|k_n'\|_{L^\infty(\mu_\infty;B)} \leq \sup_n \|k_n\|_{\mathrm{Lip}_0(X_\infty;B)} \leq 2L^2\|f\|_{\mathrm{Lip}_0(X_\infty;B)}$ and $\|k_n'\| \overset{n\to\infty}{\to} 0$ μ_∞-almost everywhere (again by Theorem 2.6. This means k_n' *boundedly converges* to 0, and we will apply the DCT theorem later). It suffices to prove

$$\limsup_{n\to\infty} \limsup_{i\to\infty} \sup_{y\in B_{Rr_i(x)}(x)} \frac{\|k_n(y)-k_n(x)\|}{r_i(x)} = 0$$

Define $y_i := \pi_i^\infty(y)$ and $x_i := \pi_i^\infty(x)$. then by Theorem 2.5,

$$[\mathbb{E}_i^\infty(k_n)](y_i) - [\mathbb{E}_i^\infty(k_n)](x_i) = \int_{(\pi_i^\infty)^{-1}(y_i)} k_n d\mu_\infty^{y_i} - \int_{(\pi_i^\infty)^{-1}(x_i)} k_n d\mu_\infty^{x_i}$$

Furthermore, for any $y \in (\pi_i^\infty)^{-1}(y_i)$ and $x \in (\pi_i^\infty)^{-1}(x_i)$, we have $\frac{d(y,y_i)}{r_i(x)}, \frac{d(y,y_i)}{r_i(x)} \leq 2\delta_i'$ by (2.11), which, together with the previous equation (and triangle inequality) gives us

$$\frac{\|k_n(y)-k_n(x)\|}{r_i(x)} \leq \frac{1}{r_i(x)} \int_{(\pi_i^\infty)^{-1}(y_i)} \|k_n - k_n(y)\| d\mu_\infty^{y_i}$$

$$+\frac{1}{r_i(x)} \left\| \int_{(\pi_i^\infty)^{-1}(y_i)} k_n d\mu_\infty^{y_i} - \int_{(\pi_i^\infty)^{-1}(x_i)} k_n d\mu_\infty^{x_i} \right\| + \frac{1}{r_i(x)} \int_{(\pi_i^\infty)^{-1}(x_i)} \|k_n - k_n(x)\| d\mu_\infty^{x_i}$$

$$\leq \frac{\|[\mathbb{E}_i^\infty(k_n)](y_i) - [\mathbb{E}_i^\infty(k_n)](x_i)\|}{r_i(x)} + 4\|k_n\|_{\mathrm{Lip}_0(X_\infty;B)}\delta_i'$$

$$\leq \frac{\|[\mathbb{E}_i^\infty(k_n)](y_i) - [\mathbb{E}_i^\infty(k_n)](x_i)\|}{r_i(x)} + 4(2L^2)\|f\|_{\mathrm{Lip}_0(X_\infty;B)}\delta_i'$$

so it suffices to prove that

$$\limsup_{n\to\infty} \limsup_{i\to\infty} \sup_{y_i\in B^i_{2Rr_i(x)}(x_i)} \frac{\|[\mathbb{E}_i^\infty(k_n)](y_i) - [\mathbb{E}_i^\infty(k_n)](x_i)\|}{r_i(x)} = 0$$

for almost every x, where B^i indicates a ball in the space (X_i, d) (since $2Rr_i(x) \geq Rr_i(x) + \delta_i' r_i(x)$).

By Lemma 2.4, for almost every x, if i is sufficiently large (depending on R and x), then $B^i_{2Rr_i(x)}(x_i)$ is completely contained in $(\pi^i_{i-1})^{-1}(e_{i-1}(x))$. Thus, by Theorem 2.7, for such i and any $y_i \in B^i_{2Rr_i(x)}(x_i)$,

$$\frac{\|[\mathbb{E}_i^\infty(k_n)](y_i) - [\mathbb{E}_i^\infty(k_n)](x_i)\|}{r_i(x)} \leq 2\frac{|y_i - x_i|}{r_i(x)} \fint_{[x_i,y_i]} \mathbb{E}_i^\infty(\|k_n'\|) d\mu_i =: (**)$$

By (2.14), since $y_i \in B^i_{2Rr_i(x)}(x_i) \subseteq (\pi^i_{i-1})^{-1}(e_{i-1}(x))$, $|y_i - x_i| \leq 2Rr_i(x)$, and so

$$(**) \leq 4R \fint_{[x_i,y_i]} \mathbb{E}_i^\infty(\|k_n'\|) d\mu_i =: (***)$$

by the definition of M, the maximal operator defined by (2.21), we get

$$(***) \leq 4R[M(\|k_n'\|)](x)$$

Then it suffices to show that

$$\limsup_{n\to\infty} [M(\|k_n'\|)](x) = 0$$

for almost every $x \in X_\infty$.

For this, it suffices to show that

$$\left\| \limsup_{n\to\infty} M(\|k_n'\|) \right\|_{L^{1,w}(\mu_\infty)} = 0$$

We have, by the DCT and Theorem 2.8,

$$\left\| \limsup_{n\to\infty} M(\|k_n'\|) \right\|_{L^{1,w}(\mu_\infty)} \overset{\text{DCT}}{=} \limsup_{n\to\infty} \|M(\|k_n'\|)\|_{L^{1,w}(\mu_\infty)} \overset{\text{Theorem 2.8}}{\leq} \limsup_{n\to\infty} 2C \|k_n'\|_{L^2(\mu_\infty;B)} \overset{\text{DCT}}{=} 0$$

□

2.7 Application to Non-BiLipschitz Embeddability

In this section we apply Theorem 1.14 to prove a new negative biLipschitz embeddability result.

Corollary 2.1. *A complete metric space M containing a thick family of geodesics does not biLipschitz embed into the product metric space $G \times B$, where G is a Carnot group and B is an RNP space.*

Proof. We'll proceed by contradiction. Let G be a Carnot group, B an RNP space, M a metric space containing a thick family of geodesics, and $f = (f_1, f_2) : M \to G \times B$ a biLipschitz map. We may assume M

is complete. Then we let $X_\infty \subseteq M$, $S_\infty \subseteq X_\infty$, and μ_∞ be as in Theorem 1.14, and from here on consider f to be restricted to X_∞.

Let $\psi : G : \mathbb{R}^k$ be the abelianization map. The map ψ satisfies a well known unique lifting property: given any Lipschitz map $\gamma : \mathbb{R} \to \mathbb{R}^k$, there exists a unique Lipschitz lift $\tilde{\gamma} : \mathbb{R} \to G$, meaning $\psi \circ \tilde{\gamma} = \gamma$. Precomposing with f gives a Lipschitz map $(\psi, id_B) \circ (f_1, f_2) = (\psi \circ f_1, f_2) : X_\infty \to \mathbb{R}^k \oplus B$ into an RNP space.

By Theorem 1.14, $(\psi \circ f_1, f_2)$ satisfies the weak form of differentiability μ_∞-almost everywhere. Pick a point $x \in S_\infty$ of differentiability (which exists since $\mu_\infty(S_\infty) > 0$) and an ultrafilter $\mathcal{U}(x)$ given to us by Theorem 1.14. This means the blowup $(\psi \circ f_1, f_2)_x : T_x^{r_i(x),\mathcal{U}(x)} X_\infty \to \mathbb{R}^k \times B$ exists and factors as $(\psi \circ f_1, f_2)_x = ((\psi \circ f_1)'(x), f_2'(x)) \circ \pi_x$, where $\pi_x : T_x^{r_i(x),\mathcal{U}(x)} X_\infty \to \mathbb{R}$ is the blowup of π and $(\psi \circ f_1)'(x) : \mathbb{R} \to \mathbb{R}^k$ and $f_2'(x) : \mathbb{R} \to B$ are linear. Breaking these into components gives us the two factorizations

$$(\psi \circ f_1)_x = (\psi \circ f_1)'(x) \circ \pi_x$$

$$(f_2)_x = f_2'(x) \circ \pi_x \tag{2.23}$$

Let us consider the blowup map $(\psi \circ f_1)_x$. It turns out that both blowups $(f_1)_x : T_x^{r_i(x),\mathcal{U}(x)} X_\infty \to T_{f_1(x)}^{r_i(x),\mathcal{U}(x)} G = G$ and $\psi = \psi_{f_1(x)} : G = T_{f_1(x)}^{r_i(x),\mathcal{U}(x)} G \to \mathbb{R}^k$ exist and thus $(\psi \circ f_1)_x = \psi \circ (f_1)_x$. The blowup $(f_1)_x$ exists because the target space G is proper, and $T_{f_1(x)}^{r_i(x),\mathcal{U}(x)} G = G$ because G is proper and self-similar. Similar reasoning implies $\psi_{f_1(x)} : G \to \mathbb{R}^k$ exists and $\psi_{f_1(x)} = \psi$. Thus our first factorization can be re-expressed as

$$\psi \circ (f_1)_x = (\psi \circ f_1)'(x) \circ \pi_x \tag{2.24}$$

By Remark 2.6, there are two geodesics $\gamma, \gamma' : \mathbb{R} \to T_x^{r_i(x),\mathcal{U}(x)} X_\infty$ whose combined image forms a circle of height α, that coincide with each other outside that circle, and satisfy $\pi_x \circ \gamma = \pi_x \circ \gamma' = id_{\mathbb{R}}$. Using these equations, (2.24), and (2.23) yields

$$\psi \circ (f_1)_x \circ \gamma = (\psi \circ f_1)'(x) = \psi \circ (f_1)_x \circ \gamma'$$

$$(f_2)_x \circ \gamma = f_2'(x) = (f_2)_x \circ \gamma'$$

Since f_1, γ, γ' are Lipschitz, the unique lifting property of ψ implies

$$(f_1)_x \circ \gamma = (f_1)_x \circ \gamma'$$

Combining these yields

$$(f_1, f_2)_x \circ \gamma = (f_1, f_2)_x \circ \gamma'$$

Since (f_1, f_2) is biLipschitz, so is $(f_1, f_2)_x$. Thus, $\gamma = \gamma'$. This is a contradiction since the combined image of two equal geodesics would be a line and could not contain (even topologically) a circle. $\square$

2.8 Inverse Limit of Graphs in nonRNP Spaces

In this section we modify the thick family of geodesics construction in [Ost14b] to obtain an embedding of an inverse limit of an admissible system of graphs into any nonRNP Banach space. To do so, we use the following characterization of nonRNP spaces (see Theorem 2.7 of [Pis16]): for any nonRNP space B, there exist a $\delta > 0$ and an open, convex subset C of the unit ball of B such that for every $c \in C$, $c \in \mathrm{co}(C \setminus B_{4\delta}(c))$.

2.8.1 Generalized Diamond Systems

Definition 2.13. A **generalized diamond system** is an inverse system of connected metric graphs, $\ldots \overset{\pi_2^3}{\to} X_2 \overset{\pi_1^2}{\to} X_1 \overset{\pi_0^1}{\to} X_0$ satisfying:

(D1) X_0 has two vertices and one edge of length 1. We identify X_0 with $I := [0, 1]$.

(D2) For any vertex $v \in V(X_i)$, $(\pi_i^{i+1})^{-1}(\{v\})$ consists of a single vertex of X_{i+1}. We identify this vertex with v and consider $V(X_i)$ as a subset of $V(X_{i+1})$.

(D3) There exist an m_i and a subdivision X_i' of X_i so that:

- (i) For vertex $v \in V(X_i')$, $(\pi_i^{i+1})^{-1}(\{v\})$ consists of one or two vertices of X_{i+1}. If u, v are adjacent vertices in X_i', then at most one of $(\pi_i^{i+1})^{-1}(\{u\})$, $(\pi_i^{i+1})^{-1}(\{v\})$ consists of two vertices.
- (ii) Each edge $e \in E(X_i)$ is subdivided into 2^{m_i} edges of X_i' of equal length.
- (iii) $\pi_i^{i+1} : X_{i+1} \to X_i'$ is open, simplicial, and an isometry on every edge.
- (iv) For any edge $e' \in E(X_i')$, $(\pi_i^{i+1})^{-1}(e')$ consists of one or two edges, and if e' is a terminal subedge of e (meaning it shares a vertex with e), then $(\pi_i^{i+1})^{-1}(e')$ consists of only one edge.

A generalized diamond system admits a canonical sequence of Borel probability measures $(\mu_i)_{i=0}^{\infty}$ satisfying

(D4) μ_0 is Lebesgue measure on I.

(D5) Restricted to each edge of X_i, μ_i is a constant multiple of length measure.

(D6) For each $e' \in E(X'_i)$, if $(\pi_i^{i+1})^{-1}(e')$ consists of two edges, then the μ_{i+1} measure of each of these edges equals $\frac{1}{2}\mu_i(e')$, and if $(\pi_i^{i+1})^{-1}(e')$ consists of one edge, then the μ_{i+1} measure of this edge equals $\mu_i(e')$.

Remark 2.7. With a small adjustment, these axioms imply the axioms of an "admissible" inverse system from [CK15]. The only problem is that in [CK15], each edge of X_i is subdivided into m edges of X'_i, where m is independent of i, and our subdivisions are into 2^{m_i} subedges, where m_i can depend on i. To conform to the [CK15] axiom, we can augment our inverse system by inserting extra graphs X_i^j between X_i and X_{i+1} that are simply subdivisions of X_i into 2^j subedges, for $1 \leq j \leq m_i$. The maps between them are identity maps. This new system will now be an admissible inverse system with subdivision parameter 2, and the inverse limit of the original system and augmented system will be the same. Thus, by Theorem 1.1 of [CK15], the inverse limit $(X_\infty, d_\infty, \mu_\infty)$ of a generalized diamond system is a PI space.

There is one last axiom for a generalized diamond system which implies (10.3) from [CK15] holds μ_∞-almost everywhere.

(D7) For any edge $e \in E(X_i)$, every point in $(\pi_i^{i+1})^{-1}(e_{1/2})$ is at most 2 edge lengths (of X_{i+1}) away from a vertex of degree 4, where $e_{1/2}$ denotes the middle half of e.

Theorem 2.10. *Every nonRNP Banach space contains a biLipschitz copy of a metric measure space satisfying the differentiation nonembeddability criterion. The metric measure space is an inverse limit of admissible graphs, as in [CK15], with nonEuclidean tangent cones at almost every point.*

Proof. We begin by making some reductions. First, notice that it suffices to embed into $B \oplus_\infty \mathbb{R}$ for any nonRNP space B. This is because we may pick any closed, codimension-1 subspace $B' \subseteq B$, which is also necessarily a nonRNP space, and get $B \cong B' \oplus_\infty \mathbb{R}$.

Let B be a nonRNP space (in a slight abuse of notation, we'll use $\|\cdot\|$ to stand for both the norm on B and the norm on $B \oplus_\infty \mathbb{R}$, but this shouldn't cause any confusion). We'll construct a sequence of subsets $(X_i)_{i=0}^\infty$ of $B \oplus_\infty \mathbb{R}$ and maps $\pi_i^{i+1} : X_{i+1} \to X_i$ such that (X_i, d_i) is a connected metric graph and $\dots \overset{\pi_2^3}{\to} X_2 \overset{\pi_1^2}{\to} X_1 \overset{\pi_0^1}{\to} X_0$ is a generalized diamond system, where d_i denotes the intrinsic metric on X_i (shortest path metric, where path length is measured with respect to ambient Banach space). The construction will be such that there exist a $\delta > 0$ and $\delta_i > \delta$ such that X_i is δ_i^{-1}-quasiconvex in $B \oplus \mathbb{R}$, meaning $\delta_i d_i(x,y) \leq \|x-y\| \leq d_i(x,y)$ for all $x, y \in X_i$. Furthermore, the construction will be such that for any $v \in V(X_i) \subseteq V(X_{i+1})$, $\pi_i^{i+1}(v) = v$ (see Axiom (D2) for the identification of $V(X_i)$ as a subset of $V(X_{i+1})$). By density of the the vertices in the inverse limit space, this implies that the closure of $\cup_i V(X_i)$ in $B \oplus \mathbb{R}$ is δ^{-1}-biLipschitz equivalent to the inverse limit of $\dots \overset{\pi_2^3}{\to} X_2 \overset{\pi_1^2}{\to} X_1 \overset{\pi_0^1}{\to} X_0$.

Previously, we introduced geodesics as isometric maps on intervals, but in this proof it will be more convenient to consider the image of these maps instead of the map itself. For this reason, we use the term *geodesic path* to mean the image of a geodesic map. Additionally, if p and q are points in a graph, we previously used the notation $|p-q|$ to denote the distance between p and q with respect to the length metric, but such notation will cause problems in this proof since we are working in a normed space. Instead, we will use the term *intrinsic metric* which has the same meaning as length metric, and notation for this distance will be set subsequently.

Model Graph

Let $\delta > 0$ and let C be an open, convex subset of the unit ball of B such that $0 \in C$ and $c \in \text{co}(C \setminus B_{4\delta}(c))$ for every $c \in C$, where $B_r(x)$ is the closed unit ball of radius r centered at x. We describe how to construct a graph, for each $c \in C$, that will serve as a building block for the graphs X_i.

Let $c \in C$. We'll form two piecewise affine, geodesic paths from $(0,0)$ to $(c,1)$, denoted $\gamma_0(c)$ and $\gamma_1(c)$. The reader should refer to Figure 2.5 for a helpful visual of the construction. Since $c \in C$, $c = \alpha_1 c_1 + \ldots \alpha_k c_k$ for some $\alpha_j \in (0,1)$ and $c_1, \ldots c_k \in C$ with $\alpha_1 + \ldots \alpha_k = 1$ and $\|c - c_j\| \geq 4\delta_c > 4\delta$ (note that since c, c_j belong to the unit ball of B, $\delta_c \leq \frac{1}{2}$). Since C is open, we may assume each α_j is a dyadic rational with common denominator 2^n, by density of dyadic rationals in $[0,1]$. Additionally, by "splitting" up terms of the form $\frac{m}{2^n} c_j$ into the m-fold sum $\frac{1}{2^n} c_j + \frac{1}{2^n} c_j + \ldots \frac{1}{2^n} c_j$, we may assume $\alpha_j = 2^{-n_c}$ and $k = 2^{n_c}$ for some $n_c \geq 1$, independent of j (of course we do not have that $\{c_j\}$ are distinct, but that is no issue). The path $\gamma_0(c)$ consists of a piecewise affine interpolation between $2 \cdot 2^{n_c} + 1$ vertices, $v_0, v_1', v_1, v_2', v_2, \ldots v_{2^{n_c}}', v_{2^{n_c}}$. These vertices are such that $v_0 = (0,0)$, and for each j, $v_j' - v_{j-1} = 2^{-(n_c+1)}(c,1)$ and $v_j - v_j' = 2^{-(n_c+1)}(c_j,1)$. Likewise, $\gamma_1(c)$ consists of a piecewise affine interpolation between $2 \cdot 2^{n_c} + 1$ vertices, $w_0, w_1', w_1, w_2', w_2, \ldots w_{2^{n_c}}', w_{2^{n_c}}$. These vertices are such that $w_0 = (0,0)$, and for each j, $w_j' - w_{j-1} = 2^{-(n_c+1)}(c_j,1)$ and $w_j - w_j' = 2^{-(n_c+1)}(c,1)$ (notice the flipping of primed and unprimed terms). It follows that $v_j = w_j$ for each j, and that $v_{2^{n_c}} = (c,1) = w_{2^{n_c}}$. These are indeed geodesic paths because the vectors c, c_j all have norm 1 in B, and we take an ∞-norm direct sum. An isometry from these geodesics paths onto the interval $[0,1]$ is provided by projection onto the second coordinate.

$\gamma_0(c)$ is equipped with a graph structure. The vertex set is the ordered set $(v_0, v_1', v_1, v_2', v_2, \ldots v_{2^{n_c}}', v_{2^{n_c}})$ and there is one edge between consecutive vertices consisting of the line segment between them. The path $\gamma_1(c)$ is similarly equipped with a graph structure. We let $\Gamma(c) = \gamma_0(c) \cup \gamma_1(c)$. Since $\gamma_0(c)$ and $\gamma_1(c)$ intersect only on their vertices, $\Gamma(c)$ inherits an induced graph structure. The vertex set is $\{v_0 = w_0 = (0,0), v_1', w_1', v_1 = w_1, \ldots v_{2^{n_c}}', w_{2^{n_c}}', v_{2^{n_c}} = w_{2^{n_c}} = (c,1)\}$. See Figure 2.5 for an example of $\Gamma(c)$ for $2^{n_c} = 4$. Loosely, $\Gamma(c)$ is made up of a sequence of parallelograms increasing in the "$\mathbb{R}$ direction" of $B \oplus \mathbb{R}$ such that

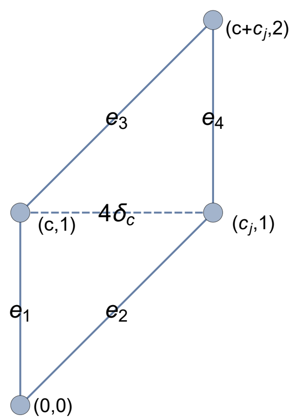

Figure 2.4: The parallelogram with vertices $(0,0)$, $(c,1)$, $(c_j,1)$, and $(c+c_j,2)$. The horizontal axis is in the "B direction" of $B \oplus \mathbb{R}$, and the vertical axis is in the "$\mathbb{R}$ direction". The extrinsic and intrinsic distance between any two points on $e_1 \cup e_3$ or any two points on $e_2 \cup c_4$ agree. The extrinsic distance between the two vertices $(c,1)$, $(c_j,1)$ is $4\delta_c$. All edge lengths are 1.

adjacent parallelograms share a common vertex. Because of this, for any two points of $x, y \in \Gamma(c)$ belonging to distinct parallelograms, the extrinsic distance $\|x-y\|$ and intrinsic distance $d_{in}(x,y)$ agree. We claim that each of these parallelograms is δ_c^{-1}-quasiconvex. Then this claim together with the preceding sentence imply that $\Gamma(c)$ is δ_c^{-1}-quasiconvex.

Proof of Claim. Consider one of the parallelograms of $\Gamma(c)$. It has vertices $v_{j-1} = w_{j-1}, v_j', w_j', v_j = w_j$ for some j. First notice that translations and dilations don't change the quasiconvexity constant of parallelograms, so we may perform such modifications to ours to obtain one that is easier to calculate with. Translate the parallelogram by $-v_{j-1}$ $(= -w_{j-1})$ so that one of the vertices is $(0,0)$, and the other vertices are $2^{-(n_c+1)}(c,1)$, $2^{-(n_c+1)}(c_j,1)$, and $2^{-(n_c+1)}(c+c_j,2)$. Then scale by 2^{n_c+1} so that the vertices are $(0,0)$, $(c,1)$, $(c_j,1)$, and $(c+c_j,2)$. Now we label the edges: let e_1 be the edge between $(0,0)$ and $(c,1)$, e_2 the edge between $(0,0)$ and $(c_j,1)$, e_3 the edge between $(c,1)$ and $(c+c_j,2)$, and e_4 be the edge between $(c_j,1)$ and $(c+c_j,2)$. Figure 2.4 shows an example of this parallelogram, and it will be helpful to keep this picture in mind while reading the remaining proof of the claim.

Note that $e_1 \cup e_3$ is a subpath of the geodesic path corresponding to $\gamma_0(c)$, and $e_2 \cup e_4$ is a subpath of the geodesic path corresponding to $\gamma_1(c)$, so the intrinsic and extrinsic distance agree on these subsets. Let x and y be elements of the parallelogram. As just mentioned, if x and y belong to $e_1 \cup e_3$, or both belong to $e_2 \cup e_3$, then the intrinsic and extrinsic distance between x and y agree. Suppose then that x belongs to e_1 and y belongs to e_2. Then $x = \alpha(c,1)$ for some $\alpha \in [0,1]$, $y = \beta(c_j,1)$ for some $\beta \in [0,1]$, and the intrinsic distance between x and y is $\alpha + \beta$. Without loss of generality, assume $\beta \geq \alpha$, so that the intrinsic distance between x and y, $d_{in}(x,y)$ is bounded by 2β. Then the extrinsic distance between x and y is

$$\|x-y\| = \|\alpha(c,1) - \beta(c_j,1)\| = \|(\beta(c-c_j) + (\alpha-\beta)c, \alpha-\beta)\|$$

$$= \max(\|\beta(c-c_j) + (\alpha-\beta)c\|, |\alpha-\beta|)$$

$$\geq \max(\|\beta(c-c_j)\| - \|(\alpha-\beta)c\|, |\alpha-\beta|)$$

$$\geq \max(\|\beta(c-c_j)\| - |\alpha-\beta|, |\alpha-\beta|)$$

$$\geq \max(\beta 4\delta_c - |\alpha-\beta|, |\alpha-\beta|)$$

$$\geq (2\beta)\delta_c \geq \delta_c d_{in}(x,y)$$

showing that the quasiconvexity constant is bounded above by δ_c^{-1} in this case. By symmetry, we get the same upper bound if x belongs to e_3 and y belongs to e_4. There is one remaining case (since the rest of the cases follow from this one by symmetry), in which x belongs to e_1 and y belongs to e_4. In this case, $x = \alpha(c,1)$ for some $\alpha \in [0,1]$, $y = (c_j,1) + \beta(c,1)$ for some $\beta \in [0,1]$, and we use the trivial bound $d_{in}(x,y) \leq 2$ for the intrinsic distance. Then for the extrinsic distance, we have

$$\|x-y\| = \|\alpha(c,1) - ((c_j,1) + \beta(c,1))\|$$

$$= \|((\alpha-\beta-1)\,c + (c-c_j), \alpha-\beta-1)\|$$

$$= \max\left(\|(\alpha-\beta-1)\,c + (c-c_j)\|, |\alpha-\beta-1|\right)$$

$$\geq \max\left(\|c-c_j\| - |\alpha-\beta-1|, |\alpha-\beta-1|\right)$$

$$\geq \max\left(4\delta_c - |\alpha-\beta-1|, |\alpha-\beta-1|\right)$$

$$\geq 2\delta_c \geq \delta_c d_{in}(x,y)$$

This completes the proof of the δ_c^{-1}-quasiconvexity of the parallelogram.

End Proof of Claim.

Since $\gamma_0(c)$, $\gamma_1(c)$, and $[(0,0),(c,1)]$ are all geodesics with endpoints $(0,0)$ and $(c,1)$, there are unique isometries $\gamma_0(c) \to [(0,0),(c,1)]$ and $\gamma_1(c) \to [(0,0),(c,1)]$ fixing the endpoints. If we let $[(0,0),(c,1)]'$ denote the subdivision of $[(0,0),(c,1)]$ into subedges of length $2^{-(n_c+1)}$, the maps are graph isomorphisms. Combining these gives us a map $\pi_c : \Gamma(c) \to [(0,0),(c,1)]'$ which is open, simplicial, and an isometry on

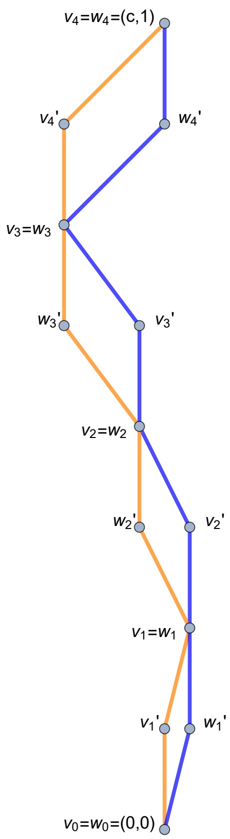

Figure 2.5: The model graph $\Gamma(c)$ The geodesic path $\gamma_0(c)$ is shown in orange, and the geodesic path $\gamma_1(c)$ is shown in blue. The horizontal axis is in the "B direction" of $B \oplus \mathbb{R}$, and the vertical axis is in the "$\mathbb{R}$ direction".

every edge. Furthermore, the preimage of any edge in $[(0,0),(c,1)]'$ consists of two edges of $\Gamma(c)$. Let $((0,0) = t_0, t_1', t_1, t_2', t_2, \ldots t_{2^{n_c}}', t_{2^{n_c}} = (c,1))$ be the ordered vertex set of $[(0,0),(c,1)]'$. Then $(\pi_c^{-1})(\{t_j\}) = \{v_j\} = \{w_j\}$, a single vertex, and $(\pi_c^{-1})(\{t_j'\}) = \{v_j, w_j'\}$, a set of two vertices. Finally, if $j \neq 0, 2^{n_c}$, the vertex $v_j = w_j$ has degree four, so every point in $\Gamma(c)$ is at most two edge lengths away from a vertex of degree four. Thus, $\pi_c : \Gamma(c) \to [(0,0),(c,1)]'$ satisfies the conditions listed for π_i^{i+1} in Axioms (D2), (D3), and (D7).

For any $\alpha_0 \in \mathbb{R} \setminus \{0\}$, $b_0 \in B$, and $A \subseteq B$, we let $\alpha_0 A + b_0$ be the image of A under the invertible similarity $b \mapsto \alpha_0 b + b_0$. The sets $\alpha_0\Gamma(c) + b_0$ and $(\alpha_0[(0,0),(c,1)] + b_0)'$ inherit graph structures from $\Gamma(c)$ and $[(0,0),(c,1)]'$, respectively, and there is also an induced map $\alpha_0\pi_c + b_0 : \alpha_0\Gamma(c) + b_0 \to (\alpha_0[(0,0),(c,1)] + b_0)'$ that, like π_c, satisfies Axioms (D2), (D3), and (D7).

Inductive Construction of X_i

For the base case, let $X_0 = \{0\} \times I \subseteq B \oplus \mathbb{R}$. For the inductive hypothesis, assume that the inverse system $X_i \overset{\pi_{i-1}^i}{\to} X_{i-1} \ldots \overset{\pi_0^1}{\to} X_0$ and X_{i-1}' have been constructed and satisfy Axioms (D1)-(D7) from Definition 2.13. For $e \in E(X_i)$, let $v_0(e)$ and $v_1(e)$ denote the terminal vertices of e. Assume that the inverse system satisfies the additional properties:

(P1) For all $e \in E(X_i)$, e equals the line segment joining $v_0(e)$ to $v_1(e)$. That is, $e = [v_0(e), v_1(e)] := \{(1-t)v_0(e) + tv_1(e) : t \in [0,1]\}$.

(P2) For all $e \in E(X_i)$, e is parallel to an associated vector $(c,1) \in C \times \{1\}$. That is, $v_1(e) - v_0(e) = \alpha(c,1)$ for some $\alpha \in \mathbb{R}$ and $c \in C$. Furthermore, $\alpha = 2^{-n_i}$ for some $n_i \geq 1$. The number n_i depends on i but not on e. It follows that every edge of X_i has length 2^{-n_i}.

Now we need to construct X_{i+1}, X_i', and $\pi_i^{i+1} : X_{i+1} \to X_i'$. Let $e \in E(X_i)$, and $c \in C$ and $n_i \geq 1$ such that $v_1(e) - v_0(e) = 2^{-n_i}(c,1)$. Subdivide e into into 3 subedges, the middle one having length $\frac{1}{2}|e|$, and the terminal ones having length $\frac{1}{4}|e|$. Let e_0 and e_1 denote the terminal subedges, and $e_{1/2}$ the middle subedge. Note that, for any $x \in e_{1/2}$ and $y \in X_i \setminus e$,

$$d_i(x,y) \geq \frac{|e|}{4} = 2^{-(n_i+2)} \tag{2.25}$$

Let $\delta' = \frac{\delta + \delta_i}{2}$, so that $\delta < \delta' < \delta_i$. Choose N to be large enough so that

$$2^{-N} \leq (\delta_i - \delta')2^{-2} \tag{2.26}$$

Subdivide $e_{1/2}$ into 2^N edges of equal length. So now e is divided into a total of $2^N + 2$ subedges, and two of them, e_0 and e_1, are marked as terminal subedges. Doing this for every $e \in E(X_i)$ gives us a subdivision $\tilde{X}_i$ of X_i. Let f be a subedge of $e_{1/2}$. Then $v_1(f) - v_0(f) = 2^{-(n_i+1+N)}(c, 1)$. We create X_{i+1} by replacing f with the graph $2^{-(n_i+1+N)}\Gamma(c) + v_0(f)$, which has the same vertices as f. Thus, X_{i+1} consists of the union of $e_0, e_1, 2^{-(n_i+1+N)}\Gamma(c) + v_0(f)$ over all $f \subseteq e_{1/2}$ and $e \in E(X_i)$, with each e_0 and e_1 subdivided into subedges so that every edge of X_{i+1} has equal length. The graph X_{i+1} satisfies (P1) and (P2). Since there are only finitely many $e \in E(X_i)$, and thus finitely many $c \in C$ associated to e, we may choose the subdivision parameter n_c of Section 2.8.1 independent of c.

X_i' is simply the subdivision of $\tilde{X}_i$ into subedges all having length the same as any edge of X_{i+1}. For any e_0, e_1, and $f \subseteq e_{1/2}$, let e_0', e_1', and f' denote the subdivisions in X_i'. Let $2^{-(n_i+1+N)}\pi_c + v_0(f)$: $2^{-(n_i+1+N)}\Gamma(c) + v_0(f) \to f'$ be the map defined in Section 2.8.1. We paste all these maps along with all the identity maps $e_0 \to e_0'$, $e_1 \to e_1'$ together to obtain the quotient map $\pi_i^{i+1} : X_{i+1} \to X_i'$. Then π_i^{i+1} satisfies Axioms (D2), (D3), and (D7) because each map $2^{-(n_i+1+N)}\pi_c + v_0(f)$ does.

The map π_i^{i+1} is a 1-Lipschitz quotient with respect to the metrics d_{i+1} and d_i. Furthermore, it has the property that, if $x, y \in X_{i+1}$ and $\pi_i^{i+1}(x)$ and $\pi_i^{i+1}(y)$ do not belong to the same edge of $\tilde{X}_i$, then

$$d_{i+1}(x, y) = d_i(\pi_i^{i+1}(x), \pi_i^{i+1}(y)) \tag{2.27}$$

Set $\delta_{i+1} := \min_c(\delta_c, \delta') > \delta$, where the minimum is over each $(c, 1)$ associated to an edge e of X_i. Since there are only finitely many edges of X_i, the minimum is well-defined and $\delta_{i+1} > \delta$. We now check that X_{i+1} is δ_{i+1}^{-1}-quasiconvex.

Let $x, y \in X_{i+1}$. First consider the case $\pi_i^{i+1}(x)$ and $\pi_i^{i+1}(y)$ belong to the same edge f of $\tilde{X}_i$, with $v_1(f) - v_0(f) = 2^{-n_i}(c, 1)$ for some $c \in C$. Then x and y both belong to $2^{-(n_i+1+N)}\Gamma(c) + v_0(f)$, on which the intrinsic distance is δ_c^{-1}-quasiconvex, so the desired conclusion holds in this case.

Now assume $\pi_i^{i+1}(x)$ and $\pi_i^{i+1}(y)$ do not belong to the same edge of $\tilde{X}_i$ but do belong to the same edge of X_i. Then the intrinsic and extrinsic distance between x and y, and the intrinsic and extrinsic distance between $\pi_i^{i+1}(x)$ and $\pi_i^{i+1}(y)$ are all equal.

Finally, assume $\pi_i^{i+1}(x)$ and $\pi_i^{i+1}(y)$ do not belong to the same edge of X_i. We consider two subcases: both x and y belong to terminal subedges of $e, f \in E(X_i)$, or one does not belong to a terminal subedge. In the first case, if both x and y belong to terminal subedges of X_i, then π_i^{i+1} acts identically, on x and y, and so

$$d_{i+1}(x, y) = d_i(\pi_i^{i+1}(x), \pi_i^{i+1}(y)) \leq \delta_i^{-1}\|\pi_i^{i+1}(x) - \pi_i^{i+1}(y)\| = \|x - y\|$$

by the inductive hypothesis and so the conclusion holds. Now assume, without loss of generality, that $\pi_i^{i+1}(x) \in e_{1/2}$ for some $e \in E(X_i)$ and $y \in X_{i+1} \setminus (\pi_i^{i+1})^{-1}(e)$. Then we get

$$d_i(\pi_i^{i+1}(x), \pi_i^{i+1}(y)) \geq \|\pi_i^{i+1}(x) - \pi_i^{i+1}(y)\| \geq \delta_i d_i(x,y) \overset{(2.25)}{\geq} \delta_i \frac{|e|}{4} = \delta_i 2^{-(n_i+2)} \tag{2.28}$$

Since π_i^{i+1} acts identically on the vertices of $\tilde{X}_i$, the d_{i+1} diameter of any fiber of π_i^{i+1} is at most the length of an edge of $\tilde{X}_i$, which is $2^{-(n_i+1+N)}$. This implies

$$\|\pi_i^{i+1}(x) - x\|, \|\pi_i^{i+1}(y) - y\| \leq 2^{-(n_i+1+N)} \tag{2.29}$$

Thus,

$$\begin{aligned}
\|x-y\| &\geq \|\pi_i^{i+1}(x) - \pi_i^{i+1}(y)\| - \|\pi_i^{i+1}(x) - x\| - \|y - \pi_i^{i+1}(y)\| \\
&\overset{(2.29)}{\geq} \delta_i d_i(\pi_i^{i+1}(x), \pi_i^{i+1}(y)) - 2^{-(n_i+N)} \\
&\overset{(2.26)}{\geq} \delta_i d_i(\pi_i^{i+1}(x), \pi_i^{i+1}(y)) - (\delta_i - \delta')2^{-(n_i+2)} \\
&\overset{(2.28)}{\geq} \delta_i d_i(\pi_i^{i+1}(x), \pi_i^{i+1}(y)) - (\delta_i - \delta') d_i(\pi_i^{i+1}(x), \pi_i^{i+1}(y)) \\
&= \delta' d_i(\pi_i^{i+1}(x), \pi_i^{i+1}(y)) \overset{(2.27)}{=} \delta' d_{i+1}(x,y) \geq \delta_{i+1} d_{i+1}(x,y)
\end{aligned}$$

□

Chapter 3

Essentially Uniformly Discrete Metric Spaces

3.1 Introduction

This chapter is devoted to proving the results stated in Section 1.2.2.

3.1.1 Outline

Section 3.2 contains the proof of Theorem 3.1 and the following section contains the proof of Theorem 3.2. The final section contains examples of essentially uniformly discrete metric spaces which do not obviously biLipschitz embed into RNP spaces without the aid of Theorem 1.15.

3.1.2 Notation

If $\lambda \in \mathrm{Lip}_{\mathrm{fin}}(X)^*$, we assign to it the real numbers $(c_p)_{p\in X\setminus\{0\}} \in \mathbb{R}^{X\setminus\{0\}}$ and write $\lambda = \sum_{p\in X} c_p e_p$ (even though this infinite sum doesn't necessarily have a usual meaning as a limit of finite sums) if $\lambda\left(1_{\{p\}}\right) = c_p$ for all $p \in X \setminus \{0\}$ (and interpret $c_0 = 0$). The assignment $\lambda \mapsto (c_p)_{p\in X\setminus\{0\}}$ is injective.

Note that $\mathrm{res}(\delta_p) = e_p$ (Recall the definition of res from Section 1.2.2). If $v \in \mathrm{LF}(X)$, we assign to it the real numbers $(c_p)_{p\in X\setminus\{0\}} \in \mathbb{R}^{X\setminus\{0\}}$ and write $v = \sum_{p\in X} c_p\delta_p$ (even though this infinite sum doesn't necessarily have a usual meaning as a limit of finite sums) if $\mathrm{res}(v) = \sum_{p\in X} c_p e_p$ (and interpret $c_0 = 0$). The assignment $v \mapsto (c_p)_{p\in X\setminus\{0\}}$ is injective if and only if res is injective.

3.2 Embedding Properties of res

Throughout the rest of this chapter, we **assume that** X **is countable** and bounded. Scaling the metric ρ by a nonzero factor scales the norm $\|\cdot\|_{\mathrm{LF}(X)}$ by the same factor, hence resulting in an equivalent norm. Thus, we shall additionally **assume that** $\mathrm{diam}(X) = 1$. Of course, all our results hold under the weaker assumption that X is bounded, but with perhaps different constants, which is inconsequential as far as Theorems 1.15 and 3.2 are concerned. Finally, by adding a new point to the space and declaring it the basepoint 0, we may assume that $\rho(0,p)=1$ for every $p \in X \setminus \{0\}$, and hence that $\|\delta_p\|=1$. Under these standing assumptions on X, we characterize the metric properties of X under which res is an isomorphic embedding.

First we need a proposition that should clarify the role of essrad in our study of res.

Proposition 3.1. *For each $p \in X$, $\|e_p\|_{\mathrm{Lip}_{\mathrm{fin}}(X)^*} = \mathrm{essrad}\,(p)$.*

Proof. Let $p \in X$. Let $f \in B_{\mathrm{Lip}_{\mathrm{fin}}(X)}$ and $\epsilon > 0$ be arbitrary. By definition of $\mathrm{essrad}\,(p)$, $|B_{\mathrm{essrad}(p)+\epsilon}(p)| = \infty$. Thus, f must vanish at some point in $B_{\mathrm{essrad}(p)+\epsilon}(p)$. This implies $|f(p)| \leq \mathrm{essrad}\,(p)+\epsilon$. Since $f \in B_{\mathrm{Lip}_{\mathrm{fin}}(X)}$, $\epsilon > 0$ were arbitrary, this in turn implies $\|e_p\|_{\mathrm{Lip}_{\mathrm{fin}}(X)^*} \leq \mathrm{essrad}\,(p)$.

Now consider the function $f : X \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 - \rho(x,p) & x \in B_{\mathrm{essrad}(p)-\epsilon}(p) \\ 0 & x \notin B_{\mathrm{essrad}(p)-\epsilon}(p) \end{cases}$$

Then $f \in \mathrm{Lip}_{\mathrm{fin}}\,(X)$ by definition of $\mathrm{essrad}\,(p)$, $\|f\|_{\mathrm{Lip}_{\mathrm{fin}}(X)} \leq (\mathrm{essrad}\,(p)-\epsilon)^{-1}$, and $e_p(f)=1$. It follows that $\|e_p\|_{\mathrm{Lip}_{\mathrm{fin}}(X)^*} \geq \mathrm{essrad}\,(p)-\epsilon$. Since $\epsilon > 0$ was arbitrary, $\|e_p\|_{\mathrm{Lip}_{\mathrm{fin}}(X)^*} \geq \mathrm{essrad}\,(p)$. □

Theorem 3.1. res *is a θ^{-1}-isomorphic embedding if and only if X is θ'-essentially uniformly discrete for some $\theta' \geq \theta$.*

Proof. Assume X is not θ'-essentially uniformly discrete for all $\theta' \geq \theta$. This implies that there is some $p \in X$ with $\mathrm{essrad}\,(p) < \theta$. Then

$$\|\delta_p\|_{\mathrm{LF}(X)} = 1 > \theta^{-1}\mathrm{essrad}\,(p) \overset{\text{Prop 3.1}}{=} \theta^{-1}\|e_p\|_{\mathrm{Lip}_{\mathrm{fin}}(X)^*} = \theta^{-1}\|\mathrm{res}(\delta_p)\|_{\mathrm{Lip}_{\mathrm{fin}}(X)^*}$$

Now assume X is θ'-essentially uniformly discrete for some $\theta' \geq \theta$. It suffices to assume $\theta' = \theta$. Let $\epsilon > 0$ be arbitrary. Let $F \subseteq X$ be finite and $\sum_{p\in F} c_p\delta_p \in \mathrm{LF}_{\mathrm{fin}}\,(X)$. Let $f \in B_{\mathrm{Lip}_0(F)}$ such that

$$\left| f\left(\sum_{p\in F} c_p\delta_p\right)\right| = \left\|\sum_{p\in F} c_p\delta_p\right\|_{\mathrm{LF}(F)}$$

We may assume that $0 \in F$ so that f vanishes at some point in F. Thus $\|f\|_{L^\infty(F)} \leq \mathrm{diam}(X)\|f\|_{\mathrm{Lip}_0(F)} \leq 1$. We'll now extend f to a function $\tilde{f} \in \mathrm{Lip}_{\mathrm{fin}}\,(X)$ with $\|\tilde{f}\|_{\mathrm{Lip}_{\mathrm{fin}}(X)} \leq (\theta-\epsilon)^{-1}$. For $x \in F$, set $\tilde{f}(x) := f(x)$, and for $x \in X \setminus (\cup_{p\in F} B_{\theta-\epsilon}(p))$, set $\tilde{f}(x) := 0$. Since $\|f\|_{\mathrm{Lip}_0(F)}, \|f\|_{L^\infty(F)} \leq 1$, the Lipschitz constant of $\tilde{f}$ on its domain of definition is $\leq \theta^{-1}$, and $\tilde{f}$ is supported on the finite set F. The set of points where $\tilde{f}$ remains undefined is $\cup_{p\in F}\,(B_{\theta-\epsilon}(p) \setminus \{p\})$, which is finite by definition of θ-essentially uniformly discreteness. We apply the McShane extension theorem to extend $\tilde{f}$ to all of X without increasing the Lipschitz constant. Then the following estimate concludes the proof of the desired implication.

$$\left\|\sum_{p\in F} c_p\delta_p\right\|_{\mathrm{LF}(X)} = \left|f\left(\sum_{p\in F} c_p\delta_p\right)\right| = \left|\tilde{f}\left(\mathrm{res}\left(\sum_{p\in F} c_p\delta_p\right)\right)\right|$$

$$\leq \left\|\tilde{f}\right\|_{\mathrm{Lip_{fin}}(X)}\left\|\mathrm{res}\left(\sum_{p\in F} c_p\delta_p\right)\right\|_{\mathrm{Lip_{fin}}(X)^*} \leq (\theta-\epsilon)^{-1}\left\|\mathrm{res}\left(\sum_{p\in F} c_p\delta_p\right)\right\|_{\mathrm{Lip_{fin}}(X)^*}$$

□

3.3 Separability of $\mathrm{Lip_{fin}}(X)^*$

We recall the (slightly modified) definition of De Leeuw's map, Φ, from [Wea99, Definition 2.1.1]. Let $\Delta \subseteq X \times X$ denote the diagonal and set $\tilde{X} := X \times X \setminus \Delta$. Then $\rho : \tilde{X} \to \mathbb{R}$ is non-vanishing. Let $\ell^\infty(\tilde{X})\rho$ denote the vector space of all real-valued functions on $\tilde{X}$ of the form $f\rho$, where $f \in \ell^\infty(\tilde{X})$. Equip $\ell^\infty(\tilde{X})\rho$ with the unique norm so that $f \mapsto f\rho$ is a linear isometry from $\ell^\infty(\tilde{X})$ onto $\ell^\infty(\tilde{X})\rho$. Define $\Phi : \mathrm{Lip_{fin}}(X) \to \ell^\infty(\tilde{X})\rho$ by $\Phi(f)((x,y)) := f(y) - f(x)$. Then Φ is a linear isometric embedding. The Riesz-Markov representation theorem implies $(\ell^\infty(\tilde{X})\rho)^*$ can be identified with the Banach space of measures μ on $\beta\tilde{X}$ equipped with the norm $\|\mu\| = \int \rho d|\mu|$, denoted $\mathcal{M}\left(\beta\tilde{X}\right)/\rho$. Here, $\beta\tilde{X}$ denotes the Stone-Čech compactification of the discrete topological space $\tilde{X}$, and we've implicitly extended ρ (in the unique way) to a continuous function on $\beta\tilde{X}$ and will continue to do so for all functions in $\ell^\infty(\tilde{X})\rho$. Under this identification, Φ^* maps $\mathcal{M}\left(\beta\tilde{X}\right)/\rho$ onto $\mathrm{Lip_{fin}}(X)^*$, with the action given by $\Phi^*(\mu)(f) = \int(f(y) - f(x))d\mu(x,y)$. As is well-known, $\partial\tilde{X} := \beta\tilde{X}\setminus\tilde{X}$ can be identified with the set of nonprincipal ultrafilters on $\tilde{X}$. We wish to identify special subsets of $\partial\tilde{X}$. For each $p \in X$, let $\mathcal{U}_p$ denote the set of all nonprincipal ultrafilters $\mathcal{U}$ on $\tilde{X}$ such that $\{p\} \times (X \setminus \{p\}) \in \mathcal{U}$ or $(X \setminus \{p\}) \times \{p\} \in \mathcal{U}$. Note that $\mathcal{U}_p$ is closed, and that $\rho(\mathcal{U}_p) \subseteq [\mathrm{essrad}(p), \mathrm{diam}(X)]$, by definition of $\mathrm{essrad}(p)$. Also by definition of $\mathrm{essrad}(p)$, there is a $\mathcal{U}_p^* \in \mathcal{U}_p$ such that $\rho(\mathcal{U}_p^*) = \mathrm{essrad}(p)$ (there could be many; $\mathcal{U}_p^*$ is chosen arbitrarily using AoC). Furthermore, we can, and do, chose $\mathcal{U}_p^*$ so that $\{p\} \times (X \setminus \{p\}) \in \mathcal{U}_p^*$ (and so $(X \setminus \{p\}) \times \{p\} \notin \mathcal{U}_p^*$). $\mathcal{M}\left(\beta\tilde{X}\right)/\rho$ splits into the (internal) ℓ^1-direct sum $\ell^1(\tilde{X})/\rho \oplus_1 \mathcal{M}\left(\cup_{p\in X}\mathcal{U}_p\right)/\rho \oplus_1 \mathcal{M}\left(\partial\tilde{X} \setminus \cup_{p\in X}\mathcal{U}_p\right)/\rho$.

Proposition 3.2. *The following are true.*

1. Φ^* *restricted to* $\ell^1(\tilde{X})/\rho$ *equals* $\mathrm{res} \circ \pi$ *(recall the definition of* $\pi : \ell^1(\tilde{X})/\rho \to \mathrm{LF}(X)$ *from Section 3.1.2).*

2. *For any $p \in X$ and $\mu \in \mathcal{M}(\mathcal{U}_p)/\rho$, $\Phi^*(\mu) \in span(e_p)$. Additionally, $\Phi^*(\delta_{\mathcal{U}_p^*}) = e_p$.*

3. *Φ^* vanishes on $\mathcal{M}\left(\partial\tilde{X} \setminus \cup_{p\in X}\mathcal{U}_p\right)/\rho$.*

Proof. (1) follows immediately from the definitions of Φ and π.

To prove (2), since the span of the point mass measures in $\mathcal{M}(\mathcal{U}_p)/\rho$ is weak*-dense in $\mathcal{M}(\mathcal{U}_p)/\rho$ and Φ^* is weak*-weak* continuous, it suffices to show $\Phi^*(\delta_{\mathcal{U}}) = \pm e_p$ for each $\mathcal{U} \in \mathcal{U}_p$. Let $\mathcal{U} \in \mathcal{U}_p$. We'll assume $\{p\} \times (X \setminus \{p\}) \in \mathcal{U}$ and show $\Phi^*(\delta_{\mathcal{U}}) = e_p$ (a similar argument shows $\Phi^*(\delta_{\mathcal{U}}) = -e_p$ if $(X \setminus \{p\}) \times \{p\} \in \mathcal{U}$). Let $f \in \mathrm{Lip}_{\mathrm{fin}}(X)$. Then $CF := \{(x,y) \in \{p\} \times (X \setminus \{p\}) : f(x) - f(y) = f(p)\}$ is a cofinite subset of $\{p\} \times (X \setminus \{p\})$. Then by definition of nonprincipal ultrafilter, $CF \in \mathcal{U}$. Hence,

$$\Phi^*(\delta_{\mathcal{U}})(f) = f(\mathcal{U}) = \mathcal{U}\text{-}\lim_{(x,y)}(f(x) - f(y)) = f(p) = e_p(f)$$

This shows (2).

Finally we prove (3). By the same density and continuity reasoning as before, it suffices to show $\Phi^*(\delta_{\mathcal{U}}) = 0$ for each $\mathcal{U} \in \partial\tilde{X} \setminus \cup_{p\in X}\mathcal{U}_p$. Let $\mathcal{U} \in \partial\tilde{X} \setminus \cup_{p\in X}\mathcal{U}_p$. Let $p \in X$, and let 1_p denote the indicator function of $\{p\}$. Then $\Phi^*(\delta_{\mathcal{U}})(1_p) = \mathcal{U}\text{-}\lim_{(x,y)}(1_p(x) - 1_p(y))$. For any $(x,y) \in \tilde{X}$, $1_p(x) - 1_p(y) \in \{-1,0,1\}$, and thus $\mathcal{U}\text{-}\lim_{(x,y)}(1_p(x) - 1_p(y)) \in \{-1,0,1\}$. Now, $1_p(x) - 1_p(y) = -1$ if and only if $(x,y) \in (X \setminus \{p\}) \times \{p\}$ and $1_p(x) - 1_p(y) = 1$ if and only if $(x,y) \in \{p\} \times (X \setminus \{p\})$. From this it's clear that $\mathcal{U}\text{-}\lim_{(x,y)}(1_p(x) - 1_p(y)) \notin \{-1,1\}$ since $\mathcal{U} \notin \cup_{p\in X}\mathcal{U}_p$, and thus the only remaining option is $\mathcal{U}\text{-}\lim_{(x,y)}(1_p(x) - 1_p(y)) = 0$. Since $p \in X$ was arbitrary, linearity implies $\Phi^*(\delta_{\mathcal{U}})(f) = 0$ for all $f \in \mathrm{Lip}_{\mathrm{fin}}(X)$. □

Theorem 3.2. *For every $\lambda \in \mathrm{Lip}_{\mathrm{fin}}(X)^*$, there exist $w \in \mathrm{LF}(X)$ and $\sum_{p\in X} c_p e_p \in \mathrm{Lip}_{\mathrm{fin}}(X)^*$ such that $\lambda = \mathrm{res}(w) + \sum_{p\in X} c_p e_p$ and $\|\lambda\| \le \|w\| + \sum_{p\in X} |c_p| \mathrm{essrad}(p)$. Consequently, $\mathrm{Lip}_{\mathrm{fin}}(X)^*$ is separable.*

Proof. Let $\lambda \in \mathrm{Lip}_{\mathrm{fin}}(X)^*$. By Hahn-Banach, there exists $\mu_1 + \mu_2 + \mu_3 \in \ell^1(\tilde{X})/\rho \oplus_1 \mathcal{M}(\cup_{p\in X}\mathcal{U}_p)/\rho \oplus_1 \mathcal{M}\left(\partial\tilde{X} \setminus \cup_{p\in X}\mathcal{U}_p\right)/\rho$ such that $\|\mu_1\| + \|\mu_2\| + \|\mu_3\| = \|\lambda\|$ and $\Phi^*(\mu_1 + \mu_2 + \mu_3) = \lambda$. Note that this is equivalent to saying $\mu_1 + \mu_2 + \mu_3$ has minimal norm among all elements in $(\Phi^*)^{-1}(\lambda)$. This minimal norm property and Proposition 3.2(3) imply $\mu_3 = 0$. Set $c_p := \mu_2(\mathcal{U}_p)$ and $\mu_2' := \sum_{p\in X} c_p \delta_{u_p^*}$. Then Proposition 3.2(2) implies $\Phi^*(\mu_2) = \sum_{p\in X} c_p e_p$ and $\|\mu_2\| = \int \rho d|\mu_2| \ge \sum_{p\in X} |c_p| \mathrm{essrad}(p)$. Set $w := \pi(\mu_1)$, and observe that Proposition 3.2(1) implies $\lambda = \mathrm{res}(w) + \sum_{p\in X} c_p e_p$. □

3.4 Examples

In this section, we construct a family of countable, essentially uniformly discrete metric spaces that do not obviously biLipschitz embed into RNP spaces without the aid of Theorem 1.15.

First we discuss how to construct, in any Banach space V, an essentially uniformly discrete subset that is not uniformly discrete. Let V be a Banach space and $X' \subseteq V$ an infinite subset such that every bounded subset of X' is finite. Let $f : [0, \infty) \to [1, \infty)$ be any function such that

1. $\limsup_{t\to\infty} f(t) = \infty$.
2. $\lim_{t\to\infty} \frac{f(t)}{t} = 0$.

Define the radial stretch map $r : V \to V$ by $r(x) := \frac{x}{f(\|x\|)}$. set $X := r(X')$. (1) and the infinitude of X' imply X is not a uniformly discrete subset of V, and (2) and the finiteness of X' on bounded subsets imply X is ∞-essentially uniformly discrete. In general, different choices of f will result in metric spaces X that are not canonically biLipschitz equivalent.

Now let (X, ρ) be an essentially uniformly discrete, non-uniformly discrete metric space that isometrically embeds into a Banach space V, such as the one constructed above. Let $h : [0, \infty) \to [0, \infty)$ be a function such that

- h is concave and increasing.
- $h(t) > 0$ for $t > 0$.
- $\lim_{t\to 0} h(t) = h(0) = 0$.

Then $h \circ \rho$ is another metric on X that is not biLipschitz equivalent to ρ (via the identity map) unless $t/C \le h(t) \le Ct$ for some $C < \infty$. Transforming the metric by h is not compatible in any obvious way with RNP biLipschitz embeddability. In other words, even if V is an RNP space, it is not clear that the metric space $(V, h \circ \|\cdot\|)$ should biLipschitz embed into any RNP space. However, this metric transform $\rho \mapsto h \circ \rho$ does preserve essentially uniform discreteness, so $(X, h \circ \rho)$ isometrically embeds into an RNP space by Theorem 1.15.

Chapter 4

Markov Convexity of Carnot Groups

4.1 Introduction

This chapter is devoted to proving the results stated in Section 1.2.3.

4.1.1 Outline

We begin with an informal discussion of the main theorems in Section 4.1.2. Section 4.2 contains some preliminary inequalities, and Sections 4.3 and 4.4 contain the proofs of Theorems 4.1 and 4.2, respectively.

4.1.2 Discussion of Proof Methods

We engage here in informal discussion of the proofs of Theorem 4.1 and 4.2. This discussion is intended to give a brief overview of the proofs for readers with a sufficient background in the relevant topics. For Theorem 4.1, the relevant topics are graded nilpotent Lie algebras, the group structure they inherit via the Baker-Campbell Hausdorff formula, and their graded-homogeneous group quasi-norms. For Theorem 4.2, the relevant topics are Markov convexity of diamond-type graphs, jet space Carnot groups, and Khintchine's inequality. Readers unfamiliar with these topic may find this section unuseful.

Discussion of Proof of Theorem 4.1

The method employed by Mendel-Naor to prove that p-convexity of Banach spaces implies Markov p-convexity is to:

1. Invoke the well-known result that p-convex Banach spaces have equivalent norms $\|\cdot\|$ satisfying the parallelogram inequality $(\|x\|^p + \|x-y\|^p)/2 - \|y/2\|^p \gtrsim \|x - y/2\|^p$.

2. Prove 4-point p-convexity: $(2d(y,x)^p + d(z,y)^p + d(y,w)^p)/2 - (d(x,w)/2)^p - (d(x,z)/2)^p \gtrsim d(z,w)^p$, where $d(x,y) = \|x-y\|$.

3. Prove the Markov p-convexity inequality, Definition 1.10.

We prove the analogous inequalities for graded nilpotent Lie groups:

1. Lemma 4.13. Construct a group quasi-norm N satisfying $(N(x)^p + N(y^{-1}x)^p)/2 - (N(y)/2)^p \gtrsim N(\delta_{1/2}(y)^{-1}x)^p$.

2. Lemma 4.14. Prove 4-point p-convexity: $(2d(y,x)^p + d(z,y)^p + d(y,w)^p)/2 - (d(x,w)/2)^p - (d(x,z)/2)^p \gtrsim d(z,w)^p$, where $d(x,y) = N(y^{-1}x)$.

3. Prove Theorem 4.1. The Markov p-convexity inequality.

The passage from (1) to (2) and from (2) to (3) is exactly the same as in Banach space case. To prove (1), we recursively construct a sequence of homogeneous quasi-norms on the group, and prove that they satisfy (1) inductively. Actually, the following stronger version of (1) (with $p = 2s$, the case $p \geq 2s$ is taken care of later) is needed for the induction to close, this is Lemma 4.12.

$$(N_s(x)^{2s} + N_s(y^{-1}x)^{2s})/2 - (N_s(y)/2)^{2s} \gtrsim SN_s(x,y)^{2s} + D_s(x,y) + N_s(\delta_{1/2}(y)^{-1}x)^{2s}$$

There are two extra terms that appear in this inequality, $SN_s(x,y)$ and $D_s(x,y)$, defined in Definitions 4.3 and 4.4. $D_s(x,y)$ is designed to bound (up to constants) the square of any BCH polynomial of degree s (see Definition 4.1), so one may guess how it would be useful to prove (1).

$SN_s(x,y)$ is nearly a positive definite quasi-norm of $(x_1, \ldots x_s, y_1, \ldots y_s)$ (the name SN is meant to suggest that it is a seminorm instead of a norm, since it is not positive definite), but not quite as it vanishes when $x_1 = y_1/2$ and $x_i = y_i = 0$ for $i \geq 2$. However, this is not an issue as we will have an extra $\|y_1\|$ term in the induction, so that $\|y_1\| + SN_s(x,y)$ is genuinely a quasi-norm of $(x_1, \ldots x_s, y_1, \ldots y_s)$. Here are D_s and SN_s for some small s:

$$D_3(x,y) = \|(x_3,y_3)\|^2 + \|(x_1,y_1)\|^2\|(x_2,y_2)\|^2 + \|(x_1,y_1)\|^2\boldsymbol{\tau}^2(x,y)$$

$$D_4(x,y) = \|(x_4,y_4)\|^2 + \|(x_1,y_1)\|^2\|(x_3,y_3)\|^2 + \|(x_2,y_2)\|^4$$

$$+\|(x_1,y_1)\|^4\|(x_2,y_2)\|^2 + \|(x_2,y_2)\|^2\boldsymbol{\tau}^2(x,y) + \|(x_1,y_1)\|^4\boldsymbol{\tau}^2(x,y)$$

$$SN_3(x,y) = \max(\|x_1 - y_1/2\|, \|(x_2,y_2)\|^{1/2}, \|(x_3,y_3)\|^{1/3})$$

The polynomial $\boldsymbol{\tau}^2(x,y)$ is designed to bound the squares of terms coming from the bracket between two vectors from the horizontal layer. For example, in the second Heisenberg group,

$$\boldsymbol{\tau}^2(x,y) = (x_{11}y_{12} - x_{12}y_{11})^2 + (x_{13}y_{14} - x_{14}y_{13})^2$$

We recursively construct the quasi-norms N_{s+1} given all the previous quasi-norms by defining $N_{s+1}(x)$ to be an $\ell^{2(s+1)}$ sum of $\lambda_{s+1}\|x_{s+1}\|^{1/(s+1)}$ and the top half of the previously defined quasi-norms, where λ_{s+1} is a positive constant chosen small enough (depending on the product structure of the group in question) to make the inequality of Lemma 4.12(1) hold. Specifically, from (4.1),

$$N_2(x) = \sqrt[4]{\|x_1\|^4 + \lambda_2\|x_2\|^2}$$

$$N_{s+1}(x) = \sqrt[2(s+1)]{\lambda_{s+1}\|x_{s+1}\|^2 + \sum_{s'=\lceil (s+1)/2\rceil}^{s} N_{s'}^{2(s+1)}(x)}$$

The reason why we add the top half of the previously defined norms, and the reason for the inclusion $SN_s(x, y)$ term in the inequality, is to help pass from $D_s(x, y)$ to $D_{s+1}(x, y)$ during the proof of the inductive step. When proving the inductive step, we have terms like
$(SN_{s'}(x, y)^{2s'} + D_{s'}(x, y))^{(s+1)/s'}$, $s' \le s$, appearing to which we apply Lemma 4.1 and obtain a term like $SN_{s'}(x, y)^{2(s+1-s')}D_{s'}(x, y)$. This term bounds $\|(x_{s+1-s'}, y_{s+1-s'})\|^2 D_{s'}(x, y)$ exactly when $\lceil (s + 1)/2\rceil \le s' \le s$. Then summing $\|(x_{s+1-s'}, y_{s+1-s'})\|^2 D_{s'}(x, y)$ over this range of s' accounts for all the terms in $D_{s+1}(x, y)$, except for the top-layer term $\|(x_{s+1}, y_{s+1})\|^2$ (since any other term in $D_{s+1}(x, y)$ contains as a factor a variable from one of the lower half layers, see Lemma 4.7 for details), which is accounted for later.

Discussion of Proof of Theorem 4.2

We recursively construct a sequence of directed graphs Γ_m and maps from them into the jet space of step r ($J^{r-1}(\mathbb{R})$) to show that it is not Markov p-convex for any $p < 2r$. The Markov processes we use are standard directed random walks on the graphs. This is very similar to the method used in [Li16], where something akin to the Laakso-Lang-Plaut diamond graphs were used. The main feature of those graphs G_m is that G_{m+1} is obtained from G_i by replaced each edge of G_1 with a copy of G_m. Roughly speaking, Li recursively maps G_{m+1} into $\mathbb{R}^2$ by replacing each edge of a distorted image of G_1 by a rotated, distorted copy of the image of G_i. The distortion is done in such a way that the coLipschitz constant (the Lipschitz constant of the inverse map) is on the order of $\sqrt[4]{m}\sqrt{\ln(m + 1)}$, and the fact that rotations are isometries of the Heisenberg group affords one uniform control on the Lipschitz constants. One can conclude from this that the Heisenberg group is not Markov p-convex for $p < 4$ (the 4 coming from the fourth root of m).

Our graphs differ from those in [Li16] in that, to obtain Γ_{m+1} from Γ_m, we first glue together *many* copies of Γ_m together with a small number of copies of a single edge I in series to get a new graph Γ'_{m+1}, and then replace each edge of Γ_1 with a copy of Γ'_{m+1} (this isn't exactly how our construction is defined, but is close enough to get the main idea). See Definition 4.5 for the full details. We will explain the reasoning for this after describing our maps of Γ_m into $J^{r-1}(\mathbb{R})$.

Our maps differ from those in [Li16] in that we *do not* rotate the image of Γ_m before using it to replace the edges of the image of Γ_1, as rotations are not Lipschitz maps in higher step groups like they are in the Heisenberg group. Refer to Figure 4.2 throughout this discussion to get an idea of the construction of these maps. Instead of rotating, we simply add (many copies of) the image of Γ_m to a distorted copy of the image of Γ_1 to obtain the mapping of Γ_{m+1} into $\mathbb{R}^2$. More specifically, we map each directed path γ in Γ_{m+1} to the jet of a function ϕ_γ - a horizontal curve in $J^{r-1}(\mathbb{R})$. The Lipschitz constant of this map is controlled by $\left\|\frac{d^r}{d^r x}\phi_\gamma\right\|_\infty$. We still distort the graphs Γ_m with the same asymptotics as in [Li16], so that the coLipschitz constant is on the order of $\sqrt[2r]{m}\sqrt[r]{\ln(m+1)}$ (at least on the pairs of random walks $(X_t^m, \tilde{X}_t^m(t-2^k))$. That we get the $2r^{\text{th}}$ root of m instead of the fourth root of m comes from the fact that $J^{r-1}(\mathbb{R})$ is of step r and the Heisenberg group is of step 2. One potential problem is that the absence of isometric rotations and the fact that $(\sqrt{m}\ln(m))^{-1}$ isn't summable means $\left\|\frac{d^r}{d^r x}\phi_\gamma\right\|_\infty$ blows up along some paths, and thus we do not have uniform control on the Lipschitz constant of the map, unlike [Li16]. However, $(\sqrt{m}\ln(m))^{-1}$ *is square-summable*, and together with the nature of the image of the random walk X_t^m in $J^{r-1}(\mathbb{R})$, this allows us to control $\mathbb{E}[d_{CC}(X_{t+1}^m, X_t^m)^p]$ uniformly in m, t. Loosely, along the random walk in the horizontal layer (which has x- and u_{r-1}-coordinates), every time one is confronted with a choice of direction to walk in, the choice is to walk 1 unit in the x-direction and $+(\sqrt{i}\ln(i+1))^{-1}$ units in the u_{r-1}-direction with probability 1/2, or 1 unit in the x-direction and $-(\sqrt{i}\ln(i+1))^{-1}$ units in the u_{r-1}-direction with probability 1/2 (for some i depending on how far one has walked). Thus, one might expect $d_{CC}(X_{t+1}^m, X_t^m)$ to be bounded by a random variable distributed like $1 + |\sum_{i=1}^t \epsilon_i(\sqrt{i}\ln(i+1))^{-1}|$, where $\{\epsilon_i\}_i$ are iid Rademachers, and then Khintchine's inequality implies we should have a uniform bound on $\mathbb{E}[d_{CC}(X_{t+1}^m, X_t^m)^p]$ (which is the real quantity of interest, recall Definition 1.10). Of course, the random walk is not distributed like this, but it turns out that this intuition is correct nonetheless, see Lemmas 4.4 and 4.18(4) for the specifics.

Finally, the reason we use many copies of Γ_m in creating Γ_{m+1} is so that, compared to the diameter of Γ_{m+1}, the diameter of the copies of Γ_m is very small, and thus those that replaced opposite edges of Γ_1 don't get too close together, which would ruin the coLipschitz constant. Morally, this "decouples" any interaction between different scales in Γ_{m+1}.

4.2 Probabilistic and Convexity Inequalities

In this article, we will often justify an inequality with the phrase "by convexity" or "by the parallelogram law". The convexity inequalities we refer to are almost always of the form

$$\frac{a^p + b^p}{2} \geq \left(\frac{a+b}{2}\right)^p$$

or

$$a^p + b^p \leq (a+b)^p$$

for $p \geq 1$ and $a, b \geq 0$. The form of the parallelogram law we most often use is

$$\frac{\|u\|^2 + \|u-v\|^2}{2} = \|v/2\|^2 + \|u - v/2\|^2$$

for u, v in a Hilbert space, which implies the inequality

$$\|u\|^2 + \|u-v\|^2 \geq \frac{\|v\|^2}{2}$$

We may also use either of these inequalities without explicitly mentioning convexity or the parallelogram law.

We collect here some basic inequalities related to convexity and an additional one on L^p-norms of random variables.

Lemma 4.1. *For all $a, b \geq 0$ and $q \geq 1$,*

$$(a+b)^q \geq a^q + qa^{q-1}b$$

Proof. Let a, b, q be as above. The inequality is obviously true if $a = 0$. Then if $a > 0$, after dividing each side by a^q and replacing b/a with t, it suffices to prove $(1+t)^q \geq 1 + qt$. This inequality is true since the right hand is the linearization of the left hand side at $t = 0$, and the left hand side is a convex function of t. □

Lemma 4.2. *For each $p > 0$ and $k \geq 1$,*

$$\sum_{t=1}^{k} (2t)^p > k^{p+1}/2$$

Proof. Let $p > 0$ and $k \geq 1$. Since the function $t \mapsto (2t)^p$ is increasing,

$$\sum_{t=1}^{k}(2t)^p > \int_0^k (2t)^p dt = \frac{2^p}{p+1}k^{p+1} \geq k^{p+1}/2$$

□

The following two lemmas are frequently used in tandem to prove Khintchine's inequality (for example, Proposition 4.5 of [Wol03]). We will need them for a similar inequality used in Section 4.4.2.

Lemma 4.3. *For all* $y \in \mathbb{R}$, $\cosh(y) \leq \exp(y^2/2)$.

Proof. Let $y \in \mathbb{R}$.

$$\cosh(y) = \frac{e^y + e^{-y}}{2} = \frac{1}{2}\sum_{k=0}^{\infty}\frac{y^k + (-y)^k}{k!} = \sum_{k=0}^{\infty}\frac{y^{2k}}{(2k)!} \leq \sum_{k=0}^{\infty}\frac{(y^2/2)^k}{k!} = \exp\left(\frac{y^2}{2}\right)$$

□

Lemma 4.4. *For each* $p \geq 1$ *and* $0 < A, B < \infty$, *there is a constant* $C = C(p, A, B) < \infty$ *such that any real-valued random variable* Y *satisfying the moment generating function subgaussian bound*

$$\mathbb{E}[\exp(yY)] \leq Ae^{By^2}$$

also satisfies the L^p*-norm bound*

$$\mathbb{E}[|Y|^p] \leq C$$

Proof. This is a standard result from the theory of subgaussian random variables whose proof appears in any text on measure concentration. For the sake of completeness we'll include the proof, roughly following the proof of Proposition 4.5 from [Wol03]. Let p, A, B, Y be as above. For any $t > 0$, Markov's inequality and our assumption imply

$$\mathbb{P}(Y \geq t) = \mathbb{P}\left(\exp\left(\frac{t}{2B}Y\right) \geq \exp\left(\frac{t^2}{2B}\right)\right) \leq \exp\left(-\frac{t^2}{2B}\right)\mathbb{E}\left[\exp\left(\frac{t}{2B}Y\right)\right]$$

$$\leq A\exp\left(-\frac{t^2}{2B} + \frac{t^2}{4B}\right) = A\exp\left(-\frac{t^2}{4B}\right)$$

Likewise,

$$\mathbb{P}(Y \leq -t) \leq A\exp\left(-\frac{t^2}{4B}\right)$$

giving us

$$\mathbb{P}(|Y| \geq t) \leq 2A \exp\left(-\frac{t^2}{4B}\right)$$

We then use the layer cake principle to calculate $\mathbb{E}[|Y|^p]$:

$$\mathbb{E}[|Y|^p] = p\int_0^\infty t^{p-1}\mathbb{P}(|Y| \geq t)dt \leq p\int_0^\infty t^{p-1}2A\exp\left(-\frac{t^2}{4B}\right)dt = C(p, A, B) < \infty$$

□

4.3 Upper Bound on Markov Convexity of Graded Nilpotent Lie Groups

Throughout this section, fix a graded nilpotent Lie algebra $(\mathfrak{g}, [\cdot, \cdot])$ of step $r \geq 2$ with grading $\oplus_{i=1}^r \mathfrak{g}_i$ and $\dim(\mathfrak{g}_i) = k_i$. Choose an ordered basis $U_{i,1}, \ldots U_{i,k_i}$ for each $\mathfrak{g}_i$ and equip $\mathfrak{g}$ with a Hilbert norm $\|\cdot\|$ such that these vectors form an orthonormal basis. We also use $\|\cdot\|$ to denote the Euclidean norm on any $\mathbb{R}^n$. Given $x \in \mathfrak{g}$, let $x_i \in \mathfrak{g}_i$ denote its $\mathfrak{g}_i$-component. Given $x_i \in \mathfrak{g}_i$, let $x_{i,j} \in \mathbb{R}$ denote its $U_{i,j}$-component. Thus,

$$\|x\|^2 = \sum_{i=1}^r \|x_i\|^2 \qquad \text{and} \qquad \|x_i\|^2 = \sum_{j=1}^{k_i} |x_{i,j}|^2$$

Consider $\mathfrak{g}$ as a graded nilpotent Lie group as in Section 1.3.3. It's easy to see that 0 is the group identity element and $x^{-1} = -x$. Whenever $u, v \in \mathfrak{g}$ or $u, v \in \mathbb{R}^n$, we use the notation $\|(u, v)\|^2$ to mean $\|u\|^2 + \|v\|^2$.

4.3.1 BCH Polynomials

Definition 4.1. For $s \geq 0$, a function $P : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ that is a monomial(polynomial) in the variables $x_{n,m}, y_{n,m}$ is a *graded-homogeneous monomial(polynomial) of degree* s if $P(\delta_t(x), \delta_t(y)) = t^s P(x, y)$ for all $x, y \in \mathfrak{g}$ and $t \in \mathbb{R}_{>0}$. Clearly, any graded-homogeneous polynomial of degree s must be a sum of graded-homogeneous monomials of degree s.

In this section, a *multiset* is a finite sequence of positive integers modulo permutations. Disjoint unions $I_1 \sqcup I_2$ of multisets are defined in the obvious way. Given a multiset I, $\|I\|_1$ denotes the sum of the elements and $\|I\|_\infty$ the maximum of the elements. Given a nonzero graded-homogeneous monomial M of degree s, we associate to it a multiset $I(M)$ defined recursively on the number of variables in the monomial by $I(M) = \{i\} \sqcup I(M')$ if $M(x, y) = x_{i,n}M'(x, y)$ or $M(x, y) = y_{i,n}M'(x, y)$ for some $n \leq k_i$ and graded-homogeneous polynomial M' of degree $s - i$ (the base case is $I(1) = \emptyset$). By the homogeneity property, it must hold that if M is nonzero and graded-homogeneous of degree s, $\|I(M)\|_1 = s$.

For $s \geq 1$, let $\mathbf{1}_s$ denote the unique multiset with $\|\mathbf{1}_s\|_1 = s$ and $\|\mathbf{1}_s\|_\infty = 1$ (and $\mathbf{1}_0 = \emptyset$). For each $n, m \leq k_1$, let $\tau_{n,m}(x,y) := x_{1,n}y_{1,m} - x_{1,m}y_{1,n}$. A graded-homogeneous polynomial P of degree $s \geq 2$ is of *τ-type* if $P(x,y) = \tau_{n,m}(x,y)M'(x,y)$ for some $n, m \leq k_1$ and graded-homogeneous monomial M' with $I(M') = \mathbf{1}_{s-2}$.

A graded homogeneous polynomial of degree $s \geq 2$ of the form $\sum_j Q_j$ (the sum is finite), where each Q_j is of τ-type or a graded-homogeneous monomial of degree s with $1 < \|I(Q_j)\|_\infty < s$ is called a *BCH polynomial of degree s*.

Remark 4.1. Obviously a sum of BCH polynomials of degree s is another such polynomial. If P is a BCH polynomial of degree s, $1 \leq i \leq r$, and $1 \leq j \leq k_i$, $x_{i,j}P(x,y)$ and $y_{i,j}P(x,y)$ are BCH polynomials of degree $s+i$. If $P(x,y)$ is a BCH polynomial of degree s, then so is $P(x, \delta_t(y))$ for any $t \in \mathbb{R}_{>0}$.

Example 4.1. Let $M(x,y) = 6x_{1,6}x_{1,1}^2y_{4,3}, P(x,y) = -y_{1,2}(x_{1,1}y_{1,2} - x_{1,2}y_{1,1})$, $Q(x,y) = x_{1,1}y_{1,1}$, and $R(x,y) = y_{3,2}$. M is a graded-homogeneous monomial of degree 7 with $I(M) = \{1,1,1,4\}$, P is a graded homogeneous polynomial of degree 3 of τ-type, Q is a graded-homogeneous monomial of degree 2 with $I(Q) = \{1,1\}$, and R is a graded homogeneous monomial of degree 3 with $I(r) = \{3\}$. M and P are BCH polynomials, but Q and R are not because they are monomials with $\|I(Q)\|_\infty = 1$ and $\|I(R)\|_\infty = \|I(R)\|_1$.

We now arrive at a key structural lemma for the group product on graded nilpotent Lie algebras. The rest of this subsection is dedicated to its proof.

Lemma 4.5. *For all $x, y \in \mathfrak{g}$ and $2 \leq s \leq r$,*

$$
\begin{array}{lrcl}
(1) & (y^{-1}x)_1 & = & x_1 - y_1 \\
(2) & (y^{-1}x)_s & = & x_s - y_s + \sum_{j=1}^{k_s} P_{s,j}(x,y)U_{s,j}
\end{array}
$$

where each $P_{s,j}$ is a BCH polynomial of degree s.

A trusting reader familiar with the group structure of graded nilpotent Lie algebras induced by the Baker-Campbell-Hausdorff formula may safely skip the rest of this subsection. Before proving the lemma, we need to set some useful notation that allows us to work with nested Lie brackets, and then prove a lemma about these brackets.

Definition 4.2. Given $x, y \in \mathfrak{g}$, $i \geq 1$, and $\epsilon \in \{1,2\}^i$, we recursively define $(x,y)^\epsilon$ as follows: for $i = 1$, $(x,y)^\epsilon := x$ if $\epsilon = 1$ and $(x,y)^\epsilon := y$ if $\epsilon = 2$. Assume $(x,y)^\epsilon$ has been defined for all $\epsilon \in \{1,2\}^i$ for some $i \geq 1$. Let $\epsilon \in \{1,2\}^{i+1}$. Then ϵ equals $(1,\epsilon')$ or $(2,\epsilon')$ for some $\epsilon' \in \{1,2\}^i$. We define $(x,y)^\epsilon := [x,(x,y)^{\epsilon'}]$ if $\epsilon = (1,\epsilon')$ and $(x,y)^\epsilon := [y,(x,y)^{\epsilon'}]$ if $\epsilon = (2,\epsilon')$.

Example 4.2. $(x, y)^{(1,2,2,1)} = [x, [y, [y, x]]]$. The 1 or 2 in the superscript should be thought of as indicating the first or second component of (x, y) in the nested Lie bracket.

Lemma 4.6. *For all $x, y \in \mathfrak{g}$, $2 \le i_1, i_2 \le r$, and $\epsilon \in \{1, 2\}^{i_1}$,*

$$((x, y)^{\epsilon})_{i_2} = \sum_{j=1}^{k_{i_2}} Q_{i_2,j}(x, y) U_{i_2,j}$$

where each $Q_{i_2,j}$ is a BCH polynomial of degree i_2 if $i_1 \le i_2$ 0 if $i_1 > i_2$.

Proof. Let $x, y \in \mathfrak{g}$. By the grading property, $((x, y)^{\epsilon})_{i_2} = 0$ if $\epsilon \in \{1, 2\}^{i_1}$ and $i_1 > i_2$. We'll prove the remaining case by induction on i_1.

Proof of base case. The base case is $i_1 = 2$. Let $\epsilon \in \{1, 2\}^2$. Then ϵ equals $(1, 1)$, $(1, 2)$, $(2, 1)$, or $(2, 2)$. Since $(x, y)^{(1,1)} = (x, y)^{(2,2)} = 0$ and $(x, y)^{(2,1)} = -(x, y)^{(1,2)}$, it suffices to only consider $\epsilon = (1, 2)$, in which case $(x, y)^{\epsilon} = [x, y]$. Let $i_2 \ge 2$. We treat the two cases $i_2 = 2$ and $i_2 > 2$. First assume $i_2 = 2$. Then we have

$$[x, y]_2 = [x_1, y_1] = \left[\sum_{j=1}^{k_1} x_{1,j} U_{1,j}, \sum_{j'=1}^{k_1} y_{1,j'} U_{1,j'}\right] = \sum_{j=1}^{k_1}\sum_{j'=1}^{k_1} x_{1,j} y_{1,j'} \left[U_{1,j}, U_{1,j'}\right]$$

$$= \frac{1}{2}\left(\sum_{n,m=1}^{k_1} x_{1,n} y_{1,m} \left[U_{1,n}, U_{1,m}\right] + \sum_{n,m=1}^{k_1} x_{1,m} y_{1,n} \left[U_{1,m}, U_{1,n}\right]\right)$$

$$= \frac{1}{2}\sum_{n,m=1}^{k_1} (x_{1,n} y_{1,m} - x_{1,m} y_{1,n}) \left[U_{1,n}, U_{1,m}\right] = \frac{1}{2}\sum_{n,m=1}^{k_1} \tau_{n,m}(x, y) \left[U_{1,n}, U_{1,m}\right]$$

$$= \frac{1}{2}\sum_{n,m=1}^{k_1} \tau_{n,m}(x, y) \sum_{j=1}^{k_2} c_{j,n,m} U_{2,j} = \sum_{j=1}^{k_2}\left(\sum_{n,m=1}^{k_1} \frac{c_{j,n,m}}{2} \tau_{n,m}(x, y)\right) U_{2,j}$$

for some $c_{j,n,m} \in \mathbb{R}$. The inner sum is a sum of polynomials of degree 2 of τ-type, and thus a BCH polynomial of degree i_2.

Now we consider the case $i_2 > 2$.

$$[x, y]_{i_2} = \sum_{n=1}^{i_2-1} [x_n, y_{i_2-n}] = \sum_{n=1}^{i_2-1}\left[\sum_{j=1}^{k_n} x_{n,j} U_{n,j}, \sum_{j'=1}^{k_{i_2-n}} y_{i_2-n,j'} U_{i_2-n,j'}\right]$$

$$= \sum_{n=1}^{i_2-1}\sum_{j=1}^{k_n}\sum_{j'=1}^{k_{i_2-n}} x_{n,j} y_{i_2-n,j'} \left[U_{n,j}, U_{i_2-n,j'}\right] = \sum_{n=1}^{i_2-1}\sum_{j=1}^{k_n}\sum_{j'=1}^{k_{i_2-n}} x_{n,j} y_{i_2-n,j'} \sum_{m=1}^{k_{i_2}} c_{m,n,j,j'} U_{i_2,m}$$

$$= \sum_{m=1}^{k_{i_2}}\left(\sum_{n=1}^{i_2-1}\sum_{j=1}^{k_n}\sum_{j'=1}^{k_{i_2-n}} c_{m,n,j,j'} x_{n,j} y_{i_2-n,j'}\right) U_{i_2,m}$$

for some $c_{m,n,j,j'} \in \mathbb{R}$. Notice that, for each n, j, j', $I(x_{n,j} y_{i_2-n,j'}) = \{n, i_2 - n\}$, and so since $i_2 > 2$ and $1 \leq n \leq i_2 - 1$, $1 < \|I(x_{n,j} y_{i_2-n,j'})\|_\infty < i_2$, and thus $x_{n,j}, y_{i_2-n,j'}$ is a BCH polynomial of degree i_2. This completes the proof of the base case.

Proof of inductive step. Now assume the lemma holds for some $2 \leq i_1 < r$. Let $\epsilon \in \{1,2\}^{i_1+1}$. Then ϵ equals $(1, \epsilon')$ or $(2, \epsilon')$ for some $\epsilon' \in \{1,2\}^{i_1}$. Without loss of generality, assume $\epsilon = (1, \epsilon')$. Let $i_2 \geq i_1 + 1$. Then

$$\begin{aligned}
((x,y)^\epsilon)_{i_2} = [x, (x,y)^{\epsilon'}]_{i_2} &= \sum_{n=1}^{i_2-1} [x_n, ((x,y)^{\epsilon'})_{i_2-n}] \\
&\overset{\text{ind hyp}}{=} \sum_{n=1}^{i_2-1} \left[\sum_{j=1}^{k_n} x_{n,j} U_{n,j}, \sum_{j'=1}^{k_{i_2-n}} P_{i_2-n,j'}(x,y) U_{i_2-n,j'} \right] \\
&= \sum_{n=1}^{i_2-1} \sum_{j=1}^{k_n} \sum_{j'=1}^{k_{i_2-n}} x_{n,j} P_{i_2-n,j'}(x,y) \, [U_{n,j}, U_{i_2-n,j'}] \\
&= \sum_{n=1}^{i_2-1} \sum_{j=1}^{k_n} \sum_{j'=1}^{k_{i_2-n}} x_{n,j} P_{i_2-n,j'}(x,y) \sum_{m=1}^{k_{i_2}} c_{m,n,j,j'} U_{i_2,m} \\
&= \sum_{m=1}^{k_{i_2}} \left(\sum_{n=1}^{i_2} \sum_{j=1}^{k_n} \sum_{j'=1}^{k_{i_2-n}} c_{m,n,j,j'} x_{n,j} P_{i_2-n,j'}(x,y) \right) U_{i_2,m}
\end{aligned}$$

for some $c_{m,n,j,j'} \in \mathbb{R}$ and BCH polynomials $P_{i_2-n,j',\ell}$ of degree $i_2 - n$. This implies $x_{n,j} P_{i_2-n,j'}(x,y)$ is a BCH polynomial of degree i_2, as desired. $\square$

Proof of Lemma 4.5. The Baker-Campbell-Hausdorff formula, (1.1), implies that there are constants (many can be taken to be 0) $\{\alpha_\epsilon\}_{\epsilon \in \cup_{i=2}^r \{1,2\}^i} \subseteq \mathbb{R}$ such that

$$y^{-1}x = x - y + \sum_{i=2}^{r} \sum_{\epsilon \in \{1,2\}^i} \alpha_\epsilon (x,y)^\epsilon$$

Since

$$(y^{-1}x)_i = x_i - y_i + \sum_{i=2}^{r} \sum_{\epsilon \in \{1,2\}^i} \alpha_\epsilon ((x,y)^\epsilon)_i$$

the desired conclusion follows by appealing to Lemma 4.6. $\square$

4.3.2 Convex Metrics

The goal of this subsection is to prove Theorem 4.1. To do so, we construct a left invariant homogeneous quasi-metric on $\mathfrak{g}$ that satisfies a certain 4-point inequality. This is the content of Lemma 4.14. All the lemmas and definitions preceding Lemma 4.14 exist to prove it.

We next define a graded-homogeneous polynomial of degree $2s$ that dominates the square of any BCH polynomial of degree s, Lemma 4.7. As a consequence of this we get two domination inequalities involving norms of group products, Lemmas 4.9 and 4.10. These types of domination are what will ultimately allow us to prove Lemma 4.12, the key lemma used in the proof of Lemma 4.14.

Definition 4.3. Let

$$\boldsymbol{\tau}(x,y) := \sqrt{\sum_{n,m\le k_1} \tau_{n,m}^2(x,y)}$$

so that $\boldsymbol{\tau}(x,y)^2 \ge \tau_{n,m}(x,y)^2$ for every n and m. For each $2 \le s \le r$, define $D_s : \mathfrak{g}\times\mathfrak{g} \to \mathbb{R}_{\ge 0}$ recursively by

$$D_2(x,y) := \boldsymbol{\tau}^2(x,y) + \|(x_2,y_2)\|^2$$

$$D_{s+1} := \|(x_{s+1},y_{s+1})\|^2 + \sum_{s'=1}^{\lfloor (s+1)/2\rfloor} \|(x_{s'},y_{s'})\|^2 D_{s+1-s'}(x,y)$$

Lemma 4.7. *For any $2 \le s \le r$ and BCH polynomial P of degree s, there exists $0 < c \le 1$ such that for all $x, y \in \mathfrak{g}$,*

$$D_s(x,y) - \|(x_s,y_s)\|^2 \ge cP^2(x,y)$$

Proof. The proof is by induction on s. The base case $s = 2$ is clear from the definition of D_2 and BCH polynomial of degree 2. Assume the inequality holds for all $s_0 \le s$ for some $s < r$. Let P be a BCH polynomial of degree $s+1$. By definition of BCH polynomial, it suffices to prove the inequality assuming P is a monomial with $1 < \|I(P)\|_\infty < s+1$ or P is of τ-type. First assume P is a monomial with $1 < \|I(P)\|_\infty < s+1$. There are two subcases to consider: $1 \in I(P)$ and $1 \notin I(P)$. Assume the first subcase holds. Then $P = x_{1,n}M(x,y)$ or $P = y_{1,n}M(x,y)$ for some $n \le k_1$ and monomial M of degree s with $1 < \|I(M)\|_\infty < s+1$. Then

$$D_{s+1}(x,y) - \|(x_{s+1},y_{s+1})\|^2 = \sum_{s'=1}^{\lfloor (s+1)/2\rfloor} \|(x_{s'},y_{s'})\|^2 D_{s+1-s'}(x,y)$$

$$\ge \|(x_1,y_1)\|^2 D_s(x,y) \overset{\text{ind hyp}}{\ge} c\|(x_1,y_1)\|^2 M^2(x,y) \ge cP^2(x,y)$$

Now assume the second subcase holds. Then $P(x,y) = x_{i,j}M(x,y)$ or $P(x,y) = y_{i,j}M(x,y)$ for some $1 < i \le \lfloor (s+1)/2\rfloor$, $j \le k_i$, and monomial M of degree $s+1-i$ with $1 < \|I(M)\|_\infty < s+1$. Then

$$D_{s+1}(x,y) - \|(x_{s+1},y_{s+1})\|^2 = \sum_{s'=1}^{\lfloor (s+1)/2\rfloor} \|(x_{s'},y_{s'})\|^2 D_{s+1-s'}(x,y)$$

$$\ge \|(x_i,y_i)\|^2 D_{s+1-i}(x,y) \overset{\text{ind hyp}}{\ge} c\|(x_i,y_i)\|^2 M^2(x,y) \ge cP^2(x,y)$$

Now assume P is of τ-type. By definition, since P has degree $s+1$, this means $P(x,y) = \tau_{n,m}(x,y)M'(x,y)$ for some $n,m \leq k_1$ and graded-homogeneous monomial M' with $I(M') = \mathbf{1}_{s-1}$. This implies $P(x,y) = x_{1,\ell}P'(x,y)$ or $P(x,y) = y_{1,\ell}P'(x,y)$ for some $\ell \leq k_1$, and degree s polynomial P' of τ-type. Then

$$D_{s+1}(x,y) - \|(x_{s+1}, y_{s+1})\|^2 = \sum_{s'=1}^{\lfloor (s+1)/2 \rfloor} \|(x_{s'}, y_{s'})\|^2 D_{s+1-s'}(x,y)$$

$$\geq \|(x_1,y_1)\|^2 D_s(x,y) \overset{\text{ind hyp}}{\geq} c\|(x_1,y_1)\|^2 (P')^2(x,y) \geq cP^2(x,y)$$

□

Lemma 4.8. *Let $2 \leq s \leq r$. For any $t > 0$, there is a constant $c > 0$ such that for all $x, y \in \mathfrak{g}$,*

$$D_s(x,y) - \|(x_s,y_s)\|^2 \geq c\|(\delta_t(y)^{-1}x)_s - (x_s - t^s y_s)\|^2$$

Proof. Let $t > 0$. By Lemma 4.5,

$$\|(\delta_t(y)^{-1}x)_s - (x_s - t^s y_s)\|^2 \overset{\text{Lem 4.5}}{=} \sum_j |P_{s,j}(x, \delta_t(y))|^2 = \sum_j |P'_{s,j,t}(x,y)|^2$$

where each $P_{s,j}$ is a BCH polynomial of degree s, and by Remark 4.1, each $P'_{s,j,t}$ is a BCH polynomial of degree s. Then the desired inequality follows from Lemma 4.7. □

Lemma 4.9. *Let $2 \leq s \leq r$ and $c > 0$. For all sufficiently small $\lambda > 0$ (depending on c), for all $x, y \in \mathfrak{g}$,*

$$c(D_s(x,y) - \|(x_s,y_s)\|^2) + \lambda\|(y^{-1}x)_s\|^2 \geq \frac{\lambda}{2}\|x_s - y_s\|^2$$

Proof. Let $\lambda > 0$. By Lemma 4.8, there is a constant $c' > 0$ (independent of x, y) such that

$$c(D_s(x,y) - \|(x_s,y_s)\|^2) + \lambda\|(y^{-1}x)_s\|^2 \geq c'\|(y^{-1}x)_s - (x_s - y_s)\|^2 + \lambda\|(y^{-1}x)_s\|^2 =: (*)$$

Thus, if $\lambda \leq c'$,

$$(*) \geq \lambda\|(y^{-1}x)_s - (x_s - y_s)\|^2 + \lambda\|(y^{-1}x)_s\|^2 \geq \frac{\lambda}{2}\|x_s - y_s\|^2$$

where the last inequality follows from the parallelogram law. □

Lemma 4.10. *Let $2 \leq s \leq r$. There is a constant $c > 0$ such that for all $x, y \in \mathfrak{g}$,*

$$D_s(x,y) \geq c\|(\delta_{1/2}(y)^{-1}x)_s\|^2$$

Proof. By Lemma 4.8, it suffices to show

$$\|(\delta_{1/2}(y)^{-1}x)_s - (x_s - 2^{-s}y_s)\|^2 + \|(x_s, y_s)\|^2 \geq \frac{1}{8}\|(\delta_{1/2}(y)^{-1}x)_s\|^2$$

Since

$$\|x_s - 2^{-s}y_s\|^2 \leq 4\|(x_s, y_s)\|^2$$

it suffices to show

$$\|(\delta_{1/2}(y)^{-1}x)_s - (x_s - 2^{-s}y_s)\|^2 + \frac{1}{4}\|x_s - 2^{-s}y_s\|^2 \geq \frac{1}{8}\|(\delta_{1/2}(y)^{-1}x)_s\|^2$$

This inequality is true by the parallelogram law. □

Lemma 4.11. *There is a constant $c > 0$ such that for all $x, y \in \mathfrak{g}$,*

$$\|y_1\|\|x_1 - y_1/2\| \geq c\boldsymbol{\tau}(x, y)$$

Proof. It suffices to show, for each fixed $n, m \leq k_1$, $\|y_1\|\|x_1 - y_1/2\| \geq |\tau_{n,m}(x, y)|$. By Cauchy-Schwarz,

$$\|y_1\|\|x_1 - y_1/2\| \geq \|(y_{1,m}, -y_{1,n})\|\|(x_{1,n}, x_{1,m}) - (y_{1,n}, y_{1,m})/2\|$$

$$\overset{\text{C-S}}{\geq} |y_{1,m}(x_{1,n} - y_{1,n}/2) - y_{1,n}(x_{1,m} - y_{1,m}/2)| = |x_{1,n}y_{1,m} - x_{1,m}y_{1,n}| = |\tau_{n,m}(x, y)|$$

□

Definition 4.4. For $2 \leq s \leq r$, define $SN_s : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ by

$$SN_s(x, y) := \max\{\|x_1 - y_1/2\|, \|(x_2, y_2)\|^{1/2}, \|(x_3, y_3)\|^{1/3}, \ldots \|(x_s, y_s)\|^{1/s}\}$$

Remark 4.2. Using the maximum of the terms is not important here; it could be replaced by any ℓ^p-sum or other such norm. If a different choice of norm was used, the rest of the section would proceed the exact same way except with possibly different values of constants (but still independent of x, y).

Lemma 4.12. *For each $2 \leq s \leq r$, there exists a homogeneous quasi-norm N_s and a constant $c > 0$ such that for all $x, y \in \mathfrak{g}$,*

1.

$$(N_s(x)^{2s} + N_s(y^{-1}x)^{2s})/2 - (N_s(y)/2)^{2s} \geq cSN_s^{2s}(x, y) + cD_s(x, y) + cN_s(\delta_{1/2}(y)^{-1}x)^{2s}$$

2. $N_s(y) \geq \|y_1\|$ *for all* $s \geq 2$. *Consequently,* $N_s(y) + SN_s(x,y) \geq b\|(x_{s'}, y_{s'})\|^{1/s'}$ *for some* $b > 0$ *and all* $1 \leq s' \leq s$.

3. *If* $N_s(y) = 0$, $y_i = 0$ *for all* $1 \leq i \leq s$. *In particular,* N_r *is a positive definite homogeneous quasi-norm.*

Proof. The proof is by induction on s. The functions N_s we construct will clearly be homogeneous quasi-norms and satisfy (2) and (3), so we will only concern ourselves with proving (1).

Proof of base case: The base case is $s = 2$. Throughout the proof of the base case, c', c'', c''' denote (small) positive constants that depend on $\mathfrak{g}$ but not on x, y. Each of the constants may depend on the ones previously appearing, but of course this is compatible with the fact that they are all independent of x, y. Define

$$N_2(x) := \sqrt[4]{\|x_1\|^4 + \lambda\|x_2\|^2}$$

where $\lambda > 0$ is to be chosen later. Recalling that $SN_2(x,y)^4 = \max(\|x_1 - y_2/2\|^4, \|(x_2, y_2)\|^2) \leq \|x_1 - y_2/2\|^4 + \|(x_2, y_2)\|^2$ and $D_2(x,y) = \boldsymbol{\tau}^2(x,y) + \|(x_2, y_2)\|^2$, we need to show

$$(N_2(x)^4 + N_2(y^{-1}x)^4)/2$$

$$\geq (N_2(y)/2)^4 + c\|x_1 - y_1/2\|^4 + c\boldsymbol{\tau}^2(x,y) + c\|(x_2,y_2)\|^2 + cN_2(\delta_{1/2}(y)^{-1}x)^4$$

for some $\lambda, c > 0$. First let's write out the definitions of some of the terms in the inequality.

$$\begin{aligned}
N_2(x)^4 &= \|x_1\|^4 + \lambda\|x_2\|^2 \\
N_2(y^{-1}x)^4 &= \|x_1 - y_1\|^4 + \lambda\|(y^{-1}x)_2\|^2 \\
N_2(y)^4 &= \|y_1\|^4 + \lambda\|y_2\|^2
\end{aligned}$$

By convexity, parallelogram law, and Lemma 4.11,

$$(\|x_1\|^4 + \|x_1 - y_1\|^4)/2 \geq ((\|x_1\|^2 + \|x_1 - y_1\|^2)/2)^2$$

$$= (\|y_1/2\|^2 + \|x_1 - y_1/2\|^2)^2 = (\|y_1\|/2)^4 + 2\|y_1/2\|^2\|x_1 - y_1/2\|^2 + \|x_1 - y_1/2\|^4$$

$$\overset{\text{Lem 4.11}}{\geq} (\|y_1\|/2)^4 + c'\boldsymbol{\tau}^2(x,y) + \|x_1 - y_1/2\|^4$$

For some $c' > 0$. Thus, it suffices to show that for sufficiently small $\lambda, c > 0$,

$$\frac{c'}{2}\boldsymbol{\tau}^2(x,y) + \lambda\|x_2\|^2 + \lambda\|(y^{-1}x)_2\|^2 + \frac{1}{2}\|x_1 - y_1/2\|^4$$

$$\geq 2^{-4}\lambda\|y_2\|^2 + c\boldsymbol{\tau}^2(x,y) + c\|(x_2,y_2)\|^2 + cN_2(\delta_{1/2}(y)^{-1}x)^4$$

By Lemma 4.9, the following inequality is true for sufficiently small $\lambda > 0$:

$$\frac{c'}{4}\boldsymbol{\tau}^2(x,y) + \lambda\|(y^{-1}x)_2\|^2 = \frac{c'}{4}(D_2(x,y) - \|(x_2,y_2)\|\|^2) + \lambda\|(y^{-1}x)_2\|^2 \overset{\text{Lem 4.9}}{\geq} \frac{\lambda}{2}\|x_2 - y_2\|^2$$

Thus it suffices for the following inequality to hold for $\lambda, c > 0$ sufficiently small:

$$\frac{c'}{4}\boldsymbol{\tau}^2(x,y) + \lambda\|x_2\|^2 + \frac{\lambda}{2}\|x_2 - y_2\|^2 + \frac{1}{2}\|x_1 - y_1/2\|^4$$

$$\geq 2^{-4}\lambda\|y_2\|^2 + c\|(x_2,y_2)\|^2 + cN_2(\delta_{1/2}(y)^{-1}x)^4$$

We have

$$\lambda\|x_2\|^2 + \frac{\lambda}{2}\|x_2 - y_2\|^2 \geq \frac{\lambda}{2}(\|x_2\|^2 + \|x_2 - y_2\|^2)$$

$$= \frac{\lambda}{4}(\|x_2\|^2 + \|x_2 - y_2\|^2) + \frac{\lambda}{4}(\|x_2\|^2 + \|x_2 - y_2\|^2) \geq 2^{-4}\lambda\|y_2\|^2 + c''\|(x_2,y_2)\|^2$$

Thus it remains to show

$$\frac{c'}{4}\boldsymbol{\tau}^2(x,y) + \frac{c''}{2}\|(x_2,y_2)\|^2 + \frac{1}{2}\|x_1 - y_1/2\|^4 \geq cN_2(\delta_{1/2}(y)^{-1}x)^4$$

for $c > 0$ sufficiently small. By Lemma 4.10 we have

$$\frac{c'}{4}\boldsymbol{\tau}^2(x,y) + \frac{c''}{2}\|(x_2,y_2)\|^2 \geq c'''(\boldsymbol{\tau}^2(x,y) + \|(x_2,y_2)\|^2) = c'''D_s(x,y) \overset{\text{Lem 4.10}}{\geq} c\lambda\|(\delta_{1/2}(y)^{-1}x)_2\|^2$$

$c > 0$ sufficiently small, and thus it remains to show

$$c\lambda\|(\delta_{1/2}(y)^{-1}x)_2\|^2 + \frac{1}{2}\|x_1 - y_1/2\|^4 \geq cN_2(\delta_{1/2}(y)^{-1}x)^4$$

This is true by definition of N_2. This completes the proof of the base case.

Proof of inductive step: Now assume the statement holds for all $2 \leq s' \leq s$ some $2 \leq s \leq r-1$. Define N_{s+1} by

$$N_{s+1}(x) := \sqrt[2(s+1)]{\lambda\|x_{s+1}\|^2 + \sum_{s'=\lceil (s+1)/2 \rceil}^{s} N_{s'}^{2(s+1)}(x)} \tag{4.1}$$

where λ is a (small) positive constant (different λ than in the base case) to be chosen later (independent of x, y). Throughout the remainder of the proof, $c_1 - c_7$ denote (small) positive constants that depend on $\mathfrak{g}$

but not on x, y. Each of the constants may depend on the ones previously appearing, but of course this is compatible with the fact that they are all independent of x, y. The constant λ will end up depending on c_2 (which in turn depends on c_1), and the subsequent constants will depend on λ.

We now prove the inductive step. In what follows, we adopt some conventions to help make the proof more readable. There are two types of equalities/inequalities we use relating each of the expressions below. The first type is simply using a lemma, definition, inductive hypothesis, or convexity or trivial numerical inequality. Whenever an equality/inequality of this type is used, the particular terms in the expression that change from one to the next are **bolded**. No other terms change, except for the bolded ones to which the particular lemma, definition, inductive hypothesis, or convexity or trivial numerical inequality apply. Apart from the trivial numerical inequalities, the name of the lemma or definition, "ind hyp", or "convexity" decorates the equality/inequality symbol. The second type of equality/inequality used is always an equality and the equality symbol is decorated with the word "rearrange". This means we use trivialities like commutivity of addition or multiplication, reindexing of a sum, or no symbolic changes at all. Importantly, we also use equalities decorated with "rearrange" to change which terms are bolded in the expression, in preparation for the use of another equality/inequality of the first type.

$$\boldsymbol{(N_{s+1}(x)^{2(s+1)} + N_{s+1}(y^{-1}x)^{2(s+1)})/2}$$

$$\stackrel{(4.1)}{=} \quad \begin{aligned} &\boldsymbol{\Big(\lambda\|x_{s+1}\|^2 + \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} N_{s'}^{2(s+1)}(x) + \|(y^{-1}x)_{s+1}\|^2} \\ &\quad \boldsymbol{+ \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} N_{s'}^{2(s+1)}(y^{-1}x)\Big)/2} \end{aligned}$$

$$\stackrel{\text{rearrange}}{=} \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2) + \boldsymbol{\Big(\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} N_{s'}^{2(s+1)}(x) + N_{s'}^{2(s+1)}(y^{-1}x)\Big)/2}$$

$$\stackrel{\text{convexity}}{\geq} \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2) + \boldsymbol{\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} \left(\frac{N_{s'}^{2s'}(x) + N_{s'}^{2s'}(y^{-1}x)}{2}\right)^{\frac{s+1}{s'}}}$$

$$\stackrel{\text{ind hyp (1)}}{\geq} \quad \begin{aligned} &\frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2) \\ &+ \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} \boldsymbol{((N_{s'}(y)/2)^{2s'} + c_1 S N_{s'}^{2s'}(x,y)} \\ &\quad \boldsymbol{+ c_1 D_{s'}(x,y) + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2s'})^{\frac{s+1}{s'}}} \end{aligned}$$

$$\stackrel{\text{Lem 4.1}}{\geq} \quad \begin{aligned} &\frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2) \\ &+ \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} \boldsymbol{((N_{s'}(y)/2)^{2s'} + c_1 S N_{s'}^{2s'}(x,y) + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2s'})^{\frac{s+1}{s'}}} \\ &\quad \boldsymbol{+ ((N_{s'}(y)/2)^{2s'} + c_1 S N_{s'}^{2s'}(x,y))^{\frac{s+1-s'}{s'}} c_1 D_{s'}(x,y)} \end{aligned}$$

$$\overset{\text{rearrange}}{=} \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2)$$
$$+\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} \boldsymbol{((N_{s'}(y)/2)^{2s'} + c_1 SN_{s'}^{2s'}(x,y) + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2s'})^{\frac{s+1}{s'}}}$$
$$+((N_{s'}(y)/2)^{2s'} + c_1 SN_{s'}^{2s'}(x,y))^{\frac{s+1-s'}{s'}} c_1 D_{s'}(x,y)$$

$$\overset{\text{convexity}}{\geq} \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2)$$
$$+\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} \boldsymbol{(N_{s'}(y)/2)^{2(s+1)} + c_1 SN_{s'}^{2(s+1)}(x,y) + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}}$$
$$+((N_{s'}(y)/2)^{2s'} + c_1 SN_{s'}^{2s'}(x,y))^{\frac{s+1-s'}{s'}} c_1 D_{s'}(x,y)$$

$$\overset{\text{rearrange}}{=} \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2)$$
$$+\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} (N_{s'}(y)/2)^{2(s+1)} + c_1 SN_{s'}^{2(s+1)}(x,y) + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}$$
$$\boldsymbol{+((N_{s'}(y)/2)^{2s'} + c_1 SN_{s'}^{2s'}(x,y))^{\frac{s+1-s'}{s'}} c_1} D_{s'}(x,y)$$

$$\overset{\text{ind hyp (2)}}{\geq} \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2)$$
$$+\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} (N_{s'}(y)/2)^{2(s+1)} + c_1 SN_{s'}^{2(s+1)}(x,y) + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}$$
$$\boldsymbol{+c_2\|(x_{s+1-s'}, y_{s+1-s'})\|^2} D_{s'}(x,y)$$

$$\overset{\text{rearrange}}{=} \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2)$$
$$\boldsymbol{+\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s}} (N_{s'}(y)/2)^{2(s+1)} + c_1 \boldsymbol{SN_{s'}^{2(s+1)}(x,y)} + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}$$
$$+\textstyle\sum_{s'=1}^{\lfloor (s+1)/2\rfloor} c_2\|(x_{s'}, y_{s'})\|^2 D_{s+1-s'}(x,y)$$

$$\geq \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2)$$
$$+c_1 \boldsymbol{SN_{s}^{2(s+1)}(x,y)} + \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} (N_{s'}(y)/2)^{2(s+1)} + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}$$
$$+\textstyle\sum_{s'=1}^{\lfloor (s+1)/2\rfloor} c_2\|(x_{s'}, y_{s'})\|^2 D_{s+1-s'}(x,y)$$

$$\overset{\text{rearrange}}{=} \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2)$$
$$+c_1 SN_{s}^{2(s+1)}(x,y) + \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} (N_{s'}(y)/2)^{2(s+1)} + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}$$
$$\boldsymbol{+\textstyle\sum_{s'=1}^{\lfloor (s+1)/2\rfloor} c_2\|(x_{s'}, y_{s'})\|^2 D_{s+1-s'}(x,y)}$$

$$\overset{\text{Def 4.3}}{=} \quad \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2)$$
$$+c_1 SN_s^{2(s+1)}(x,y) + \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} (N_{s'}(y)/2)^{2(s+1)} + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}$$
$$+\boldsymbol{c_2(D_{s+1}(x,y) - \|(x_{s+1}, y_{s+1})\|^2)}$$

$$\overset{\text{rearrange}}{=} \quad c_1 SN_s^{2(s+1)}(x,y) + \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} (N_{s'}(y)/2)^{2(s+1)} + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}$$
$$+\frac{c_2}{2}(D_{s+1}(x,y) - \|(x_{s+1}, y_{s+1})\|^2)$$
$$\boldsymbol{\frac{c_2}{2}(D_{s+1}(x,y) - \|(x_{s+1}, y_{s+1})\|^2) + \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2)} =: (*)$$

By Lemma 4.9, we can choose $\lambda > 0$ sufficiently small so that

$$\frac{c_2}{2}(D_{s+1}(x,y) - \|(x_{s+1}, y_{s+1})\|^2) + \frac{\lambda}{2}(\|x_{s+1}\|^2 + \|(y^{-1}x)_{s+1}\|^2) \overset{\text{Lem 4.9}}{\geq} \frac{\lambda}{4}(\|x_{s+1}\|^2 + \|x_{s+1} - y_{s+1}\|^2)$$

$$= \frac{\lambda}{8}(\|x_{s+1}\|^2 + \|x_{s+1} - y_{s+1}\|^2) + \frac{\lambda}{8}(\|x_{s+1}\|^2 + \|x_{s+1} - y_{s+1}\|^2)$$

$$\geq \frac{\lambda}{16}\|y_{s+1}\|^2 + c_3\|(x_{s+1}, y_{s+1})\|^2 \geq 2^{-(s+1)}\lambda\|y_{s+1}\|^2 + c_3\|(x_{s+1}, y_{s+1})\|^2$$

And thus we get

$$(*) \geq \quad c_1 SN_s^{2(s+1)}(x,y) + \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} (N_{s'}(y)/2)^{2(s+1)} + c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}$$
$$+\frac{c_2}{2}(D_{s+1}(x,y) - \|(x_{s+1}, y_{s+1})\|^2)$$
$$\boldsymbol{2^{-(s+1)}\lambda\|y_{s+1}\|^2 + c_3\|(x_{s+1}, y_{s+1})\|^2}$$

$$\overset{\text{rearrange}}{=} \quad c_1 SN_s^{2(s+1)}(x,y) + \boldsymbol{2^{-(s+1)}\lambda\|y_{s+1}\|^2 + \textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} (N_{s'}(y)/2)^{2(s+1)}}$$
$$+\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)} + \frac{c_2}{2}(D_{s+1}(x,y) - \|(x_{s+1}, y_{s+1})\|^2)$$
$$+c_3\|(x_{s+1}, y_{s+1})\|^2$$

$$\overset{(4.1)}{=} \quad c_1 SN_s^{2(s+1)}(x,y) + \boldsymbol{(N_{s+1}(y)/2)^{2(s+1)}}$$
$$+\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)} + \frac{c_2}{2}(D_{s+1}(x,y) - \|(x_{s+1}, y_{s+1})\|^2)$$
$$+c_3\|(x_{s+1}, y_{s+1})\|^2$$

$$\overset{\text{rearrange}}{=} \quad \boldsymbol{c_1 SN_s^{2(s+1)}(x,y) + \frac{c_3}{2}\|(x_{s+1}, y_{s+1})\|^2} + (N_{s+1}(y)/2)^{2(s+1)}$$
$$+\textstyle\sum_{s'=\lceil (s+1)/2\rceil}^{s} c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}$$
$$+\frac{c_2}{2}(D_{s+1}(x,y) - \|(x_{s+1}, y_{s+1})\|^2) + \frac{c_3}{2}\|(x_{s+1}, y_{s+1})\|^2$$

$$\overset{\text{Def 4.4}}{\geq} \quad \boldsymbol{c_4 S N_{s+1}^{2(s+1)}(x,y)} + (N_{s+1}(y)/2)^{2(s+1)} + \sum_{s'=\lceil (s+1)/2 \rceil}^{s} c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)} + \frac{c_2}{2}(D_{s+1}(x,y) - \|(x_{s+1},y_{s+1})\|^2) + \frac{c_3}{2}\|(x_{s+1},y_{s+1})\|^2$$

$$\overset{\text{rearrange}}{=} \quad (N_{s+1}(y)/2)^{2(s+1)} + c_4 S N_{s+1}^{2(s+1)}(x,y) + \sum_{s'=\lceil (s+1)/2 \rceil}^{s} c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)} + \boldsymbol{\frac{c_2}{2}(D_{s+1}(x,y) - \|(x_{s+1},y_{s+1})\|^2) + \frac{c_3}{2}\|(x_{s+1},y_{s+1})\|^2}$$

$$\geq \quad (N_{s+1}(y)/2)^{2(s+1)} + c_4 S N_{s+1}^{2(s+1)}(x,y) + \sum_{s'=\lceil (s+1)/2 \rceil}^{s} c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)} + \boldsymbol{c_5 D_{s+1}(x,y)}$$

$$\overset{\text{rearrange}}{=} \quad (N_{s+1}(y)/2)^{2(s+1)} + c_4 S N_{s+1}^{2(s+1)}(x,y) + \frac{c_5}{2} D_{s+1}(x,y) + \sum_{s'=\lceil (s+1)/2 \rceil}^{s} c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)} + \boldsymbol{\frac{c_5}{2} D_{s+1}(x,y)}$$

$$\overset{\text{Lem 4.10}}{\geq} \quad (N_{s+1}(y)/2)^{2(s+1)} + c_4 S N_{s+1}^{2(s+1)}(x,y) + \frac{c_5}{2} D_{s+1}(x,y) + \sum_{s'=\lceil (s+1)/2 \rceil}^{s} c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)} + \boldsymbol{c_6 \|(\delta_{1/2}(y)^{-1}x)_{s+1}\|^2}$$

$$\overset{\text{rearrange}}{=} \quad (N_{s+1}(y)/2)^{2(s+1)} + c_4 S N_{s+1}^{2(s+1)}(x,y) + \frac{c_5}{2} D_{s+1}(x,y) + \boldsymbol{\sum_{s'=\lceil (s+1)/2 \rceil}^{s} c_1 N_{s'}(\delta_{1/2}(y)^{-1}x)^{2(s+1)} + c_6 \|(\delta_{1/2}(y)^{-1}x)_{s+1}\|^2}$$

$$\overset{(4.1)}{\geq} \quad (N_{s+1}(y)/2)^{2(s+1)} + c_4 S N_{s+1}^{2(s+1)}(x,y) + \frac{c_5}{2} D_{s+1}(x,y) + \boldsymbol{c_7 N_{s+1}(\delta_{1/2}(y)^{-1}x)^{2(s+1)}}$$

□

Lemma 4.13. *There exists a positive definite homogeneous quasi-norm N_r on $\mathfrak{g}$ and a constant $c > 0$ (depending on $\mathfrak{g}$ but not on x, y) such that for all $p \geq r$ and all $x, y \in \mathfrak{g}$,*

$$(N_r(x)^{2p} + N_r(y^{-1}x)^{2p})/2 - (N_r(y)/2)^{2p} \geq c^{p/r} N_r(\delta_{1/2}(y)^{-1}x)^{2p}$$

Proof. Let N_r, c be as in the conclusion of Lemma 4.12. Let $p \geq r$. Then by convexity and that lemma,

$$(N_r(x)^{2p} + N_r(y^{-1}x)^{2p})/2 \geq ((N_r(x)^{2r} + N_r(y^{-1}x)^{2r})/2)^{p/r}$$

$$\overset{\text{Lem 4.12}}{\geq} ((N_r(y)/2)^{2r} + c N_r(\delta_{1/2}(y)^{-1}x)^{2r})^{p/r} \geq (N_r(y)/2)^{2p} + c^{p/r} N_r(\delta_{1/2}(y)^{-1}x)^{2p}$$

□

Lemma 4.14. *There exists a left invariant, homogeneous, positive definite quasi-metric d_{N_r} on $\mathfrak{g}$ and a constant $c > 0$ (depending on $\mathfrak{g}$ but not on w, x, y, z) such that for all $p \geq r$ and $w, x, y, z \in \mathfrak{g}$,*

$$(2d_{N_r}(y,x)^{2p} + d_{N_r}(y,w)^{2p} + d_{N_r}(y,z)^{2p})/2 - (d_{N_r}(x,w)/2)^{2p} - (d_{N_r}(x,z)/2)^{2p} \geq c' d_{N_r}(w,z)^{2p}$$

Proof. Let N_r, c be as in the previous lemma. Let d_{N_r} be the metric derived from N_r; $d_{N_r}(x,y) := N_r(y^{-1}x)$. By left invariance of the metric, we may assume $x = 0$. Then by applying the previous lemma to each of the pairs (y, w) and (y, z), we obtain

$$(d_{N_r}(y,0)^{2p} + d_{N_r}(y,w)^{2p})/2 - (d_{N_r}(0,w)/2)^{2p} \geq c^{p/r} d_{N_r}(\delta_{1/2}(w), 0)^{2p}$$

$$(d_{N_r}(y,0)^{2p} + d_{N_r}(y,z)^{2p})/2 - (d_{N_r}(0,z)/2)^{2p} \geq c^{p/r} d_{N_r}(\delta_{1/2}(z), 0)^{2p}$$

Adding these and then using using Hölder, the quasi-triangle inequality, and homogeneity gives

$$(2d_{N_r}(y,0)^{2p} + d_{N_r}(y,w)^{2p} + d_{N_r}(y,z)^{2p})/2 - (d_{N_r}(0,w)/2)^{2p} - (d_{N_r}(0,z)/2)^{2p}$$

$$\geq c^{p/r}(d_{N_r}(\delta_{1/2}(w),0)^{2p} + d_{N_r}(\delta_{1/2}(z),0)^{2p}) \geq 2^{-2p+1} c^{p/r}(d_{N_r}(\delta_{1/2}(w),0) + d_{N_r}(\delta_{1/2}(z),0))^{2p}$$

$$\geq c' d_{N_r}(\delta_{1/2}(w), \delta_{1/2}(z))^{2p} = 2^{-2p} c' d_{N_r}(w,z)^{2p}$$

for some $c' > 0$. □

Theorem 4.1. *Every graded nilpotent Lie group G of step r, equipped with a left invariant metric homogeneous with respect to the dilations induced by the grading, is 4-point p-convex - and consequently Markov p-convex - for every $p \in [2r, \infty)$.*

Proof. Let G be as above. Lemma 4.14 exactly states that G is 4-point p-convex for every $p \in [2r, \infty)$. Then Theorem 1.11 implies G is Markov p-convex for every $p \in [2r, \infty)$. □

4.4 Lower Bound on Markov Convexity of $J^{r-1}(\mathbb{R})$

The goal of this section is to prove Theorem 4.2, which occurs at the conclusion. The strategy is to construct a sequence of directed graphs (see Definition 4.5) with bad Markov convexity properties. These bad properties are manifested by the dispersive nature of random walks on the graphs. This is the content of Lemma 4.16.

We then map these graphs into $J^{r-1}(\mathbb{R})$ with sufficient control over the distortion (Lemma 4.18) to prove Theorem 4.2.

4.4.1 Directed Graphs and Random Walks

Let $(N_m)_{m=0}^{\infty}$ be any sequence of integers with $N_0 = 0$ and $N_{m+1} \geq \max(1, N_m + \lceil 2\log_2(m+1)\rceil)$. We'll define a sequence of directed graphs $(\Gamma_m)_{m=0}^{\infty}$. The graphs will be directed from unique source vertex to unique and sink vertex, which we will denote by 0_m and 1_m, respectively. Let $\operatorname{diam}(\Gamma_m)$ be the number of edges in a directed edge path from 0 to 1, which is also equal to the diameter of Γ_m with respect to the shortest path metric. The construction will be such that $\operatorname{diam}(\Gamma_m) = 2^{N_m}$.

Definition 4.5. We'll perform the construction and also prove that $\operatorname{diam}(\Gamma_m) = 2^{N_m}$ by induction. Let Γ_0 be the interval I, that is, a graph with two vertices $0, 1$ and a single edge connecting them, directed from 0 to 1. Suppose Γ_m has been constructed for some $m \geq 0$. We define an intermediate graph Γ'_{m+1} by gluing together $a := 2^{N_{m+1}-\lceil 2\log_2(m+1)\rceil-1}$ copies of I, then $A := 2^{N_{m+1}-N_m} - 2^{N_{m+1}-N_m-\lceil 2\log_2(m+1)\rceil} = 2^{-N_m}(2^{N_{m+1}} - 2a) = 2^{N_{m+1}-N_m}(1 - 2^{-\lceil 2\log_2(m+1)\rceil})$ copies of Γ_m, then a more copies of I again together in series. The source vertex of this graph is the source vertex of the first copy of I, and the sink vertex is the sink vertex of the last copy of I. The diameter of this graph is

$$a \cdot \operatorname{diam}(I) + A \cdot \operatorname{diam}(\Gamma_m) + a \cdot \operatorname{diam}(I) \overset{\text{ind hyp}}{=} 2a + 2^{N_m}A = 2^{N_{m+1}}$$

We then define Γ_{m+1} to be two copies of Γ'_{m+1}, denoted $+\Gamma'_{m+1}$ and $-\Gamma'_{m+1}$, glued together in parallel. Denote the common source vertex 0_m and sink vertex 1_m. The diameter of Γ_{m+1} is the same as the diameter of Γ'_{m+1}. We note that each copy of Γ_m in Γ_{m+1} is isometrically embedded; any shortest path between two points in a copy of $\Gamma_m \subseteq \Gamma_{m+1}$ completely belongs to Γ_m.

By swapping $+\Gamma'_{m+1}$ and $-\Gamma'_{m+1}$ in Γ_{m+1}, we obtain a directed graph involution $\iota : \Gamma_{m+1} \to \Gamma_{m+1}$.

For $q_1, q_2 \in \Gamma_m$, (q_1, q_2) is called a *vertical pair* if $d_m(q_1, 0_m) = d_m(q_2, 0_m)$.

For each $m \geq 0$, let $(X_t^m)_{t=0}^{2^{N_m}}$ be the standard directed random walk on Γ_m. Let d_m denote the shortest path metric on Γ_m. With full probability, $d(X_t^m, 0_m) = t$ for $0 \leq t \leq 2^{N_m}$.

See the two right-hand graphs of Figure 4.2 for what Γ_1 and Γ_2 look like when $N_0 = 0$, $N_1 = 2$, and $N_2 = 4$. The graphs are drawn in such a way that the direction is from left to right, $+\Gamma'_m$ lies above the x-axis, and $-\Gamma'_m$ lies below the x-axis. The source vertices 0_m are both drawn at $(0,0)$, and the sink vertices 1_2 are both drawn at $(1,0)$.

Lemma 4.15. *For all $p > 0$ and $m \geq 0$,*

$$\sum_{k=0}^{N_m} \sum_{t=1}^{2^{N_m}} \frac{\mathbb{E}[d_m(X_t^m, \tilde{X}_t^m(t-2^k))^p]}{2^{kp}} \geq \frac{m}{8} 2^{N_m} \Pi_{i=1}^{m-1}(1-(i+1)^{-2})$$

Proof. Let $p \geq 1$. The proof is by induction on m. The base case $m = 0$ is trivially true. Assume the inequality holds for some $m \geq 0$. Now we consider the standard random walk X_t^{m+1} on Γ_{m+1}. Consider k and t in the range $a + 1 \leq t \leq 2^{N_{m+1}} - a$, $0 \leq k \leq N_m$, where $a = 2^{N_{m+1} - \lceil 2\log_2(m+1) \rceil - 1}$. Then $t - 2^k \geq 2^{N_{m+1} - \lceil 2\log_2(m+1)\rceil - 1} + 1 - 2^{N_m} \geq 1$, so X_1^{m+1} and $\tilde{X}_1^{m+1}(t-2^k)$ agree. Then for all subsequent times, with full probability, X_t^{m+1} and $\tilde{X}_t^{m+1}(t-2^k)$ belong to the same copy of Γ'_{m+1} in Γ_{m+1}. Then, after recalling the construction of Γ'_{m+1} as a number of copies of Γ_m and I glued together, it can be seen that for the range of t in interest, X_t^{m+1} and $\tilde{X}_t^{m+1}(t-2^k)$ are standard random walks across $A = 2^{N_{m+1}-N_m}(1 - 2^{-\lceil 2\log_2(m+1)\rceil})$ consecutive copies of Γ_m, which we denote as $A \cdot X_t^{m+1}$ and $A \cdot \tilde{X}_t^{m+1}(t-2^k)$. Thus, under our assumptions on k and t, $d_{m+1}(X_t^{m+1}, \tilde{X}_t^{m+1}(t-2^k))$ has the same distribution as $d_m(A \cdot X_t^m, A \cdot \tilde{X}_t^m(t-2^k))$. Hence we obtain by the inductive hypothesis

$$\sum_{k=0}^{N_m} \sum_{t=a+1}^{2^{N_{m+1}}-a} \frac{\mathbb{E}[d_{m+1}(X_t^{m+1}, \tilde{X}_t^{m+1}(t-2^k))^p]}{2^{kp}} = \sum_{k=0}^{N_m} \sum_{t=a+1}^{2^{N_{m+1}}-a} \frac{\mathbb{E}[d_m(A \cdot X_t^m, A \cdot \tilde{X}_t^m(t-2^k))^p]}{2^{kp}}$$

$$= \sum_{k=0}^{N_m} \sum_{T=1}^{A} \left(\sum_{t=a+(T-1)2^{N_m}+1}^{a+T2^{N_m}} \frac{\mathbb{E}[d_m(A \cdot X_t^m, A \cdot \tilde{X}_t^m(t-2^k))^p]}{2^{kp}} \right)$$

$$= \sum_{k=0}^{N_m} \sum_{T=1}^{A} \sum_{t=1}^{2^{N_m}} \frac{\mathbb{E}[d_m(X_t^m, \tilde{X}_t^m(t-2^k))^p]}{2^{kp}} \overset{\text{ind hyp}}{\geq} \sum_{T=1}^{A} \frac{m}{8} 2^{N_m} \Pi_{i=1}^{m-1}(1-(i+1)^{-2})$$

$$= 2^{N_m} A \frac{m}{8} \Pi_{i=1}^{m-1}(1-(i+1)^{-2}) = 2^{N_{m+1}}(1 - 2^{-\lceil 2\log_2(m+1)\rceil}) \frac{m}{8} \Pi_{i=1}^{m-1}(1-(i+1)^{-2})$$

$$\geq 2^{N_{m+1}}(1-(m+1)^{-2}) \frac{m}{8} \Pi_{i=1}^{m-1}(1-(i+1)^{-2}) = \frac{m}{8} 2^{N_{m+1}} \Pi_{i=1}^{m}(1-(i+1)^{-2})$$

In summary,

$$\sum_{k=0}^{N_m} \sum_{t=a+1}^{2^{N_{m+1}}-a} \frac{\mathbb{E}[d_{m+1}(X_t^{m+1}, \tilde{X}_t^{m+1}(t-2^k))^p]}{2^{kp}} \geq \frac{m}{8} 2^{N_{m+1}} \Pi_{i=1}^{m}(1-(i+1)^{-2}) \tag{4.2}$$

Now consider k and t in the range $0 \leq k \leq N_{m+1} - 1$, $1 \leq t \leq 2^k$, so that $t - 2^k \leq 0$. Note that this means this range is disjoint from the one previously considered. Since $t - 2^k \leq 0$, the random walks X^{m+1} and $\tilde{X}^{m+1}(t-2^k)$ evolved independently immediately. Thus, with probability $1/2$, X^{m+1} and $\tilde{X}^{m+1}(t-2^k)$ belong to different copies of Γ'_{m+1} in Γ_{m+1}. This implies that, with probability $1/2$,

$d_{m+1}(X_t^{m+1}, \tilde{X}_t^{m+1}(t-2^k)) = 2t$. Thus,

$$\sum_{k=0}^{N_{m+1}-1} \sum_{t=1}^{2^k} \frac{\mathbb{E}[d_{m+1}(X_t^{m+1}, \tilde{X}_t^{m+1}(t-2^k))^p]}{2^{kp}} \geq \sum_{k=0}^{N_{m+1}-1} \sum_{t=1}^{2^k} \frac{(2t)^p}{2^{kp+1}}$$

$$\overset{\text{Lem 4.2}}{>} \sum_{k=0}^{N_{m+1}-1} \frac{2^{k(p+1)}}{2^{kp+2}} = \sum_{k=0}^{N_{m+1}-1} 2^{k-2} = 2^{N_{m+1}-2} - \frac{1}{4} \geq \frac{1}{8} 2^{N_{m+1}}$$

In summary,

$$\sum_{k=0}^{N_{m+1}-1} \sum_{t=1}^{2^k} \frac{\mathbb{E}[d_{m+1}(X_t^{m+1}, \tilde{X}_t^{m+1}(t-2^k))^p]}{2^{kp}} > \frac{1}{8} 2^{N_{m+1}} \tag{4.3}$$

Again, notice that in (4.2) and (4.3), the range of t, k we consider are disjoint from each other and are subsets of the range $0 \leq k \leq N_{m+1}$, $1 \leq t \leq 2^{N_{m+1}}$. Thus, by adding (4.2) and (4.3), we obtain

$$\sum_{k=0}^{N_{m+1}} \sum_{t=1}^{2^{N_{m+1}}} \frac{\mathbb{E}[d_m(X_t^m, \tilde{X}_t^m(t-2^k))^p]}{2^{kp}} > \frac{m}{8} 2^{N_{m+1}} \Pi_{i=1}^m (1-(i+1)^{-2}) + \frac{1}{8} 2^{N_{m+1}}$$

$$> \left(\frac{m+1}{8}\right) 2^{N_{m+1}} \Pi_{i=1}^m (1-(i+1)^{-2})$$

completing the inductive step. □

Lemma 4.16.

$$\sum_{k=0}^{\infty} \sum_{t=1}^{2^{N_m}} \frac{\mathbb{E}[d_m(X_t^m, \tilde{X}_t^m(t-2^k))^p]}{2^{kp}} \gtrsim m 2^{N_m}$$

for all $p > 0$.

Proof. This follows from Lemma 4.15 and the fact that $\Pi_{i=1}^{m-1}(1-(i+1)^{-2}) > \Pi_{i=1}^{\infty}(1-(i+1)^{-2}) > 0$ for all $m \geq 0$. □

4.4.2 Mapping the Graphs into $J^{r-1}(\mathbb{R})$

Lemma 4.17. *There exists $\phi \in C^{r-1,1}([0,1])$ such that*

1. *ϕ is symmetric across the line $x = \frac{1}{2}$, that is, $\phi(x) = \phi(1-x)$ for all $x \in [0, \frac{1}{2}]$.*

2. *$\phi(x) \geq (2x)^r$ for all $x \in [0, \frac{1}{2}]$.*

3. *$[j^{r-1}(0)](\phi) = (0,0)$, and thus by (1), $[j^{r-1}(1)](\phi) = (1,0)$.*

4. *For every integer $0 \leq i < 2^r$ and every $x \in [i2^{-r}, (i+1)2^{-r})$, $\phi^{(r)}(x) = \phi^{(r)}(i2^{-r})$ (so $\phi^{(r)}$ is constant on intervals of this form).*

Since $\phi \in C^{r-1,1}([0,1])$, $\phi^{(r)} \in L^\infty([0,1])$. We also remark here that whenever dealing with L^∞ functions, we choose representatives that are everywhere (not just almost everywhere) bounded by their norm.

Proof. The proof is by induction on r. For the base case $r = 1$, define

$$\phi(x) := \begin{cases} 2x & x \in [0, \frac{1}{2}] \\ 2 - 2x & x \in [\frac{1}{2}, 1] \end{cases}$$

ϕ satisfies (1) - (4).

Now suppose such a function ϕ exists for some $r \geq 1$. We'll construct a function ψ that satisfies (1) - (4) for $r + 1$. Define $\overline{\phi} \in C^{r-1,1}([0,1])$ by

$$\overline{\phi}(x) := \begin{cases} \phi(2x) & x \in [0, \frac{1}{2}] \\ -\phi(2 - 2x) & x \in [\frac{1}{2}, 1] \end{cases}$$

Then define $\overline{\Phi} \in C^{r,1}([0,1])$ by

$$\overline{\Phi}(x) := \int_0^x \overline{\phi}(\xi) d\xi$$

$\overline{\Phi}$ satisfies (1), (3), and (4) by the inductive hypothesis. Note that the inductive hypothesis applied to (2) implies $\overline{\phi}(x) \geq 2^r (2x)^r$ for every $x \in [0, \frac{1}{4}]$, and hence

$$\overline{\Phi}(x) \geq \frac{2^{r-1}}{r+1}(2x)^{r+1} \geq \frac{1}{2}(2x)^{r+1}$$

Also, since $\phi \geq 0$, (which follows from the inductive hypothesis applied to (1) and (2)),

$$\overline{\Phi}(x) \geq \overline{\Phi}\left(\frac{1}{4}\right) \geq \left(\frac{1}{2}\right)^{r+2}$$

for all $x \in [\frac{1}{4}, \frac{1}{2}]$. Together, these two inequalities imply

$$\psi(x) := 2^{r+2}\overline{\Phi}(x) \geq (2x)^{r+1}$$

for all $x \in [0, \frac{1}{2}]$. Thus, ψ satisfies (1)-(4), completing the inductive step. $\square$

See Figure 4.1 for graphs of ϕ and its first two derivatives when $r = 3$. Note that these graphs are not on the same scale.

Lemma 4.18. *Let ϕ be the function from Lemma 4.17. Set $N_0 = 0$, and for $m \geq 1$, set $N_m := \lceil Cm \log_2(m+$*

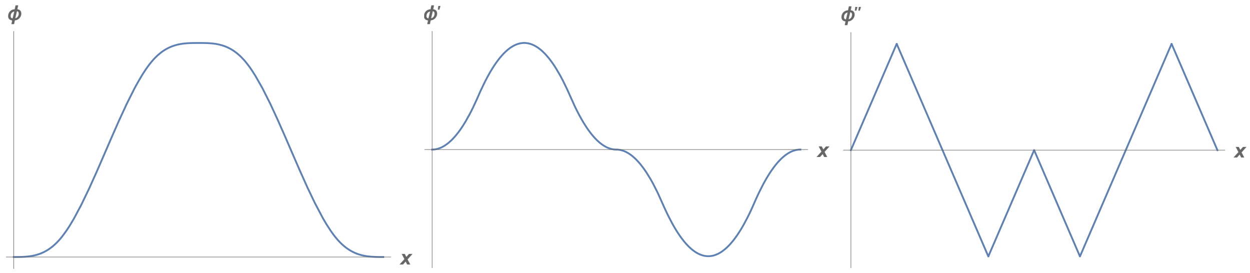

Figure 4.1: Graphs of the function ϕ from Lemma 4.17 and its first two derivatives when $r = 3$. Note that these are not shown to the same scale.

$1)\rceil$, where C is a sufficiently large constant to be chosen later, so that $N_m \geq r$ and $N_{m+1} \geq \max(1, N_m + \lceil 2\log_2(m+1)\rceil)$. Then there exists a sequence of maps $F_m : \Gamma_m \to J^{r-1}(\mathbb{R})$ such that, for all $m \geq 0$ and all directed paths γ from 0_m to 1_m in Γ_m, there is a function $\phi_\gamma \in C^{r-1,1}([0, 2^{N_m}])$ such that

1. *$[j^{r-1}(0)](\phi_\gamma) = (0, 0)$ and $[j^{r-1}(2^{N_m})](\phi_\gamma) = (2^{N_m}, 0)$.*

2. *After isometrically identifying γ with $[0, 2^{N_m}]$ via $q \mapsto d_m(q, 0_m)$, F_m restricted to γ equals the jet of ϕ_γ; $F_m(t) = [j^{r-1}(t)](\phi_\gamma)$.*

3. *For all vertical pairs $(q_1, q_2) \in \Gamma_m \times \Gamma_m$,*

$$\sqrt{m}\ln(m+1)|\pi_0(F_m(q_1)) - F_m(q_2))| \geq d_m(q_1, q_2)^r$$

4. *Let $\gamma(X^m)$ denote the directed path followed by the random walk X^m (so $\gamma(X^m)$ is itself a path-valued random variable). For all $y \in \mathbb{R}$, and $0 \leq t < 2^{N_m}$,*

$$\mathbb{E}\left[\exp\left(y\left(\sup_{[t,t+1]} \phi^{(r)}_{\gamma(X^m)}\right)\right)\right] \leq \exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \sum_{n=1}^{m} \frac{1}{n\ln(n+1)^2}\right)$$

and

$$\mathbb{E}\left[\exp\left(y\left(\inf_{[t,t+1]} \phi^{(r)}_{\gamma(X^m)}\right)\right)\right] \leq \exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \sum_{n=1}^{m} \frac{1}{n\ln(n+1)^2}\right)$$

and thus there exists a constant $B < \infty$ (not depending on y, t, or m) such that

$$\mathbb{E}\left[\exp\left(y\left\|\phi^{(r)}_{\gamma(X^m)}\right\|_{L^\infty[t,t+1]}\right)\right] \leq 2e^{By^2}$$

5. *$\|\phi^{(r)}_\gamma\|_\infty \leq 2\sqrt{m}\|\phi^{(r)}\|_\infty$.*

6. *$\|\pi_0 \circ F_m\|_\infty \leq 2^r(m+1)^{Crm+1}\|\phi\|_\infty$.*

Proof. The proof is by induction on m. The base case $m = 0$ is easy, we simply define F_0 to be the jet of the 0 function on $\Gamma_0 = I$. Then (1) - (6) hold. Assume such a sequence of maps $F_0, \dots F_m$ exist for some $m \geq 0$. Set

$$K := \|\pi_0 \circ F_m\|_\infty \tag{4.4}$$

Since $N_{m+1} \geq C(m+1)\log_2(m+2)$, we may (and do) choose C sufficiently large so that

$$K \overset{\text{ind hyp (6)}}{\leq} \|\phi\|_\infty 2^r (m+1)^{Crm+1} \leq \frac{2^{r(N_{m+1} - \lceil 2\log_2(m+1)\rceil - 1) - 1}}{\sqrt{m+1}\ln(m+2)} \tag{4.5}$$

Define $\tilde{\phi} \in C^{r-1,1}([0, 2^{N_{m+1}}])$ by

$$\tilde{\phi}(x) := \frac{2^{rN_{m+1}}}{\sqrt{m+1}\ln(m+2)}\phi(2^{-N_{m+1}}x)$$

Note that since $N_{m+1} \geq r$, Lemma 4.17(4) tells us:

$$\tilde{\phi}^{(r)}(x) = \tilde{\phi}^{(r)}(i) \tag{4.6}$$

for every integer $0 \leq i < 2^{N_m}$ and every $x \in [i, i+1)$. We also have by the chain rule

$$\left\|\tilde{\phi}^{(r)}\right\|_\infty = \frac{\left\|\phi^{(r)}\right\|_\infty}{\sqrt{m+1}\ln(m+2)} \tag{4.7}$$

and additionally

$$\left\|\tilde{\phi}\right\|_\infty \leq 2^{rN_{m+1}}\|\phi\|_\infty \leq 2^{r(C(m+1)\log_2(m+2)+1)}\|\phi\|_\infty = 2^r(m+2)^{Cr(m+1)}\|\phi\|_\infty \tag{4.8}$$

We will now define the function F_{m+1} on $\Gamma_{m+1} = +\Gamma'_{m+1} \cup -\Gamma'_{m+1}$. Let us first work with $+\Gamma'_{m+1}$. Let γ be a directed path from 0_m to 1_m in $+\Gamma'_{m+1}$. Then by definition of $+\Gamma'_{m+1}$, γ consists of $a = 2^{N_{m+1} - \lceil 2\log_2(m+1)\rceil - 1}$ copies of I, then $A = 2^{-N_m}(2^{N_{m+1}} - 2a)$ copies of different directed paths γ_i, $1 \leq i \leq A$, each belonging to Γ_m and connecting 0_m to 1_m, then a more copies of I glued together in series. Identify γ isometrically with $[0, 2^{N_{m+1}}]$ via $q \mapsto d_{m+1}(q, 0_{m+1})$. Under this identification, the first set of copies of I gets identified with the subinterval $[0, a]$, each γ_i gets identified with the subinterval $[a + (i-1)2^{N_m}, a + i2^{N_m}]$, and the last set of copies of I gets identified with the subinterval $[2^{N_{m+1}} - a, 2^{N_{m+1}}]$. We then define

$$\phi_\gamma := \tilde{\phi} + f_\gamma \tag{4.9}$$

where f_γ is defined as follows: f_γ is identically 0 on $[0,a] \cup [2^{N_{m+1}} - a, 2^{N_{m+1}}]$, and $f_\gamma(x) = \phi_{\gamma_i}(x - a - (i-1)2^{N_m})$ on $[a + (i-1)2^{N_m}, a + i2^{N_m}]$ (ϕ_{γ_i} is given to us by the inductive hypothesis). By the inductive hypothesis applied to (1) and Lemma 4.17(3), $\phi_\gamma \in C^{r-1,1}([0, 2^{N_{m+1}}])$ and satisfies (1). It is also clear from this definition, (4.7), and the inductive hypothesis applied to (5) that

$$\left\|\phi_\gamma^{(r)}\right\|_\infty \overset{(4.9)}{\leq} \left\|\tilde{\phi}^{(r)}\right\|_\infty + \max_{1\leq i\leq A}\left\|\phi_{\gamma_i}^{(r)}\right\|_\infty \overset{(4.7)}{\leq} \frac{\left\|\phi^{(r)}\right\|_\infty}{\sqrt{m+1}\ln(m+2)} + \max_{1\leq i\leq A}\left\|\phi_{\gamma_i}^{(r)}\right\|_\infty$$

$$\overset{\text{ind hyp (5)}}{\leq} \frac{\left\|\phi^{(r)}\right\|_\infty}{\sqrt{m+1}\ln(m+2)} + 2\sqrt{m}\left\|\phi^{(r)}\right\|_\infty \leq 2\sqrt{m+1}\left\|\phi^{(r)}\right\|_\infty$$

verifying (5). We can finally define F_{m+1} on $+\Gamma'_{m+1}$ by declaring it to be the jet of ϕ_γ on γ. We need to check that F_{m+1} is well-defined. Since every point of $+\Gamma_{m+1}$ is contained in some directed path from 0_m to 1_m, we only need to check what happens when one point belongs to two different paths. Let $q \in +\Gamma'_{m+1}$ and suppose $q \in \gamma \cap \gamma'$ for some directed paths γ, γ' from 0_{m+1} to 1_{m+1} in $+\Gamma'_{m+1}$. Set $t := d(q, 0_{m+1})$. There are two cases: $t \in [0,a] \cup [2^{N_{m+1}} - a, 2^{N_{m+1}}]$ or $t \in [a + (i-1)2^{N_m}, a + i2^{N_m}]$ for some i. Assume the first case holds. Then our definition of $F_{m+1}(q)$ based on either $q \in \gamma$ or $q \in \gamma'$ is

$$F_{m+1}(q) = [j^{r-1}(t)](\tilde{\phi})$$

so well-definedness holds in this case. In the other case, our definition of $F_{m+1}(q)$ based on $q \in \gamma$ is, by the inductive hypothesis applied to (2),

$$F_{m+1}(q) = [j^{r-1}(t)](\tilde{\phi}) + ([j^{r-1}(t - a - (i-1)2^{N_m})](\phi_{\gamma_i}) + (a + (i-1)2^{N_m} - t, 0)$$

$$\overset{\text{ind hyp}}{=} [j^{r-1}(t)](\tilde{\phi}) + F_m(q) + (a + (i-1)2^{N_m} - t, 0)$$

and likewise based on $q \in \gamma'$,

$$F_{m+1}(q) = [j^{r-1}(t)](\tilde{\phi}) + ([j^{r-1}(t - a - (i-1)2^{N_m})](\phi_{\gamma'_i}) + (a + (i-1)2^{N_m} - t, 0)$$

$$\overset{\text{ind hyp}}{=} [j^{r-1}(t)](\tilde{\phi}) + F_m(q) + (a + (i-1)2^{N_m} - t, 0)$$

(note that the term $(a+(i-1)2^{N_m} - t, 0)$ is present so that the x-coordinate of the entire expression will be t, and that we identify q as belonging to a copy of Γ_m so that $F_m(q)$ makes sense) so well-definedness holds in this case as well. Thus F_{m+1} is well-defined on $+\Gamma'_{m+1}$. We define F_{m+1} on $-\Gamma_{m+1}$ by $F_{m+1}(q) = -F_{m+1}(\iota(q))$, where $\iota : +\Gamma'_{m+1} \to -\Gamma'_{m+1}$ is the involution. It follows from this that if γ is a directed 0_{m+1}-1_{m+1} path in $-\Gamma'_{m+1}$, then $\phi_\gamma = -\phi_{\iota(\gamma)}$. Thus, (1) and (2) are satisfied. It remains to show (3), (4), and (6). Before doing so, let us summarize the discussion on F_{m+1} of this paragraph: for $q \in \Gamma_{m+1}$ and $t = d_{m+1}(q, 0_{m+1})$,

$$F_{m+1}(q) = \begin{cases} [j^{r-1}(t)](\tilde{\phi}) & t \in [0,a] \cup [2^{N_{m+1}} - a, 2^{N_{m+1}}] \\ & q \in +\Gamma'_{m+1} \\ \\ [j^{r-1}(t)](\tilde{\phi}) & \\ +F_m(q) + (a + (i-1)2^{N_m} - t, 0) & t \in [a + (i-1)2^{N_m}, a + i2^{N_m}] \\ & q \in +\Gamma'_{m+1} \\ \\ [j^{r-1}(t)](-\tilde{\phi}) & t \in [0,a] \cup [2^{N_{m+1}} - a, 2^{N_{m+1}}] \\ & q \in -\Gamma'_{m+1} \\ \\ [j^{r-1}(t)](-\tilde{\phi}) & \\ -F_m(q) - (a + (i-1)2^{N_m} - t, 0) & t \in [a + (i-1)2^{N_m}, a + i2^{N_m}] \\ & q \in -\Gamma'_{m+1} \end{cases} \tag{4.10}$$

See Figure 4.2 for the images of Γ_1 and Γ_2, based on $N_0 = 0$, $N_1 = 2$, $N_2 = 4$, in $J^1(\mathbb{R})$. Using (4.10), we can quickly verify (6):

$$\|\pi_0 \circ F_{m+1}\|_\infty \overset{(4.10)}{\leq} \left\|\tilde{\phi}\right\|_\infty + \|\pi_0 \circ F_m\|_\infty \overset{\text{ind hyp (6)}}{\leq} \left\|\tilde{\phi}\right\|_\infty + 2^r(m+1)^{Crm+1}\|\phi\|_\infty$$

$$\overset{(4.8)}{\leq} 2^r(m+2)^{Cr(m+1)}\|\phi\|_\infty + 2^r(m+1)^{Crm+1}\|\phi\|_\infty \leq 2^r(m+2)^{Cr(m+1)+1}\|\phi\|_\infty$$

(3) and (4) require more involved arguments.

Proof of (3). Let $(q_1, q_2) \in \Gamma_{m+1} \times \Gamma_{m+1}$ be a vertical pair. By definition of vertical pair, $d_{m+1}(q_1, 0_{m+1}) = d_{m+1}(q_2, 0_{m+1})$. Let t denote this common value. There are two cases, q_1, q_2 belong to the same copy of Γ'_{m+1}, or they belong to different copies. First assume they belong to the same copy. Without loss of generality say $+\Gamma'_{m+1}$. Then there are two subcases for t: $t \in [0,a] \cup [2^{N_{m+1}} - a, 2^{N_{m+1}}]$ or $t \in [a+(i-1)2^{N_m}, a+i2^{N_m}]$ for some $1 \leq i \leq A$. Assume the first subcase holds. Then by construction of $+\Gamma'_{m+1}$, q_1, q_2 belong to a copy of I, and thus the equality $d_{m+1}(q_1, 0_{m+1}) = d_{m+1}(q_2, 0_{m+1})$ implies $q_1 = q_2$, so (3) trivially holds. Assume the second subcase for t. Then

$$|\pi_0(F_{m+1}(q_1) - F_{m+1}(q_2))| \overset{(4.10)}{=} |\pi_0(F_m(q_1) - F_m(q_2))|$$

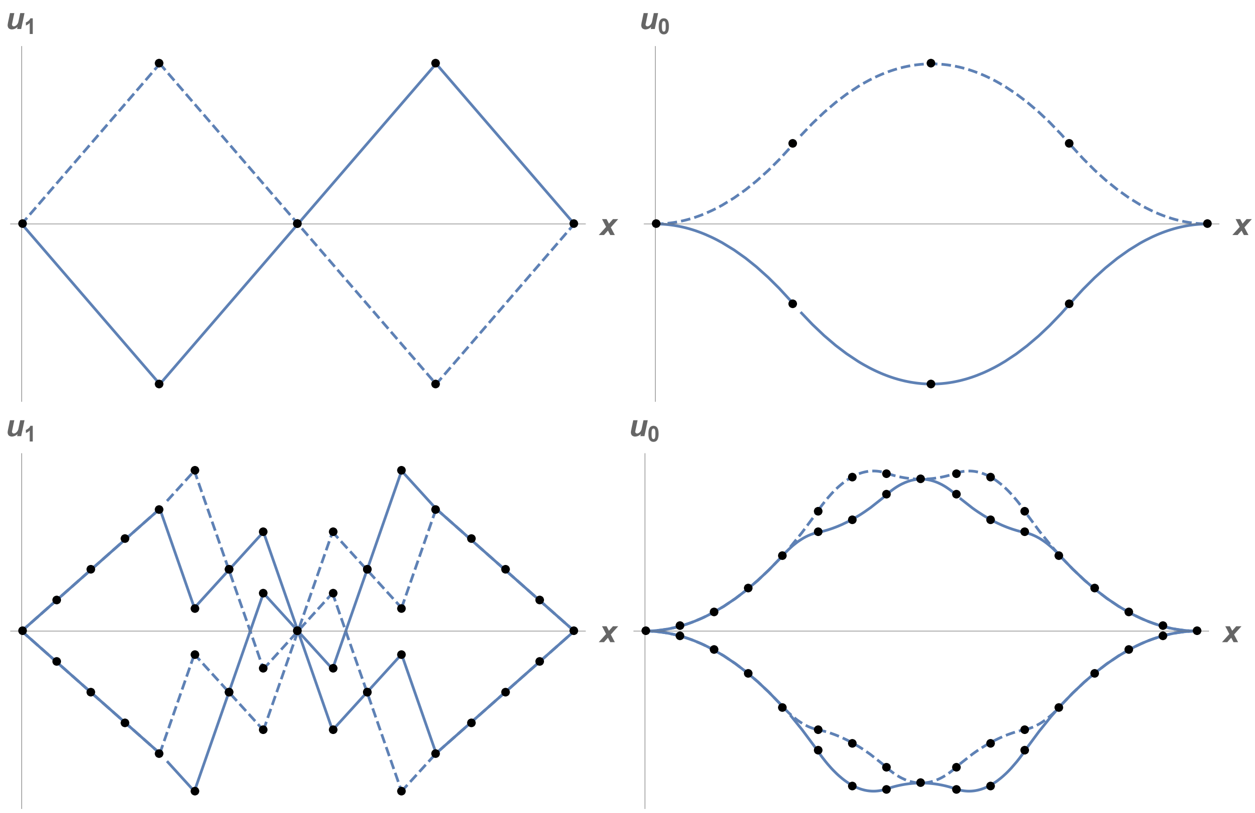

Figure 4.2: Above, the image of Γ_1, and below, the image of Γ_2, based on $N_0 = 0$, $N_1 = 2$, $N_2 = 4$, in $J^1(\mathbb{R})$ under the map F_2. $J^1(\mathbb{R})$ is identified with $\mathbb{R}^3$ via the coordinates x, u_1, u_0. These are not drawn to the same scale. The two images on the right are respectively graph isomorphic to Γ_1 and Γ_2.

and so (3) holds by the inductive hypothesis.

Now assume we are in the second case where q_1, q_2 belong to different copies of Γ'_{m+1}. Without loss of generality, assume $q_1 \in +\Gamma'_{m+1}$ and $q_2 \in -\Gamma'_{m+1}$. Observe that under this assumption, $d_{m+1}(q_1, q_2) = 2t$ if $t \leq 2^{N_{m+1}-1}$ and $d_{m+1}(q_1, q_2) = 2(2^{N_{m+1}} - t)$ if $t \geq 2^{N_{m+1}-1}$. Because of the symmetry of $\tilde{\phi}$ about the line $x = 2^{N_{m+1}-1}$, it suffices to only check the case $t \leq 2^{N_{m+1}-1}$. Let us first record the following inequality:

$$\pi_0([j^{r-1}(t)](\tilde{\phi})) \geq \frac{(2t)^r}{\sqrt{m+1}\ln(m+2)} \tag{4.11}$$

which can be proven by

$$\pi_0([j^{r-1}(t)](\tilde{\phi})) = \tilde{\phi}(t) = \frac{2^{rN_{m+1}}}{\sqrt{m+1}\ln(m+2)}\phi(2^{-N_{m+1}}t) \overset{\text{Lem 4.17(2)}}{\geq} \frac{(2t)^r}{\sqrt{m+1}\ln(m+2)}$$

Again split into two subcases: $t \in [0, a]$ or $t \in [a, 2^{N_{m+1}-1}]$. In the first subcase we have

$$\pi_0(F_{m+1}(q_1)) \overset{(4.10)}{=} \pi_0([j^{r-1}(t)](\tilde{\phi})) \overset{(4.11)}{\geq} \frac{(2t)^r}{\sqrt{m+1}\ln(m+2)}$$

and

$$\pi_0(F_{m+1}(q_2)) \overset{(4.10)}{=} \pi_0([j^{r-1}(t)](-\tilde{\phi})) \overset{(4.11)}{\leq} -\frac{(2t)^r}{\sqrt{m+1}\ln(m+2)}$$

and thus

$$|\pi_0(F_{m+1}(q_1) - F_{m+1}(q_2))| \geq \frac{2(2t)^r}{\sqrt{m+1}\ln(m+2)} = \frac{2d_{m+1}(q_1, q_2)^r}{\sqrt{m+1}\ln(m+2)}$$

proving (3) in this subcase.

Now assume the second subcase, $t \in [a, 2^{N_{m+1}-1}]$. Then

$$\begin{aligned}
\pi_0(F_{m+1}(q_1)) &\overset{(4.10)}{=} \pi_0([j^{r-1}(t)](\tilde{\phi}) + F_m(q_1) + (a + (i-1)2^{N_m} - t, 0)) = \pi_0([j^{r-1}(t)](\tilde{\phi})) + \pi_0(F_m(q_1)) \\
&\overset{(4.4)}{\geq} \pi_0([j^{r-1}(t)](\tilde{\phi})) - K \overset{(4.5)}{\geq} \pi_0([j^{r-1}(t)](\tilde{\phi})) - \frac{2^{r(N_{m+1}-\lceil 2\log_2(m+1)\rceil)-1}}{\sqrt{m+1}\ln(m+2)} \\
&\overset{(4.11)}{\geq} \frac{(2t)^r - 2^{r(N_{m+1}-\lceil 2\log_2(m+1)\rceil)-1}}{\sqrt{m+1}\ln(m+2)} = \frac{(2t)^r - (2a)^r/2}{\sqrt{m+1}\ln(m+2)} \geq \frac{(2t)^r - (2t)^r/2}{\sqrt{m+1}\ln(m+2)} = \frac{(2t)^r}{2\sqrt{m+1}\ln(m+2)}
\end{aligned}$$

Similarly,

$$\pi_0(F_{m+1}(q_2)) \leq -\frac{(2t)^r}{2\sqrt{m+1}\ln(m+2)}$$

and thus

$$|\pi_0(F_{m+1}(q_1) - F_{m+1}(q_2))| \geq \frac{(2t)^r}{\sqrt{m+1}\ln(m+2)} = \frac{d_{m+1}(q_1, q_2)^r}{\sqrt{m+1}\ln(m+2)}$$

proving (3) in this final subcase.

Proof of (4). Let $0 \leq t < 2^{N_{m+1}}$ be an arbitrary integer. Again we consider two cases for t: $t \in [0, a) \cup [2^{N_{m+1}} - a, 2^{N_{m+1}})$ or $t \in [a, 2^{N_{m+1}} - a)$. Assume the first case holds. There are two subcases to consider for $\gamma(X^{m+1})$: $\gamma(X^{m+1})$ belongs to $+\Gamma'_{m+1}$ or $\gamma(X^{m+1})$ belongs to $-\Gamma'_{m+1}$. These are complementary events each occuring with probability 1/2. Restricted to the first event, for every $x \in [t, t+1]$,

$$\phi^{(r)}_{\gamma(X^{m+1})}(x) \overset{(4.9)}{=} \tilde{\phi}^{(r)}(x) + f_{\gamma(X^{m+1})}(x) = \tilde{\phi}^{(r)}(x) \overset{(4.6)}{=} \tilde{\phi}^{(r)}(t)$$

where the second equality holds by the definition of f succeeding (4.9). Thus,

$$\sup_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})} = \inf_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})} = \tilde{\phi}^{(r)}(t)$$

Likewise, for the second subcase where we restrict to the event that $\gamma(X^{m+1})$ belongs to $-\Gamma'_{m+1}$,

$$\sup_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})} = \inf_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})} = -\tilde{\phi}^{(r)}(t)$$

Combining these yields

$$\mathbb{E}\left[\exp\left(y\left(\sup_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})}\right)\right)\right] = \frac{1}{2}\left(\exp\left(y\tilde{\phi}^{(r)}(t)\right) + \exp\left(-y\tilde{\phi}^{(r)}(t)\right)\right)$$

$$= \cosh\left(y\tilde{\phi}^{(r)}(t)\right) \leq \cosh\left(y\left\|\tilde{\phi}^{(r)}\right\|_\infty\right) \overset{(4.7)}{=} \cosh\left(y\left\|\phi^{(r)}\right\|_\infty \frac{1}{\sqrt{m+1}\ln(m+2)}\right)$$

$$\overset{\text{Lem 4.3}}{\leq} \exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|^2_\infty \frac{1}{(m+1)\ln(m+2)^2}\right)$$

and the same estimate holds for the essential infimum, verifying (4) in this case.

Now consider the second case, $t \in [a + (i-1)2^{N_m}, a + i2^{N_m}]$ for some $1 \leq i \leq A$. Again, there are two subcases to consider for $\gamma(X^{m+1})$: $\gamma(X^{m+1})$ belongs to $+\Gamma'_{m+1}$ or $\gamma(X^{m+1})$ belongs to $-\Gamma'_{m+1}$. Restricted to the first event, and for the range of t under consideration, X^{m+1} is equal in distribution to a copy of X^m (after an appropriate shift in the time parameter), by definition of $+\Gamma'_{m+1}$. Thus, for every $x \in [t, t+1]$,

$$\phi^{(r)}_{\gamma(X^{m+1})}(x) \overset{(4.9)}{=} \tilde{\phi}^{(r)}(x) + f_{\gamma(X^{m+1})}(x) = \tilde{\phi}^{(r)}(x) + \phi_{\gamma(X^m)}(x') \overset{(4.6)}{=} \tilde{\phi}^{(r)}(t) + \phi_{\gamma(X^m)}(x')$$

where $x' = x - a - (i-1)2^{N_m}$, and the second equality holds by the definition of f succeeding (4.9). Thus,

$$\sup_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})} = \tilde{\phi}^{(r)}(t) + \sup_{[t',t'+1]} \phi_{\gamma(X^m)}$$

$$\inf_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})} = \tilde{\phi}^{(r)}(t) + \inf_{[t',t'+1]} \phi_{\gamma(X^m)}$$

where $t' = t - a - (i-1)2^{N_m}$. Likewise, for the second subcase where we restrict to the event that $\gamma(X^{m+1})$ belongs to $-\Gamma'_{m+1}$,

$$\sup_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})} = -\tilde{\phi}^{(r)}(t) - \inf_{[t',t'+1]} \phi_{\gamma(X^m)}$$

$$\inf_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})} = -\tilde{\phi}^{(r)}(t) - \sup_{[t',t'+1]} \phi_{\gamma(X^m)}$$

Combining these and using the inductive hypothesis applied to (4) and some basic monotonicity and symmetry properties of cosh yields

$$\mathbb{E}\left[\exp\left(y\left(\sup_{[t,t+1]} \phi^{(r)}_{\gamma(X^{m+1})}\right)\right)\right]$$

$$= \frac{1}{2}\exp\left(y\tilde{\phi}^{(r)}(t)\right)\mathbb{E}\left[\exp\left(y\left(\sup_{[t',t'+1]} \phi^{(r)}_{\gamma(X^{m+1})}\right)\right)\right]$$

$$+\frac{1}{2}\exp\left(-y\tilde{\phi}^{(r)}(t)\right)\mathbb{E}\left[\exp\left(-y\left(\inf_{[t',t'+1]} \phi^{(r)}_{\gamma(X^{m+1})}\right)\right)\right]$$

$$\overset{\text{ind hyp}}{\leq} \frac{1}{2}\exp\left(y\tilde{\phi}^{(r)}(t)\right)\exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \sum_{n=1}^{m} \frac{1}{n\ln(n+1)^2}\right)$$

$$+\frac{1}{2}\exp\left(-y\tilde{\phi}^{(r)}(t)\right)\exp\left(\frac{(-y)^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \sum_{n=1}^{m} \frac{1}{n\ln(n+1)^2}\right)$$

$$= \cosh\left(y\tilde{\phi}^{(r)}(t)\right)\exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \sum_{n=1}^{m} \frac{1}{n\ln(n+1)^2}\right)$$

$$\leq \cosh\left(y\left\|\tilde{\phi}^{(r)}\right\|_\infty\right)\exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \sum_{n=1}^{m} \frac{1}{n\ln(n+1)^2}\right)$$

$$\overset{(4.7)}{=} \cosh\left(y\left\|\phi^{(r)}\right\|_\infty \frac{1}{\sqrt{m+1}\ln(m+2)}\right)\exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \sum_{n=1}^{m} \frac{1}{n\ln(n+1)^2}\right)$$

$$\overset{\text{Lem 4.3}}{\leq} \exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \frac{1}{(m+1)\ln(m+2)^2}\right)\exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \sum_{n=1}^{m} \frac{1}{n\ln(n+1)^2}\right)$$

$$= \exp\left(\frac{y^2}{2}\left\|\phi^{(r)}\right\|_\infty^2 \sum_{n=1}^{m+1} \frac{1}{n\ln(n+1)^2}\right)$$

and the same estimate holds for the infimum, verifying (4) in this case. This completes the inductive step and the proof of the lemma. □

Theorem 4.2. *For every $p > 0$, $r \geq 1$, coarsely dense set $N \subseteq J^{r-1}(\mathbb{R})$, and $R \geq 3$, let $B_N(R) := \{x \in N : d_{CC}(0,x) \leq R\}$. Then*

$$\Pi_p(B_N(R)) \gtrsim \frac{\ln(R)^{\frac{1}{p}-\frac{1}{2r}}}{\ln(\ln(R))^{\frac{1}{p}+\frac{1}{2r}}}$$

where the implicit constant can depend on r, p but not on N, R.

Proof. Let p, r, N be as above. Since the Markov convexity constant Π_p is scale-invariant, then by applying a dilation we may assume without loss of generality that every point of $J^{r-1}(\mathbb{R})$ is at a distance of at most 1 away from a point of N. Let $F_m : \Gamma_m \to J^{r-1}(\mathbb{R})$ be the sequence of maps from Lemma 4.18. Extend the domain of t for the random walks on Γ_m by $X_t^m := X_0^m$ if $t \leq 0$, and $X_t^m := X_{2^{N_m}}^m$ if $t \geq 2^{N_m}$. Each $\{X_t^m\}_{t\in\mathbb{Z}}$ is a Markov process on the state space Γ_m.

With full probability, $d_{CC}(X_t^m, 0_m) = \min(\max(0,t), 2^{N_m})$. Since $\tilde{X}_t^m(t-2^k)$ equals X_t^m in distribution, $(X_t^m, \tilde{X}_t^m(t-2^k))$ is a vertical pair with full probability. Then Lemma 4.18(3) applies, and we get the following lower bound for the left hand side of the Markov convexity inequality in Definition 1.10:

$$\sum_{k=0}^{\infty}\sum_{t\in\mathbb{Z}} \frac{\mathbb{E}[d_{CC}(F_m(X_t^m), F_m(\tilde{X}_t^m(t-2^k)))^p]}{2^{kp}} \overset{\text{Lem 1.2}}{\geq} \sum_{k=0}^{\infty}\sum_{t\in\mathbb{Z}} \frac{\mathbb{E}[|\pi_0(f_m(X_t^m) - f_m(\tilde{X}_t^m(t-2^k)))|^{p/r}]}{2^{kp}}$$

$$\overset{\text{Lem 4.18(3)}}{\geq} \frac{m^{-\frac{p}{2r}}}{\ln(m+1)^{\frac{p}{s}}} \sum_{k=0}^{\infty}\sum_{t\in\mathbb{Z}} \frac{\mathbb{E}[d_m(X_t^m, \tilde{X}_t^m(t-2^k))^p]}{2^{kp}} \overset{\text{Lem 4.16}}{\gtrsim} m^{-\frac{p}{2r}}\ln(m+1)^{-\frac{p}{r}} m 2^{N_m} = \frac{m^{1-\frac{p}{2r}}2^{N_m}}{\ln(m+1)^{\frac{p}{r}}}$$

In summary,

$$\sum_{k=0}^{\infty}\sum_{t\in\mathbb{Z}} \frac{\mathbb{E}[d_{CC}(F_m(X_t^m), F_m(\tilde{X}_t^m(t-2^k)))^p]}{2^{kp}} \gtrsim \frac{m^{1-\frac{p}{2r}}2^{N_m}}{\ln(m+1)^{\frac{p}{r}}} \tag{4.12}$$

Now we upper bound the right hand side of the Markov convexity inequality. Since $d_{CC}(F_m(X_{t+1}^m), F_m(X_t^m)) = 0$ whenever $t \leq 0$ or $t \geq 2^{N_m}$,

$$\sum_{t\in\mathbb{Z}} \mathbb{E}[d_{CC}(F_m(X_{t+1}^m), F_m(X_t^m))^p] = \sum_{t=0}^{2^{N_m}-1} \mathbb{E}[d_{CC}(F_m(X_{t+1}^m), F_m(X_t^m))^p] =: (*) \tag{4.13}$$

Then

$$(*) \overset{\text{Lem 4.18(2)}}{=} \sum_{t=0}^{2^{N_m}-1} \mathbb{E}\left[d_{CC}([j^{r-1}(t+1)](\phi_{\gamma(X^m)})([j^{r-1}(t)](\phi_{\gamma(X^m)}))^p\right]$$

$$\overset{\text{Lem 1.1}}{\leq} \sum_{t=0}^{2^{N_m}-1} \mathbb{E}\left[\left(1+\left\|\phi^{(r)}_{\gamma(X^m)}\right\|_{L^\infty[t,t+1]}\right)^p\right]$$

$$\lesssim \sum_{t=0}^{2^{N_m}-1} 1+\mathbb{E}\left[\left\|\phi^{(r)}_{\gamma(X^m)}\right\|^p_{L^\infty[t,t+1]}\right] \overset{\text{Lems 4.4,4.18(4)}}{\lesssim} \sum_{t=0}^{2^{N_m}-1} 1 = 2^{N_m}$$

In summary,

$$\sum_{t\in\mathbb{Z}} \mathbb{E}[d_{CC}(F_m(X^m_{t+1}), F_m(X^m_t))^p] \lesssim 2^{N_m} \tag{4.14}$$

Let $\pi_N : J^{r-1}(\mathbb{R}) \to N$ be any map so that

$$d_{CC}(x, \pi_N(x)) \leq 1 \tag{4.15}$$

which exists by our initial assumption. We'll use π_N to transfer inequalities (4.12) and (4.14) to corresponding inequalities on N. Consider the maps $\bar{F}_m : \Gamma_m \to N$ defined by $\bar{F}_m := \pi_N \circ \delta_{2m} \circ F_m$. By Lemma 4.18(3),

$$d_{CC}(\delta_{2m}(F_m(q_1)), \delta_{2m}(F_m(q_2))) \geq 2d_m(q_1, q_2) \geq 4$$

for any vertical pair $(q_1, q_2) \in \Gamma_m \times \Gamma_m$. Combining this with (4.15) yields

$$d_{CC}(\bar{F}_m(q_1), \bar{F}_m(q_2)) \overset{(4.15)}{\geq} d_{CC}(\delta_{2m}(F_m(q_1)), \delta_{2m}(F_m(q_2))) - 2$$

$$\geq \frac{1}{2} d_{CC}(\delta_{2m}(F_m(q_1)), \delta_{2m}(F_m(q_2))) = m d_{CC}(F_m(q_1), F_m(q_2))$$

for any vertical pair (q_1, q_2). Combining this with (4.12) yields

$$\sum_{k=0}^{\infty} \sum_{t\in\mathbb{Z}} \frac{\mathbb{E}[d_{CC}(\bar{F}_m(X^m_t), \bar{F}_m(\tilde{X}^m_t(t-2^k)))^p]}{2^{kp}} \gtrsim \frac{m^{p+1-\frac{p}{2r}} 2^{N_m}}{\ln(m+1)^{\frac{p}{r}}} \tag{4.16}$$

Next,

$$d_{CC}(\bar{F}_m(X^m_{t+1}), \bar{F}_m(X^m_t)) \overset{(4.15)}{\leq} d_{CC}(\delta_{2m}(F_m(X^m_{t+1})), \delta_{2m}(F_m(X^m_t))) + 2$$

$$= 2m d_{CC}(F_m(X^m_{t+1}), F_m(X^m_t)) + 2$$

Combining this with (4.14) and (4.13) yields

$$\sum_{t\in\mathbb{Z}} \mathbb{E}[d_{CC}(\bar{F}_m(X^m_{t+1}), \bar{F}_m(X^m_t))^p] \lesssim m^p 2^{N_m} \tag{4.17}$$

For each $R \geq 1$, let $m(R)$ denote the largest m so that $\bar{F}_{m(R)}(\Gamma_{m(R)}) \subseteq B_N(R)$. Then (4.16) and (4.17) imply

$$\Pi_p(B_N(R)) \gtrsim \frac{m(R)^{\frac{1}{p}-\frac{1}{2r}}}{\ln(m(R)+1)^{\frac{1}{r}}} \tag{4.18}$$

Now we wish to estimate the quantity $m(R)$. Let $m \geq 0$ be arbitrary. Since any two points of Γ_m are connected by a geodesic that is a piecewise directed path, the Lipschitz constant of any map on Γ_m is the maximum of the Lipschitz constants of the map restricted to directed paths. Thus, by Lemmas 4.18(2), 4.18(5), and 1.1, $\mathrm{Lip}(F_m) \lesssim \sqrt{m}$. Since $\mathrm{diam}(\Gamma_m) = 2^{N_m} \leq 2^{Cm\log_2(m+1)+1}$ and $F_m(0_m) = 0$, this implies $F_m(\Gamma_m) \subseteq B_{J^{r-1}(\mathbb{R})}(R')$ with $R' \lesssim (m+1)^{Cm+\frac{1}{2}}$. Then $\delta_{2m}(F_m(\Gamma_m)) \subseteq B_{J^{r-1}(\mathbb{R})}(R'')$ with $R'' \lesssim (m+1)^{Cm+\frac{3}{2}}$. Then $\bar{F}_m(\Gamma_m) = \pi_N(\delta_{2m}(F_m(\Gamma_m))) \subseteq B_{J^{r-1}(\mathbb{R})}(R''+1)$. This implies, for any $R \geq 1$, $R \lesssim (m(R)+1)^{Cm(R)+\frac{3}{2}}$, where the implied constant is independent of R. This implies $m(R) \gtrsim \frac{\ln(R)}{\ln(\ln(R))}$ for $R \geq 3$. Plugging this into (4.18) yields

$$\Pi_p(B_N(R)) \gtrsim \frac{\ln(R)^{\frac{1}{p}-\frac{1}{2r}}}{\ln(\ln(R))^{\frac{1}{p}+\frac{1}{2r}}}$$

$\square$

Chapter 5

References

[ANT13] T. Austin, A. Naor, and R. Tessera, *Sharp quantitative nonembeddability of the Heisenberg group into superreflexive Banach spaces*, Groups Geom. Dyn. **7** (2013), no. 3, 497–522. MR 3095705

[Aro76] N. Aronszajn, *Differentiability of Lipschitzian mappings between Banach spaces*, Studia Math. **57** (1976), no. 2, 147–190. MR 0425608

[Ass83] P. Assouad, *Plongements lipschitziens dans* $\mathbf{R}^n$, Bull. Soc. Math. France **111** (1983), no. 4, 429–448. MR 763553

[Bat15] D. Bate, *Structure of measures in Lipschitz differentiability spaces*, J. Amer. Math. Soc. **28** (2015), no. 2, 421–482. MR 3300699

[BL00] Y. Benyamini and J. Lindenstrauss, *Geometric nonlinear functional analysis. Vol. 1*, American Mathematical Society Colloquium Publications, vol. 48, American Mathematical Society, Providence, RI, 2000. MR 1727673

[BLU07] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007. MR 2363343

[Bou86] J. Bourgain, *The metrical interpretation of superreflexivity in Banach spaces*, Israel J. Math. **56** (1986), no. 2, 222–230. MR 880292

[Che99] J. Cheeger, *Differentiability of Lipschitz functions on metric measure spaces*, Geom. Funct. Anal. **9** (1999), no. 3, 428–517. MR 1708448

[Chr73] J. P. R. Christensen, *Measure theoretic zero sets in infinite dimensional spaces and applications to differentiability of Lipschitz mappings*, Publ. Dép. Math. (Lyon) **10** (1973), no. 2, 29–39, Actes du Deuxième Colloque d'Analyse Fonctionnelle de Bordeaux (Univ. Bordeaux, 1973), I, pp. 29–39. MR 0361770

[CK06] J. Cheeger and B. Kleiner, *On the differentiability of Lipschitz maps from metric measure spaces to Banach spaces*, Inspired by S. S. Chern, Nankai Tracts Math., vol. 11, World Sci. Publ., Hackensack, NJ, 2006, pp. 129–152. MR 2313333

[CK09] ______, *Differentiability of Lipschitz maps from metric measure spaces to Banach spaces with the Radon-Nikodým property*, Geom. Funct. Anal. **19** (2009), no. 4, 1017–1028. MR 2570313

[CK10] ______, *Differentiating maps into* L^1*, and the geometry of BV functions*, Ann. of Math. (2) **171** (2010), no. 2, 1347–1385. MR 2630066

[CK15] ______, *Inverse limit spaces satisfying a Poincaré inequality*, Anal. Geom. Metr. Spaces **3** (2015), 15–39. MR 3300718

[Dal15] A. Dalet, *Free spaces over some proper metric spaces*, Mediterr. J. Math. **12** (2015), no. 3, 973–986. MR 3376824

[dLPP19] B. de Mendonça Braga, G. Lancien, C. Petitjean, and A. Procházka, *On Kalton's interlaced graphs and nonlinear embeddings into dual Banach spaces*, arXiv e-prints (2019), arXiv:1909.12132.

[God15] G. Godefroy, *A survey on Lipschitz-free Banach spaces*, Comment. Math. **55** (2015), no. 2, 89–118. MR 3518958

[HK98] J. Heinonen and P. Koskela, *Quasiconformal maps in metric spaces with controlled geometry*, Acta Math. **181** (1998), no. 1, 1–61. MR 1654771

[HS90] W. Hebisch and A. Sikora, *A smooth subadditive homogeneous norm on a homogeneous group*, Studia Math. **96** (1990), no. 3, 231–236. MR 1067309

[Jen68] T. M. Jenkins, *Banach Spaces of Lipschitz Functions of an Abstract Metric Space*, ProQuest LLC, Ann Arbor, MI, 1968, Thesis (Ph.D.)–Yale University. MR 2617138

[JL84] W. B. Johnson and J. Lindenstrauss, *Extensions of Lipschitz mappings into a Hilbert space*, Conference in modern analysis and probability (New Haven, Conn., 1982), Contemp. Math., vol. 26, Amer. Math. Soc., Providence, RI, 1984, pp. 189–206. MR 737400

[Jun17] D. Jung, *A variant of Gromov's problem on Hölder equivalence of Carnot groups*, J. Math. Anal. Appl. **456** (2017), no. 1, 251–273. MR 3680967

[Jun19] ______, *Bilipschitz embeddings of spheres into jet space Carnot groups not admitting Lipschitz extensions*, Ann. Acad. Sci. Fenn. Math. **44** (2019), no. 1, 261–280. MR 3919136

[Kal04] N. J. Kalton, *Spaces of Lipschitz and Hölder functions and their applications*, Collect. Math. **55** (2004), no. 2, 171–217. MR 2068975

[KM16] B. Kleiner and J. M. Mackay, *Differentiable structures on metric measure spaces: a primer*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **16** (2016), no. 1, 41–64. MR 3524664

[Laa00] T. J. Laakso, *Ahlfors Q-regular spaces with arbitrary $Q > 1$ admitting weak Poincaré inequality*, Geom. Funct. Anal. **10** (2000), no. 1, 111–123. MR 1748917

[LD17] E. Le Donne, *A primer on Carnot groups: homogenous groups, Carnot-Carathéodory spaces, and regularity of their isometries*, Anal. Geom. Metr. Spaces **5** (2017), no. 1, 116–137. MR 3742567

[LDLM18] E. Le Donne, S. Li, and T. Moisala, *Gâteaux differentiability on infinite-dimensional carnot groups*, arxiv (2018).

[LDRW18] E. Le Donne, T. Rajala, and E. Walsberg, *Isometric embeddings of snowflakes into finite-dimensional Banach spaces*, Proc. Amer. Math. Soc. **146** (2018), no. 2, 685–693. MR 3731701

[Li14] S. Li, *Coarse differentiation and quantitative nonembeddability for Carnot groups*, J. Funct. Anal. **266** (2014), no. 7, 4616–4704. MR 3170215

[Li16] ______, *Markov convexity and nonembeddability of the Heisenberg group*, Ann. Inst. Fourier (Grenoble) **66** (2016), no. 4, 1615–1651. MR 3494180

[LN06] J. R. Lee and A. Naor, *Lp metrics on the heisenberg group and the goemans-linial conjecture*, 2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06), Oct 2006, pp. 99–108.

[LN14] V. Lafforgue and A. Naor, *Vertical versus horizontal Poincaré inequalities on the Heisenberg group*, Israel J. Math. **203** (2014), no. 1, 309–339. MR 3273443

[LNP09] J. R. Lee, A. Naor, and Y. Peres, *Trees and Markov convexity*, Geom. Funct. Anal. **18** (2009), no. 5, 1609–1659. MR 2481738

[LP01] U. Lang and C. Plaut, *Bilipschitz embeddings of metric spaces into space forms*, Geom. Dedicata **87** (2001), no. 1-3, 285–307. MR 1866853

[Man73] P. Mankiewicz, *On the differentiability of Lipschitz mappings in Fréchet spaces*, Studia Math. **45** (1973), 15–29. MR 0331055

[MN13] M. Mendel and A. Naor, *Markov convexity and local rigidity of distorted metrics*, J. Eur. Math. Soc. (JEMS) **15** (2013), no. 1, 287–337. MR 2998836

[MO80] P. W. McCartney and R. C. O'Brien, *A separable Banach space with the Radon-Nikodým property that is not isomorphic to a subspace of a separable dual*, Proc. Amer. Math. Soc. **78** (1980), no. 1, 40–42. MR 548080

[Mos62] G. D. Mostow, *Homogeneous spaces with finite invariant measure*, Ann. of Math. (2) **75** (1962), 17–37. MR 145007

[Mos73] ______, *Strong rigidity of locally symmetric spaces*, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1973, Annals of Mathematics Studies, No. 78. MR 0385004

[Nao12] A. Naor, *An introduction to the Ribe program*, Jpn. J. Math. **7** (2012), no. 2, 167–233. MR 2995229

[Nao18] ______, *Metric dimension reduction: A snapshot of the ribe program*, Proc. Int. Cong. of Math., vol. 1, 2018, pp. 759–838.

[NY18] A. Naor and R. Young, *Vertical perimeter versus horizontal perimeter*, Ann. of Math. (2) **188** (2018), no. 1, 171–279. MR 3815462

[Ost13] M. I. Ostrovskii, *Metric embeddings*, De Gruyter Studies in Mathematics, vol. 49, De Gruyter, Berlin, 2013, Bilipschitz and coarse embeddings into Banach spaces. MR 3114782

[Ost14a] ______, *Metric spaces nonembeddable into Banach spaces with the Radon-Nikodým property and thick families of geodesics*, Fund. Math. **227** (2014), no. 1, 85–96. MR 3247034

[Ost14b] ______, *Radon-Nikodým property and thick families of geodesics*, J. Math. Anal. Appl. **409** (2014), no. 2, 906–910. MR 3103207

[Ost16] ______, *Metric characterizations of some classes of Banach spaces*, Harmonic analysis, partial differential equations, complex analysis, Banach spaces, and operator theory. Vol. 1, Assoc. Women Math. Ser., vol. 4, Springer, [Cham], 2016, pp. 307–347. MR 3627726

[Pan89] P. Pansu, *Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un*, Ann. of Math. (2) **129** (1989), no. 1, 1–60. MR 979599

[Pis16] G. Pisier, *Martingales in Banach spaces*, Cambridge Studies in Advanced Mathematics, vol. 155, Cambridge University Press, Cambridge, 2016. MR 3617459

[Rag72] M. S. Raghunathan, *Discrete subgroups of Lie groups*, Springer-Verlag, New York-Heidelberg, 1972, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68. MR 0507234

[Rib76] M. Ribe, *On uniformly homeomorphic normed spaces*, Ark. Mat. **14** (1976), no. 2, 237–244. MR 0440340

[RW10] S. Rigot and S. Wenger, *Lipschitz non-extension theorems into jet space Carnot groups*, Int. Math. Res. Not. IMRN (2010), no. 18, 3633–3648. MR 2725507

[Sch14] A. Schioppa, *Derivations and Alberti representations*, ProQuest LLC, Ann Arbor, MI, 2014, Thesis (Ph.D.)–New York University. MR 3279153

[Sem96] S. Semmes, *On the nonexistence of bi-Lipschitz parameterizations and geometric problems about A_∞-weights*, Rev. Mat. Iberoamericana **12** (1996), no. 2, 337–410. MR 1402671

[Vem04] S. S. Vempala, *The random projection method*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 65, American Mathematical Society, Providence, RI, 2004, With a foreword by Christos H. Papadimitriou. MR 2073630

[War05] B. Warhurst, *Jet spaces as nonrigid Carnot groups*, J. Lie Theory **15** (2005), no. 1, 341–356. MR 2115247

[Wea99] N. Weaver, *Lipschitz algebras*, World Scientific Publishing Co., Inc., River Edge, NJ, 1999. MR 1832645

[Wol03] T. H. Wolff, *Lectures on harmonic analysis*, University Lecture Series, vol. 29, American Mathematical Society, Providence, RI, 2003, With a foreword by Charles Fefferman and a preface by Izabella Łaba, Edited by Łaba and Carol Shubin. MR 2003254

[Yu00] G. Yu, *The coarse Baum-Connes conjecture for spaces which admit a uniform embedding into Hilbert space*, Invent. Math. **139** (2000), no. 1, 201–240. MR 1728880