MINICOURSE: QUANTITATIVE DIFFERENTIABILITY AND RECTIFIABILITY

ROBERT YOUNG

CONTENTS

Differentiability measures how well a function can be approximated by affine functions. Rectifiability likewise measures how well a set can be approximated by planes. But both of these notions operate at *infinitesimal* scales; they deal with the properties of limits. In these notes, we will study quantitative versions of these notions that operate at *local* scales; scales that are small but bounded away from zero. How well can a function or a set be approximated by affine functions or sets at local scales? How often can a function or set fail to be approximated by an affine function or set? How can we use questions like these to study the geometry and analysis of sets and functions?

1. COARSE DIFFERENTIATION OF CURVES

We'll start with the simplest case: Let X be a metric space, say $\mathbb{R}^n$, Hilbert space, or a Banach space. How can we describe the infinitesimal and local structure of maps $f : \mathbb{R} \to X$?

For maps from $\mathbb{R}$ to $\mathbb{R}$, the Lebesgue Differentiation Theorem says:

Theorem 1.1. *If $f : \mathbb{R} \to \mathbb{R}$ is an absolutely continuous map, then f is differentiable almost everywhere. That is, for almost every $x \in \mathbb{R}$, there is an $f'(x)$ such that*

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = f'(x).$$

In particular, if f is Lipschitz, f is differentiable almost everywhere. But this is an infinitesimal result – what about local? What's the largest segment we can expect to find where f is ϵ–close to affine?

To answer this, we need a notion of coarse differentiation (see works of Jones [Jon90], Eskin–Fisher–Whyte [EFW12], Cheeger [Che99], et al.).

Date: June 10, 2025.

Definition 1.2. Let X be a metric space and let $f: \mathbb{R} \to X$. Let $I = [a, b]$ be an interval of length L, let $\epsilon = \frac{1}{n}$ for some $n \in \mathbb{N}$. We say that f is ϵ*–efficient* on I if

$$\sum_{i=0}^{n-1} d_X(f(a+(i+1)\frac{L}{n}), f(a+i\frac{L}{n}))| \leq d_X(f(b), f(a)) + \epsilon L.$$

If $X = \mathbb{R}$, this is also called ϵ–monotone – the inequality implies that graph can't backtrack by more than about ϵ.

Then the following lemma is simple:

Lemma 1.3. *Let $f: [0,1] \to X$ be 1–Lipschitz, let $\epsilon > 0$, and let $n = \epsilon^{-1} \in \mathbb{Z}$. There is an interval $I \subset [0,1]$ of length $L > \epsilon^n$ such that f is ϵ–efficient on I.*

Proof. Let

$$\ell_k = \sum_{i=0}^{\epsilon^{-k}} d(f(i\epsilon^k), f((i+1)\epsilon^k)).$$

By the Lipschitz property, $\ell_k \leq 1$. Let $x_{i,k} = i\epsilon^k$. If $I_{i,k} = [x_{i,k}, x_{i+1,k}]$ is not ϵ–efficient for any i, then for any i,

$$\sum_{j=0}^{n-1} d(f(x_{ni+j,k+1}), f(x_{ni+j+1,k+1})) \geq d(f(x_{i,k}), f(x_{i+1,k})) + \epsilon^{k+1},$$

and

$$\ell_{k+1} = \sum_{i=0}^{n^k-1} \sum_{j=0}^{n-1} d(f(x_{ni+j,k+1}), f(x_{ni+j+1,k+1})) > \ell_k + \epsilon.$$

By induction, $\ell_k > \epsilon k$. When $k > n$, this is a contradiction, so f is ϵ–efficient on some interval of length at most ϵ^n. □

The basic idea behind this proof is one that we'll see again and again: If we can break down some finite geometric quantity into contributions from many different intervals, then there must be intervals that contribute less than ϵ. In this case, we break the length of f into contributions

$$\ell_{i,k} = d(f(x_{i,k}), f(x_{i+1,k})) - \sum_{j=0}^{n-1} d(f(x_{ni+j,k+1}), f(x_{ni+j+1,k+1}))$$

associated with each interval. Intervals where f is ϵ–*in*efficient contribute a nonzero quantity to the length of f, so the number of ϵ–inefficient intervals is bounded in terms of the length of f.

Ideally, the quantity should be *coercive*: intervals on which the quantity is zero should lie in some nice class, and intervals on which the quantity is ϵ should be close to that class. The notion of ϵ–efficiency is mildly coercive: if X is a length space and $f: I \to X$ is ϵ–efficient on every interval and for all ϵ, then f is a length-minimizing curve parameterized monotonically. We can improve this by adding a dimension.

For $\epsilon > 0$, we say that $f: I \to \mathbb{R}^k$ is ϵ*–coarsely differentiable* on I if there is an affine function such that $|f(t) - \lambda(t)| \leq \epsilon \ell(I)$ for every $t \in I$.

Proposition 1.4. *Let $f\colon [0,1] \to \mathbb{R}^k$ be 1–Lipschitz, $\epsilon > 0$. There is an interval $I \subset [0,1]$ such that f is ϵ*–coarsely differentiable *on I.*

Proof. Consider the graph $g\colon [0,1] \to \mathbb{R}^{k+1}$, $g(t) = (t, f(t))$. This is Lipschitz, so for any $\delta > 0$, there is an interval I on which g is δ–efficient. By the Pythagorean Theorem, $g(I)$ is in the $\sim \sqrt{\delta}\ell(I)$–neighborhood of a line segment $L \in \mathbb{R}^{k+1}$, say $L = \{(t, \lambda(t)\}$, where $\lambda(t) = mt + b$. Since f was 1–Lipschitz, we can take $|m| \le 1$. Then $d((x, y), L) \approx |y - \lambda(x)|$, so for $t \in I$, we have

$$|f(t) - \lambda(t)| \approx d((t, f(t)), L) \lesssim \sqrt{\delta}\ell(I).$$

□

In fact, intervals like this are abundant. The notions of ϵ–efficient and ϵ–coarse differentiable are scale-invariant, so Lemma 1.3 and Proposition 1.4 both apply to a Lipschitz function on any interval. So any subinterval of $[0,1]$ contains a smaller interval on which f is ϵ–efficient or ϵ–coarse differentiable.

More quantitatively, f is ϵ–efficient on "most" subintervals of $[0,1]$ in the following sense.

Definition 1.5. Let $D \subset \mathbb{R} \times \mathbb{R}^+$ be a measurable set. Let $D_r = \{x \mid (x, r) \in D\}$. We say that D is a *(C-)Carleson set* if there is a $C > 0$ such that for every bounded interval $I \subset \mathbb{R}$ of length $L = |I|$,

$$\int_0^L \int_I \mathbf{1}_D(x, r)\,\mathrm{d}x \frac{\mathrm{d}r}{r} = \int_0^L |D_r \cap I| \frac{\mathrm{d}r}{r} = \int_{-\infty}^{\log L} |D_{e^t} \cap I|\,\mathrm{d}t \le C|I|, \tag{1}$$

where $|D_r \cap I|$ is the Lebesgue measure of $D_r \cap I$.

In many cases, including the results below, D represents the set of intervals on which a function f is poorly behaved, so that $(x, r) \in D$ if and only if f is "bad" on $[x - r, x + r]$. Equation (1) then says that for any interval I, f is "good" on most subintervals of I (i.e., the set of "bad" subintervals has finite measure with respect to $\mathrm{d}x \frac{\mathrm{d}r}{r}$).

Exercise 1. Technically, the integral in (1) is computed over intervals centered in I with radius at most L. Show that most of these intervals are subintervals of I.

Exercise 2. Let $S \subset [0,1]$. What are some conditions under which most subintervals of $[0,1]$ are disjoint from S?

Theorem 1.6. *Let $f : \mathbb{R} \to X$ be 1–Lipschitz, $\epsilon > 0$ and let $S_\epsilon \subset \mathbb{R} \times \mathbb{R}^+$ be the set*

$$S_\epsilon = \{(x, r) \mid f \textit{ is not } \epsilon\textit{–efficient on } [x - \frac{r}{2}, x + \frac{r}{2}]\}.$$

Then S_ϵ is Carleson.

We'll see a number of theorems that look like this, and I suggest considering a few examples. For example, $f(x) = |x|$ fails to be ϵ–efficient on intervals containing 0, and one can check that the corresponding set $\{(x, r) \mid |x| < r\}$ is Carleson. Similarly, for any $\epsilon > 0$, there are $0 < a < b$ such that $f(x) = \sin x$ is ϵ–efficient on intervals with length less than a or greater than b, and $S_\epsilon \subset \mathbb{R} \times [a, b]$, which likewise a Carleson set.

Exercise 3. Check the details above for $f(x) = |x|$ and $f(x) = \sin x$.

By the Pythagorean theorem argument in Proposition 1.4, when $X = \mathbb{R}^k$, we can replace ϵ–efficient with ϵ–coarse differentiable:

Corollary 1.7. *Let $f : \mathbb{R} \to \mathbb{R}^k$ be 1–Lipschitz, $\epsilon > 0$ and let $D_\epsilon \subset [0,1] \times (0,1]$ be the set*

$$D_\epsilon = \{(x,r) \mid f \textit{ is not } \epsilon\textit{–coarse differentiable on } [x - \frac{r}{2}, x + \frac{r}{2}]\}. \tag{2}$$

Then D_ϵ is Carleson.

Corollary 1.7 is an example of a *weak geometric lemma*. Weak geometric lemmas state that a function or set is well-behaved (up to error ϵ) except on a C_ϵ–Carleson set, where C_ϵ depends on ϵ.

Proof of Theorem 1.6. By scale-invariance, it suffices to check the Carleson condition for the interval $[0,1]$.

First, the proof of Lemma 1.3 gives a discrete version of the Carleson condition. As before, let

$$\ell_k = \sum_{i=0}^{\epsilon^{-k}-1} d(f(i\epsilon^k), f((i+1)\epsilon^k)|.$$

Then, letting $S_\epsilon = S_\epsilon^f$,

$$\ell_{k+1} \geq \ell_k + \epsilon^{k+1} \sum_{i=0}^{\epsilon^{-k}-1} \mathbf{1}_{S_\epsilon}((i + \frac{1}{2})\epsilon^k, \epsilon^k),$$

so

$$\sum_{k=0}^{\infty} \sum_{i=0}^{\epsilon^{-k}-1} \epsilon^{k+1} \mathbf{1}_{S_\epsilon}((i + \frac{1}{2})\epsilon^k, \epsilon^k) \leq 1. \tag{3}$$

Let $f_t(x) = f(x+t)$. Then (3) applies to f_t, so

$$\sum_{k=0}^{\infty} \sum_{i=0}^{\epsilon^{-k}-1} \epsilon^{k+1} \mathbf{1}_{S_\epsilon^t}((i + \frac{1}{2})\epsilon^k, \epsilon^k) = \sum_{k=0}^{\infty} \sum_{i=0}^{\epsilon^{-k}-1} \epsilon^{k+1} \mathbf{1}_{S_\epsilon}(t + (i + \frac{1}{2})\epsilon^k, \epsilon^k) \leq 1.$$

We integrate over t to get

$$\begin{aligned} &\int_{-1}^{1} \sum_{k=0}^{\infty} \sum_{i=0}^{\epsilon^{-k}-1} \epsilon^{k+1} \mathbf{1}_{S_\epsilon}(t + (i + \frac{1}{2})\epsilon^k, \epsilon^k)\, \mathrm{d}t \\ &= \sum_{k=0}^{\infty} \sum_{i=0}^{\epsilon^{-k}-1} \int_{-1}^{1} \epsilon^{k+1} \mathbf{1}_{S_\epsilon}(t + (i + \frac{1}{2})\epsilon^k, \epsilon^k)\, \mathrm{d}t \leq 2. \end{aligned} \tag{4}$$

Let $S_{\epsilon,r} = S_{\epsilon,r}^f = \{x \mid (x,r) \in S_\epsilon\}$. Then for any $x_0 \in [0,1]$ and $r > 0$,

$$\int_{-1}^{1} \mathbf{1}_{S_\epsilon}(t + x_0, r)\, \mathrm{d}t = |S_{\epsilon,r} \cap [x_0 - 1, x_0 + 1]| \geq |S_{\epsilon,r} \cap [0,1]|.$$

Combining this with (4), we find

$$\begin{aligned}2 &\geq \sum_{k=0}^{\infty} \sum_{i=0}^{\epsilon^{-k}-1} \epsilon^{k+1} |S_{\epsilon,\epsilon^k} \cap [0,1]| \\ &\geq \sum_{k=0}^{\infty} \epsilon |S_{\epsilon,\epsilon^k} \cap [0,1]|.\end{aligned}$$

A similar argument holds for rescalings. For $\rho > 0$, let $f^\rho(t) = \rho^{-1} f(\rho t)$. Then f^ρ is 1–Lipschitz and

$$S_\epsilon^{f^\rho} = \{(x,r) \mid (\rho x, \rho r) \in S_\epsilon\}.$$

It follows that

$$\sum_{k=0}^{\infty} |S^{f^\rho}_{\epsilon,\epsilon^k} \cap [0,1]| = \sum_{k=0}^{\infty} \rho^{-1} |S_{\epsilon,\rho\epsilon^k} \cap [0,\rho]| \leq 2\epsilon^{-1}$$

for all $\rho > 0$ and thus

$$\sum_{k=0}^{\infty} |S_{\epsilon,\rho\epsilon^k} \cap [0,\rho]| \leq 2\rho\epsilon^{-1}.$$

Therefore,

$$\begin{aligned}\int_0^1 |S_{\epsilon,r} \cap [0,1]| \frac{\mathrm{d}r}{r} &= \sum_{k=1}^{\infty} \int_{\epsilon^k}^{\epsilon^{k-1}} |S_{\epsilon,r} \cap [0,1]| \frac{\mathrm{d}r}{r} \\ &\leq \int_1^{\epsilon^{-1}} \sum_{k=0}^{\infty} |S_{\epsilon,r\epsilon^k} \cap [0,r]| \frac{\mathrm{d}r}{r} \\ &\leq \int_1^{\epsilon^{-1}} 2r\epsilon^{-1} \frac{\mathrm{d}r}{r} \\ &\leq 2\epsilon^{-2}.\end{aligned}$$

□

Remark. These coarse differentiability results aren't quite enough to imply that f is genuinely differentiable. For example, Corollary 1.7 implies that for almost every x, there is an affine λ and a sequence of radii r_k such that $\frac{\|f-\lambda\|_{L_\infty([x-r_k,x+r_k])}}{r_k} \to 0$, but λ need not be unique.

Remark. It is natural to consider versions of Proposition 1.4 and Theorem 1.6 for maps to other spaces, such as Banach spaces. When $f\colon [0,1] \to X$ is a map to a Banach space, the coarse differentiability of f depends on the convexity of $\|\cdot\|_X$. For example, let $f\colon [0,1] \to L_1([0,1])$, $f(t) = \mathbf{1}_{[0,t]}$ be the map that sends t to the characteristic function $\mathbf{1}_{[0,t]}$ of $[0,t]$. For $s < t$,

$$\|f(t) - f(s)\|_1 = \|\mathbf{1}_{(s,t]}\|_1 = |s-t|,$$

so f is a geodesic and thus δ–efficient for any δ, but f cannot be approximated by an affine map on any interval. This arises partly from the fact that the unit ball in L_1 is convex but not strictly convex, so there can be many geodesics between two points.

2. CURVES, SURFACES, AND METRIC SPACES

We now pass from curves (maps from $\mathbb{R}$ to X) to maps between metric spaces (maps from X to Y). The first case we consider is functions of multiple variables. In this case, Rademacher's Theorem generalizes Theorem 1.1 to Lipschitz functions $f\colon \mathbb{R}^k \to \mathbb{R}$.

Theorem 2.1. *If $f\colon \mathbb{R}^k \to \mathbb{R}$ is a Lipschitz map, then f is differentiable almost everywhere. That is, for almost every $x \in \mathbb{R}$, there is a linear function $Df_x\colon \mathbb{R}^k \to \mathbb{R}$ such that*

$$\lim_{h\to 0} \frac{f(x+h)-f(x)-Df_x(h)}{h} = 0.$$

Again, this is an infinitesimal result, but, as before, we can show that f is coarse differentiable on "most" balls.

To formalize this, we first generalize the definition of a Carleson set to $\mathbb{R}^k$.

Definition 2.2. Let $D \subset \mathbb{R}^k \times \mathbb{R}^+$ be a measurable set. Let $D_r = \{x \mid (x,r) \in D\}$. We say that D is a *(C–)Carleson set* if there is a $C > 0$ such that for every ball $B \subset \mathbb{R}^k$ of radius r,

$$\int_0^r \int_B \mathbf{1}_D(x,t)\,\mathrm{d}x \frac{\mathrm{d}t}{t} = \int_0^r |D_t \cap B| \frac{\mathrm{d}t}{t} \le C|B|. \tag{5}$$

Similarly to Definition 1.5, D often represents the set of balls on which some function f is "bad", with $(x,r) \in D$ if and only if f is "bad" on $B_r(x)$. As before, if D is Carleson, then for any ball B, f is "good" on most subballs of B.

We claim that Lipschitz functions $f\colon \mathbb{R}^k \to \mathbb{R}$ satisfy a weak geometric lemma. For $\epsilon > 0$, we say that $f\colon \mathbb{R}^k \to \mathbb{R}$ is *ϵ–coarse differentiable* on a ball $B_r(x)$ if there is an affine function such that $|f(v) - \lambda(v)| \le \epsilon r$ for every $v \in B_r(x)$.

Proposition 2.3 (Coarse Rademacher)**.** *Let $f\colon \mathbb{R}^k \to \mathbb{R}$ be 1–Lipschitz, $\epsilon > 0$ and let $D_\epsilon \subset \mathbb{R}^k \times \mathbb{R}^+$ be the set*

$$D_\epsilon = \{(x,r) \mid f \text{ is not } \epsilon\text{–coarsely differentiable on } B_r(x)\}.$$

Then D_ϵ is Carleson.

Our strategy to prove Proposition 2.3 has two parts. First, we use Corollary 1.7 to show that f is coarsely differentiable on "most" line segments in $\mathbb{R}^k$. Then, we show that a function which is coarsely differentiable on "most" line segments is coarsely differentiable on balls.

For the first part, we will need some lemmas on transformations and combinations of Carleson sets. The first is straightforward and we omit the proof.

Lemma 2.4. *Let $\rho > 0$ and suppose that $D \subset \mathbb{R}^k \times \mathbb{R}^+$ is a Carleson set. Then*

$$D' = \{(x,r) \in \mathbb{R}^k \times \mathbb{R}^+ \mid (x,\rho r) \in D$$

is also a Carleson set.

The second is more subtle.

Lemma 2.5. *Let $C > 0$, let A be a set, and let μ be a probability measure on A. For $\alpha \in A$, let $D_\alpha \subset \mathbb{R}^k \times \mathbb{R}^+$ be a C–Carleson set so that D_α depends measurably on α.*

Let μ be a measure on A with $\mu(A) = 1$. For any $\gamma \in [0,1]$, the set

$$D' = \{(x,r) \in \mathbb{R}^k \times \mathbb{R}^+ \mid \mu(\{\alpha \in A \mid (x,r) \in D_\alpha\}) \geq \gamma\}$$

is $\gamma^{-1}C$–Carleson.

For example, if $D_1,\dots,D_n$ are C–Carleson, we can apply Lemma 2.5 with $A = \{1,\dots,n\}$, μ the uniform measure on A, and $\gamma = \frac{1}{n}$ to show that $\bigcup_{i=1}^n D_i$ is nC–Carleson.

Proof. Let $\mathbf{1}_{D_\alpha}$ be the characteristic function of D_α and let

$$g(x,r) = \mu(\{\alpha \in A \mid (x,r) \in D_\alpha\}) = \int_A \mathbf{1}_{D_\alpha}(x,r)\,\mathrm{d}\mu(\alpha).$$

Then $D' = g^{-1}([\gamma,1])$, i.e.,

$$\mathbf{1}_{D'}(x,r) \leq \gamma^{-1} \int_A \mathbf{1}_{D_\alpha}(x,r)\,\mathrm{d}\mu(\alpha). \tag{6}$$

Let $B \subset \mathbb{R}^k$ be a ball of radius r. Integrating (6), we get

$$\int_0^r \int_B g(x,t)\,\mathrm{d}x \frac{\mathrm{d}t}{t} \leq \gamma^{-1} \int_0^r \int_B \int_A \mathbf{1}_D(x,t)\,\mathrm{d}\mu(\alpha)\,\mathrm{d}x \frac{\mathrm{d}t}{t} \leq \gamma^{-1} C|B|,$$

as desired, where the second inequality follows from Fubini's Theorem and the Carleson condition. □

Finally, we leave the following lemma as an exercise for the reader.

Lemma 2.6. *Let $S \subset \mathbb{R}^k \times \mathbb{R}^+$ be a C–Carleson set and let $\delta > 0$. Then there is a $C' = C'(C,\delta,k)$ such that*

$$T = \{(x,r) \in \mathbb{R}^k \times \mathbb{R}^+ \mid |B_r(x) \cap S_r| > \delta |B_r(x)|\}.$$

is C'–Carleson.

Exercise 4. For $x, v \in \mathbb{R}^k$ and $r > 0$, let $\eta_v(x,r) = (x + rv, r)$. Show that if S is C–Carleson, then $\eta_v(S)$ is also Carleson, with constant depending on C, k, and $\|v\|$. Use these sets and Lemma 2.5 to prove Lemma 2.6.

Proof of Proposition 2.3. Since f is Lipschitz, the restriction of f to any line is Lipschitz. We thus consider foliations of $\mathbb{R}^k$ by lines. For any vector $v \in \mathbb{R}^k$, let

$$D_{\epsilon,v}(f) = \{(x,r) \mid f \text{ is not } \epsilon\text{–coarsely differentiable on the line segment } [x - rv, x + rv].$$

By Corollary 1.7 and Fubini's Theorem, each set $D_{\epsilon,v}$ is Carleson, with constant depending on $\|v\|$.

Let ν be the uniform probability measure on the unit sphere S^{k-1}. By Lemma 2.4, Lemma 2.5 and Lemma 2.6, for any $\delta > 0$, the set

$$T_\delta(f) = \{(x,r) \mid \int_{S^{k-1}} \int_{B_{2r}(x)} \mathbf{1}_{D_{\delta,v}(f)}(y, 4r)\,\mathrm{d}y\,\mathrm{d}\nu(v) > \delta |B_r|\}$$

is Carleson.

If $(x,r) \not\in T_\delta(f)$, then f is δ–coarsely differentiable on all but a small fraction of lines through the ball $B_{2r}(x)$. We claim that if δ is sufficiently small, then f is coarsely differentiable on $B_r(x)$.

There are many ways to prove this, but the following argument by compactness is particularly slick. Suppose by contradiction that there is an $\epsilon > 0$ such that for any $n > 0$, there is a 1–Lipschitz function $f_n : \mathbb{R}^k \to \mathbb{R}$ such that $(x_n, r_n) \not\in T_{\frac{1}{n}}(f_n)$, but f_n is not ϵ–coarsely differentiable on $B_{r_n}(x_n)$. After rescaling and translating, we may suppose that $x_n = \mathbf{0}$, $r_n = 1$, and $f_n(\mathbf{0}) = 0$.

By the Arzelà–Ascoli theorem, we can pass to a subsequence so that f_n converges uniformly to a limit f. Furthermore, $(\mathbf{0}, 1) \not\in T_\delta(f)$ for all $\delta > 0$, which means that f is affine on every line segment in B_1. Such a function is affine on B_1. When n is sufficiently large, we have $\|f_n - f\|_{L_\infty(B_1)} < \epsilon$, so f_n is ϵ–coarsely differentiable, which is a contradiction.

Therefore, for any $\epsilon > 0$, there is a $\delta > 0$ such that if $(x,r) \not\in T_\delta(f)$, then f is ϵ–coarsely differentiable on $B_r(x)$. Since $T_\delta(f)$ is Carleson, this proves the proposition. □

2.1. **Application: Embeddings of the Heisenberg group.** These ideas work for other spaces too, notably Carnot groups. Let $\mathbb{H}$ be the Heisenberg group and consider a Lipschitz map $f : \mathbb{H} \to \mathbb{R}^k$. Can f be a bilipschitz embedding?

Let X, Y and Z be the standard left-invariant frame field of $\mathbb{H}$, so that Z is vertical and X and Y span the horizontal distribution of $\mathbb{H}$. Let

$$S^1_H = \{X \cos\theta + Y \sin\theta \mid \theta \in \mathbb{R}\}$$

be the set of unit horizontal vectors. Then for every left-invariant $V \in S^1_H$, there is a one-parameter subgroup generated by V, which we denote by $\langle V \rangle$, and the cosets of $\langle V \rangle$ form a foliation of $\mathbb{H}$ by lines. These lines are horizontal curves, so we call them *horizontal lines.* We write the elements of $\langle V \rangle$ as V^t, $t \in \mathbb{R}$

Every horizontal line L is embedded isometrically in $\mathbb{H}$, so if $f : \mathbb{H} \to \mathbb{R}^k$ is Lipschitz, then so is $f|_L$. Hence, similarly to the proof of Proposition 2.3, for any $\delta > 0$, the set of (x,r) such that

$$\int_{B_{\delta^{-1} r}(x)} \int_{S^1_H} \mathbf{1}_{f \text{ is } \delta\text{–coarse diff on } [yV^{4r}, yV^{-4r}]} \, \mathrm{d}V \, \mathrm{d}y > \delta |B_r(x)|$$

is Carleson. Call this set T_δ.

As before, if $(x,r) \in T_\delta$ for all $\delta > 0$, then f is affine on $B_r(x)$, but this is a little more difficult to prove. (To start, show that Vf is constant for any $V \in S^1_H$.) Once this is shown, the same argument by contradiction used in Proposition 2.3 proves the following generalization of Proposition 2.3.

Proposition 2.7 (Coarse Pansu)**.** *Let $f : \mathbb{H} \to \mathbb{R}^k$ be 1–Lipschitz and $\epsilon > 0$. We say that f is* coarse Pansu-differentiable *if there is a homomorphism $\lambda : \mathbb{H} \to \mathbb{R}^k$ and a $w \in \mathbb{R}^k$ such that $|f(v) - w - \lambda(x^{-1}v)| \le \epsilon r$ for every $v \in B_r(x)$.*

Let $D_\epsilon \subset \mathbb{R}^k \times \mathbb{R}^+$ be the set

$$D_\epsilon = \{(x,r) \mid f \textit{ is not } \epsilon\textit{–coarse Pansu-differentiable on } B_r(x)\}.$$

Then D_ϵ is Carleson.

In particular, this implies that $\mathbb{H}$ does not have a bilipschitz embedding into $\mathbb{R}^k$. Suppose that $f\colon \mathbb{H} \to \mathbb{R}^k$ is Lipschitz. Then for every $\epsilon > 0$, there is some ball $B_r(x)$ such that f is ϵ–coarse Pansu-differentiable on $B_r(x)$. Let $w \in \mathbb{R}^k$ and let $\lambda\colon \mathbb{H} \to \mathbb{R}^k$ be a homomorphism such that $|f(v) - w - \lambda(x^{-1}v)| \le \epsilon r$ for every $v \in B_r(x)$. Since $\mathbb{R}^k$ is abelian, we have $[\mathbb{H},\mathbb{H}] = \langle Z\rangle \subset \ker\lambda$. Therefore,

$$|f(x) - f(xZ^{\frac{1}{25}r^2})| \le |w + \lambda(x) - (w + \lambda(xZ^{\frac{1}{25}r^2})| + 2\epsilon r = 2\epsilon r,$$

but

$$d_{\mathbb{H}}(x, xZ^{\frac{1}{25}r^2}) \ge \frac{r}{5}.$$

It follows that for any $\epsilon > 0$, we can find $g = x, g' = xZ^{\frac{1}{25}r^2}$ such that

$$|f(g) - f(g')| \le 5\epsilon d_{\mathbb{H}}(g, g'),$$

i.e., f is not a bilipschitz map.

3. Rectifiability and uniform rectifiability

References

[Che99] J Cheeger. Differentiability of lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3):428–517, August 1999.

[EFW12] Alex Eskin, David Fisher, and Kevin Whyte. Coarse differentiation of quasi-isometries i: Spaces not quasi-isometric to cayley graphs. *Ann. Math.*, 176(1):221–260, July 2012.

[Jon90] Peter W. Jones. Rectifiable sets and the traveling salesman problem. *Invent. Math.*, 102(1):1–15, 1990.