RADEMACHER'S THEOREM IN METRIC MEASURE SPACES AND ALBERTI REPRESENTATIONS

DAVID BATE

1. Introduction

Rademacher's theorem states that any Lipschitz $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is differentiable Lebesgue almost everywhere. It is a fundamental result in geometric measure theory. For example, it shows that any rectifiable set possesses a weak tangent plane at almost every point. Moreover the statement itself is very interesting in its own right; the fact that a seemingly rather simple condition can impose such strong regularity is quite remarkable. Rademacher's theorem leads to many more questions that can be answered.

Recently there has been a wealth of interest in generalising results of classical analysis to the setting of metric spaces, possibly with an underlying measure. Naturally, Rademacher's theorem is a candidate for such a generalisation. This course will focus on a new proof of Cheeger's generalisation which replaces the domain with a doubling metric measure space (X, d, μ) that satisfies a Poincaré inequality [Che99]. This proof uses modern techniques developed in [Bat14] which consider a rich structure of Lipschitz curves in the metric space (known as "Alberti representations") which allow us to form a partial derivative of any Lipschitz function $f\colon X \to \mathbb{R}$. By considering many such families of curves, we are able to form a derivative of f and hence deduce Cheeger's theorem.

At the end of the course we will also discuss Kirchheim's generalisation of Rademacher's theorem for Lipschitz functions $f\colon \mathbb{R}^n \to X$, with X an arbitrary metric space [Kir94].

This constitutes the third proof of Cheeger's theorem. In addition to the original, Keith [Kei04] gave an independent proof (see also [KM11]). One common element to all three of these proofs is the use of weak tangent spaces (defined via Gromov-Hausdorff convergence), which is only natural considering the nature of the derivative. Essentially, after an analysis of a large collection of Lipschitz functions in the metric space X, one takes a weak limit and the tangential behaviour of these Lipschitz functions is significantly more rigid (i.e. linear in the classical theory). The doubling condition then gives an upper bound on the number of different (i.e. linearly independent) tangential behaviours the large set of Lipschitz functions can posses, which ultimately leads to the existence of a derivative.

Except for standard concepts in measure theory, this course will be self contained.

2. Basic concepts of metric measure spaces

Definition 2.1. A *metric measure space* (X, d, μ) consists of a complete metric space (X, d) equipped with a σ-finite Borel measure μ. That is, all Borel subsets of X are measurable and there exists a decomposition

$$X = \bigcup_{i \in \mathbb{N}} X_i$$

into Borel sets such that $\mu(X_i) < \infty$ for each $i \in \mathbb{N}$.

For many purposes, this definition includes spaces that are simply too large to work in. Often, it is natural to impose the following mild "finite dimensional" condition.

Given a metric space, unless otherwise stated, we will denote by $B(x, r)$ the closed ball centred at x with radius $r > 0$. Sometimes it may be necessary to consider open balls, but we will use the same notation and mention it explicitly. Given an (open or closed) ball B and $\lambda > 0$, we will denote by λB the ball with the same centre and with radius λ times bigger.

Definition 2.2. A metric measure space (X, d, μ) is *doubling* if there exists a $C \geq 1$ (the *doubling constant* of μ) such that, for each ball $B \subset X$,

$$0 < \mu(2B) \leq C\mu(B) < \infty.$$

In particular, by induction,

$$\mu(2^n B) \leq C^n \mu(B) \tag{2.1}$$

for each $n \in \mathbb{N}$.

A standard property of doubling measures is that they satisfy the Vitali covering theorem and hence the Lebesgue differentiation theorem, which we now establish.

The following lemma is often called the "$5r$" or "Vitali covering" lemma.

Lemma 2.3 (5r covering lemma)**.** *Let (X, d) be a metric space and $\mathcal{B}$ a collection of balls in X such that*

$$M := \sup\{\mathrm{rad}(B) : B \in \mathcal{B}\} < \infty.$$

Then there exists a disjoint sub-collection $\mathcal{C} \subset \mathcal{B}$ such that every ball B of $\mathcal{B}$ intersects a ball from $\mathcal{C}$ with radius at least half that of B. In particular

$$\bigcup_{C \in \mathcal{C}} 5C \supset \bigcup_{B \in \mathcal{B}} B. \tag{2.2}$$

Proof. For each $n \in \mathbb{N}$ let

$$\mathcal{B}_n = \{B \in \mathcal{B} : 2^{-n}M < \mathrm{rad}(B) \leq 2^{-n-1}M\}$$

and let $\mathcal{C}_1$ be a maximal disjoint sub-collection of $\mathcal{B}_1$. Inductively, if we have chosen $\mathcal{C}_1, \ldots \mathcal{C}_n$, we define

$$\mathcal{C}'_{n+1} = \{B \in \mathcal{B}_n : B \cap B' = \emptyset \ \forall B' \in \mathcal{C}_1 \cup \ldots \mathcal{C}_n\}$$

and let $\mathcal{C}_{n+1}$ be a maximal disjoint sub-collection of $\mathcal{C}'_{n+1}$. That is, if $B \in \mathcal{C}'_{n+1} \setminus \mathcal{C}_{n+1}$, then there exists $B' \in \mathcal{C}_{n+1}$ with $B \cap B' \neq \emptyset$.

We claim that

$$\mathcal{C} := \bigcup_{n \in \mathbb{N}} \mathcal{C}_n$$

is the required sub-collection. By construction, this sub-collection is disjoint, and so we only need to prove eq. (2.2). To see this, suppose that $B \in \mathcal{B}$ and $B \notin \mathcal{C}$. This implies that $B \notin \mathcal{C}_n$ with $2^{-n}M < \mathrm{rad}(B) \leq 2^{-n-1}M$. However, $B \in \mathcal{B}_n$ and so either $B \notin \mathcal{C}'_n$, or $B \in \mathcal{C}'_n$ but $B \notin \mathcal{C}_n$.

In the first case there exists $B' \in C_i$, for some $1 \leq i < n$ with $B \cap B' \neq \emptyset$. In the second case, since $\mathcal{C}_n$ is maximal, there must exist $B' \in \mathcal{C}_n$ with $B \cap B' \neq \emptyset$. In either case, $\mathrm{rad}(B') \geq \mathrm{rad}(B)/2$ by the definition of the $\mathcal{B}_i$. In particular, by the triangle inequality, $B \subset 5B'$. Indeed, if $x \in B \cap B'$, $B' = B(x_0, r)$ and $y \in B$, then $d(x, y) \leq 2\,\mathrm{rad}(B) \leq 4r$ and $d(x, x_0) < r$, so that $d(y, x_0) < 5r$ and hence $y \in 5B'$ (and similarly if B' is a closed ball). □

This easily leads to the Vitali covering theorem.

Theorem 2.4 (Vitali covering theorem)**.** *Let (X,d,μ) be a doubling metric measure space, $S \subset X$ and $\mathcal{B}$ a collection of closed balls such that, for each $x \in S$,*

$$\inf\{r > 0 : B(x,r) \in \mathcal{B}\} = 0.$$

Then there exists a countable disjoint sub-collection $\mathcal{C} \subset \mathcal{B}$ such that

$$\mu\left(S \setminus \bigcup_{B \in \mathcal{C}} B\right) = 0.$$

Proof. We will prove the case when S is bounded. The unbounded case is an exercise. By first discarding balls from $\mathcal{B}$ if necessary, we may suppose that the balls in $\mathcal{B}$ have radii at most 1.

Let $\mathcal{C}$ be a $5r$ sub-covering of $\mathcal{B}$ as obtained from Lemma 2.3. Since X is separable (see the exercises) and the elements of $\mathcal{C}$ are disjoint, $\mathcal{C}$ must be countable. Enumerate $\mathcal{C}$ as $B_1, B_2, \ldots$. Since S is bounded and each B_i has radius at most 1, each B_i is contained in a fixed ball B. In particular, since the B_i are disjoint, eq. (2.1) gives

$$\sum_{i \in \mathbb{N}} \mu(5B_i) \leq C \sum_{i \in \mathbb{N}} \mu(B_i) = C\mu\left(\bigcup_{i \in \mathbb{N}} B_i\right) \leq C\mu(B) < \infty$$

so that

$$\sum_{i>n} \mu(5B_i) \to 0 \tag{2.3}$$

as $n \to \infty$.

To complete the proof we will show that

$$S \setminus \bigcup_{i=1}^{n} B_i \subset \bigcup_{i>n} 5B_i \tag{2.4}$$

for each $n \in \mathbb{N}$. Once we have done this, eq. (2.3) completes the proof. To see eq. (2.4), fix $n \in \mathbb{N}$ and let $x \in S \setminus \cup_{i=1}^{n} B_i$. Because the balls B_i are closed, by the main assumption of the theorem, there exists a $\delta > 0$ such that $B(x,\delta)$ is disjoint from the B_i with $1 \leq i \leq n$ and such that $B(x,\delta) \in \mathcal{B}$. However, since $\mathcal{C}$ is a $5r$ sub-cover of $\mathcal{B}$, $B(x,\delta)$ intersects a ball $B_i \in \mathcal{C}$ with $\mathrm{rad}(B') \geq \mathrm{rad}(B)/2$. In particular, $i > n$ and $B \subset 5B'$, proving eq. (2.4), as required. □

Finally, we prove the Lebesgue differentiation theorem.

Theorem 2.5 (Lebesgue differentiation theorem)**.** *Let (X,d,μ) be a doubling metric measure space and $f\colon X \to \mathbb{R}$ a positive integrable function. Then*

$$\lim_{r \to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f \, d\mu = f(x) \tag{2.5}$$

for μ-a.e. $x \in X$.

Proof. First suppose that $S \subset X$ consists only of points $x \in X$ for which

$$\limsup_{r \to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f \, d\mu > t.$$

Then let $U \supset S$ with $\int_U f\,d\mu < \int_S f\,d\mu + \epsilon$ and let $\mathcal{B}$ be the collection of balls satisfying $B \subset U$ and $\int_B f\,d\mu > t\mu(B)$. By hypotheses, $\mathcal{B}$ satisfies the hypotheses of Theorem 2.4 and so has a countable disjoint sub-cover $\mathcal{B}$ covering almost all of S. Therefore

$$\int_S f\,d\mu + \epsilon \geq \int_U f\,d\mu \geq \sum_{B \in \mathcal{C}} \int_B f\,d\mu > t \sum_{B \in \mathcal{B}} \mu(B) = t\mu\left(\bigcup_{B \in \mathcal{B}} B\right) \geq t\mu(S).$$

Since $\epsilon > 0$ is arbitrary, we have

$$t\mu(S) \leq \int_S f d\mu. \tag{2.6}$$

A similar argument shows that, if $S \subset X$ consists only of points $x \in X$ for which

$$\liminf_{r\to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f d\mu < s,$$

then

$$\int_S f d\mu \leq s\mu(S). \tag{2.7}$$

We first show that the limit in eq. (2.5) exists almost everywhere. For $s < t \in \mathbb{R}$ let $S_{s,t}$ be the set of those $x \in X$ for which

$$\liminf_{r\to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f d\mu < s < t < \limsup_{r\to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f d\mu.$$

Then eqs. (2.6) and (2.7) imply

$$t\mu(S_{s,t}) \leq \int_{S_{s,t}} f d\mu \leq s\mu(S_{s,t}),$$

so that $\mu(S_{s,t}) = 0$. Therefore, the limit in eq. (2.5) exists exists almost everywhere. We will note the value of this limit by $g(x)$.

To see that $g = f$ almost everywhere, let $\epsilon > 0$, $S \subset X$ and for each $n \in \mathbb{Z}$ let

$$S_n = \{x \in S : (1+\epsilon)^n < g(x) \leq (1+\epsilon)^{n+1}\}.$$

Then by eq. (2.7)

$$\int_S g d\mu = \sum_{n\in\mathbb{N}} \int_{S_n} g d\mu \leq \sum_{n\in\mathbb{Z}} (1+\epsilon)^{n+1} \mu(S_n) \leq (1+\epsilon) \sum_{n\in\mathbb{Z}} \int_{S_n} f d\mu = (1+\epsilon) \int_S f d\mu.$$

Similarly, by eq. (2.7),

$$\int_S g d\mu \geq (1+\epsilon)^{-1} \int_S f d\mu.$$

Since $\epsilon > 0$ is arbitrary, we must have

$$\int_S f d\mu = \int_S g d\mu.$$

Since $S \subset X$ is arbitrary, this implies that $f = g$ almost everywhere. □

We also get the Lebesgue density theorem.

Corollary 2.6. *Let (X, d, μ) be a doubling metric measure space and $S \subset X$ measurable. Then for μ-a.e. $x \in S$,*

$$\lim_{r\to 0} \frac{\mu(B(x,r) \cap S)}{\mu(B(x,r))} = 1.$$

3. The classical Rademacher theorem and Poincaré inequality

Definition 3.1. A function $f : (X, d) \to (Y, \rho)$ is *Lipschitz* if there exists an $L \geq 0$ such that

$$\rho(f(x), f(y)) \leq Ld(x, y)$$

for each $x, y \in X$. The smallest such L is called the *Lipschitz constant* of f.

For us, a fundamental result regarding Lipschitz functions is the following.

Theorem 3.2 (Lebesgue)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be Lipschitz. Then f is differentiable (Lebesgue) almost everywhere.*

Proof. Observe that it suffices to prove the result for f is defined on $[0,1]$.

For each $x \in [0,1]$ let

$$V(x) = \sup \sum_{i=1}^{n} |f(x_{i+1}) - f(x_i)|,$$

where the supremum is taken over all $0 < x_1 < x_2 < \ldots < x_n < 1$. This V is called the *variation* of f, and since f is Lipschitz, its variation is bounded. Moreover, V is increasing as is $V - f$. Thus, it suffices to prove that any increasing function is differentiable almost everywhere.

To this end, let $g\colon [0,1] \to \mathbb{R}$ be increasing and $s \in \mathbb{R}$. If S consists only of points $x \in [0,1]$ for which

$$\liminf_{y \to x} \frac{g(y) - g(x)}{y - x} < s, \tag{3.1}$$

then by the Vitali covering theorem, we may cover almost all of S by disjoint intervals $[a_i, b_i]$ for which $g(b_i) - g(a_i) < s(b_i - a_i)$. By summing over such i and *using the fact that g is increasing*, this implies that $\mathcal{L}^1(g(S)) \leq s\mathcal{L}^1(S)$. Similarly, if S consists only of points $x \in [0,1]$ for which

$$\limsup_{y \to x} \frac{g(y) - g(x)}{y - x} > t, \tag{3.2}$$

then $\mathcal{L}^1(g(S)) \geq t\mathcal{L}^1(S)$. Therefore, for $s < t$, if S consists of all those $x \in [0,1]$ satisfying both eqs. (3.1) and (3.2), $t\mathcal{L}^1(S) \leq s\mathcal{L}^1(S)$, so that $\mathcal{L}^1(S) = 0$. By considering all rational $s < t$, this shows that g is differentiable almost everywhere. □

Following this we prove Rademacher's theorem. There are many possible ways of proceeding once we have proved Lebesgue's theorem.

Theorem 3.3 (Rademacher)**.** *Let $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz. Then f is differentiable almost everywhere.*

Proof. Note that it suffices to assume that $m = 1$. We will prove the statement for $n = 2$. The general proof uses exactly the same idea, but with more notation and a little more care at the end.

Let

$$G = \{x \in \mathbb{R}^n : \partial_1 f(x), \partial_2 f(x) \text{ exist}\}.$$

Note that, by Fubini's theorem and Theorem 3.2, G has full measure. For $x \in G$ define $Df(x) = (\partial_1 f(x), \partial_2 f(x))$ and let

$$B = \{x \in G : Df(x) \text{ is not the derivative of } f \text{ at } x\}.$$

We need to show that $\mathcal{L}(B) = 0$.

To do this, we need to "filter" B into subsets that we can handle. Observe that, for each $x \in B$ there exists an $\epsilon > 0$ such that

$$\limsup_{y \to x} \frac{|f(y) - f(x) - Df(x)(y - x)|}{\|y - x\|} > \epsilon. \tag{3.3}$$

For each $\epsilon > 0$ let B_ϵ be those $x \in B$ which satisfy eq. (3.3), so that $B = \cup_{\epsilon > 0} B_\epsilon$. However, note that we also have $B = \cup_{\mathbb{Q} \ni \epsilon > 0} B_\epsilon$. This small modification is important for showing that B has measure zero: because this is a countable union, we only have to show that each B_ϵ has measure zero.

Now fix an $\epsilon > 0$ and, for each $D \in \mathbb{Q}^2$ let

$$B_{\epsilon, D} = \{x \in B_\epsilon : \|Df(x) - D\| < \epsilon/10\}.$$

Note that, by the triangle inequality, for any $x \in B_{\epsilon,D}$,

$$\limsup_{y\to x} \frac{|f(y)-f(x)-D(x)(y-x)|}{\|y-x\|} > 9\epsilon/10. \tag{3.4}$$

Moreover, because $\mathbb{Q}^2$ is dense in $\mathbb{R}^2$, $B_\epsilon = \cup_{D\in\mathbb{Q}^2} B_{\epsilon,D}$, so that it suffices to prove each $B_{\epsilon,D}$ has measure zero.

We make one final filter. For $R>0$ let $B_{\epsilon,D,R}$ be those $x\in B_{\epsilon,D}$ such that

$$\frac{|f(x+he_i)-f(x)-D_i h|}{|h|} < \epsilon/10 \tag{3.5}$$

for each $i=1,2$ and each $0<|h|<R$. Because $\partial_i f(x)$ exists and $|\partial_i f(x) - D_i| < \epsilon/10$ for each $x\in B$, $B_{\epsilon,D} = \cup_{\mathbb{Q}\ni R} B_{\epsilon,D,R} > 0$, and so it suffices to prove that each one has measure zero.

To do this, fix $0<R,\epsilon<1/10$, $D\in\mathbb{Q}^2$ and $x\in B_{\epsilon,D}$. We will show that there exists a $\lambda>0$ and a sequence $x_j\to x$ such that $B(x_j,\lambda\|x_j-x\|)$ is disjoint from $B_{\epsilon,D,R}$ for each $j\in\mathbb{N}$. The Lebesgue density theorem then implies that $B_{\epsilon,D,R}$ has measure zero.

To this end, write $g(y)=f(y)-D\cdot y$ and pick $L\geq 1$ greater than the Lipschitz constant of g. For each $j\in\mathbb{N}$ let y be a point as in eq. (3.4) with $\|y-x\|<1/j$. We set $x_j=y$ and $\lambda=\epsilon/10L$ and let $z\in B(x_j,\lambda\|x-x_j\|)$. Let z' be the projection of z onto the line $x+\mathbb{R}e_1$, so that $\|x-z'\|\leq\|x-x_j\|<R$ and $\|z-z'\|\leq\|z-x\|$. Observe that by eq. (3.4)

$$\frac{|g(z)-g(x)|}{\|z-x\|} > \frac{8\epsilon}{10}\frac{1}{1+\epsilon/10L} \geq \frac{80\epsilon}{101} \tag{3.6}$$

and by eq. (3.5)

$$\frac{|g(z')-g(x)|}{\|z-x\|} \leq \frac{|g(z')-g(x)|}{\|z'-x\|} < \frac{\epsilon}{10}.$$

Thus, by the triangle inequality,

$$|g(z)-g(z')| > \frac{79\epsilon\|z-x\|}{101}.$$

Since $z'\in z+\mathbb{R}e_2$, and $\|z-z'\|<R$, $z\notin B_{\epsilon,D,R}$, as required. □

Together with the Lebesgue and Rademacher differentiation theorems, we have the fundamental theorem of calculus.

Theorem 3.4 (Fundamental theorem of calculus). *Let $f\colon\mathbb{R}\to\mathbb{R}$ be Lipschitz. Then for any $x\leq y\in\mathbb{R}$*

$$f(y)-f(x) = \int_x^y Df\,\mathrm{d}\mathcal{L}.$$

Proof. First suppose that $Df(t)=0$ for almost every $t\in\mathbb{R}$ and let $\epsilon>0$. By the Vitali covering theorem, we may cover $[x,y]$ by a null set N and a countable number of disjoint intervals $[a_i,b_i]$ such that $|f(t)-f(a_i)|\leq\epsilon|t-a_i|\leq\epsilon|b_i-a_i|$ for each $t\in[a_i,b_i]$ and each $i\in\mathbb{N}$. Consequently,

$$\mathcal{L}(f([x,y])) \leq \mathcal{L}(f(N)) + \sum_{i\in\mathbb{N}} \epsilon(b_i-a_i) \leq \mathcal{L}(f(N)) + \epsilon(y-x).$$

Since $\epsilon>0$ is arbitrary, this gives $\mathcal{L}(f([x,y]))\leq f(N)$. Since f is Lipschitz and $\mathcal{L}(N)=0$, we have $\mathcal{L}(f([x,y]))=0$. Since $f([x,y])$ is an interval, it must be a single point, and so f must be constant.

Now suppose that f is as in the statement of the theorem and define

$$g(t) = \int_x^t Df\,\mathrm{d}\mathcal{L}$$

for each $t \in [x, y]$. Then g is a Lipschitz function (because $|Df|$ is bounded by $\operatorname{Lip} f$) and, by Theorem 2.5, $Dg = Df$ almost everywhere in $[x, y]$. Thus $f - g$ has derivative zero almost everywhere in $[x, y]$ and so by the first part of the proof is constant. That is

$$f(y) - f(x) = g(y) - g(x) = \int_x^y Df \, \mathrm{d}\mathcal{L}.$$

□

Remark 3.5. In the previous proof, we may replace the hypothesis that f is Lipschitz by f being totally bounded (so that its derivative exists almost everywhere) and the fact that f maps sets of measure zero to sets of measure zero. Such a function is said to be *absolutely continuous.*

We now state a version of the *Poincaré inequality* for Lipschitz functions defined on Euclidean space. This version is much weaker than the "standard" Poincaré inequality but, as we will see later, has a suitable interpretation in the metric measure space setting.

We introduce some standard notation. For an integrable $f\colon (X, d, \mu) \to \mathbb{R}$ and $B \subset X$ of positive and finite measure, let

$$\fint_B f \, \mathrm{d}\mu = \frac{1}{\mu(B)} \int_B f \, \mathrm{d}\mu$$

and $f_B = \fint_B f \, \mathrm{d}\mu$.

Theorem 3.6. *For any Lipschitz $f\colon \mathbb{R}^n \to \mathbb{R}$ and any ball $B \subset \mathbb{R}^n$,*

$$\fint_B |f - f_B| \, \mathrm{d}\mathcal{L}^n \leq C_n \operatorname{rad}(B) \fint_B \|Df\| \, \mathrm{d}\mathcal{L}^n$$

Proof. For $x \in B$, we calculate $f(x) - f_B$ using polar coordinates: for $\omega \in \mathbb{S}^{n-1}$ let $R(\omega)$ be the largest $R \geq 0$ such that $x + R\omega \in B$. Then

$$|f(x) - f_B| \leq \frac{1}{\mathcal{L}^n(B)} \int_{\mathbb{S}^{n-1}} \int_0^{R(\omega)} r^{n-1} |f(x) - f(x + r\omega)| \, \mathrm{d}r \, \mathrm{d}\omega.$$

Applying Theorem 3.4 to the integrand gives

$$|f(x) - f_B| \leq \frac{1}{\mathcal{L}^n(B)} \int_{\mathbb{S}^{n-1}} \int_0^{R(\omega)} r^{n-1} \int_0^r \|Df(x + t\omega)\| \, \mathrm{d}t \, \mathrm{d}r \, \mathrm{d}\omega$$

and so, since $r \leq R(\omega) \leq 2 \operatorname{rad}(B)$,

$$|f(x) - f_B| \leq \frac{C_n R(\omega)^n}{\mathcal{L}^n(B)} \int_{\mathbb{S}^{n-1}} \int_0^{R(\omega)} \|Df(x + t\omega)\| \, \mathrm{d}t \, \mathrm{d}\omega.$$

By making the change of variables back to B, this integral becomes

$$|f(x) - f_B| \leq C_n \int_B \frac{\|Df(y)\|}{\|x - y\|^{n-1}} \, \mathrm{d}y.$$

Finally, by integrating over $x \in B$ and applying Fubini we get

$$\int_B |f - f_B| \, \mathrm{d}\mathcal{L}^n \leq C_n \int_B \|Df(y)\| \int_B \frac{1}{\|x - y\|^{n-1}} \, \mathrm{d}x \, \mathrm{d}y \leq C_n \operatorname{rad}(B) \int_B \|Df\| \, \mathrm{d}\mathcal{L}^n.$$

Dividing by the measure of B completes the proof. □

4. The Poincaré inequality in metric measure spaces and Cheeger's theorem

The proof of Poincaré inequality relies on the behaviour of Lipschitz functions along lines in Euclidean space. Namely that they satisfy the fundamental theorem of calculus. This motivates our definition of a Poincaré inequality in the setting of metric measure spaces, but we must replace a "line" by a rectifiable curve.

First we introduce some standard notation of how we integrate a function on a rectifiable curve.

Definition 4.1. A *rectifiable curve* in a metric space X is a Lipschitz function $\gamma\colon [a,b] \to X$.

The *length* of a rectifiable curve is

$$\operatorname{len}(\gamma) = \sup \left\{ \sum_{i=1}^{n} d(\gamma(t_i), \gamma(t_{i-1})) : a = t_0 < t_1 < \ldots < t_n = b \right\}.$$

If γ is injective, then this agrees with $\mathcal{H}^1(\gamma([a,b]))$.

Whilst formally, we have defined a rectifiable curve to be a Lipschitz function, we will usually think of one as a subset of a metric space. This requires us to canonically define the integral of a function on a rectifiable curve, which we now do.

Given a rectifiable curve $\gamma\colon [a,b] \to X$, let $l_\gamma\colon [a,b] \to \mathbb{R}_+$ be defined by $l_\gamma(t) = \operatorname{len}(\gamma|_{[a,t]})$. There exists a 1-Lipschitz $\gamma^*\colon [0, \operatorname{len}(\gamma)] \to X$ such that $\gamma = \gamma^* \circ l_\gamma$. Such a γ^* is the *arc-length parametrisation* of γ, and if $\gamma = \gamma^*$ then we say that γ is *parametrised by arc-length.* Moreover, if γ is injective, then so is γ^* and, for any Borel $S \subset [0, \operatorname{len}(\gamma)]$, $\mathcal{H}^1(S) = \mathcal{H}^1(\gamma^*(S))$ (see the exercises).

Definition 4.2. Given a Borel function $\rho\colon X \to \mathbb{R}$ we define the *line integral* of ρ over a rectifiable curve γ to be

$$\int_\gamma \rho \,\mathrm{d}s = \int_0^{\operatorname{len}(\gamma)} \rho \circ \gamma^*(t) \,\mathrm{d}t.$$

If γ is injective, this agrees with

$$\int_{\gamma([a,b])} \rho \,\mathrm{d}\mathcal{H}^1.$$

We can now define a replacement for the fundamental theorem of calculus in a metric space.

Definition 4.3. Let X be a metric space and $f\colon X \to \mathbb{R}$ a Lipschitz function. A Borel function $\rho\colon X \to [0,\infty)$ is an *upper gradient* of f if, for every $x, y \in X$ and every rectifiable curve γ joining x to y,

$$|f(x) - f(y)| \leq \int_\gamma \rho \,\mathrm{d}s.$$

Observe that, if $X = \mathbb{R}^n$, then by Lebesgue's theorem, $\rho = \|Df\|$ is an upper gradient of any Lipschitz $f\colon \mathbb{R}^n \to \mathbb{R}$. More generally, for any metric space X and any Lipschitz $f\colon X \to \mathbb{R}$, $\operatorname{Lip} f$ is an upper gradient of f. See also the exercises.

Observe also that this definition is only of value when X has a rich structure of rectifiable curves. For example, if X contains no non-trivial rectifiable curves, as is the case when $X = \{0,1\}$, then $\rho = 0$ is an upper-gradient of *any* Lipschitz function.

The notion of a Poincaré inequality in a metric measure space says that any upper gradient of a Lipschitz function must control the behaviour in a very precise way. The Poincaré inequality in this setting was first introduced by Heinonen and Koskela [HK98]. We give a slightly different formulation that is equivalent whenever the measure is doubling.

Definition 4.4. For $p \geq 1$, a metric measure space (X, d, μ) satisfies a *p-Poincaré inequality* if there exists a $C \geq 1$ such that, for any ball $B \subset X$, any Lipschitz $f \colon X \to \mathbb{R}$ and any upper-gradient ρ of f,

$$\fint_B |f - f_B| \leq C \operatorname{rad}(B) \left(\fint_B \rho^p \, \mathrm{d}\mu \right)^{\frac{1}{p}}.$$

We say that (X, d, μ) is a *p-PI space* if it is doubling and satisfies a p-Poincaré inequality.

By Holder's inequality, the Poincaré inequality becomes weaker as p increases. Also, Theorem 3.6 shows that Euclidean space is a 1-PI space.

Laakso gave several very interesting, non-trivial, and highly non-Euclidean examples of 1-PI spaces [Laa00; Laa02].

We now move on to considering the derivatives of real valued Lipschitz functions defined on a metric space. The first way to generalise a derivative to this setting was introduced by Cheeger [Che99] and was later refined by Keith [Kei04]. It simply replaces to coordinate functions in Euclidean space by an arbitrary vector valued Lipschitz function.

Definition 4.5. Let $\phi \colon X \to \mathbb{R}^n$ be a fixed Lipschitz function and $x_0 \in X$.

A function $f \colon X \to \mathbb{R}$ is *differentiable with respect to* ϕ at x_0 if there exists a unique $Df(x_0) \in \mathbb{R}^n$ such that

$$\limsup_{x \to x_0} \frac{|f(x) - f(x_0) - Df(x_0) \cdot (\phi(x) - \phi(x_0))|}{d(x, x_0)} = 0.$$

Note that we must *assume* that the derivative is unique as a part of the definition.

We can now say what it means for a metric measure space to satisfy a generalisation of Rademacher's theorem.

Definition 4.6. A metric measure space (X, d, μ) is a *Lipschitz differentiability space* if there exists a countable Borel decomposition

$$X = \bigcup_{i \in \mathbb{N}} U_i$$

and countably many Lipschitz functions $\phi_i \colon X \to \mathbb{R}^{n_i}$ such that the following is true: For every Lipschitz $f \colon X \to \mathbb{R}$ and every $i \in \mathbb{N}$, f is differentiable at μ-a.e. point in U_i with respect to ϕ_i.

We call the pair (U_i, ϕ_i) a *chart*.

This is all the required concepts to state Cheeger's theorem [Che99].

Theorem 4.7 (Cheeger)**.** *Any PI space is a Lipschitz differentiability space.*

5. Alberti representations as an alternative generalisation of Rademacher's theorem

We now introduce an alternative way to generalise Rademacher's theorem to metric measure spaces. However, it turns out that this in an equivalent formulation to Cheeger's. This is nice because it allows us to have two very different descriptions of the same phenomenon.

This alternative formulation is based on the idea of forming a *partial derivative* of any Lipschitz function along a rectifiable curve.

Definition 5.1. Given a metric space X, define the set of *rectifiable curves* in X to be

$$\Gamma(X) = \{\gamma\colon [0,1] \to X : \gamma \text{ 1-Lipschitz}\}$$

equipped with the supremum metric

$$d(\gamma,\gamma') = \sup_{t\in[0,1]} d(\gamma(t),\gamma'(t)).$$

Definition 5.2. An *Alberti representation* of a metric measure space (X,d,μ) consists of a probability measure $\mathbb{P}$ on $\Gamma(X)$ and for each $\gamma \in \Gamma(X)$ a measure $\mu_\gamma \ll \mathcal{H}^1|_{\gamma([0,1])}$ such that

$$\mu(B) = \int_{\Gamma(X)} \mu_\gamma(B)\, \mathrm{d}\mathbb{P}(\gamma)$$

for every Borel $B \subset X$.

The motivation for considering Alberti representations is the following. Suppose that $f\colon X \to \mathbb{R}$ is Lipschitz and that $\gamma \in \Gamma(X)$. Then the composition $f\circ\gamma\colon [0,1] \to X$ is Lipschitz and so is differentiable almost everywhere. That is, for $\mathcal{H}^1$ almost every $x \in \gamma([0,1])$, there exists a *partial derivative* of f at x given by $(f\circ\gamma)'(\gamma^{-1}(x))$. It is then not hard to see that, if (X,d,μ) has an Alberti representation, then such a partial derivative exists for μ almost every $x \in X$.

You should compare this to our proof of Rademacher's theorem; the first step is to use the Fubini and Lebesgue theorems to find a partial derivative of any Lipschitz function. Here, we are replacing the use of Fubini's theorem with the *definition* of an Alberti representation. Of course, Fubini's theorem tells us that Lebesgue measure has an Alberti representation.

Of course, in the proof of Rademacher's theorem, it is important that we can find all n partial derivatives of a Lipschitz $f\colon \mathbb{R}^n \to \mathbb{R}$. Our next step is to interpret this for Alberti representations.

In Euclidean space, we can distinguish different families of curves by considering their tangents: some curves travel in a direction close to the e_1, direction, some others close to the e_2 direction and so on. In metric spaces there is no natural way to assign a tangent. However, suppose we fix a Lipschitz $\phi\colon X \to \mathbb{R}^n$. Then, as above, for any $\gamma \in \Gamma(X)$, $(\phi\circ\gamma)'(t)$ exists for $\mathcal{H}^1$ almost every $t \in [0,1]$. We can use this tangent to give a direction to γ (with respect to ϕ).

Definition 5.3. Let (X,d) be a metric space, $\phi\colon X \to \mathbb{R}^n$ Lipschitz and $C \subset \mathbb{R}^n$ a cone. We say that $\gamma \in \Gamma(X)$ is a C-curve (with respect to ϕ) if $(\phi\circ\gamma)'(t) \in C \setminus \{0\}$ for $\mathcal{L}^1$ almost every $t \in [0,1]$.

Similarly, an Alberti representation $(\mathbb{P},\{\mu_\gamma\})$ of a metric measure space is a C-Alberti representation (with respect to ϕ) if $\mathbb{P}$ almost every curve is a C curve (with respect to ϕ).

Finally, we say that a collection of Alberti representations

$$(\mathbb{P}_1,\{\mu_\gamma^1\}),\dots,(\mathbb{P}_n,\{\mu_\gamma^n\})$$

is *independent* if there exists a Lipschitz $\phi\colon X \to \mathbb{R}^n$ and independent cones $C_1,\dots,C_n \subset \mathbb{R}^n$ such that $(\mathbb{P}_i,\{\mu_\gamma^i\})$ is a C_i-Alberti representation with respect to ϕ for each $1 \le i \le n$.

We say that cones $C_1,\dots,C_n \subset \mathbb{R}^n$ are *independent* if any choice $v_i \in C_i \setminus \{0\}$ forms a linearly independent set.

Often the underlying function ϕ will be clear from the context and we will not explicitly mention it.

We see from Fubini's theorem that Lebesgue measure has a collection of n independent Alberti representations (and no more), and that this is precisely the property we required in our proof of Rademacher's theorem.

The first step of our proof of Cheeger's theorem will be to prove that a PI space has a large collection of independent Alberti representations. Before we prove this, we will develop a general technique for deciding when a measure has many independent Alberti representations.

Definition 5.4. Let X be a metric space, $\phi\colon X \to \mathbb{R}^n$ Lipschitz and $C \subset \mathbb{R}^n$ a cone. A Borel $S \subset X$ is *C-null (with respect to ϕ)* if $\mathcal{H}^1(\gamma \cap S) = 0$ for each C-curve γ.

Observe that, if a metric measure space is C-null then it cannot support a C-Alberti representation. This is in fact a characterisation of those measures with C-Alberti representations. We will use this fact without proof.

Lemma 5.5. *Let (X, d, μ) be a metric measure space, $\phi\colon X \to \mathbb{R}^n$ Lipschitz and $C \subset \mathbb{R}^n$ a cone. There exists a decomposition $X = A \cup S$ such that $\mu \llcorner A$ has a C-Alberti representation and S is C-null. This decomposition is unique up to μ null sets.*

By applying this lemma to a collection of independent cones $C_1, C_2, \ldots, C_n$, we obtain a decomposition $X = A \cup S_1 \cup S_2 \cup \ldots \cup S_n$ where each S_i is C_i null. and $\mu|_A$ has n independent Alberti representations. This is a good start, but, in a PI space, we will only be able to prove that the S_i have measure zero when θ is sufficiently wide. Of course, we cannot simply apply the lemma to very wide cones, because they will not be independent. Therefore, we must do something more involved. The first step is to show that we can combine the cones of C-null sets. However, for this we require a very rich collection of curves provided by an ambient Banach space.

Lemma 5.6. *Let X be a metric space and $f\colon X \to \mathbb{R}^n$ Lipschitz. There exists a Banach space B, a biLipschitz $\iota\colon X \to B$ and a linear $T\colon X \to \mathbb{R}^n$ such that $\phi(x) = T(\iota(x))$ for every $x \in X$. If X is separable, then we may find such a B that is also separable.*

Proof. Let $\iota^*\colon X \to B^*$ be an isometric embedding of X into a Banach space. If X is separable we may suppose that B^* is separable too. The map

$$\iota(x) = (\phi(x), \iota^*(x)) \in \mathbb{R}^n \times B^* := B$$

is the desired embedding. □

Since a PI space is preserved under biLipschitz equivalences, from now on we may suppose that our metric measure space is a subset of some Banach space and that the map ϕ is linear.

We define our cones as follows. For $w \in \mathbb{S}$ and $0 < \theta < 1$ let $C(w, \theta)$ be the cone of $v \in \mathbb{R}^n$ such that $w \cdot v \geq \theta \|v||$. Note that this cone becomes very wide as $\theta \to 0$.

Proposition 5.7. *Let B be a Banach space, $\phi\colon B \to \mathbb{R}^n$ linear and $C \subset \mathbb{R}^n$ a cone. Suppose that $S \subset B$ is C-null with respect to ϕ. Then for any $\gamma \in \Gamma(B)$, $(\phi \circ \gamma)'(t) \notin \operatorname{interior}(C)$ for almost every $t \in \gamma^{-1}(S)$.*

Proof. Let $C = C(w, \theta)$ for some $w \in \mathbb{S}^{n-1}$ and $0 < \theta < 1$. Suppose that the conclusion is false and that for some $\gamma \in \Gamma(B)$,

$$B = \{t \in \gamma^{-1}(S) : (\phi \circ \gamma)'(t) \in \operatorname{interior}(C)\}$$

has positive measure. Since

$$\text{interior}(C) = \bigcup_{j\in\mathbb{N}} C(w,\theta+1/j)\setminus B(0,1/j),$$

there exists some $j\in\mathbb{N}$ such that

$$B_1 = \{t\in\gamma^{-1}(S) : (\phi\circ\gamma)'(t)\in C(w,\theta+1/j)\setminus U(0,1/j)\}$$

satisfies $\mathcal{H}^1(\gamma(B_1))>0$. Further, since $C(w,\theta+1/j)\setminus U(0,1/j)$ is a closed subset of interior(C), there exists an $R>0$ such that the set

$$B_2 = \{t\in B_1 : \frac{\phi(\gamma(t+r))-\phi(\gamma(t))}{r}\in C\ \forall\ 0<|r|<R\}$$

also satisfies $\mathcal{H}^1(\gamma(B_2))>0$.

By dividing B_2 into finitely many subsets of diameter R, at least one of these subsets B_3 must satisfy $\mathcal{H}^1(\gamma(B_3))>0$. Note that in particular, for any $s,t\in B_3$,

$$\frac{\phi(\gamma(t))-\phi(\gamma(s))}{t-s}\in C.$$

Finally, there exists a compact $K\subset B_3$ such that $\mathcal{H}^1(\gamma(K))>0$.

Let $\min K=a$ and $\max K=b$. Note that $[a,b]\setminus K$ is open and so is a countable union of disjoint intervals $\cup_i(c_i,d_i)$. We extend γ linearly on the complement of K: for any $i\in\mathbb{N}$ and $t\in(c_i,d_i)$ let $0<\lambda<1$ such that $t=\lambda c_i+(1-\lambda)d_i$ and set $\gamma(t)=\lambda\gamma(c_i)+(1-\lambda)\gamma(d_i)$. Since $\phi(\gamma(d_i))-\phi(\gamma(c_i))\in C$ and ϕ is linear, $(\phi\circ\gamma)'(t)\in C$ for all $t\in(c_i,d_i)$. □

From the previous Proposition we can deduce the following method to "refine" the directions of an Alberti representation. Of course, this makes most sense when the C_i are very thin.

Lemma 5.8. *Let B be a Banach space, $\phi\colon B\to\mathbb{R}^n$ linear and $C\subset\mathbb{R}^n$ a cone. Suppose that a measure μ has a C-Alberti representation with respect to ϕ. Then, for any collection of cones $C_1,\ldots,C_m\subset\mathbb{R}^n$ such that*

$$C\setminus\{0\}\subset\bigcup_{i=1}^m \text{interior}(C_i),$$

there exists a Borel decomposition $B=B_1\cup\ldots\cup B_m$ such that each $\mu|_{B_i}$ has a C_i-Alberti representation.

Proof. By Lemma 5.5 there exists a decomposition $B=B_1\cup S_1$ such that B_1 has the required form and S_1 is C_1-null. By applying Lemma 5.5 again (to $\mu|_{S_1}$) we obtain a decomposition $B=B_1\cup B_2\cup S_2$ where B_2 has the required form and S_2 is both C_1-null and C_2-null. By repeating, we obtain a decomposition $B=B_1\cup\ldots\cup B_m\cup S$ such that each B_i has the required form and S is C_i-null for each $1\leq i\leq m$. By Lemma 5.8, we see that S is C-null. Since μ has a C-Alberti representation, we must have $\mu(S)=0$. Therefore, $B_1'=B_1\cup S$ also has the required properties of B_1, and $B_1'\cup B_2\cup\ldots\cup B_m$ is the required decomposition. □

Finally we can prove the decomposition that we are after; we will later show that the 'singular sets' appearing have measure zero whenever the measure supports a Poincaré inequality.

Proposition 5.9. *Let (X,d,μ) be a metric measure space and $\phi\colon X\to\mathbb{R}^n$ Lipschitz. For any $0<\theta<1$ there exists a decomposition*

$$X=\bigcup_{i=1}^m A_i\cup\bigcup_{i=1}^m S_i \tag{5.1}$$

and $w_1, \ldots, w_m \in \mathbb{S}^{n-1}$ such that each $\mu|_{A_i}$ has n independent Alberti representations and each S_i is $C(w_i,\theta)$-null.

Proof. We will iteratively find the Alberti representations one-by-one.

First we apply Lemma 5.5 to an arbitrary cone $C(w,\theta)$ to obtain a decomposition $X = A \cup S$ where $\mu|_A$ has a $C(w,\theta)$-Alberti representation and S is $C(w,\theta)$-null.

Now suppose that, for $1 \leq d < n$ we have a decomposition as in eq. (5.1) where each $\mu|_{A_i}$ has d independent Alberti representations. By applying Lemma 5.8 (and increasing m) we may suppose that these Alberti representations are $C(w_j^i,\alpha)$-Alberti representations for $1 \leq j \leq d$ and $0 < \alpha < 1$ as small as we wish. Fix a $1 \leq i \leq m$ and pick w_{d+1}^i to be orthogonal to $w_1^i, \ldots, w_d^i$, which is possible because by assumption $1 \leq d < n$. We pick α so small that $C(w_{d+1}^i,\theta)$ and each $C(w_j^i,\alpha)$ are disjoint for $1 \leq j \leq d$ (this is independent of the choice of the w_j^i and only depends on n). Then the cones $C(w_1^i,\alpha), \ldots, C(w_d^i,\alpha)$ and $C(w_{d+1}^i,\theta)$ are independent.

Finally, we apply Lemma 5.5 to obtain a decomposition $A_i = A_i' \cup S'$ where $\mu|_{A_i'}$ has $d+1$ independent Alberti representations and S' is $C(w_{d+1}^i,\theta)$-null.

Repeating this for each $1 \leq d < n$ completes the proof. □

6. Alberti representations and the Poincaré inequality

The next step is to demonstrate that PI spaces possess many independent Alberti representations. Of course, we cannot use Proposition 5.9 to deduce that there arbitrarily large collections of independent Alberti representations; $\mathbb{R}^n$ can have at most n independent Alberti representations. Thus we must find some reasonable conditions on ϕ for which that proposition is useful (see also the exercises for when it is vacuous).

To understand what such a condition on ϕ should look like, it is important to think of how we will eventually deduce the conclusion of Cheeger's theorem. We start with the following elementary lemma of measure theory.

Lemma 6.1. *Let μ be a σ-finite measure on a set X and let $\mathcal{S}$ be a collection of μ measurable sets that is closed under taking countable unions. Then there exists a decomposition*

$$X = S \cup A$$

such that $S \in \mathcal{S}$ and, for every $B \subset A$, if $B \in \mathcal{S}$ then $\mu(B) = 0$. This decomposition is unique up to sets of μ measure zero.

Let $\mathcal{T}$ be a collection of μ measurable subsets of X such that, for any measurable $Y \subset X$ with $\mu(Y) > 0$ there exists a $T \in \mathcal{T}$ with $T \subset Y$ and $\mu(T) > 0$. Then there exists a decomposition

$$X = N \cup \bigcup_{i\in\mathbb{N}} T_i$$

where $\mu(N) = 0$ and $T_i \in \mathcal{T}$ for each $i \in \mathbb{N}$.

Proof. It suffices to prove the lemma for μ finite.

For the first part of the lemma, consider

$$m = \sup\{\mu(S) : S \in \mathcal{S}\} < \infty.$$

First observe that there exists $S \in \mathcal{S}$ with $\mu(S) = m$. Indeed, for each $i \in \mathbb{N}$ let $S_i \in \mathcal{S}$ with $\mu(S_i) \geq m - 1/i$ and set $S = \cup_i S_i$. Then $S \in \mathcal{S}$ (since $\mathcal{S}$ is closed under countable unions) and

$$m \geq \mu(S) \geq \mu(S_i) \geq m - 1/i$$

for each $i \in \mathbb{N}$, so that $\mu(S) = m$. Set $A = X \setminus S$.

Now suppose that $B \subset A$, $B \in \mathcal{S}$ and $\mu(B) > 0$. Then $B \cup S \in \mathcal{S}$ and $\mu(B \cup S) > \mu(S) > m$, a contradiction. Thus, for any $B \in \mathcal{S}$ with $B \subset A$, we must have $\mu(B) = 0$, as required.

For the second part, given such a set $\mathcal{T}$, define

$$\mathcal{S} = \{\bigcup_{i \in \mathbb{N}} T_i : T_i \in \mathcal{T}\}$$

which is closed under taking countable unions. Thus, by the first part of the lemma, there exists a decomposition

$$X = A \cup \bigcup_{i \in \mathbb{N}} T_i$$

where $T_i \in \mathcal{T}$ for each $i \in \mathbb{N}$ and, if $T \in \mathcal{T}$, $T \subset A$, then $\mu(T) = 0$. By assumption, this implies that $\mu(A) = 0$, as required. □

To construct the chart maps in the definition of a Lipschitz differentiability space we will use the following.

Proposition 6.2. *Let (X, d, μ) be a σ-finite metric measure space and suppose that there exists an $N \in \mathbb{N}$ for which the following is true. When ever $\phi \colon X \to \mathbb{R}^n$ is Lipschitz with the property that*

$$\operatorname{Lip}(D \cdot \phi, x) > 0 \quad \textit{for every } D \in \mathbb{R}^n \setminus \{0\}, \tag{6.1}$$

for a set of positive measure $x \in X$, then $n \leq N$. Then (X, d, μ) is a Lipschitz differentiability space.

Proof. Let $U \subset X$ be a Borel set of positive measure. Either every Lipschitz $\phi \colon X \to \mathbb{R}$ satisfies $\operatorname{Lip}(\phi, x) = 0$ for μ almost every $x \in U$ or there exists a $U_1 \subset U$ of positive measure and a Lipschitz $\phi_1 \colon X \to \mathbb{R}$ with $\operatorname{Lip}(\phi_1, x) > 0$ for every $x \in U_1$. Given the first option we stop; in this case X is a Lipschitz differentiability space with respect to the zero function. Otherwise we proceed iteratively.

Suppose that we have, for some $n \in \mathbb{N}$, a Lipschitz $\phi_n \colon X \to \mathbb{R}^n$ and a $U_n \subset U$ with $\mu(U_n) > 0$ such that $\operatorname{Lip}(D \cdot \phi, x) > 0$ for every $D \in \mathbb{S}^{n-1}$ for almost every $x \in U_n$. Then either

- (U_n, ϕ) is a chart with respect to which every Lipschitz $f \colon X \to \mathbb{R}$ is differentiable almost everywhere;
- or there exists a $U_{n+1} \subset U_n$ of positive measure and a Lipschitz $\phi^{n+1} \colon X \to \mathbb{R}$ such that $\phi' = (\phi, \phi^{n+1})$ satisfies $\operatorname{Lip}(D \cdot \phi', x) > 0$ for every $D \in \mathbb{R}^{n+1} \setminus \{0\}$.

Note that, for the first case, the uniqueness of the derivative is equivalent to the hypotheses on ϕ (see the exercises).

By hypothesis, the second option cannot hold for sufficiently large n, and so at some point we find a subset of U of positive measure which forms a chart. Since $U \subset X$ is an arbitrary set of positive measure, Lemma 6.1 completes the proof. □

The way forward is now more clear; given a Lipschitz $\phi \colon X \to \mathbb{R}^n$ satisfying eq. (6.1) for all x in a subset U of positive measure, Proposition 5.9 gives a decomposition of U into some subsets which have n independent Alberti representations, and other sets which are $C(w, \theta)$-null for θ as small as we like. To show that these $C(w, \theta)$-null subsets of U have measure zero, we begin the following Lemma.

Lemma 6.3. *Let (X, d, μ) be a doubling metric measure space and $f \colon X \to \mathbb{R}$ 1-Lipschitz. There exists a constant $\eta = \eta(\operatorname{doub}(\mu)) \geq 1$ such that, for any ball $B \subset X$,*

$$⨍_{2B} |f - f_{2B}| \, \mathrm{d}\mu \geq C(\mu) \operatorname{rad} B \sup \left\{ \frac{|f(x) - f(y)|}{\operatorname{rad} B} : x, y \in B \right\}^{\eta}.$$

In particular, if (X,d,μ) is a (weak) PI space and ρ is an upper gradient of f,

$$\operatorname{Lip}(f,x) \leq C(\mathrm{PI}(\mu))\rho(x)^{1/\eta}$$

for μ almost every $x \in X$.

Remark 6.4. The 'in particular' statement of the previous lemma is the only part in the proof that we require the Poincaré inequality.

Proof. Fix a ball $B \subset X$ and let

$$M = \sup\left\{\frac{|f(x)-f(y)|}{\operatorname{rad} B} : x,y \in B\right\} \leq 1$$

If $M=0$ there is nothing to prove and so we may suppose $M>0$. Since B is compact, there exist $x_1, x_2 \in B$ such that

$$|f(x_1)-f(y_2)| = M \operatorname{rad} B.$$

For $i=1,2$ let $B_i = B(x_i, M\operatorname{rad} B/10) \subset 2B$. Then, since f is 1-Lipschitz,

$$|f(y_1)-f(y_2)| \geq 8M\operatorname{rad} B/10$$

for each $y_i \in B_i$ and each $i=1,2$. In particular, there exists $i \in 1,2$ such that

$$|f(y_i)-f_{2B}| \geq 4M\operatorname{rad} B/10$$

for each $y_i \in B_i$. Therefore,

$$\fint_{2B} |f-f_{2B}| \,\mathrm{d}\mu \geq \frac{1}{\mu(2B)}\int_{B_i} |f-f_{2B}|\,\mathrm{d}\mu \geq \frac{2M\operatorname{rad} B\mu(B_i)}{5\mu(2B)} \tag{6.2}$$

Now, $\operatorname{rad} B_i = M\operatorname{rad} B/10$ and so, since μ is doubling,

$$C(\mu)^n \mu(B_i) \geq \mu(2B)$$

for $n = -\log_2 M/10 + 2$. That is, if $\delta = \log_2 C > 0$ depending on $C(\mu)$ such that

$$\frac{\mu(B_i)}{\mu(2B)} \geq CM^{\delta}$$

Combining this with eq. (6.2) completes the first part of the proof.

The second part of the lemma simply follows from the Lebesgue differentiation theorem. □

Proposition 6.5. *Let (X,d,μ) be a (weak) PI space and $\phi\colon X \to \mathbb{R}^n$ Lipschitz satisfying eq. (6.1) for all $x \in U \subset X$. There exists a decomposition $U = \cup_i U_i$ such that, for each $i \in \mathbb{N}$ and any $w \in \mathbb{S}^{n-1}$, any $C(w,1/i)$-null subset of U_i has μ measure zero.*

Proof. Since the definition of a $C(w,\theta)$-null set is invariant under scaling ϕ, we may suppose that ϕ is 1-Lipschitz. Secondly, for each $x \in U$, the function

$$D \in \mathbb{S}^{n-1} \mapsto \operatorname{Lip}(D\cdot\phi, x)$$

is Lipschitz and positive by assumption. Therefore, it is bounded below by some $\lambda_x > 0$ for each $x \in U$. The map $x \mapsto \lambda_x$ is Borel and so we may decompose U into sets $U_i = \{x : \lambda_x > 1/i\}$, $i \in \mathbb{N}$.

For $0<\theta<1$ and $w \in \mathbb{S}^{n-1}$, fix a $C(w,\theta)$-null set $S \subset U$. Let $f = w\cdot\phi\colon X \to \mathbb{R}$. We claim that

$$\rho = 1_{S^c} + \theta 1_S$$

is an upper gradient of f. Indeed, if $\gamma\colon [0,l] \to X$ is parametrised by arc-length, then by Proposition 5.7, $(\phi\circ\gamma)' \notin \operatorname{int}(C(w,\theta))$ for almost every $t \in \gamma^{-1}(S)$. That is,

$$|(f\circ\gamma)'(t)| = |w\cdot(\phi\circ\gamma)'(t)| \leq \theta\|(\phi\circ\gamma)'(t)\| \leq \theta\operatorname{Lip}\phi = \theta$$

for almost every $t \in \gamma^{-1}(S)$. Therefore, by the fundamental theorem of calculus,

$$|f(\gamma(l)) - f(\gamma(0))| \leq \int_0^l |(f \circ \gamma)'(t)| \leq \int_{\gamma^{-1}(S)} \theta + \int_{\gamma^{-1}(X \setminus S)} 1 = \int_0^l \rho(t)\,\mathrm{d}t,$$

as required.

By applying Lemma 6.3, we find a $C, \eta \geq 1$ depending only on $\mathrm{PI}(\mu)$ such that, for μ almost every $x \in S$,

$$\frac{1}{i} \leq \lambda_x \leq \mathrm{Lip}(f, x) \leq C\rho(x)^{1/\eta} = C\theta^{1/\eta}.$$

This is impossible if θ is sufficiently small, and so we must have $\mu(S) = 0$. Precisely, any $C(w, i^{-\eta}/C)$-null subset of U_i has μ measure zero. Re-indexing the U_i completes the proof. □

We complete this section by summarising the main result.

Theorem 6.6. *Let (X, d, μ) be a PI space. Suppose that $U \subset X$ and $\phi \colon X \to \mathbb{R}^n$ is Lipschitz such that eq.* (6.1) *holds for all $x \in U$. Then there exists a decomposition $U = \cup_i A_i$ such that each $\mu|_{A_i}$ has n independent Alberti representations.*

In particular, if there exists an $N \in \mathbb{N}$ such that, for any $U \subset X$ with $\mu(U) > 0$, $\mu|_U$ can have at most N independent Alberti representations, then X is a Lipschitz differentiability space.

Proof. Let $U = \cup_i U_i$ be the decomposition given by Proposition 5.7 and fix an $i \in \mathbb{N}$. Apply Proposition 5.9 to $\mu|_{U_i}$ with $\theta = 1/i$. Since any $C(w, 1/i)$-null subset of U_i has μ measure zero, this gives a decomposition $U_i = \cup_j A_i^j$ where each $\mu|_{A_i^j}$ has n independent Alberti representations. After re-indexing the sequence, this is the required decomposition.

For the second part, by combining the hypotheses with the first part of the theorem, we precisely satisfy the hypotheses of Proposition 6.2. The conclusion follows. □

7. Gromov–Hausdorff convergence and tangents of metric spaces

An important property of the set of rectifiable curves is that it is compact. This follows from the Arzelà–Ascoli theorem.

Theorem 7.1 (Arzelà–Ascoli)**.** *Let X, Y be compact metric spaces and $L \geq 0$. The set of all L-Lipschitz $f \colon X \to Y$ equipped with the supremum metric is compact.*

Proof. The set of all L-Lipschitz $f \colon X \to Y$ is a metric space and so, given a sequence f_n of such functions, it suffices to prove that f_n has a subsequence that converges to some L-Lipschitz $f \colon X \to Y$.

Since X is totally bounded, for each $m \in \mathbb{N}$, there exists $x_1, \ldots, x_{M_m} \in X$ such that

$$X = B(x_1^m, 1/m) \cup \ldots \cup B(x_{M_m}^m, 1/m). \tag{7.1}$$

First we define the function f. Consider the sequence $f_n(x_1^1) \in Y$. Since Y is compact, there exists a subsequence $f_{n_1(j)}(x_1^1)$ and some $y_1 \in Y$ such that $f_{n_1(j)}(x_1^1) \to y_1$ as $j \to \infty$. We define $f(x_1^1) = y_1$.

Now we move onto $f_{n_1(j)}(x_2^1)$. Again, since Y is compact there exists $y_2 \in Y$ and a subsequence $f_{n_2(j)_k}(x_2^1)$ such that $f_{n_2(j)_k}(x_2^1) \to y_2$ as $k \to \infty$. We define $f(x_2^1) = y_2$ and note that, because this is a subsequence of the previous subsequence, we also have $f_{n_2(j)_k}(x_1^1) \to f(x_1^1)$ as $k \to \infty$.

If we continue this for each $x_1^1, \ldots, x_{M_1}^1$ we obtain a subsequence $f_{n^1(j)}$ of f_n such that

$$f_{n^1(j)}(x_i^1) \to f(x_i^1) \quad \text{as } j \to \infty \ \forall 1 \leq i \leq M_1.$$

We now repeat this process for each $x_1^2, x_2^2, \dots, x_{M_2}^2$, but using $f_{n^1(j)}$ as the starting sequence. This gives a subsequence $f_{n^2(j)}$ such that

$$f_{n^2(j)}(x_i^2) \to f(x_i^2) \quad \text{as } j \to \infty \ \forall 1 \leq i \leq M_2.$$

Also, since this is a subsequence of f_{n^1}, we also have

$$f_{n^2(j)}(x_i^1) \to f(x_i^1) \quad \text{as } j \to \infty \ \forall 1 \leq i \leq M_1.$$

By repeating this over each scale and taking a diagonal subsequence, we get a subsequence, which for simplicity we will call f_n, such that

$$f_n(x_i^m) \to f(x_i^m) \quad \text{as } n \to \infty$$

for each $m \in \mathbb{N}$ and each $1 \leq i \leq M_m$.

By the exercises,

$$f \colon \bigcup_{m \in \mathbb{N}} \{x_1^m, \dots, x_{M_m}^m\} \to Y$$

is L-Lipschitz and may be extended to an L-Lipschitz $f \colon X \to Y$. To prove that $f_n \to f$ uniformly, let $m \in \mathbb{N}$ be arbitrary and $N \in \mathbb{N}$ such that, for all $n \geq N$, $|f_n(x_i^m) - f(x_i^m)| < 1/m$ for each $1 \leq i \leq M_m$. Then, for any $x \in X$, by eq. (7.1), there exists a $1 \leq i \leq M_m$ such that $d(x, x_i^m) < 1/m$. In particular, since f_n and f are L-Lipschitz,

$$\begin{aligned} |f_n(x) - f(x)| &\leq |f_n(x) - f_n(x_i^m)| + |f_n(x_i^m) - f(x_i^m)| + |f(x_i^m) - f(x)| \\ &\leq L/m + 1/m + L/m = (2L+1)/m. \end{aligned}$$

Thus, $f_n \to f$ uniformly, as required. □

We want to consider a notion of convergence of metric spaces. There are many variations of a core idea developed by Gromov. Collectively they are all known as "Gromov–Hausdorff convergence". The application we are interested in (defining the tangent of a doubling metric space) requires us to consider a distinguished point of the metric space. A *pointed metric space* (X, d, x) is a metric space (X, d) and a point $x \in X$.

Definition 7.2 (Gromov–Hausdorff)**.** A sequence (X_n, d_n, x_n) of pointed metric spaces *Gromov–Hausdorff converges* to a pointed metric space (X, d, x) if there exists a sequence of maps (called *Hausdorff approximations*) $\iota_n \colon X \to X_n$ with $\iota_n(x) = x_n$ such that, for every $\epsilon, R > 0$, there exists an $N \in \mathbb{N}$ such that, for all $n \geq N$,

$$|d_n(\iota_n(y), \iota_n(z)) - d(y, z)| < \epsilon \tag{7.2}$$

for all $y, z \in B(x, R)$ and

$$B(\iota_n(B(x, R)), \epsilon) \supset B(x_i, R - \epsilon). \tag{7.3}$$

Remark 7.3. Note that we do not assume any regularity on the Hausdorff approximations.

The first condition says that the Hausdorff approximations become more and more isometric on compact sets, whilst the second says they become more and more surjective.

Using similar techniques to the Arzelà–Ascoli theorem, we have a compactness theorem for Gromov–Hausdorff convergence.

Theorem 7.4. *Let (X_n, d_n, x_n) be a sequence of uniformly doubling pointed metric spaces. Then there exists a doubling pointed metric space (X, d, x) (with the same doubling constant) such that, after possibly taking a subsequence, $(X_n, d_n, x_n) \to (X, d, x)$.*

Proof. Fix $R>0$. For each $k,n\in\mathbb{N}$ let

$$Y_k^n=\{y_0^n,y_1^n,\ldots,y_{N(k)}^n\}$$

be a $1/k$-net of $B(x_n,R)$. Note that since the X_n are uniformly doubling, such a net, where the number of elements is independent of n, exists. We may also require that $y_0^n=x_n$ for each $n\in\mathbb{N}$. Note that, by our choice of notation, we are also requiring that the $1/k$-nets of each X_n are nested as k increases.

Consider the tuples

$$D_k^n:=(d_n(y_i^n,y_j^n):1\le i,j\le N(k))\in[0,R]^{N(k)^2}.$$

By the compactness of $[0,R]$, by taking a subsequence if necessary, we may suppose that there exists a $D_k\in[0,R]^{N(k)^2}$ such that $D_k^n\to D_k$ as $n\to\infty$. Let $\hat{B}_k$ be an arbitrary set of $N(k)$ elements, say

$$\hat{B}_k=\{x_1,\ldots,x_{N(k)}\}$$

and define $\hat{d}_k(x_i,x_j)=(D_k)_{i,j}$, so that

$$\hat{d}_k(x_i,x_j)=\lim_{n\to\infty}d_n(y_i^n,y_j^n). \tag{7.4}$$

Thus, $\hat{d}_k$ defines a *pseudo metric* on $\hat{B}_k$. That is, $\hat{d}_k$ satisfies all of the properties of a metric except possibly we may have some $i\neq j$ for which $\hat{d}_k(x_i,x_j)=0$.

We now repeat this for each $k\in\mathbb{N}$ and take a diagonal subsequence such that eq. (7.4) holds for each $k\in\mathbb{N}$. Since the D_k^n are nested, the $\hat{B}_k$ are also nested and $\hat{d}_k|_{\hat{X}_j}=\hat{d}_j$ whenever $j\le k$. Thus we may define $\hat{B}=\cup_k\hat{B}_k$ and $\hat{d}=\lim\hat{d}_k$ on $\hat{B}$, so that $\hat{d}$ is also a pseudo metric. We let $(\tilde{B},\tilde{d})$ be the quotient metric space of $(\hat{B},\hat{d})$. That is, let $\tilde{B}$ be the set of equivalence classes of $\hat{B}$ under $x\sim y\Leftrightarrow\hat{d}(x,y)=0$ and $\tilde{d}([x],[y])=\hat{d}(x,y)$ (you can check that this is a well defined metric space). Finally, we let (B,d) be the completion of $(\tilde{B},\tilde{d})$.

Now observe that, for any $n\in\mathbb{N}$ and $k\le j\in\mathbb{N}$, D_k^n is a $1/k$-net of D_j^n. It follows that $\hat{B}_k$ is a $1/k$-net of $\hat{B}_j$ and hence of $\hat{B}$. Thus

$$N_k:=\{[x]:x\in\hat{B}_k\}$$

is a $1/k$-net of $\tilde{B}$ and hence of B. We define $\iota_n\colon N_n\to B_n$ as follows. For any $y\in N_n$, choose the smallest $1\le i\le N(k)$ such that $z=[x_i]$ and define $\iota_n(z)=y_i^n$. We extend ι_n to B by mapping any $z\in B$ to an arbitrarily chosen $\iota_n(x)$ with $x\in N_n$ and $x'\in B(x,1/n)$.

In this construction, Y_k^n is a $1/k$-net of $B(x_n,R)$. It follows that $B=B(x_0,R)$. The final step of the construction is to take a final diagonal convergent subsequence as $R\to\infty$.

It is now a matter of checking the conclusion of the theorem is satisfied. For simplicity, we use the notation for $B=B(x,R)$ as above. Since

$$\iota_n(B(x,R))\supset\iota_n(N_n)=B_n,$$

and B_n is a $1/n$-net of $B(x_n,R)$, we have

$$B(\iota_n(B(x_0,R)),1/n)\supset B(x_n,R),$$

and hence eq. (7.3). To see eq. (7.2), observe that for any $k\in\mathbb{N}$ and $x,y\in N_k$,

$$d_n(\iota_n(x),\iota_n(y))\to d(x,y)$$

by the definition of d. Since N_k contains only finitely many points, we can ensure that

$$|d_n(\iota_n(x),\iota_n(y))-d(x,y)|<1/k$$

for all $x, y \in N_k$ and all n greater than some N. Therefore, for any $x', y' \in B$ and $n \geq N$, there exists $x, y \in N_k$ with $d(x, x'), d(y, y') \leq 1/k$ and $\iota_n(x) = \iota_n(x')$ and $\iota_n(y) = \iota_n(y')$. Thus

$$\begin{aligned} |d_n(\iota_n(x'), \iota_n(y')) - d(x', y')| &\leq |d_n(\iota_n(x), \iota_n(y)) - d(x, y)| + |d(x, y) - d(x', y')| \\ &\leq 1/k + 2/k. \end{aligned}$$

□

In what follows, we will not need to consider the specific metric in a metric space (X, d) and so will simply refer to it as X. Given $\lambda > 0$, we will write λX for the metric space $(X, \lambda d)$.

Definition 7.5. Let (X, x) be a pointed metric space. A *Gromov–Hausdorff tangent* of X at x is any limit of a sequence of the form $(\lambda X, x)$ for $\lambda \to \infty$.

We denote by $\operatorname{Tan}(X, x)$ the set of all Gromov–Hausdorff tangents of X at x.

If X is doubling, then for any $\lambda > 0$, λX is also doubling with the same constant. Thus, for any $x \in X$ and $\lambda_i \to \infty$, by the previous compactness theorem, there exists a subsequence λ_{i_k} such that $(\lambda_{i_k} X, x)$ converges to some tangent space. Therefore, $\operatorname{Tan}(X, x) \neq \emptyset$. Note, however, that tangents are rarely unique. Certainly, if $(Y, y) \in \operatorname{Tan}(X, x)$ then also $(\lambda Y, y) \in \operatorname{Tan}(X, x)$ for any $\lambda > 0$. However, much more non uniqueness properties can occur: for n_j an increasing sequence of integers, consider the tangents to the graph of the following function at the origin:

$$f(x) = x\chi_A$$

for

$$A = \{x \in \mathbb{R} : \exp(\exp(-n_j)) \leq x \leq \exp(\exp(-n_{j-1})),\ j \text{ odd}\}.$$

Of course, much stranger phenomena can occur.

We will use the following fact.

Fact. *Gromov–Hausdorff convergence of doubling metric spaces is metrisable. That is, there exists a metric d_{GH} on the set $\mathcal{M}$ of all pointed doubling metric spaces such that (X_n, x_n) converges to (X, x) if and only if*

$$d_{GH}((X_n, x_n), (X, x)) \to 0.$$

By Theorem 7.4, $(\mathcal{M}, d_{GH})$ is a countable union of compact metric spaces and hence is separable.

As we will see in the following results, the precise value of this metric is not important to us, only the fact that Gromov–Hausdorff convergence is metrisable.

Recall that a point $a \in A \subset (X, d)$ is a *porosity point* of A if there exists a $\lambda > 0$ and $x_n \in X$ such that $x_n \to a$ and

$$A \cap B(x_n, \lambda d(x_n, a)) = \emptyset.$$

Therefore, if a is *not* a porosity point,

$$\lim_{R \to 0} \sup_{y \in B(a,R)} \frac{d(y, A \cap B(a, R))}{R} = 0. \tag{7.5}$$

We know, by the Lebesgue density theorem, if X is equipped with a doubling measure μ, then μ almost every point of A is not a porosity point.

Lemma 7.6. *Suppose that $A \subset (X, d)$ and that $a \in A$ is not a porosity point of A. Then $\operatorname{Tan}(A, a) = \operatorname{Tan}(X, a)$.*

Proof. Let $(Y,t) \in \mathrm{Tan}(A,a)$. Then there exists $\lambda_i \to \infty$ and Hausdorff approximations $\iota_n\colon (Y,y) \to (\lambda_n A, a)$. We define Hausdorff approximations, also called ι_n, into $(\lambda_n X, a)$ by simply post composing with the inclusion of A into X. Certainly eq. (7.2) remains true for these Hausdorff approximations, we just need to check eq. (7.3). Given $R, \epsilon > 0$, since eq. (7.3) applies for the Hausdorff approximations into $(\lambda_n A, a)$, there exists $N_1 \in \mathbb{N}$ such that, for all $n \geq N_1$,

$$B(\iota_n(B(y,R)),\epsilon) \supset B(a,(R-\epsilon)/\lambda_n) \cap A.$$

By eq. (7.5), there exists $N_2 \in \mathbb{N}$ such that, for all $n \geq N_2$,

$$B(B(a,(R-\epsilon)/\lambda_n) \cap A, \epsilon/\lambda_n) \supset B(a,(R-\epsilon)/\lambda_n).$$

Therefore, for all $n \geq N_1, N_2$,

$$B(\iota_n(B(t,R)),2\epsilon) \supset B(a,(R-2\epsilon)/\lambda_n),$$

as required.

Now suppose that $(Y,y) \in \mathrm{Tan}(X,a)$. Then there exists $\lambda_n \to 0$ and Hausdorff approximations $\iota_n\colon (Y,y) \to (\lambda_n X, a)$. We define Hausdorff approximations $\tilde{\iota}_n\colon (Y,y) \to (\lambda_n A, a)$ as follows. For $z \in Y$, if $\iota_n(z) \in A$ let $\tilde{\iota}_n(z) = \iota_n(z)$. Otherwise, pick $p \in A$ with

$$d(\iota_n(z),p) \leq d(\iota_n(z),A) + \frac{1}{n\lambda_n} \tag{7.6}$$

and set $\tilde{\iota}_n(z) = p$. Note that eq. (7.3) is automatically true for $\tilde{\iota}_n$. To see eq. (7.2), observe that, for any $R > 0$, eq. (7.5) and and eq. (7.6) show that

$$\lambda_n d(\iota_n(z), \tilde{\iota}_n(z)) \to 0$$

uniformly on $B(t,R)$. Therefore, eq. (7.2) for the $\tilde{\iota}_n$ follows from eq. (7.2) for the ι_n. □

The previous observation allows us to show that moving the base point of a tangent gives a space that is also a tangent.

Theorem 7.7. *Let (X,μ) be a doubling metric measure space. Then for μ almost every $x \in X$ the following is true. If $(Y,y) \in \mathrm{Tan}(X,x)$ and $y' \in Y$, then $(Y,y') \in \mathrm{Tan}(X,x)$.*

Proof. We must show that the following set has μ measure zero:

$$\{x \in X : \exists\ (Y,y) \in \mathrm{Tan}(X,x),\ y' \in Y \text{ s.t. } (Y,y') \notin \mathrm{Tan}(X,x)\}.$$

If $(Y,y') \notin \mathrm{Tan}(X,x)$ then there exists $k \in \mathbb{N}$ such that

$$d((Y,y'),(\lambda X,x)) > 1/k \quad \forall\ \lambda > k.$$

Thus, it suffices to prove, for a $\delta > 0$ which we now fix, that the following set has μ measure zero:

$$\{x \in X : \exists\ (Y,y) \in \mathrm{Tan}(X,x),\ y' \in Y \text{ s.t. } d((Y,y'),(\lambda X,x)) > \delta\ \forall\ \lambda > \delta\}.$$

Since the set of all doubling metric spaces is separable with respect to Gromov–Hausdorff convergence, there exist a countable cover by sets of diameter $\delta/4$. Let B be one such set. It suffices to prove that the set A defined to be those $x \in X$ for which

$$\exists\ (Y,y) \in B \cap \mathrm{Tan}(X,x),\ y' \in Y \text{ s.t. } d((Y,y'),(\lambda X,x)) > \delta\ \forall\ \lambda > \delta$$

has measure zero. Finally, for μ almost every $x \in A$, $\mathrm{Tan}(X,x) = \mathrm{Tan}(A,x)$. Therefore, it suffices to prove that the set A' of those $x \in A$ for which

$$\exists\ (Y,y) \in B \cap \mathrm{Tan}(A,x),\ y' \in Y \text{ s.t. } d((Y,y'),(X,d/r,x)) > \delta\ \forall\ \lambda > \delta$$

has measure zero. In fact, we will show that it is empty.

Indeed, suppose that $x \in A'$. Then there exist $\lambda_n \to \infty$ and Hausdorff approximations $\iota_n : (Y,y) \to (\lambda_n A, x)$. Then, for $a_n = \iota_n(y') \in A$, $(\lambda_n A, a_n) \to (Y, y')$, and so there exists $N \in \mathbb{N}$ such that

$$d((\lambda_n A, a_n), (Y, y')) < \delta/4 \tag{7.7}$$

for each $n \geq N$.

However, each $a_n \in A$ and so there exists $(Y_n, y_n) \in \operatorname{Tan}(X, a_n)$ and a $y'_n \in Y_n$ such that $(Y_n, y'_n) \in B$ and

$$d((Y_n, y'_n), (\lambda_n A, a_n) > \delta \tag{7.8}$$

whenever $r_n < \delta$. In particular, since B has diameter $\delta/4$,

$$d((Y_n, y'_n), (Y, y')) < \delta/4 \tag{7.9}$$

for each $n, m \in \mathbb{N}$. Therefore, for any $n \geq N$ for which $r_n < \delta$,

$$\begin{aligned} \delta &\overset{(7.8)}{<} d((Y_n, y'_n), (\lambda_n A, a_n) \\ &\leq d((Y_n, y'_n), (Y, y')) + d((Y, y'), (\lambda_n A, a_n)) \\ &\overset{(7.7)(7.9)}{\leq} \delta/4 + \delta/4, \end{aligned}$$

a contradiction. □

We need to make a slight modification to our use of Gromov–Hausdorff convergence, because we must also consider the limiting behaviour of a Lipschitz function defined on the space.

Given a pointed metric space (X, x) and a Lipschitz $\phi : X \to \mathbb{R}^n$ we call the triple (X, x, f) a space function.

Definition 7.8. A space function (Y, y, ψ) is a *Gromov–Hausdorff tangent* to a space function (X, x, ϕ) if there exists $\lambda_n \to \infty$ and Hausdorff approximations $\iota_n : (Y, y) \to (\lambda_n X, x)$ (so that $(Y, y) \in \operatorname{Tan}(X, x)$) such that

$$\lambda_n(\phi(\iota_n(z)) - \phi(x)) \to g(z) - g(y)$$

uniformly on any $B(y, R)$.

The set of all Gromov–Hausdorff tangents to (X, x, ϕ) is denoted by $\operatorname{Tan}(X, x, \phi)$.

Of course, if (X, x) is doubling, we know that $\operatorname{Tan}(X, x)$ is not empty. Suppose that for some $\lambda_n \to \infty$, $(\lambda_n X, x) \to (Y, y)$. For any $\lambda > 0$ and any Lipschitz $\phi : X \to \mathbb{R}^n$, the function $\lambda\phi : \lambda X \to \mathbb{R}^n$ has the same Lipschitz constant as ϕ. Therefore, by an argument similar to the Arzelà–Ascoli theorem, there exists a subsequence $\lambda_{n_j} \to \infty$ and a Lipschitz function $\psi : Y \to \mathbb{R}^n$ such that $(Y, y, \psi) \in \operatorname{Tan}(X, x, \psi)$.

By almost the same proof, the analogue to Theorem 7.7 is also true.

Theorem 7.9. *Let (X, μ) be a doubling metric measure space and $\phi : X \to \mathbb{R}^n$ Lipschitz. Then for μ almost every $x \in X$ the following is true. If $(Y, y, \psi) \in \operatorname{Tan}(X, x, \phi)$ and $y' \in Y$, then $(Y, y', \psi - \psi(y')) \in \operatorname{Tan}(X, x, \phi)$.*

Finally, we apply the theory of Gromov–Hausdorff tangents to the setting of metric measure spaces with Alberti representations. For X a metric space and $x \in X$, a *line* passing through x is an isometry $g : \mathbb{R} \to X$ that contains x. For $v \in \mathbb{R}^n$, such a line is *in the direction of* v (with respect to some $\phi : X \to \mathbb{R}^n$) if $\phi(g(\mathbb{R})) = \mathbb{R}v$.

Lemma 7.10. *Let (X, d, μ) be a metric measure space, $\phi : X \to \mathbb{R}^n$ Lipschitz and $C \subset \mathbb{R}^n$ a cone. Suppose that, for some $A \subset X$, $\mu|_A$ has a C-Alberti representation with respect to ϕ. Then for μ almost every $x \in A$ there exists a $v \in C \setminus \{0\}$ such that, for every $(Y, y, \psi) \in (X, x, \phi)$, Y contains a line passing through y in the ψ direction of v.*

Proof. From the definition of an Alberti representation, we know for μ almost every $x \in A$, there exists a $\gamma \in \Gamma$ and $t \in [0,1]$ such that $\gamma(t) = x$ and $(\phi\circ\gamma)'(t) \in C\setminus\{0\}$. Fix such a point $x \in A$.

Now suppose that $(Y, y, \psi) \in (X, x, \phi)$ and that $\iota_n \colon (Y, y) \to (\lambda_n X, x)$ are Hausdorff approximations. Since $(\phi \circ \gamma)'(t)$ exists, by unravelling the definitions, it is straightforward (but a bit tedious) to show that there exists a map $g \colon \mathbb{R} \to Y$ containing y such that $(\psi \circ g)'(s) = (\phi \circ \gamma)'(t)$ for all $s \in \mathbb{R}$. By reparametrising g, we see that it is a line passing through y in the direction of $v := (\phi\circ\gamma)'(t) \in C\setminus\{0\}$. □

By considering many independent Alberti representations we obtain the following.

Proposition 7.11. *Let (X, d, μ) be a doubling metric measure space and $\phi \colon X \to \mathbb{R}^n$ Lipschitz. Suppose that, for some $A \subset X$, $\mu|_A$ has n independent Alberti representations with respect to ϕ. Then for μ almost every $x \in A$ and every $(Y, y, \psi) \in (X, x, \phi)$, $\psi \colon Y \to \mathbb{R}^n$ is surjective.*

Proof. By applying Lemma 7.10 to each of the Alberti representations in the hypothesis, we know that the following is true. For μ almost every $x \in A$ there exist linearly independent $v_1, \ldots, v_n \in \mathbb{R}^n$ such that, for every $(Y, y, \psi) \in (X, x, \phi)$, Y contains lines $g_1, \ldots, g_n$ each passing through y in the ψ direction of $v_1, \ldots, v_n$ respectively.

Since (X, d, μ) is doubling, we can apply Theorem 7.9. This implies that, for μ almost every $x \in A$, the following is true. There exist linearly independent $v_1, \ldots, v_n \in \mathbb{R}^n$ such that, for every $(Y, y, \psi) \in (X, x, \phi)$ and any $y' \in Y$, Y contains lines $g_1, \ldots, g_n$ each passing through y' in the ψ direction of $v_1, \ldots, v_n$ respectively. Therefore, for any $(Y, y) \in \operatorname{Tan}(X, x)$, $\psi(Y) = \mathbb{R}^n$.

Indeed, since (Y, y) contains a line passing through y in the direction of v_1, $\psi(Y)$ contains $V_1 := \psi(y) + \mathbb{R}v_1$. Now, from any point $z \in V_1$, there exists $y' \in Y$ such that $\psi(y') = z$. Since Y contains a line passing through y' in the direction of v_2, $\psi(Y)$ contains $\psi(y') + \mathbb{R}v_2$ and hence $V_2 := \psi(y) + \mathbb{R}v_1 + \mathbb{R}v_2$. Repeating iteratively, we see that $\psi(Y)$ contains

$$V_n := \psi(y) + \mathbb{R}v_1 + \mathbb{R}v_2 + \ldots + \mathbb{R}v_n = \mathbb{R}^n$$

as required. □

This gives us the required bound on the number of independent Alberti representations by the doubling constant.

Theorem 7.12. *Let (X, d, μ) be a doubling metric measure space. There exists a $N \in \mathbb{N}$ depending only on the doubling constant of X such that, for any $A \subset X$ with $\mu(A) > 0$, $\mu|_A$ can have at most N independent Alberti representations.*

Proof. Suppose that Z is an N-doubling metric space. Then we know that any ball $B(z, R)$ can be covered by at most N^n balls of radius $2^{-n}R$. In particular, for any $\alpha > \log_2 N$, $\mathcal{H}^\alpha(B(x, r)) = 0$, and hence $\dim_H Z \leq \log_2 N$.

Now suppose that $(Y, y) \in \operatorname{Tan}(X, x)$, for X as in the statement of the theorem. We know that X is N-doubling for some $N \in \mathbb{N}$ and hence so is Y. Therefore, $\dim_H Y \leq \log_2 N$. In particular, a Lipschitz function cannot map Y onto $\mathbb{R}^n$ with $n > \log_2 N$. Since (X, d, μ) is doubling, the conclusion follows from Proposition 7.11. □

To conclude we summarise the proof of Cheeger's theorem.

Proof of Theorem 4.7. Let (X, d, μ) be a PI space. By Proposition 6.2 it suffices to find an $N \in \mathbb{N}$ for which, whenever $\phi \colon X \to \mathbb{R}^n$ is Lipschitz satisfying eq. (6.1) at all points $x \in U$ with $\mu(U) > 0$, we must have $n \leq N$. Since X satisfies a Poincaré inequality, by applying Theorem 6.6, there exists a countable decomposition $U = \cup_i U_i$ such that each $\mu|_{U_i}$ has n independent Alberti representations. Thus, if we find a bound N to the number of independent of Alberti representations of any positive measure subset of X, we are done. This is given by Theorem 7.12. □

References

[Bat14] D. Bate. "Structure of measures in Lipschitz differentiability spaces". In: *J. Amer. Math. Soc.* 28.2 (2014), pp. 421–482. DOI: 10.1090/s0894-0347-2014-00810-9.

[Che99] J. Cheeger. "Differentiability of Lipschitz functions on metric measure spaces". In: *Geom. Funct. Anal.* 9.3 (1999), pp. 428–517. DOI: 10.1007/s000390050094.

[HK98] J. Heinonen and P. Koskela. "Quasiconformal maps in metric spaces with controlled geometry". In: *Acta Math.* 181.1 (1998), pp. 1–61. DOI: 10.1007/bf02392747.

[Kei04] S. Keith. "A differentiable structure for metric measure spaces". In: *Adv. Math.* 183.2 (2004), pp. 271–315. DOI: 10.1016/S0001-8708(03)00089-6.

[Kir94] B. Kirchheim. "Rectifiable metric spaces: local structure and regularity of the Hausdorff measure". In: *Proc. Amer. Math. Soc.* 121.1 (1994), pp. 113–123. DOI: 10.2307/2160371.

[KM11] B. Kleiner and J. Mackay. "Differentiable structures on metric measure spaces: A primer". In: *arXiv* 1108.1324 (2011), p. 23.

[Laa00] T. J. Laakso. "Ahlfors Q-regular spaces with arbitrary $Q > 1$ admitting weak Poincaré inequality". In: *Geom. Funct. Anal.* 10.1 (2000), pp. 111–123. DOI: 10.1007/s000390050003.

[Laa02] T. J. Laakso. "Plane with A_∞-weighted metric not bi-Lipschitz embeddable to $\mathbb{R}^N$". In: *Bull. London Math. Soc.* 34.6 (2002), pp. 667–676. DOI: 10.1112/S0024609302001200.

P.O. Box 68 (Gustaf Hällströmin katu 2b), FI-00014 University of Helsinki
Email address: david.bate@helsinki.fi