Notes on Rectifiability

Urs Lang

2007
Version of December 15, 2023

Abstract

These are the notes to a first part of a lecture course on Geometric Measure Theory I have taught at ETH Zurich a number of times in the past twenty years, lastly in the Fall Semester 2023.

Contents

1 Embeddings of metric spaces

We start with some basic and well-known isometric embedding theorems for metric spaces.

For a non-empty set S, $l_\infty(S) = (l_\infty(S), \|\cdot\|_\infty)$ denotes the Banach space of all functions $u\colon S \to \mathbb{R}$ with $\|u\|_\infty := \sup_{s\in S} |u(s)| < \infty$. Similarly, $l_\infty := l_\infty(\mathbb{N}) = \{(u_k)_{k\in\mathbb{N}} : \|(u_k)\|_\infty := \sup_{k\in\mathbb{N}} |u_k| < \infty\}$.

Proposition 1.1 (Kuratowski, Fréchet) (1) *Every metric space X admits an isometric embedding into $l_\infty(X)$.*
(2) *Every separable metric space admits an isometric embedding into l_∞.*

Proof: For (1), fix a basepoint $z \in X$ and define $f\colon X \to l_\infty(X)$,

$$x \mapsto u_x, \quad u_x(y) = d(x,y) - d(y,z).$$

Note that $\|u_x\|_\infty = \sup_y |u_x(y)| \le d(x,z)$. Moreover,

$$\|u_x - u_{x'}\|_\infty = \sup_y |d(x,y) - d(x',y)| \le d(x,x'),$$

and equality occurs for $y = x'$.

To prove (2), choose a basepoint $z \in X$ and a countable dense set $S = \{y_k : k \in \mathbb{N}\}$ in X. Define $f\colon X \to l_\infty$,

$$x \mapsto (d(x,y_k) - d(y_k,z))_{k\in\mathbb{N}}.$$

By (1), $f|_S\colon S \to l_\infty = l_\infty(S)$ is an isometric embedding. Since S is dense and f is continuous, f is an isometric embedding. □

Note that if X is bounded, there is no need for the term $-d(y,z)$ or $-d(y_k,z)$, respectively, in the definition of f. In this case, the embedding is canonical. For further reading on results of this type and detailed references we refer to [Hei2003].

Recall that a metric space X is said to be *precompact* or *totally bounded* if for every $\epsilon > 0$, X can be covered by a finite number of closed balls of radius ϵ. We call a set $Y \subset X$ *ϵ-separated* if $d(y,y') \ge \epsilon$ whenever $y, y' \in Y$, $y \ne y'$. Note that X is precompact if and only if for every $\epsilon > 0$, all ϵ-separated subsets of X are finite. A metric space is compact if and only if it is precompact and complete.

Definition 1.2 (uniformly precompact family) A family $(X_\alpha)_{\alpha\in A}$ of metric spaces is called *uniformly precompact* if for all $\epsilon > 0$ there exists a number $n = n(\epsilon) \in \mathbb{N}$ such that each X_α can be covered by n closed balls of radius ϵ. The family $(X_\alpha)_{\alpha\in A}$ is *uniformly bounded* if $\sup_{\alpha\in A} \operatorname{diam}(X_\alpha) < \infty$.

Theorem 1.3 (Gromov embedding) *Suppose that $(X_\alpha)_{\alpha \in A}$ is a uniformly precompact and uniformly bounded family of metric spaces. Then there is a compact metric space Z such that each X_α admits an isometric embedding into Z.*

We follow essentially the original proof from [Gro1981].

Proof: For $i \in \mathbb{N}$, let $\epsilon_i := 2^{-i}$ and pick $n_i \in \mathbb{N}$ such that each X_α can be covered by n_i closed balls of radius ϵ_i. Choose a partition of $\mathbb{N}$ into sets N_i, $i \in \mathbb{N}$, with cardinality $\# N_i = n_1 n_2 \dots n_i$, and define a map $\pi\colon \mathbb{N} \setminus N_1 \to \mathbb{N}$ such that $\pi^{-1}(N_i) = N_{i+1}$ and $\#(\pi^{-1}\{k\}) = n_{i+1}$ for all $i \in \mathbb{N}$ and $k \in N_i$. Now we construct in each X_α a sequence $(x_k^\alpha)_{k \in \mathbb{N}}$ according to the following inductive scheme. For $i = 1$, the points x_k^α with $k \in N_i = N_1$ are chosen such that the n_1 balls $B(x_k^\alpha, \epsilon_1)$ cover X_α. For $i \geq 1$, if the $n_1 \dots n_i$ centers x_k^α with $k \in N_i$ are selected, the $n_1 \dots n_i n_{i+1}$ points x_l^α with $l \in N_{i+1}$ are chosen such that for each $k \in N_i$, the ball $B(x_k^\alpha, \epsilon_i)$ is covered by the n_{i+1} balls

$$B(x_l^\alpha, \epsilon_{i+1}) \subset B(x_k^\alpha, 2\epsilon_i)$$

with $l \in \pi^{-1}\{k\}$. In this way we obtain for every $\alpha \in A$ a dense sequence $(x_k^\alpha)_{k \in \mathbb{N}}$ in X_α which gives rise to an isometric embedding $f_\alpha\colon X_\alpha \to l_\infty$, mapping x to $(d(x, x_k^\alpha))_{k \in \mathbb{N}}$. Whenever $i \in \mathbb{N}$, $k \in N_i$, and $l \in \pi^{-1}\{k\}$, then

$$|d(x, x_k^\alpha) - d(x, x_l^\alpha)| \leq d(x_k^\alpha, x_l^\alpha) \leq 2\epsilon_i.$$

Hence, each $f_\alpha(X_\alpha)$ lies in the set Z of all sequences $(u_k)_{k \in \mathbb{N}}$ with $0 \leq u_k \leq \sup_\alpha \operatorname{diam}(X_\alpha)$ for all $k \in \mathbb{N}$ and

$$|u_k - u_l| \leq 2\epsilon_i$$

whenever $i \in \mathbb{N}$, $k \in N_i$, and $l \in \pi^{-1}\{k\}$. Since the sequence $(\epsilon_i)_{i \in \mathbb{N}}$ is summable, it follows that Z is a compact subset of l_∞. □

2 Compactness theorems for metric spaces

For subsets A, B of a metric space X we denote by

$$N_\delta(A) = \{x \in X : d(x, A) \leq \delta\}$$

the closed δ-neighborhood of A and by

$$d_{\mathrm{H}}(A, B) = \inf\{\delta \geq 0 : A \subset N_\delta(B),\ B \subset N_\delta(A)\}$$

the *Hausdorff distance* of A and B; d_{H} defines a metric on the set C of non-empty, closed and bounded subsets of X.

Theorem 2.1 (Blaschke) *Suppose that $X = (X, d)$ is a metric space and $\mathcal{C}$ is the set of non-empty, closed and bounded subsets of X, endowed with the Hausdorff metric d_{H}.*
(1) *If X is complete, then $\mathcal{C}$ is complete.*
(2) *If X is compact, then $\mathcal{C}$ is compact.*

This was first proved by Blaschke [Bla1916] for compact convex bodies in $\mathbb{R}^3$ to settle the existence question in the isoperimetric problem.

Proof: We start with (1). Let $(C_i)_{i\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{C}$. Then the set

$$C := \bigcap_{i=1}^{\infty} \overline{\bigcup_{j\geq i} C_j}$$

is closed and bounded. We show that

$$\lim_{i\to\infty} d_{\mathrm{H}}(C_i, C) = 0.$$

Let $\epsilon > 0$. Choose i_0 such that $d_{\mathrm{H}}(C_i, C_j) < \epsilon/2$ whenever $i, j \geq i_0$. Suppose $x \in C$. Since $C \subset \overline{\bigcup_{j\geq i_0} C_j}$, there exists an index $j \geq i_0$ with $d(x, C_j) < \epsilon/2$. Hence $d(x, C_i) \leq d(x, C_j) + d_{\mathrm{H}}(C_i, C_j) < \epsilon$ for all $i \geq i_0$. This shows that $C \subset N_\epsilon(C_i)$ for $i \geq i_0$.

Now suppose $x \in C_i$ for some $i \geq i_0$. Pick a sequence $i = i_1 < i_2 < \ldots$ such that $d_{\mathrm{H}}(C_m, C_n) < \epsilon/2^k$ whenever $m, n \geq i_k$, $k \in \mathbb{N}$. Then choose a sequence $(x_k)_{k\in\mathbb{N}}$ such that $x_1 = x$, $x_k \in C_{i_k}$ and $d(x_k, x_{k+1}) < \epsilon/2^k$. As X is complete, the Cauchy sequence (x_k) converges to some point y. We have

$$d(x, y) = \lim_{k\to\infty} d(x, x_k) \leq \sum_{k=1}^{\infty} d(x_k, x_{k+1}) < \epsilon,$$

and y belongs to the closure of $C_{i_k} \cup C_{i_{k+1}} \cup \ldots$ for all k. Thus $y \in C$ and $d(x, C) < \epsilon$. This shows that $C_i \subset N_\epsilon(C)$ whenever $i \geq i_0$.

For the proof of (2), we know that $\mathcal{C}$ is complete since X is, so it suffices to show that $\mathcal{C}$ is precompact. Let $\epsilon > 0$. Since X is precompact, there exists a finite set $Z \subset X$ with $N_\epsilon(Z) = X$. We show that every $C \in \mathcal{C}$ is at Hausdorff distance at most ϵ of some subset of Z, namely $Z_C := Z \cap N_\epsilon(C)$. For every $x \in C$ there exists a point $z \in Z$ with $d(x, z) \leq \epsilon$, so $z \in Z_C$. This shows that $C \subset N_\epsilon(Z_C)$. Since also $Z_C \subset N_\epsilon(C)$, we have $d_{\mathrm{H}}(C, Z_C) \leq \epsilon$. As there are only finitely many distinct subsets of Z, we conclude that $\mathcal{C}$ is precompact. □

Definition 2.2 (Gromov–Hausdorff distance) The *Gromov–Hausdorff distance* of two metric spaces X, Y is the infimum of all $r > 0$ for which there exist a metric space (Z, d^Z) and subspaces $X' \subset Z$ and $Y' \subset Z$ isometric to X and Y, respectively, such that $d_{\mathrm{H}}^Z(X', Y') < r$.

Compare [Gro1981], [Gro1999]. Alternatively, call a metric $\bar{d}$ on the disjoint union $X \sqcup Y$ *admissible* for the given metrics $d = d^X$ and $d = d^Y$ on X and Y if $\bar{d}|_{X\times X} = d^X$ and $\bar{d}|_{Y\times Y} = d^Y$; then

$$d_{\mathrm{GH}}(X,Y) = \inf \bar{d}_{\mathrm{H}}(X,Y)$$

where the infimum is taken over all admissible metrics $\bar{d}$ on $X \sqcup Y$.

For instance, suppose that $\mathrm{diam}(X), \mathrm{diam}(Y) \le D < \infty$. Setting $\bar{d}(x,y) = D/2$ for $x \in X$ and $y \in Y$ we obtain an admissible metric on $X \sqcup Y$, in particular $d_{\mathrm{GH}}(X,Y) \le D/2$.

Proposition 2.3 (1) *d_{GH} satisfies the triangle inequality, i.e. $d_{\mathrm{GH}}(X,Z) \le d_{\mathrm{GH}}(X,Y) + d_{\mathrm{GH}}(Y,Z)$ for all metric spaces X,Y,Z.*
(2) *d_{GH} defines a metric on the set of isometry classes of compact metric spaces.*

See [BurBI2001, Proposition 7.3.16, Theorem 7.3.30]. Assertion (2) is no longer true if 'compact' is replaced with 'complete and bounded'.

Theorem 2.4 (Gromov compactness criterion) *Suppose that $(X_i)_{i\in\mathbb{N}}$ is a uniformly precompact and uniformly bounded sequence of metric spaces. Then there exist a subsequence $(X_{i_j})_{j\in\mathbb{N}}$ and a compact metric space Z such that (X_{i_j}) Gromov–Hausdorff converges to Z, i.e.* $\lim_{j\to\infty} d_{\mathrm{GH}}(X_{i_j}, Z) = 0$.

This was proved in [Gro1981].

Proof: Combine Theorems 1.3 (Gromov embedding) and 2.1(2) (Blaschke). □

3 Lipschitz maps

Let X,Y be metric spaces, and let $\lambda \in [0,\infty)$. A map $f\colon X \to Y$ is *λ-Lipschitz* if

$$d(f(x), f(x')) \le \lambda\, d(x,x')$$

for all $x, x' \in X$, and f is *Lipschitz* if

$$\mathrm{Lip}(f) := \inf\{\lambda \in [0,\infty) : f \text{ is } \lambda\text{-Lipschitz}\} < \infty$$

(where $\inf \emptyset := \infty$). We say that $f\colon X \to Y$ is *bi-Lipschitz* if f is *λ-bi-Lipschitz* for some $\lambda \in [1,\infty)$, that is,

$$\lambda^{-1} d(x,x') \le d(f(x), f(x')) \le \lambda\, d(x,x')$$

for all $x, x' \in X$.

The following basic extension result for Lipschitz maps holds, see [McS1934] and the footnote in [Whit1934].

Proposition 3.1 (McShane, Whitney) *Suppose that X is a metric space, and $A \subset X$.*

(1) *Let $n \in \mathbb{N}$. Every λ-Lipschitz map $f\colon A \to \mathbb{R}^n$ admits a $\sqrt{n}\lambda$-Lipschitz extension $\bar{f}\colon X \to \mathbb{R}^n$.*

(2) *Let S be any non-empty set. Every λ-Lipschitz map $f\colon A \to l_\infty(S)$ possesses a λ-Lipschitz extension $\bar{f}\colon X \to l_\infty(S)$.*

Proof: To prove (1), consider first the case $n = 1$. Put

$$\bar{f}(x) := \inf\{f(a) + \lambda\, d(a,x) : a \in A\}$$

for all $x \in X$. Note that for $b \in A$, since f is λ-Lipschitz,

$$\bar{f}(x) \geq \inf\{f(b) - \lambda\, d(a,b) + \lambda\, d(a,x) : a \in A\} \geq f(b) - \lambda\, d(b,x),$$

in particular $\bar{f}(x) > -\infty$ and $\bar{f}(b) \geq f(b)$. As $\bar{f}(b) \leq f(b)$ by definition, $\bar{f}\colon X \to \mathbb{R}$ is an extension of f. For $x, x' \in X$,

$$\bar{f}(x) \leq \inf\{f(a) + \lambda\, d(a,x') + \lambda\, d(x,x') : a \in A\} = \bar{f}(x') + \lambda\, d(x,x'),$$

so $\bar{f}$ is λ-Lipschitz. In the case that $n \geq 2$, note that each component of $f = (f_1, \dots, f_n)$ is λ-Lipschitz and hence admits a λ-Lipschitz extension $\bar{f}_i\colon X \to \mathbb{R}$. The function $\bar{f} = (\bar{f}_1, \dots, \bar{f}_n)$ with these components is then $\sqrt{n}\lambda$-Lipschitz.

The proof of (2) is similar. Given $f = (f_s)_{s \in S}$, each component f_s has a λ-Lipschitz extension $\bar{f}_s\colon X \to \mathbb{R}$, and then $\bar{f} = (\bar{f}_s)_{s \in S}$ is λ-Lipschitz as well. □

In (1), the factor $\sqrt{n}$ cannot be replaced with a constant $< n^{1/4}$, compare [JohLS1986] and [Lan1999]. In particular, Lipschitz maps into a Hilbert space Y cannot be extended in general. However, if X is itself a Hilbert space, one has again an optimal result:

Theorem 3.2 (Kirszbraun, Valentine) *If X, Y are Hilbert spaces, $A \subset X$, and $f\colon A \to Y$ is λ-Lipschitz, then f has a λ-Lipschitz extension $\bar{f}\colon X \to Y$.*

See [Kirs1934], [Val1945], or [Fed1969, Theorem 2.10.43]. The following argument is essentially due to Mickle (1949). A generalization to metric spaces with curvature bounds was given in [LanS1997].

Proof (sketch): It suffices to prove the result for $\lambda = 1$.

Step I. First one shows that if $A \subset X$ is finite and $x \in X \setminus A$, and $f\colon A \to Y$ is 1-Lipschitz, then there is a 1-Lipschitz extension $f_x\colon A \cup \{x\} \to Y$ of f.

Suppose that $A = \{x_1, \dots, x_n\}$, and put $r_i := \|x_i - x\|$ and $y_i := f(x_i)$. The goal is to show that $\bigcap_{i=1}^n B(y_i, r_i) \neq \emptyset$. Clearly $C_t := \bigcap_{i=1}^n B(y_i, tr_i) \neq \emptyset$ for $t > 0$ sufficiently large. Put $s := \inf\{t > 0 : C_t \neq \emptyset\}$. Use the strict convexity of balls in

Y and completeness to prove that $\mathrm{diam}(C_t) \to 0$ as $t \to s+$ and that C_s consists of a single point, $C_s = \{y\}$. It then remains to show that $s \le 1$.

Put $u_i := x_i - x$ and $v_i := y_i - y$, and note that $\|u_i\| = r_i$ and $\|v_i\| \le sr_i$. Let $I := \{i : \|v_i\| = sr_i\}$. It follows from the choice of s that y is in the convex hull of $\{y_i : i \in I\}$, so 0 can be written as a convex combination $\sum_{i \in I} \lambda_i v_i$. Since $\|v_i - v_j\|^2 = \|y_i - y_j\|^2 \le \|x_i - x_j\|^2 = \|u_i - u_j\|^2$, we have

$$s^2 r_i^2 - 2\langle v_i, v_j\rangle + s^2 r_j^2 \le r_i^2 - 2\langle u_i, u_j\rangle + r_j^2$$

for all $i, j \in I$. Now multiply this inequality by $\lambda_i \lambda_j$ and sum over $i, j \in I$. Since $\sum_{i,j} \lambda_i \lambda_j \langle v_i, v_j\rangle = \left\|\sum_i \lambda_i v_i\right\|^2 = 0$ and $\sum_{i,j} \lambda_i \lambda_j \langle u_i, u_j\rangle = \left\|\sum_i \lambda_i u_i\right\|^2 \ge 0$, this gives

$$s^2 \sum_{i,j \in I} \lambda_i \lambda_j (r_i^2 + r_j^2) \le \sum_{i,j \in I} \lambda_i \lambda_j (r_i^2 + r_j^2),$$

showing that $s \le 1$.

Step II. If $\mathcal{B}$ is a family of closed balls in Y such that every finite subfamily has non-empty intersection, then also $\bigcap \mathcal{B} \neq \emptyset$. From this (well-known) property of Hilbert spaces and the result of Step I one concludes that if $A \subset X$ is arbitrary and $x \in X \setminus A$, then every 1-Lipschitz map $f : A \to Y$ has a 1-Lipschitz extension $f_x : A \cup \{x\} \to Y$.

Step III. The theorem now follows from the result of Step II by a standard application of Zorn's Lemma. □

The next result characterizes the extendability of partially defined Lipschitz maps from $\mathbb{R}^m$ into a complete metric space Y; it is useful in connection with the definition of rectifiable sets (Definition 10.1). We call a metric space Y *Lipschitz m-connected* if there is a constant $c \ge 1$ such that for $k \in \{0, \dots, m\}$, every λ-Lipschitz map $f : S^k \to Y$ admits a $c\lambda$-Lipschitz extension $\bar{f} : B^{k+1} \to Y$; here S^k and B^{k+1} denote the unit sphere and closed ball in $\mathbb{R}^{k+1}$, endowed with the induced metric. Every Banach space is Lipschitz m-connected for all $m \ge 0$. The sphere S^n is Lipschitz $(n-1)$-connected.

Theorem 3.3 (Lipschitz maps on $\mathbb{R}^m$) *Let Y be a complete metric space, and let $m \in \mathbb{N}$. Then the following statements are equivalent:*

(1) *Y is Lipschitz $(m-1)$-connected.*

(2) *There is a constant c such that every λ-Lipschitz map $f : A \to Y$, $A \subset \mathbb{R}^m$, has a $c\lambda$-Lipschitz extension $\bar{f} : \mathbb{R}^m \to Y$.*

The idea of the proof goes back to Whitney [Whit1934]. Compare [Alm1962, Theorem (1.2)] and [JohLS1986].

Proof: It is clear that (2) implies (1). Now suppose that (1) holds, and let $f : A \to Y$ be a λ-Lipschitz map, $A \subset \mathbb{R}^m$. As Y is complete, f extends canonically to the closure of A, with the same Lipschitz constant. Hence, assume A to be closed. A *dyadic cube* in $\mathbb{R}^m$ is a set of the form $x+[0,2^k]^m$ for some $k \in \mathbb{Z}$ and $x \in (2^k\mathbb{Z})^m$. Denote by $\mathcal{C}$ the family of all dyadic cubes $C \subset \mathbb{R}^m \setminus A$ that are maximal (with respect to inclusion) subject to the condition

$$\operatorname{diam}(C) \leq 2\, d(A, C).$$

The cubes in $\mathcal{C}$ have pairwise disjoint interiors and cover $\mathbb{R}^m \setminus A$. Moreover, they satisfy

$$d(A, C) < 2 \operatorname{diam}(C),$$

for if C' is the next bigger dyadic cube containing C, then $2\, d(A, C') < \operatorname{diam}(C') = 2 \operatorname{diam}(C)$ and $d(A, C) \leq d(A, C') + \operatorname{diam}(C)$.

Let $\Sigma^k \subset \mathbb{R}^m$ denote the k-skeleton of this cubical decomposition. Choose $\pi\colon A \cup \Sigma^0 \to A$ such that $d(x, \pi(x)) = d(x, A)$ for all $x \in A \cup \Sigma^0$. If $x \in \Sigma^0$ and $a \in A$, then

$$d(\pi(x), \pi(a)) = d(\pi(x), a) \leq d(x, \pi(x)) + d(x, a) \leq 2\, d(x, a).$$

If $x, x' \in \Sigma^0$ are distinct, and $C_x \in \mathcal{C}$ is a smallest cube containing x, then

$$d(\pi(x), x) \leq d(A, C_x) + \operatorname{diam}(C_x) \leq 3 \operatorname{diam}(C_x) \leq 3\sqrt{m}\, d(x, x'),$$

and likewise $d(\pi(x'), x') \leq 3\sqrt{m}\, d(x, x')$; thus

$$d(\pi(x), \pi(x')) \leq d(\pi(x), x) + d(x, x') + d(x', \pi(x')) \leq c_0\, d(x, x')$$

for $c_0 := 6\sqrt{m} + 1$. This shows that π is c_0-Lipschitz.

Extend f to a $c_0\lambda$-Lipschitz map $f_0\colon A \cup \Sigma^0 \to Y$ by putting $f_0 := f \circ \pi$. Since Y is Lipschitz 0-connected, f_0 can be extended to a map $f_1\colon A \cup \Sigma^1 \to Y$ whose restriction to any edge of a cube in $\mathcal{C}$ is $c_1\lambda$-Lipschitz for some constant c_1. It then follows easily that the restriction of f_1 to the (relative) boundary of any 2-dimensional face of a cube in $\mathcal{C}$ is $2c_1\lambda$-Lipschitz. Since Y is Lipschitz 1-connected, f_1 admits an extension $f_2\colon A \cup \Sigma^2 \to Y$ whose restriction to any 2-face of a cube in $\mathcal{C}$ is $c_2\lambda$-Lipschitz for some constant c_2. It follows that the restriction of f_2 to the boundary of any 3-face of a cube in $\mathcal{C}$ is $2c_2\lambda$-Lipschitz. By successively constructing extensions to $A \cup \Sigma^3, \dots, A \cup \Sigma^m = \mathbb{R}^m$, using that Y is Lipschitz $(m-1)$-connected, one arrives at an extension $\bar{f} := f_m\colon \mathbb{R}^m \to Y$ of f_0 whose restriction to any cube $C \in \mathcal{C}$ is $c_m\lambda$-Lipschitz for some constant $c_m \geq 1$. Since f_0 is continuous, and $\bar{f}|_A = f$ is λ-Lipschitz, it follows easily that $\bar{f}$ is in fact $c_m\lambda$-Lipschitz on $\mathbb{R}^m$. □

A metric space Y is an *absolute (C-)Lipschitz retract* if for any isometric embedding $i\colon Y \to Z$ into another metric space Z there is a (C-)Lipschitz retraction of Z onto $i(Y)$.

Proposition 3.4 *For a metric space Y, the following are equivalent:*
(1) *For every metric space X and every Lipschitz map $f\colon A \to Y$, $A \subset Y$, there is a Lipschitz extension $\bar{f}\colon X \to Y$.*
(2) *There exists a constant $C \geq 1$ such that for every metric space X and every Lipschitz map $f\colon A \to Y$, $A \subset Y$, there is a Lipschitz extension $\bar{f}\colon X \to Y$ with $\mathrm{Lip}(\bar{f}) \leq C\,\mathrm{Lip}(f)$.*
(3) *Y is an absolute Lipschitz retract.*
(4) *Y is an absolute C-Lipschitz retract for some $C \geq 1$.*

Proof: To be completed. □

Proposition 3.5 *Let X be a metric space. Every uniformly continuous and bounded function $f\colon X \to \mathbb{R}$ is a uniform limit of a sequence of Lipschitz functions.*

This is taken from [Hei2001, Theorem 6.8].

Proof: Let $\omega(\delta) := \sup\{|f(x) - f(y)| : d(x,y) \leq \delta\}$, $\delta \geq 0$, be the modulus of continuity of f. For $k \in \mathbb{N}$, define $f_k\colon X \to \mathbb{R}$ by

$$f_k(x) := \inf\{f(y) + k\,d(x,y) : y \in X\}.$$

Then $f(x) \geq f_k(x) \geq \inf f > -\infty$, and f_k is k-Lipschitz (compare the proof of Proposition 3.1). If $d(x,y) > \delta_k := 2\sup|f|/k$, then

$$f(y) + k\,d(x,y) > f(y) + 2\sup|f| \geq f(x) \geq f_k(x),$$

so in the definition of $f_k(x)$ it suffices to take the infimum over the closed ball $B(x,\delta_k)$. In particular $f_k(x) \geq \inf\{f(y) : y \in B(x,\delta_k)\}$ and hence

$$0 \leq f(x) - f_k(x) \leq \sup\{f(x) - f(y) : y \in B(x,\delta_k)\} \leq \omega(\delta_k)$$

for all $x \in X$. We conclude that $f_k \to f$ $(k \to \infty)$ uniformly on X. □

4 Differentiability of Lipschitz maps

Recall the following definitions.

Definition 4.1 (Gâteaux and Fréchet differential) Suppose that X, Y are Banach spaces, f maps an open set $U \subset X$ into Y, and $x \in U$.
(1) The map f is *Gâteaux differentiable* at x if the directional derivative

$$D_v f(x) = \lim_{t \to 0} \frac{f(x+tv) - f(x)}{t}$$

exists for every $v \in X$ and if there is a continuous linear map $L\colon X \to Y$ such that

$$L(v) = D_v f(x)$$

for all $v \in X$. Then L is the *Gâteaux differential* of f at x.

(2) The map f is *(Fréchet) differentiable* at x if there is a continuous linear map $L\colon X \to Y$ such that

$$\lim_{v\to 0} \frac{f(x+v) - f(x) - L(v)}{\|v\|} = 0.$$

Then $L =: df_x$ is the *(Fréchet) differential* of f at x.

The map f is Fréchet differentiable at x if and only if f is Gâteaux differentiable at x and the limit $L(u) = D_u f(x)$ exists uniformly for u in the unit sphere of X, i.e., for all $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|f(x+tu) - f(x) - tL(u)\| \le \epsilon|t|$$

whenever $|t| \le \delta$ and $u \in S(0,1) \subset X$.

Lemma 4.2 (differentiable Lipschitz maps) *Suppose that Y is a Banach space, $f\colon \mathbb{R}^m \to Y$ is Lipschitz, $x \in \mathbb{R}^m$, D is a dense subset of S^{m-1}, $D_u f(x)$ exists for every $u \in D$, $L\colon \mathbb{R}^m \to Y$ is linear, and $L(u) = D_u f(x)$ for all $u \in D$. Then f is Fréchet differentiable at x with differential $df_x = L$.*

In particular, if a *Lipschitz* map $f\colon \mathbb{R}^m \to Y$ is Gâteaux differentiable at x, then f is Fréchet differentiable at x.

Proof: Let $\epsilon > 0$. Choose a finite set $D' \subset D$ such that for every $u \in S^{m-1}$ there is a $u' \in D'$ with $\|u - u'\| \le \epsilon$. Then there is a $\delta > 0$ such that

$$\|f(x+tu') - f(x) - tL(u')\| \le \epsilon|t|$$

whenever $|t| \le \delta$ and $u' \in D'$. Given $u \in S^{m-1}$, pick $u' \in D'$ with $\|u - u'\| \le \epsilon$; then

$$\begin{aligned}&\|f(x+tu) - f(x) - tL(u)\| \\ &\quad \le \epsilon|t| + \|f(x+tu) - f(x+tu')\| + |t|\|L(u-u')\| \\ &\quad \le (1 + \mathrm{Lip}(f) + \|L\|)\epsilon|t|\end{aligned}$$

for $|t| \le \delta$. By the remark preceding the lemma, this gives the result. □

Theorem 4.3 (Rademacher) *Every Lipschitz map $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at $\mathcal{L}^m$-almost all points in $\mathbb{R}^m$.*

This was originally proved in [Rad1919].

Proof: It suffices to prove the theorem for $n = 1$; in the general case, $f = (f_1, \dots, f_n)$ is differentiable at x if and only if each f_i is differentiable at x.

In the case $m = 1$ the function $f\colon \mathbb{R} \to \mathbb{R}$ is absolutely continuous and hence $\mathcal{L}^1$-almost everywhere differentiable.

Now let $m \geq 2$. Fix $u \in S^{m-1}$ for the moment, and let B_u denote the Borel set of all $x \in \mathbb{R}^m$ where $D_u f(x)$ exists. Let H_u be the linear hyperplane orthogonal to u. For $x_0 \in H_u$, the function $t \mapsto f(x_0 + tu)$ is $\mathcal{L}^1$-almost everywhere differentiable by the result for $m = 1$, hence

$$\mathcal{L}^1((x_0 + \mathbb{R}u) \setminus B_u) = 0.$$

The characteristic function of $\mathbb{R}^m \setminus B_u$ is $\mathcal{L}^m$-measurable, and an application of Fubini's Theorem shows that $\mathcal{L}^m(\mathbb{R}^m \setminus B_u) = 0$.

Now choose a dense countable subset D of S^{m-1} containing the canonical basis vectors $e_1, \dots, e_m$, and put $B := \bigcap_{u \in D} B_u$. Then $\mathcal{L}^m(\mathbb{R}^m \setminus B) = 0$, and for every $x \in B$, $D_u f(x)$ exists for all $u \in D$; in particular, the formal gradient $\nabla f(x) := (D_{e_1} f(x), \dots, D_{e_m} f(x))$ exists. It now suffices to show that for $\mathcal{L}^m$-almost all $x \in B$, the usual relation

$$D_u f(x) = \langle \nabla f(x), u \rangle$$

holds for all $u \in D$. Since the right side is linear in u, the theorem then follows from Lemma 4.2.

Let $\phi \in C_c^\infty(\mathbb{R}^m)$. By Lebesgue's bounded convergence theorem,

$$\lim_{t \to 0} \int \frac{f(x + tu) - f(x)}{t} \phi(x)\, dx = \int D_u f(x) \phi(x)\, dx,$$
$$\lim_{t \to 0} \int f(x) \frac{\phi(x - tu) - \phi(x)}{t}\, dx = -\int f(x) D_u \phi(x)\, dx.$$

Substituting $x - tu$ for x in the term $f(x + tu)\phi(x)$ we see that the two left sides coincide. Hence,

$$\int D_u f(x) \phi(x)\, dx = -\int f(x) D_u \phi(x)\, dx.$$

This holds in particular for $e_1, \dots, e_m$, and by taking linear combinations of these identities we get

$$\int \langle \nabla f(x), u \rangle \phi(x)\, dx = -\int f(x) \langle \nabla \phi(x), u \rangle\, dx.$$

Now the right sides of the last two identities coincide. As $\phi \in C_c^\infty(\mathbb{R}^m)$ is arbitrary, it follows that $D_u f(x) = \langle \nabla f(x), u \rangle$ for $\mathcal{L}^m$-almost every $x \in B$, as desired. □

Theorem 4.4 (Stepanov) *Every map $f \colon \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at $\mathcal{L}^m$-almost all points in the set*

$$L(f) := \{x : \limsup_{y \to x} \|f(y) - f(x)\| / \|y - x\| < \infty\}.$$

This generalization of Rademacher's Theorem was proved in [Step1923]. The following elegant argument is due to Malý [Mal1999].

Proof: It suffices to consider the case $n = 1$. Let $(U_k)_{k\in\mathbb{N}}$ be the family of all open balls in $\mathbb{R}^m$ with center in $\mathbb{Q}^m$ and positive rational radius such that $f|_{U_k}$ is bounded. This family covers $L(f)$. Let $a_k\colon U_k \to \mathbb{R}$ be the supremum of all k-Lipschitz functions $\le f|_{U_k}$, and let $b_k\colon U_i \to \mathbb{R}$ be the infimum of all k-Lipschitz functions $\ge f|_{U_k}$. Note that a_k, b_k are k-Lipschitz and $a_k \le f|_{U_k} \le b_k$. Let

$$B_k := \{x \in U_k : \text{both } a_k \text{ and } b_k \text{ are differentiable at } x\}.$$

By Rademacher's Theorem, $Z := \bigcup_{i=1}^{\infty} U_k \setminus B_k$ has measure zero. Let $x \in L(f) \setminus Z$. We show that there exists an index k such that $x \in B_k$ and $a_k(x) = b_k(x)$; then f is differentiable at x. Since $x \in L(f)$, there is a radius $r > 0$ such that $\|f(y) - f(x)\| \le \lambda\|y - x\|$ for all $y \in B(x, r)$ and for some $\lambda \ge 0$ independent of y. Choose k such that $k \ge \lambda$ and $x \in U_k \subset B(x, r)$. Since $x \notin Z$, $x \in B_k$. By the definition of a_k and b_k, because $k \ge \lambda$,

$$f(x) - k\|y - x\| \le a_k(y) \le f(y) \le b_k(y) \le f(x) + k\|y - x\|$$

for all $y \in U_k$. For $y = x$ this gives $a_k(x) = b_k(x)$, as desired. □

The following simple example shows that Theorem 4.3 is not valid in general for Banach space valued maps (cf. [Fed1969, 2.9.23]).

Example 4.5 Consider the map $f\colon [0, 1] \to L_1([0, 1])$ that sends s to the characteristic function $\chi_{[0,s]}$ of $[0, s]$. For $0 \le s \le s + h \le 1$ we have

$$\|f(s + h) - f(s)\|_1 = \int_0^1 |\chi_{[0,s+h]} - \chi_{[0,s]}|\, d\mathcal{L}^1 = \int_0^1 \chi_{(s,s+h]}\, d\mathcal{L}^1 = h,$$

so f is an isometric embedding, in particular f is 1-Lipschitz. However, the difference quotients $\frac{1}{h}(f(s + h) - f(s)) = \frac{1}{h}\chi_{(s,s+h]}$ have integral 1, and they form no Cauchy family for $h \to 0$. Hence, f is nowhere differentiable.

For a Banach space Y, the property that every absolutely continuous function $f\colon [0, 1] \to Y$ is differentiable almost everywhere holds if and only if Y has the so-called *Radon–Nikodym property (RNP)*, and then also every Lipschitz map $f\colon \mathbb{R}^m \to Y$ is (Fréchet) differentiable almost everywhere. The definition of the RNP postulates the existence of a Radon–Nikodym derivative in $L_1(\mu, Y)$ (a Bochner integrable function) for every Y-valued measure $\nu\colon \mathcal{M} \to Y$ whose norm is absolutely continuous with respect to a probability measure space $(X, \mathcal{M}, \mu)$. A geometric characterization is as follows: Y has the RNP if and only if every closed convex set C in the unit ball has slices of arbitrarily small diameter. Here, a *slice* is a set of the form $S(C, \lambda, \alpha) := \{y \in C : \lambda(y) \ge \sup_{x\in C} \lambda(x) - \alpha\}$ for $\lambda \in Y^*$ and $\alpha > 0$. Every separable dual space and every reflexive Banach space has the RNP. For a detailed discussion, see [BenL2000].

5 Extension of smooth functions

We state Whitney's extension theorem for C^1 functions and some applications, compare [Whit1934], [Fed1969, Sect. 3.1], and [Sim2014, Ch. 2].

Theorem 5.1 (Whitney) *Suppose $f\colon A \to \mathbb{R}$ is a function on a closed set $A \subset \mathbb{R}^m$, $g\colon A \to \mathbb{R}^m$ is continuous, and for every compact set $K \subset A$ and every $\epsilon > 0$ there is a $\delta > 0$ such that*

$$|f(y) - f(x) - \langle g(x), y - x\rangle| \le \epsilon\|y - x\|$$

whenever $x, y \in K$ and $\|y - x\| \le \delta$. Then there exists a C^1 function $\bar{f}\colon \mathbb{R}^m \to \mathbb{R}$ with $\bar{f}|_A = f$ and $\nabla\bar{f}|_A = g$.

It follows from the assumptions that f is locally Lipschitz, hence continuous: if $x, y \in K$ and $\|y - x\| \le \delta$, then $|f(y) - f(x)| \le (\epsilon + \sup_{x\in K} \|g(x)\|)\, \|y - x\|$. Note also that if $\tilde{f}\colon \mathbb{R}^m \to \mathbb{R}$ is C^1, then $f := \tilde{f}|_A$ and $g := \nabla\tilde{f}|_A$ satisfy the assumptions of the theorem. The proof of Theorem 5.1 uses a decomposition of $\mathbb{R}^m \setminus A$ as for Theorem 3.3 and a C^1 partition of unity.

A first application is the following strong approximation result.

Theorem 5.2 (C^1 approximation of Lipschitz functions) *If $f\colon \mathbb{R}^m \to \mathbb{R}$ is Lipschitz and $\epsilon > 0$, then there is a C^1 function $\bar{f}\colon \mathbb{R}^m \to \mathbb{R}$ such that*

$$\mathcal{L}^m(\{x \in \mathbb{R}^m : f(x) \ne \bar{f}(x)\}) < \epsilon.$$

Proof: Let $\epsilon > 0$. By Rademacher's Theorem, there is a Borel set $B \subset \mathbb{R}^m$ with $\mathcal{L}^m(\mathbb{R}^m \setminus B) < \epsilon/2$ such that f is differentiable at every point in B, and $g := \nabla f\colon B \to \mathbb{R}^m$ is a measurable function. For $x \in B$ and $k \in \mathbb{N}$, let

$$r_k(x) := \sup\left\{\frac{|f(y) - f(x) - \langle g(x), y - x\rangle|}{\|y - x\|} : y \in B,\ 0 < \|y - x\| \le \frac{1}{k}\right\};$$

then $r_k \to 0$ pointwise on B as $k \to \infty$. Applying both Lusin's Theorem and Egorov's Theorem, we find a closed set $C \subset B$ with $\mathcal{L}^m(B \setminus C) < \epsilon/2$ such that $g|_C$ is continuous and $r_k \to 0$ uniformly on compact subsets of C. Now extend $f|_C$ to $\mathbb{R}^m$ by means of Theorem 5.1. □

For an another application of Theorem 5.1, recall that by Sard's Theorem, the set of critical values of a C^k function $f\colon \mathbb{R}^m \to \mathbb{R}^n$ with $0 \le m-n < k$ has $\mathcal{L}^n$-measure zero. In the case that $m = 2$ and $n = 1$, the assumption that f be twice continuously differentiable seems too strong but is necessary. Indeed, using the C^1 extension result, Whitney constructed an example of a C^1 function $f\colon \mathbb{R}^2 \to \mathbb{R}$ and an injective non-rectifiable curve $\gamma\colon [0, 1] \to \mathbb{R}^2$ with $\nabla f(\gamma(t)) = 0$ for all $t \in [0, 1]$, such that $f \circ \gamma\colon [0, 1] \to [0, 1]$ is monotonic and surjective (see [Whit1935]). Thus, in this example, there is whole interval of critical values.

6 Metric differentiability

Let $Y = (Y, d)$ be a metric space. Suppose $I \subset \mathbb{R}$ is an interval (i.e. a connected set) and $\gamma\colon I \to Y$ is a curve (i.e. a continuous map). The *length* of γ is the possibly infinite number

$$L(\gamma) := \sup \sum_{i=1}^{N} d(\gamma(t_{i-1}), \gamma(t_i)),$$

where the supremum is taken over all finite sequences $t_0 \le t_1 \le \cdots \le t_N$ in I. The curve γ is called *rectifiable* if $L(\gamma) < \infty$.

Theorem 6.1 (metric derivative) *Suppose that $a < b$, Y is a metric space, and $\gamma\colon [a, b] \to Y$ is Lipschitz. Then the limit*

$$|\dot\gamma|(t) := \lim_{h\to 0} \frac{d(\gamma(t + h), \gamma(t))}{|h|}$$

exists for $\mathcal{L}^1$-almost every $t \in [a, b]$, and

$$L(\gamma) = \int_a^b |\dot\gamma|(t)\, dt.$$

The proof below follows [AmbT2004, Theorem 4.1.1] (see also [Kir1994, Proposition 1] and [BurBI2001, Theorem 2.7.6]).

Proof: Choose a dense sequence $(y_k)_{k\in\mathbb{N}}$ in $\gamma([a, b])$ and define

$$\rho_k\colon [a, b] \to \mathbb{R}, \quad \rho_k(t) = d(\gamma(t), y_k).$$

By the triangle inequality, $|\rho_k(t+h) - \rho_k(t)| \le d(\gamma(t+h), \gamma(t))$ whenever $t, t+h \in [a, b]$, in particular $\mathrm{Lip}(\rho_k) \le \mathrm{Lip}(\gamma)$. Hence, for almost every $t \in [a, b]$, the derivative $\dot\rho_k(t)$ exists for all k, and

$$M(t) := \sup_k |\dot\rho_k(t)| \le \liminf_{h\to 0} \frac{d(\gamma(t + h), \gamma(t))}{|h|}.$$

On the other hand, if $(y_{k_n})_{n\in\mathbb{N}}$ is a subsequence tending to $\gamma(t)$, then

$$d(\gamma(t + h), \gamma(t)) = \lim_{n\to\infty} |\rho_{k_n}(t + h) - \rho_{k_n}(t)| \le \mathrm{sgn}(h) \int_t^{t+h} M(s)\, ds$$

and so

$$\limsup_{h\to 0} \frac{d(\gamma(t + h), \gamma(t))}{|h|} \le \limsup_{h\to 0} \frac{1}{h} \int_t^{t+h} M(s)\, ds.$$

It follows that for every Lebesgue point t of the measurable function $s \mapsto M(s)$, the limit $|\dot\gamma|(t)$ exists and equals $M(t)$. This proves the first part of the theorem.

The above argument also shows that

$$d(\gamma(t+h),\gamma(t)) \le \int_t^{t+h} |\dot\gamma|(s)\,ds$$

whenever $a \le t < t+h \le b$, which implies that $L(\gamma) \le \int_a^b |\dot\gamma|(t)\,dt$. For the reverse inequality, fix $\epsilon > 0$, and choose $N \ge 2$ such that $h := (b-a)/N \le \epsilon$. Put $t_i := a + ih$, for $i = 0, 1, \dots, N$. Then

$$\begin{aligned}\frac{1}{h}\int_a^{b-h} d(\gamma(t),\gamma(t+h))\,dt &= \frac{1}{h}\int_0^h \sum_{i=1}^{N-1} d(\gamma(s+t_{i-1}),\gamma(s+t_i))\,ds \\ &\le \frac{1}{h}\int_0^h L(\gamma)\,ds = L(\gamma).\end{aligned}$$

Using Fatou's lemma, we conclude that

$$\begin{aligned}\int_a^{b-\epsilon} |\dot\gamma|(t)\,dt &= \int_a^{b-\epsilon} \lim_{N\to\infty} \frac{d(\gamma(t+h),\gamma(t))}{h}\,dt \\ &\le \liminf_{N\to\infty} \frac{1}{h}\int_a^{b-\epsilon} d(\gamma(t+h),\gamma(t))\,dt \le L(\gamma).\end{aligned}$$

Letting $\epsilon \to 0$, we obtain $\int_a^b |\dot\gamma|(t)\,dt \le L(\gamma)$. □

We now consider functions $f \colon \mathbb{R}^m \to Y = (Y,d)$ for $m \ge 1$. To motivate the following definition, we first remark that if Y is a Banach space and f is differentiable at x, then it is also true that

$$\lim_{\|v\|+\|w\|\to 0} \frac{f(x+v) - f(x+w) - df_x(v-w)}{\|v\|+\|w\|} = 0.$$

Indeed, given $\epsilon > 0$, there is a $\delta > 0$ such that $\|f(x+v) - f(x) - df_x(v)\| \le \epsilon\|v\|$ whenever $\|v\| \le \delta$. Hence, if also $\|w\| \le \delta$, then

$$\begin{aligned}&\|f(x+v) - f(x+w) - df_x(v-w)\| \\ &\quad = \big\|\big(f(x+v) - f(x) - df_x(v)\big) - \big(f(x+w) - f(x) - df_x(w)\big)\big\| \\ &\quad \le \epsilon(\|v\| + \|w\|)\end{aligned}$$

by the triangle inequality.

Definition 6.2 (metric differentiability) Suppose Y is a metric space, $U \subset \mathbb{R}^m$ is an open set, and $x \in U$. A map $f \colon U \to Y$ is *metrically differentiable* at x if there exists a seminorm σ on $\mathbb{R}^m$ such that

$$\lim_{\|v\|+\|w\|\to 0} \frac{d(f(x+v), f(x+w)) - \sigma(v-w)}{\|v\|+\|w\|} = 0.$$

Then we call σ the *metric differential* of f at x and denote it by $\mathrm{md}\, f_x$.

It is clear that for every x there is at most one such seminorm. In particular, if Y is a Banach space and f is differentiable at x, then it follows from the preceding remark that f is metrically differentiable at x with metric differential $\mathrm{md}\, f_x = \|df_x(\cdot)\|$.

Theorem 6.3 (metric differentiability of Lipschitz maps) *Suppose that Y is a metric space, $U \subset \mathbb{R}^m$ is an open set, and $f \colon U \to Y$ is a Lipschitz map. Then f is metrically differentiable at $\mathcal{L}^m$-almost all points in U.*

See [Kir1994] and [KorS1993]. For the proof we use the following proposition (compare [Wen2008]).

Proposition 6.4 *Let $f \colon U \to Y$ be given as in Theorem 6.3, and let B be the set of all $x \in U$ with the property that the limit*

$$\sigma_x(v) := \lim_{t\to 0} \frac{d(f(x+tv), f(x))}{|t|}$$

exists for all $v \in \mathbb{R}^m$. The following holds.

(1) *$\mathcal{L}^m(U \setminus B) = 0$, and every $x \in B$, the function $\sigma_x \colon \mathbb{R}^m \to \mathbb{R}$ satisfies $\mathrm{Lip}(\sigma_x) \le \mathrm{Lip}(f)$ and $\sigma_x(sv) = |s|\, \sigma_x(v)$ for all $v \in \mathbb{R}^m$ and $s \in \mathbb{R}$.*

(2) *There exist compact sets $K_1, K_2, \ldots \subset B$ with $\mathcal{L}^m(B \setminus \bigcup_{j=1}^\infty K_j) = 0$ and the following property: for every j and every $\epsilon > 0$ there is a $\delta > 0$ such that*

$$|d(f(x+v), f(x+w)) - \sigma_x(v-w)| \le \epsilon\|v-w\|$$

whenever $v, w \in \mathbb{R}^m$, $\|v\|, \|w\| \le \delta$, and $x, x+w \in K_j$.

Proof: To prove (1), we choose a countable dense set $D \subset S^{m-1}$. For $x \in U$, $t \in \mathbb{R} \setminus \{0\}$, and $v \in \mathbb{R}^m$, put

$$q_{x,t}(v) := \frac{d(f(x+tv), f(x))}{|t|}.$$

Using Theorem 6.1, we conclude similarly as in the first part of the proof of Theorem 4.3 (Rademacher) that the set B of all $x \in U$ with the property that the limit $\sigma_x(u) = \lim_{t\to 0} q_{x,t}(u)$ exists for all $u \in D$ is a Borel set with $\mathcal{L}^m(U \setminus B) = 0$. Let $x \in B$. For fixed t, the function $q_{x,t} \colon \mathbb{R}^m \to \mathbb{R}$ is λ-Lipschitz, $\lambda := \mathrm{Lip}(f)$. It follows that $\sigma_x \colon D \to \mathbb{R}$ is λ-Lipschitz, and since D is dense, σ_x extends uniquely to a λ-Lipschitz function on S^{m-1}, which is then the uniform limit of $q_{x,t}|_{S^{m-1}}$ for $t \to 0$ (compare Lemma 4.2). The existence of the metric derivative $\sigma_x(u)$ also implies the existence of $\sigma_x(ru)$ for all $r \in \mathbb{R}$, and it holds that $\sigma_x(sv) = |s|\, \sigma_x(v)$ for all $v \in \mathbb{R}^m$ and $s \in \mathbb{R}$. Moreover, $\sigma_x \colon \mathbb{R}^m \to \mathbb{R}$ is λ-Lipschitz.

For (2), consider the map $\sigma \colon B \to C(S^{m-1})$, $x \mapsto \sigma_x|_{S^{m-1}}$, where $C(S^{m-1})$ denotes the space of continuous real-valued functions on S^{m-1}, endowed with the supremum norm $\|\cdot\|_\infty$. This space is separable, and σ is measurable. By Lusin's

Theorem there exist closed sets $C_1, C_2, \ldots \subset B$ such that $\mathcal{L}^m(B \setminus \bigcup_{k=1}^\infty C_k) = 0$ and $\sigma|_{C_k}$ is continuous for each k. For $y \in B$ and $i \in \mathbb{N}$, let

$$r_i(y) := \sup_{0<t\le 1/i} \sup_{\|u\|=1} \left|q_{y,t}(u) - \sigma_y(u)\right|.$$

From the proof of (1) we know that $r_i(y) \to 0$ $(i \to \infty)$ for every $y \in B$. Using Egorov's Theorem we find compact sets $K_1, K_2, \ldots \subset B$ with $\mathcal{L}^m(B \setminus \bigcup_{j=1}^\infty K_j) = 0$ such that each K_j is contained in some C_k (hence $\sigma|_{K_j}$ is uniformly continuous) and $r_i \to 0$ $(i \to \infty)$ uniformly on each K_j. Now let $j \in \mathbb{N}$ and $\epsilon > 0$. Then there is an i such that

$$\sup_{\|u\|=1} |\sigma_x(u) - \sigma_y(u)| = \|\sigma_x - \sigma_y\|_\infty \le \frac{\epsilon}{2}, \quad r_i(y) \le \frac{\epsilon}{2}$$

whenever $x, y \in K_j$ and $\|x - y\| \le \delta := 1/(2i)$. Given $x \in K_j$ and $v, w \in \mathbb{R}^m$ with $\|v\|, \|w\| \le \delta$, $v \ne w$, and $y := x + w \in K_j$, put $t := \|v - w\|$ and $u := (1/t)(v - w)$. Then it follows that $0 < t \le \|v\| + \|w\| \le 2\delta = 1/i$ and

$$\begin{aligned}
&\left|d(f(x+v), f(y)) - \sigma_x(v-w)\right| \\
&\quad \le \left|d(f(y+v-w), f(y)) - \sigma_y(v-w)\right| + \left|\sigma_x(v-w) - \sigma_y(v-w)\right| \\
&\quad = t\left|q_{y,t}(u) - \sigma_y(u)\right| + t|\sigma_x(u) - \sigma_y(u)| \\
&\quad \le \epsilon t.
\end{aligned}$$

As $y = x + w$ and $t = \|v - w\|$, this completes the proof. □

Proof of Theorem 6.3: Let compact sets $K_1, K_2, \ldots \subset U$ be given as in Proposition 6.4. Suppose $x \in K_j$ is a point with Lebesgue density 1, i.e.,

$$\Theta^m(K_j, x) := \lim_{r \to 0+} \frac{\mathcal{L}^m(K_j \cap B(x,r))}{\mathcal{L}^m(B(x,r))} = 1.$$

Let $\epsilon > 0$, and let $\delta = \delta(j, \epsilon) > 0$ be given as in Proposition 6.4. By adjusting δ if necessary, we arrange that for every $w \in \mathbb{R}^m$ with $\|w\| \le \delta$ there exists a $w' = w'(w)$ such that $x + w' \in K_j$, $\|w'\| \le \|w\|$ and $\|w - w'\| \le \epsilon\|w\|$. Suppose now that $v, w \in \mathbb{R}^n$, $\|v\|, \|w\| \le \delta$, and $w' = w'(w)$. Using Proposition 6.4 we conclude that

$$\begin{aligned}
&\left|d(f(x+v), f(x+w)) - \sigma_x(v-w)\right| \\
&\quad \le \left|d(f(x+v), f(x+w')) - \sigma_x(v-w')\right| \\
&\qquad + d(f(x+w), f(x+w')) + \left|\sigma_x(v-w) - \sigma_x(v-w')\right| \\
&\quad \le \epsilon\|v - w'\| + 2\,\mathrm{Lip}(f)\|w - w'\| \\
&\quad \le \epsilon(\|v\| + \|w\|) + 2\epsilon\,\mathrm{Lip}(f)\|w\| \\
&\quad \le \epsilon(1 + 2\,\mathrm{Lip}(f))(\|v\| + \|w\|).
\end{aligned}$$

Since almost every point in U is a density point of some K_j, this shows that

$$\lim_{\|v\|+\|w\|\to 0} \frac{d(f(x+v), f(x+w)) - \sigma_x(v-w)}{\|v\| + \|w\|} = 0$$

for almost every $x \in U$.

It remains to prove that σ_x satisfies the triangle inequality. For $v, w \in \mathbb{R}^m$,

$$\begin{aligned}\sigma_x(v+w) &= \lim_{t\to 0+} \frac{d(f(x+tv), f(x-tw))}{t} \\ &\le \lim_{t\to 0+} \frac{d(f(x+tv), f(x))}{t} + \lim_{t\to 0+} \frac{d(f(x-tw), f(x))}{t} \\ &= \sigma_x(v) + \sigma_x(-w).\end{aligned}$$

Since $\sigma_x(-w) = \sigma_x(w)$, the proof is complete. □

7 Hausdorff measures

For every real number $s \ge 0$, we put

$$\alpha_s := \frac{\pi^{s/2}}{\Gamma(\frac{s}{2}+1)},$$

where $\Gamma\colon (0,\infty) \to \mathbb{R}$ is the usual gamma function, $\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}\,dx$. Note that $\Gamma(1) = 1$, $\Gamma(s+1) = s\,\Gamma(s)$ for $s > 0$, and $\Gamma(1/2) = \sqrt{\pi}$. For $m \in \mathbb{N}$, α_m equals the Lebesgue measure of the unit ball in $\mathbb{R}^m$,

$$\alpha_m = \mathcal{L}^m(B(0,1)) = \begin{cases} \dfrac{2^m \pi^{(m-1)/2} (\frac{m-1}{2})!}{m!} & \text{if } m \text{ is odd,} \\ \dfrac{\pi^{m/2}}{(\frac{m}{2})!} & \text{if } m \text{ is even.} \end{cases}$$

We recall the definition of the Hausdorff measures.

Definition 7.1 Let X be a metric space. For $s \ge 0$, $0 < \delta \le \infty$, and $A \subset X$, define

$$\mathcal{H}^s_\delta(A) := \inf \sum_{i=1}^\infty \alpha_s \big(\tfrac{1}{2} \operatorname{diam}(C_i)\big)^s,$$

where the infimum is taken over all coverings $(C_i)_{i\in\mathbb{N}}$ of A with $\operatorname{diam}(C_i) \le \delta$ for all i. (Here the conventions $\operatorname{diam}(\emptyset)^s = 0$, $0^0 = 1$ are used.) Then

$$\mathcal{H}^s(A) := \lim_{\delta\to 0+} \mathcal{H}^s_\delta(A) = \sup_{\delta>0} \mathcal{H}^s_\delta(A)$$

is the *s-dimensional Hausdorff measure* of the set A.

For every s, $\mathcal{H}^s$ is a Borel regular metric outer measure on X. As we will show below, with the chosen normalization, $\mathcal{H}^m = \mathcal{L}^m$ on $\mathbb{R}^m$. Whenever $A \subset X$ and $f\colon A \to Y$ is a Lipschitz map into another metric space Y, then

$$\mathcal{H}^s(f(A)) \le \mathrm{Lip}(f)^s\, \mathcal{H}^m(A).$$

Suppose X is a metric space, $A \subset X$, and $x \in X$. Then the s-dimensional *upper density* and *lower density* of A at x are defined by

$$\Theta^{*s}(A,x) = \limsup_{r\to 0+} \frac{\mathcal{H}^s(A \cap B(x,r))}{\alpha_s r^s},$$
$$\Theta^{s}_{*}(A,x) = \liminf_{r\to 0+} \frac{\mathcal{H}^s(A \cap B(x,r))}{\alpha_s r^s},$$

respectively. If the two coincide, then the common value $\Theta^s(A,x)$ is the s-dimensional *density* of A at x.

Theorem 7.2 (densities) *Let A be a subset of X with $\mathcal{H}^s(A) < \infty$.*
(1) *For $\mathcal{H}^s$-almost every $x \in A$, $2^{-s} \le \Theta^{*m}(A,x) \le 1$.*
(2) *If A is $\mathcal{H}^s$-measurable, then $\Theta^s(A,x) = 0$ for $\mathcal{H}^s$-almost every $x \in X \setminus A$.*
(3) *If both A and $B \subset A$ are $\mathcal{H}^s$-measurable, then $\Theta^{*s}(A,x) = \Theta^{*s}(B,x)$ and $\Theta^s_*(A,x) = \Theta^s_*(B,x)$ for $\mathcal{H}^s$-almost every $x \in B$.*

Proof: For the proof of (1) and (2) we refer to [Sim2014, Ch. 1, Theorem 3.6 and Theorem 3.26] (see also [Mat1995, Theorem 6.2 and Corollary 6.3]).

(3) follows from (2): $C := A \setminus B$ is $\mathcal{H}^s$-measurable with $\mathcal{H}^s(C) < \infty$, so $\Theta^s(C,x) = 0$ for $\mathcal{H}^s$-almost every $x \in B \subset X \setminus C$, and hence $\Theta^{*s}(A,x) = \Theta^{*s}(B,x) + \Theta^s(C,x) = \Theta^{*s}(B,x)$ and $\Theta^s_*(A,x) = \Theta^s_*(B,x) + \Theta^s(C,x) = \Theta^s_*(B,x)$. □

Theorem 7.3 (isodiametric inequality) *Let ρ be a norm on $\mathbb{R}^m$ with unit ball B_ρ. For a non-empty set $C \subset \mathbb{R}^m$, let $\mathrm{conv}(C)$ denote the convex hull of C and $\mathrm{diam}_\rho(C)$ the diameter of C with respect to ρ. Then*

$$\mathcal{L}^m(\mathrm{conv}(C)) \le \mathcal{L}^m(B_\rho)\,\big(\tfrac{1}{2}\,\mathrm{diam}_\rho(C)\big)^m.$$

In particular, for the standard Euclidean norm on $\mathbb{R}^m$, $\mathcal{L}^m(\mathrm{conv}(C)) \le \alpha_m\big(\tfrac{1}{2}\,\mathrm{diam}(C)\big)^m$.

Proof: It suffices to prove the inequality for a compact convex body C. The idea is to show that there exists a centrally symmetric convex body C^* such that $\mathcal{L}^m(C) \le \mathcal{L}^m(C^*)$ and $\mathrm{diam}_\rho(C^*) \le \mathrm{diam}_\rho(C)$. Then $C^* \subset B_\rho\big(0, \tfrac{1}{2}\,\mathrm{diam}_\rho(C^*)\big)$, hence $\mathcal{L}^m(C^*) \le \mathcal{L}^m(B_\rho)\,\big(\tfrac{1}{2}\,\mathrm{diam}_\rho(C^*)\big)^m$, and the desired inequality for C follows. As for C^*, one can put $C' := -C$ and $C^* := \frac{1}{2}(C + C')$. Then, by (a special case of) the Brunn-Minkowski inequality,

$$\mathcal{L}^m(C^*)^{1/m} \geq \mathcal{L}^m\big(\tfrac{1}{2}C\big)^{1/m} + \mathcal{L}^m\big(\tfrac{1}{2}C'\big)^{1/m} = \mathcal{L}^m(C)^{1/m}.$$

Furthermore, for every pair of points $x^* = \frac{1}{2}(x + x')$ and $y^* = \frac{1}{2}(y + y')$ in C^*, with $x, y \in C$ and $x', y' \in C'$,

$$\rho(x^* - y^*) = \frac{1}{2}\rho(x - y + x' - y') \leq \frac{1}{2}\big(\rho(x - y) + \rho(x' - y')\big),$$

thus $\mathrm{diam}_\rho(C^*) \leq \mathrm{diam}_\rho(C)$. See [BurZ1988, Theorem 11.2.1]. □

Theorem 7.4 *Let ρ be a norm on $\mathbb{R}^m$, and let $\mathcal{H}^m_\rho$ denote the Hausdorff measure with respect to the corresponding metric. Then $\mathcal{H}^m_\rho(B_\rho) = \alpha_m$. In particular, with respect to the standard Euclidean norm, $\mathcal{H}^m = \mathcal{L}^m$.*

Compare [Kir1994, Lemma 6].

Proof: The inequality $\mathcal{H}^m_\rho(B_\rho) \leq \alpha_m$ follows from the fact that the quotient $\mathcal{H}^m_\rho(B_\rho(x, r))/r^m$ is constant for all $x \in \mathbb{R}^m$ and $r > 0$ and, hence, is less than or equal to α_m since $\Theta^{*m}(\mathbb{R}^m_\rho, x) \leq 1$ by Theorem 7.2. The reverse inequality is a consequence Theorem 7.3, which implies that for every covering $(C_i)_{i\in\mathbb{N}}$ of B_ρ,

$$\sum_i \big(\tfrac{1}{2}\,\mathrm{diam}_\rho(C_i)\big)^m \geq \sum_i \frac{\mathcal{L}^m(C_i)}{\mathcal{L}^m(B_\rho)} \geq 1,$$

thus $\mathcal{H}^m_\rho(B_\rho) \geq \alpha_m$. □

8 Area formula

The next goal is to prove Theorem 8.3 below. We start with a technical lemma, compare [Kir1994, Lemma 4].

Lemma 8.1 (Borel partition) *Suppose that Y is a metric space, $f\colon \mathbb{R}^m \to Y$ is Lipschitz, and B is the Borel set of all x where f is metrically differentiable and* md f_x *is a norm. Let $\lambda > 1$. Then there exist a Borel partition $(B_i)_{i\in\mathbb{N}}$ of B and a sequence of norms ρ_i on $\mathbb{R}^m$ such that*

$$\lambda^{-1}\rho_i(x - x') \leq d(f(x), f(x')) \leq \lambda\,\rho_i(x - x'),$$
$$\lambda^{-1}\rho_i(v) \leq \mathrm{md}\, f_x(v) \leq \lambda\,\rho_i(v)$$

for all $x, x' \in B_i$ and $v \in \mathbb{R}^m$.

In the 'classical' case, when $Y = \mathbb{R}^n$ and B is the Borel set of all x where f is differentiable and df_x has rank m, all norms ρ_i may be chosen to be Euclidean, that is, induced by an inner product, compare [Fed1969, Lemma 3.2.2] and [EvaG1992, p. 94].

Proof: Choose a sequence of norms ρ_i on $\mathbb{R}^m$ such that for every norm ρ on $\mathbb{R}^m$ and for every $c > 1$ there is an $i \in \mathbb{N}$ such that

$$c^{-1}\rho_i(v) \leq \rho(v) \leq c\,\rho_i(v)$$

for all $v \in \mathbb{R}^m$. Given $\lambda > 1$, pick $\delta > 0$ such that $\lambda^{-1} + \delta < 1 < \lambda - \delta$. For $i, k \in \mathbb{N}$, denote by $B_{i,k}$ the Borel set of all $x \in B$ such that

(i) $(\lambda^{-1} + \delta)\,\rho_i(v) \leq \operatorname{md} f_x(v) \leq (\lambda - \delta)\,\rho_i(v)$ for $v \in \mathbb{R}^m$,

(ii) $|d(f(x+v), f(x)) - \operatorname{md} f_x(v)| \leq \delta\,\rho_i(v)$ for $\|v\| \leq 1/k$.

The sets $B_{i,k}$ cover B: given $x \in B$, choose $i \in \mathbb{N}$ such that (i) holds, let $c_i > 0$ be such that $\|v\| \leq c_i\,\rho_i(v)$ for all $v \in \mathbb{R}^m$, and pick $k \in \mathbb{N}$ such that (ii) holds with $(\delta/c_i)\|v\|$ in place of $\delta\,\rho_i(v)$; then $x \in B_{i,k}$. Now if $C \subset B_{i,k}$ is a set with $\operatorname{diam} C \leq 1/k$, then

$$\begin{aligned} d(f(x+v), f(x)) &\leq \operatorname{md} f_x(v) + \delta\,\rho_i(v) \leq \lambda\,\rho_i(v), \\ d(f(x+v), f(x)) &\geq \operatorname{md} f_x(v) - \delta\,\rho_i(v) \geq \lambda^{-1}\rho_i(v) \end{aligned}$$

whenever $x, x+v \in C$. By subdividing and relabeling the sets $B_{i,k}$ appropriately we obtain the result. □

Definition 8.2 (jacobian) (1) Suppose that X, Y are normed spaces, $\dim X = m \in \mathbb{N}$, and $L\colon X \to Y$ is linear. The *jacobian* $\mathbf{J}(L) \in [0, \infty)$ of L is the (unique) number satisfying

$$\mathcal{H}^m(L(A)) = \mathbf{J}(L)\,\mathcal{H}^m(A)$$

for all $A \subset X$.

(2) If σ is a seminorm on $\mathbb{R}^m$, we define the *jacobian* $\mathbf{J}(\sigma)$ of σ as the number satisfying

$$\mathcal{H}^m_\sigma(A) = \mathbf{J}(\sigma)\,\mathcal{L}^m(A)$$

for all $A \subset \mathbb{R}^m$ in case σ is a norm and $\mathbf{J}(\sigma) = 0$ otherwise.

We remark that if $A \subset \mathbb{R}^m$ is $\mathcal{L}^m$-measurable, and $f\colon A \to Y$ is a Lipschitz map into a metric space Y, then $f(A)$ is $\mathcal{H}^m$-measurable. This is because A can be written as the union of countably many compact sets and a set of measure zero, thus the same is true for $f(A)$.

Theorem 8.3 (area formula) *Suppose that Y is a metric space and $f\colon \mathbb{R}^m \to Y$ is Lipschitz.*

(1) *If* $A \subset \mathbb{R}^m$ *is* $\mathcal{L}^m$*-measurable, then*

$$\int_A \mathbf{J}(\mathrm{md}\, f_x)\, dx = \int_Y \#(f^{-1}\{y\} \cap A)\, d\mathcal{H}^m(y).$$

(2) *If* g *is a real-valued* $\mathcal{L}^m$*-integrable function on* $\mathbb{R}^m$*, then*

$$\int_{\mathbb{R}^m} g(x)\, \mathbf{J}(\mathrm{md}\, f_x)\, dx = \int_Y \sum_{x \in f^{-1}\{y\}} g(x)\, d\mathcal{H}^m(y).$$

See [Kir1994]. Note that $\# = \mathcal{H}^0$. In the case $Y = \mathbb{R}^n$ we obtain the classical area formula as $\mathbf{J}(\mathrm{md}\, f_x)$ coincides with $\mathbf{J}(df_x)$ for $\mathcal{L}^m$-almost every $x \in \mathbb{R}^m$. Compare [Fed1969, Theorem 3.2.3], [EvaG1992, Sect. 3.3]. That formula says, in particular, that the differential geometric volume of an injective C^1 immersion $f\colon U \to \mathbb{R}^n$, U an open subset of $\mathbb{R}^m$, equals $\mathcal{H}^m(f(U))$.

Proof: For (1), we may partition A into countably many measurable sets and prove the respective formula for each of these sets separately. In particular, we lose no generality in assuming $\mathcal{L}^m(A) < \infty$. Let A_0 denote the set of all $x \in A$ where f is not metrically differentiable. Then

$$\mathcal{H}^m(f(A_0)) \le \mathrm{Lip}(f)^m \mathcal{L}^m(A_0) = 0$$

by Theorem 6.3, thus A_0 does not contribute to either side of the claimed identity. Now we split $A \setminus A_0$ into the two sets A', A'', where A' consists of all x for which $\mathrm{md}\, f_x$ is a norm, that is, $\mathbf{J}(\mathrm{md}\, f_x) > 0$.

First we consider A'. Let $\lambda > 1$. Using Lemma 8.1 we find a measurable partition $(A_i)_{i \in \mathbb{N}}$ of A' and norms ρ_i on $\mathbb{R}^m$ such that $f|_{A_i}$ is injective,

$$\lambda^{-m}\, \mathcal{H}^m_{\rho_i}(A_i) \le \mathcal{H}^m(f(A_i)) \le \lambda^m\, \mathcal{H}^m_{\rho_i}(A_i),$$

and $\lambda^{-1}\rho_i \le \mathrm{md}\, f_x \le \lambda\, \rho_i$ for all $x \in A_i$. This last assertion yields

$$\lambda^{-m}\, \mathbf{J}(\rho_i) \le \mathbf{J}(\mathrm{md}\, f_x) \le \lambda^m\, \mathbf{J}(\rho_i)$$

for all $x \in A_i$. Since $\mathbf{J}(\rho_i)\, \mathcal{L}^m(A_i) = \mathcal{H}^m_{\rho_i}(A_i)$, it follows that

$$\lambda^{-m}\, \mathcal{H}^m_{\rho_i}(A_i) \le \int_{A_i} \mathbf{J}(\mathrm{md}\, f_x)\, dx \le \lambda^m\, \mathcal{H}^m_{\rho_i}(A_i)$$

and

$$\lambda^{-2m}\, \mathcal{H}^m(f(A_i)) \le \int_{A_i} \mathbf{J}(\mathrm{md}\, f_x)\, dx \le \lambda^{2m}\, \mathcal{H}^m(f(A_i)).$$

Since $f|_{A_i}$ is injective and $f(A_i)$ is $\mathcal{H}^m$-measurable, $\mathcal{H}^m(f(A_i))$ can be written as $\int_Y \#(f^{-1}\{y\} \cap A_i)\, d\mathcal{H}^m(y)$. Then, summing over i, we get that

$$\begin{aligned}\lambda^{-2m} \int_Y \#\big(f^{-1}\{y\} \cap A'\big)\, d\mathcal{H}^m(y) &\le \int_{A'} \mathbf{J}(\mathrm{md}\, f_x)\, dx \\ &\le \lambda^{2m} \int_Y \#\big(f^{-1}\{y\} \cap A'\big)\, d\mathcal{H}^m(y).\end{aligned}$$

As $\lambda > 1$ was arbitrary, this shows that (1) holds for A'.

Now we turn to the set A'' of all $x \in A$ where $\mathbf{J}(\mathrm{md}\, f_x) = 0$. We prove that $\mathcal{H}^m(f(A'')) = 0$; thus either side of the claimed identity is zero for A''. Let $\epsilon \in (0,1)$, and define

$$h\colon \mathbb{R}^m \to Y \times \mathbb{R}^m, \quad h(x) = (f(x), \epsilon x).$$

Equip $Y \times \mathbb{R}^m$ with the l^1 product metric

$$d_1((y,z),(y',z')) = d(y,y') + \|z - z'\|.$$

Clearly, h is metrically differentiable at x whenever f is, and

$$\mathrm{md}\, h_x(v) = \mathrm{md}\, f_x(v) + \epsilon\|v\|$$

for all $v \in \mathbb{R}^m$. In particular, $\mathrm{md}\, h_x$ is a norm on $\mathbb{R}^m$ for every $x \in A''$. Fix $x \in A''$ for the moment, and put $\rho := \mathrm{md}\, h_x$. Consider the norm ball $B_\rho := B_\rho(0,1) \subset \mathbb{R}^m$. Since $\mathrm{md}\, f_x$ is not a norm, there is a $v_0 \in \mathbb{R}^m$ with $\rho(v_0) = 1$ and $\mathrm{md}\, f_x(v_0) = 0$, thus $\|v_0\| = 1/\epsilon$. Moreover, if $\rho(v) = 1$, then

$$1 = \mathrm{md}\, f_x(v) + \epsilon\|v\| \le (\mathrm{Lip}(f) + 1)\|v\|,$$

thus $\|v\| \ge r := 1/(\mathrm{Lip}(f)+1)$. Hence, B_ρ contains the convex hull of $\{v_0, -v_0\} \cup B(0,r)$, where $\|v_0\| = 1/\epsilon$. It follows that $\mathcal{L}^m(B_\rho) \ge c_m r^{m-1}/\epsilon$ for some constant c_m depending only on m, hence

$$\mathbf{J}(\mathrm{md}\, h_x) = \mathbf{J}(\rho) = \frac{\mathcal{H}^m_\rho(B_\rho)}{\mathcal{L}^m(B_\rho)} \le \frac{\epsilon\, \alpha_m}{c_m r^{m-1}}$$

(recall Theorem 7.4). Applying the above result for (f, A') to (h, A''), we get that

$$\mathcal{H}^m(h(A'')) = \int_{A''} \mathbf{J}(\mathrm{md}\, h_x)\, dx \le \frac{\epsilon\, \alpha_m}{c_m r^{m-1}} \mathcal{L}^m(A'').$$

Since the canonical projection $Y \times \mathbb{R}^m \to Y$ is 1-Lipschitz and maps $h(A'')$ to $f(A'')$, we have $\mathcal{H}^m(f(A'')) \le \mathcal{H}^m(h(A''))$, and letting ϵ tend to 0 we conclude that $\mathcal{H}^m(f(A'')) = 0$. This completes the proof of (1).

Finally, (2) follows from (1), by approximating g by simple functions. □

9 Coarea formula

For the proof of the coarea formula, Theorem 9.4, we need the following general coarea inequality, which is also of independent interest.

Theorem 9.1 (coarea inequality) *Suppose that X, Y are metric spaces, $f\colon X \to Y$ is a Lipschitz map, $A \subset X$, and $m, k \ge 0$ are real numbers. Then*

$$\int_Y^* \mathcal{H}^k(f^{-1}\{y\} \cap A)\, d\mathcal{H}^m(y) \le \frac{\alpha_m \alpha_k}{\alpha_{m+k}} \mathrm{Lip}(f)^m\, \mathcal{H}^{m+k}(A).$$

Here $\int^*$ denotes the upper integral; in general, the integrand $y \mapsto \mathcal{H}^k(f^{-1}\{y\} \cap A)$ is not $\mathcal{H}^m$-measurable. However, if X is proper (that is, closed bounded subsets of X are compact) and A is $\mathcal{H}^{m+k}$-measurable with $\mathcal{H}^{m+k}(A) < \infty$, then $f^{-1}\{y\} \cap A$ is $\mathcal{H}^k$-measurable for $\mathcal{H}^m$-almost every y and $y \mapsto \mathcal{H}^k(f^{-1}\{y\} \cap A)$ is $\mathcal{H}^m$-measurable, compare [Fed1969, 2.10.26]. Theorem 9.1 is stated with some additional assumptions in [Fed1969, Theorem 2.10.25]. As remarked in [Dav1970, p. 236], these are superfluous. See also [Rei2009] and [EsmH2021].

We prove Theorem 9.1 in the case that Y is an m-dimensional normed space $(\mathbb{R}^m, \rho)$, so we will write $\mathcal{H}^m_\rho$ for the Hausdorff measure on Y.

Proof: Assume that $\mathcal{H}^{m+k}(A) < \infty$. Let $i \in \mathbb{N}$ and $\delta \in (0, 1/i]$. Choose a countable δ-bounded covering $\mathcal{C}$ of A such that

$$\sum_{C \in \mathcal{C}} \alpha_{k+m} \big(\tfrac{1}{2} \operatorname{diam}(C)\big)^{m+k} \le \mathcal{H}^{m+k}_\delta(A) + \delta.$$

For $C \in \mathcal{C}$, let D_C denote the ($\mathcal{H}^m_\rho$-measurable) closure of $f(C)$, and define $g_C : Y \to \mathbb{R}$ as $\alpha_k\big(\tfrac{1}{2} \operatorname{diam}(C)\big)^k$ times the characteristic function of D_C. For every $y \in Y$, the sets $C \in \mathcal{C}$ with $y \in D_C$ cover $f^{-1}\{y\} \cap A$, thus

$$\mathcal{H}^k_{1/i}\big(f^{-1}\{y\} \cap A\big) \le \sum_{C \in \mathcal{C}} g_C(y).$$

It follows that

$$\begin{aligned}\int^*_Y \mathcal{H}^k_{1/i}\big(f^{-1}\{y\} \cap A\big)\, d\mathcal{H}^m_\rho(y) &\le \int_Y \sum_{C \in \mathcal{C}} g_C(y)\, d\mathcal{H}^m_\rho(y) \\ &= \sum_{C \in \mathcal{C}} \int_Y g_C(y)\, d\mathcal{H}^m_\rho(y) \\ &= \sum_{C \in \mathcal{C}} \alpha_k \big(\tfrac{1}{2} \operatorname{diam}(C)\big)^k\, \mathcal{H}^m_\rho(D_C).\end{aligned}$$

By Theorems 7.3 and 7.4, and since f is Lipschitz,

$$\begin{aligned}\mathcal{H}^m_\rho(D_C) &\le \alpha_m \big(\tfrac{1}{2} \operatorname{diam}(D_C)\big)^m \\ &\le \alpha_m \operatorname{Lip}(f)^m \,\big(\tfrac{1}{2} \operatorname{diam}(C)\big)^m.\end{aligned}$$

We conclude that

$$\int^*_Y \mathcal{H}^k_{1/i}\big(f^{-1}\{y\} \cap A\big)\, d\mathcal{H}^m_\rho(y) \le \frac{\alpha_m \alpha_k}{\alpha_{m+k}} \operatorname{Lip}(f)^m \,\big(\mathcal{H}^{m+k}_\delta(A) + \delta\big).$$

Now we let first $\delta \to 0$, then $i \to \infty$. □

Lemma 9.2 (factorization) *Suppose that $f : \mathbb{R}^n \to \mathbb{R}^m$ is a Lipschitz map, where $n \ge m$, and $A \subset \mathbb{R}^n$ is an $\mathcal{L}^n$-measurable set such that df_x exists and has rank m*

for all $x \in A$. Let $\lambda > 1$. Then there exist countable families of $\mathcal{L}^n$-measurable sets $A_i \subset A$, Lipschitz maps $h_i\colon \mathbb{R}^n \to \mathbb{R}^n$, and linear maps $L_i\colon \mathbb{R}^n \to \mathbb{R}^m$ such that $\mathcal{L}^n(A \setminus \bigcup_i A_i) = 0$, $h_i|_{A_i}$ is λ-bi-Lipschitz,

$$f = L_i \circ h_i,$$

and for all $x \in A_i$, $d(h_i)_x$ exists and is λ-bi-Lipschitz.

Proof: It is easy to see that for every $x \in \mathbb{R}^n$ where df_x exists and has rank m there is a coordinate projection $p\colon \mathbb{R}^n \to \mathbb{R}^{n-m}$ such that du_x has rank n, where

$$u\colon \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^{n-m}, \quad u(x) = (f(x), p(x)).$$

According to the possible choices of p we obtain a covering of A by finitely many measurable subsets. For the proof of the lemma it suffices to consider a single such subset which, for simplicity, we denote again by A. Thus, we assume that there is a fixed projection p as above such that du_x has rank n for every $x \in A$. By applying Lemma 8.1 to u, we find a measurable partition $(B_i)_{i\in\mathbb{N}}$ of A such that each $u|_{B_i}$ is bi-Lipschitz. For every i, choose a Lipschitz extension

$$v_i\colon \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^n$$

of $(u|_{B_i})^{-1}$, thus $v_i(u(x)) = x$ for all $x \in B_i$. There is a measurable set $C_i \subset u(B_i)$ with $\mathcal{L}^n(u(B_i) \setminus C_i) = 0$ such that v_i is differentiable at every point $u(x) \in C_i$, and $d(v_i)_{u(x)} \circ du_x = \mathrm{id}_{\mathbb{R}^n}$. Let $\lambda > 1$. Applying Lemma 8.1 to v_i, we find a measurable partition $(C_{i,k})_{k\in\mathbb{N}}$ of C_i and a sequence of Euclidean norms $\rho_{i,k}$ on $\mathbb{R}^m \times \mathbb{R}^{n-m}$ such that

$$\lambda^{-1}\rho_{i,k}((y,z) - (y',z')) \le \|v_i(y,z) - v_i(y',z')\| \le \lambda\,\rho_{i,k}((y,z) - (y',z'))$$

and $\lambda^{-1}\rho_{i,k}(\cdot) \le \|d(v_i)_{(y,z)}(\cdot)\| \le \lambda\,\rho_{i,k}(\cdot)$ for all $(y,z), (y',z') \in C_{i,k}$. Now choose linear isometries $T_{i,k}\colon (\mathbb{R}^m \times \mathbb{R}^{n-m}, \rho_{i,k}) \to \mathbb{R}^n$ and put

$$h_{i,k} := T_{i,k} \circ u\colon \mathbb{R}^n \to \mathbb{R}^n.$$

For all $x, x' \in v_i(C_{i,k})$, we have $\|h_{i,k}(x) - h_{i,k}(x')\| = \rho_{i,k}(u(x) - u(x'))$ and $\|d(h_{i,k})_x(\cdot)\| = \rho_{i,k}(du_x(\cdot))$. It follows that both the restriction of $h_{i,k}$ to $v_i(C_{i,k})$ and $d(h_{i,k})_x$ are λ-bi-Lipschitz. Finally, define $q\colon \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^m$ by $q(y,z) = y$ and put

$$L_{i,k} := q \circ T_{i,k}^{-1}\colon \mathbb{R}^n \to \mathbb{R}^n.$$

Then $L_{i,k} \circ h_{i,k} = q \circ u = f$. □

Definition 9.3 (coarea factor) Suppose that X, Y are normed spaces, $\dim X = n \ge \dim Y = m$, and $L\colon X \to Y$ is linear. The m-dimensional *coarea factor* $\mathbf{C}_m(L)$ is the number satisfying

$$\mathbf{C}_m(L)\mathcal{H}^n(A) = \int_Y \mathcal{H}^{n-m}(L^{-1}\{y\} \cap A)\, d\mathcal{H}^m(y)$$

for all $\mathcal{H}^n$-measurable sets $A \subset X$ with $\mathcal{H}^n(A) < \infty$.

Compare [AmbK2000a, Sect. 9]. Note that the right side is invariant under translations of A and, by Theorem 9.1 (coarea inequality), less than or equal to $(\alpha_m \alpha_{n-m}/\alpha_n)\,\mathrm{Lip}(L)^m\,\mathcal{H}^n(A)$. Therefore $\mathbf{C}_m(L)$ is a well-defined finite number. Clearly $\mathbf{C}_m(L) = \mathbf{J}(L)$ if $n = m$. Note that $\mathbf{C}_m(L) = 0$ if L has rank $< m$, since $\mathcal{H}^m(L(X)) = 0$. Now suppose L has rank m. Choosing an m-dimensional linear subspace $V \subset X$ complementary to the kernel $\ker(L)$ and a set A of the form $A = B + C$ for $B \subset \ker(L)$ and $C \subset V$, we infer that $\mathbf{C}_m(L)\,\mathcal{H}^n(A) = \mathcal{H}^{n-m}(B)\,\mathcal{H}^m(L(A))$. Since $L(A) = L(C)$ and $\mathcal{H}^m(L(C)) = \mathbf{J}(L|_V)\,\mathcal{H}^m(C)$, this yields the identity

$$\mathbf{C}_m(L)\,\mathcal{H}^n(A) = \mathbf{J}(L|_V)\,\mathcal{H}^{n-m}(B)\,\mathcal{H}^m(C).$$

When X is a Euclidean space, it follows that

$$\mathbf{C}_m(L) = \mathbf{J}\big(L|_W\big) \geq \mathbf{J}(L|_V),$$

where W denotes the orthogonal complement of $\ker(L)$ and L is still assumed to have rank m.

Finally, we remark that if H is a linear automorphism of the Euclidean space X, and if H is λ-bi-Lipschitz, then

$$\lambda^{-m}\,\mathbf{C}_m(L) \leq \mathbf{C}_m(L \circ H) \leq \lambda^m\,\mathbf{C}_m(L).$$

This holds trivially if the rank of L is $< m$. If the rank is m, put $V := H^{-1}(W)$ for W as above. Then

$$\mathbf{C}_m(L \circ H) \geq \mathbf{J}(L \circ H|_V) = \mathbf{J}(H|_V)\,\mathbf{J}\big(L|_W\big) \geq \lambda^{-m}\,\mathbf{C}_m(L),$$

which proves the first inequality. To verify the second, apply the first with $L \circ H$ and H^{-1} in place of L and H, respectively.

Theorem 9.4 (coarea formula) *Suppose that $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is a Lipschitz map, where $n \geq m \geq 1$.*

(1) *If $A \subset \mathbb{R}^n$ is $\mathcal{L}^n$-measurable, then*

$$\int_A \mathbf{C}_m(df_x)\,dx = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}\{y\} \cap A)\,dy.$$

(2) *If g is a real-valued $\mathcal{L}^n$-integrable function on $\mathbb{R}^n$, then*

$$\int_{\mathbb{R}^n} g(x)\,\mathbf{C}_m(df_x)\,dx = \int_{\mathbb{R}^m} \int_{f^{-1}\{y\}} g(x)\,d\mathcal{H}^{n-m}(x)\,dy.$$

Proof: To prove (1), we may partition A into countably many measurable sets and prove the respective formula for each of these sets separately. In particular, we lose no generality in assuming $\mathcal{L}^n(A) < \infty$. Let A_0 denote the set of all $x \in A$ where

f is not differentiable. It follows from Theorem 9.1 that A_0, as well as any other set of $\mathcal{L}^n$ measure zero, does not contribute to either side of the claimed identity. Now we split $A \setminus A_0$ into the two sets A', A'', where A' consists of all x where $\mathbf{C}_m(df_x) > 0$, i.e. df_x has rank m.

First we consider A'. Let $\lambda > 1$. Using Lemma 9.2 we choose countable families of pairwise disjoint measurable sets $A_i \subset A'$, Lipschitz maps $h_i \colon \mathbb{R}^n \to \mathbb{R}^n$, and linear maps $L_i \colon \mathbb{R}^n \to \mathbb{R}^m$ such that $\mathcal{L}^n(A' \setminus \bigcup_i A_i) = 0$, $h_i|_{A_i}$ is λ-bi-Lipschitz, $f = L_i \circ h_i$, and $d(h_i)_x$ exists and is λ-bi-Lipschitz for all $x \in A_i$. By the definition of the coarea factor,

$$\mathbf{C}_m(L_i)\, \mathcal{L}^n(h_i(A_i)) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(L_i^{-1}\{y\} \cap h_i(A_i))\, dy.$$

Since $h_i|_{A_i}$ is λ-bi-Lipschitz and maps $f^{-1}\{y\} \cap A_i$ onto $L_i^{-1}\{y\} \cap h_i(A_i)$, it follows that

$$\begin{aligned} \lambda^{-(2n-m)}\, \mathbf{C}_m(L_i)\, \mathcal{L}^n(A_i) &\le \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}\{y\} \cap A_i)\, dy \\ &\le \lambda^{2n-m}\, \mathbf{C}_m(L_i)\, \mathcal{L}^n(A_i). \end{aligned}$$

For all $x \in A_i$, $df_x = L_i \circ d(h_i)_x$, and $d(h_i)_x$ is λ-bi-Lipschitz, hence

$$\lambda^{-m}\, \mathbf{C}_m(df_x) \le \mathbf{C}_m(L_i) \le \lambda^m\, \mathbf{C}_m(df_x).$$

We conclude that

$$\begin{aligned} \lambda^{-2n} \int_{A_i} \mathbf{C}_m(df_x)\, dx &\le \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}\{y\} \cap A_i)\, dy \\ &\le \lambda^{2n} \int_{A_i} \mathbf{C}_m(df_x)\, dx. \end{aligned}$$

Summing over i and then letting λ tend to 1 we infer that (1) holds for $\bigcup_i A_i$ and hence for A'.

Now we turn to the set A'' of all $x \in A$ where $\mathbf{C}_m(df_x) = 0$. We must show that $\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}\{y\} \cap A'')\, dy = 0$. Let $\epsilon > 0$, and define

$$\begin{aligned} &h \colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m, \quad h(x,z) = f(x) + \epsilon z, \\ &p \colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m, \quad p(x,z) = z. \end{aligned}$$

Suppose that $(x,z) \in A'' \times \mathbb{R}^m$. Then

$$dh_{(x,z)}(v,w) = df_x(v) + \epsilon w$$

for all $(v,w) \in \mathbb{R}^n \times \mathbb{R}^m$. In particular, $dh_{(x,z)}$ has rank m, and $Z := \ker(dh_{(x,z)})$ is n-dimensional. Since $Z \cap (\mathbb{R}^n \times \{0\}) = \ker(df_x) \times \{0\}$ has dimension $\ge n-(m-1)$,

$V := p(Z)$ is a proper subspace of $\mathbb{R}^m$, and so its orthogonal complement $V^\perp$ in $\mathbb{R}^m$ is non-trivial. Since $\{0\} \times V^\perp \subset W := Z^\perp$, it follows that

$$\mathbf{C}_m(dh_{(x,z)}) = \mathbf{J}\big(dh_{(x,z)}|_W\big) \le \epsilon(\mathrm{Lip}(f) + \epsilon)^{m-1}.$$

Let $C := [0,1]^m \subset \mathbb{R}^m$. Using Fubini's Theorem and Theorem 9.1 (coarea inequality) with $k = n - m$, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}\{y\} \cap A'')\,dy \\
&\quad = \int_C \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}\{y - \epsilon z\} \cap A'')\,dy\,dz \\
&\quad = \int_{\mathbb{R}^m} \int_C \mathcal{H}^{n-m}(\{x \in A'' : h(x,z) = y\})\,dz\,dy \\
&\quad = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(p^{-1}\{z\} \cap h^{-1}\{y\} \cap (A'' \times C))\,dz\,dy \\
&\quad \le \frac{\alpha_m \alpha_{n-m}}{\alpha_n} \int_{\mathbb{R}^m} \mathcal{H}^n(h^{-1}\{y\} \cap (A'' \times C))\,dy.
\end{aligned}$$

Applying the above result for (f, A') to $(h, A'' \times C)$, we get

$$\begin{aligned}
\int_{\mathbb{R}^m} \mathcal{H}^n(h^{-1}\{y\} \cap (A'' \times C))\,dy &= \int_{A'' \times C} \mathbf{C}_m(dh_{(x,z)})\,d(x,z) \\
&\le \epsilon(\mathrm{Lip}(f) + \epsilon)^{m-1} \mathcal{L}^n(A'').
\end{aligned}$$

Letting ϵ tend to 0 we conclude that $\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}\{y\} \cap A'')\,dy = 0$. This completes the proof of (1).

Finally, (2) follows from (1) by approximation. □

10 Rectifiable sets

The following notion is fundamental in geometric measure theory.

Definition 10.1 (countably rectifiable set) Let Y be a metric space. A set $E \subset Y$ is called *countably $\mathcal{H}^m$-rectifiable* if there is a countable family of Lipschitz maps $f_i \colon A_i \to Y$, where $A_i \subset \mathbb{R}^m$ is $\mathcal{L}^m$-measurable, such that

$$\mathcal{H}^m\big(E \setminus \textstyle\bigcup_i f_i(A_i)\big) = 0.$$

It is often possible to take without loss of generality $A_i = \mathbb{R}^m$, e.g. if Y is a Banach space (recall Theorem 3.3).

Lemma 10.2 (bi-Lipschitz parametrization) *Suppose that Y is a metric space and $E \subset Y$ is an $\mathcal{H}^m$-measurable and countably $\mathcal{H}^m$-rectifiable set. Then there* *exists a countable family of bi-Lipschitz maps $f_k \colon C_k \to f_k(C_k) \subset E$, with $C_k \subset \mathbb{R}^m$ compact, such that the $f_k(C_k)$ are pairwise disjoint and*

$$\mathcal{H}^m\big(E \setminus \textstyle\bigcup_k f_k(C_k)\big) = 0.$$

Compare [Fed1969, Lemma 3.2.18] and [AmbK2000b, Lemma 4.1]. When $Y = \mathbb{R}^n$ it is possible to choose all f_k to be λ-bi-Lipschitz, for any given $\lambda > 1$.

Proof: Consider first a single Lipschitz map $f \colon A \to Y$ for some $\mathcal{L}^m$-measurable set $A \subset \mathbb{R}^m$. We assume that f extends to a Lipschitz map $\bar{f}$ defined on all of $\mathbb{R}^m$; if such an extension does not exist, we may first replace f with $\iota \circ f$ for some isometric embedding $\iota \colon Y \to l_\infty(Y)$. Applying Lemma 8.1 (Borel partition) to the set of all $x \in A$ where $\operatorname{md} \bar{f}_x$ exists and is a norm, we find a sequence of measurable sets $D_j \subset A$ such that $f|_{D_j}$ is bi-Lipschitz and, by Theorem 8.3 (area formula), $\mathcal{H}^m(f(A) \setminus \bigcup_j f(D_j)) = 0$.

Suppose now that $\mathcal{H}^m(E \setminus \bigcup_i f_i(A_i)) = 0$ for some sequence of Lipschitz maps $f_i \colon A_i \to Y$, where $A_i \subset \mathbb{R}^m$ is $\mathcal{L}^m$-measurable with $\mathcal{L}^m(A_i) < \infty$. Applying the above argument to each f_i, we get (after relabeling) a sequence of bi-Lipschitz maps $g_j \colon D_j \to g_j(D_j) \subset Y$ such that $D_j \subset \mathbb{R}^m$ is $\mathcal{L}^m$-measurable with finite measure and $\mathcal{H}^m(E \setminus \bigcup_j g_j(D_j)) = 0$. Note that $E_j := E \cap g_j(D_j)$ is $\mathcal{H}^m$-measurable and $\mathcal{H}^m(E_j) < \infty$. Then there exists a sequence of pairwise disjoint F_σ sets

$$F_1 \subset E_1, \quad F_2 \subset E_2 \setminus E_1, \quad F_3 \subset E_3 \setminus (E_1 \cup E_2), \quad \ldots$$

such that $\mathcal{H}^m(E \setminus \bigcup_j F_j) = 0$. Exhausting each of the $\mathcal{L}^m$-measurable sets $g_j^{-1}(F_j)$, up to an $\mathcal{L}^m$-nullset, by a countable collection of pairwise disjoint compact sets, we obtain the result. □

Proposition 10.3 (rectifiable level sets) *Suppose that X is a metric space, $n \geq m \geq 1$, $E \subset X$ is $\mathcal{H}^n$-measurable and countably $\mathcal{H}^n$-rectifiable, and $f \colon E \to \mathbb{R}^m$ is Lipschitz. Then for $\mathcal{L}^m$-almost every $y \in \mathbb{R}^m$, $f^{-1}\{y\}$ is $\mathcal{H}^{n-m}$-measurable and countably $\mathcal{H}^{n-m}$-rectifiable.*

Proof: Consider first the case $E = X = \mathbb{R}^n$. Let B denote the set of all $x \in \mathbb{R}^n$ where df_x exists and has rank m. Choose a Borel partition $(B_i)_{i \in \mathbb{N}}$ of B and coordinate projections $p_i \colon \mathbb{R}^n \to \mathbb{R}^{n-m}$ such that each $u_i|_{B_i}$ is bi-Lipschitz, where $u_i = (f, p_i)$; compare the first part of the proof of Lemma 9.2. For all $y \in \mathbb{R}^m$,

$$f^{-1}\{y\} \cap B_i = (u_i|_{B_i})^{-1}((\{y\} \times \mathbb{R}^{n-m}) \cap u_i(B_i)).$$

For $\mathcal{L}^m$-almost every $y \in \mathbb{R}^m$, we have in addition that $\mathcal{H}^{n-m}(f^{-1}\{y\} \setminus B) = 0$ since, by Theorem 9.4 (coarea formula),

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}\{y\} \setminus B)\, dy = \int_{\mathbb{R}^n \setminus B} \mathbf{C}_m(df_x)\, dx = 0.$$

As $(u_i|_{B_i})^{-1}$ is Lipschitz, this shows that $f^{-1}\{y\}$ is countably $\mathcal{H}^{n-m}$-rectifiable for $\mathcal{L}^m$-almost every $y \in \mathbb{R}^m$.

Now consider the general case, and let $E \subset X$ and $f: E \to \mathbb{R}^m$ be given. By Lemma 10.2 there exists a sequence of Lipschitz maps $h_i: C_i \to E$, where $C_i \subset \mathbb{R}^n$ is compact, such that $\mathcal{H}^n(E \setminus D) = 0$ for $D := \bigcup_{i=1}^{\infty} h_i(C_i)$. For each i, put $f_i := f \circ h_i: C_i \to \mathbb{R}^m$ and choose a Lipschitz extension $\bar{f}_i: \mathbb{R}^n \to \mathbb{R}^m$. For $\mathcal{L}^m$-almost every $y \in \mathbb{R}^m$, the set $\bar{f}_i^{-1}\{y\}$ is countably $\mathcal{H}^{n-m}$-rectifiable by the first part of the proof, and so is

$$h_i(\bar{f}_i^{-1}\{y\} \cap C_i) = h_i(f_i^{-1}\{y\} \cap C_i) = f^{-1}\{y\} \cap h_i(C_i),$$

thus $f^{-1}\{y\} \cap D$ is a σ-compact countably $\mathcal{H}^{n-m}$-rectifiable subset of $f^{-1}\{y\}$. Moreover, $\mathcal{H}^{n-m}(f^{-1}\{y\} \setminus D) = 0$ for $\mathcal{L}^m$-almost every $y \in \mathbb{R}^m$, because

$$\int_{\mathbb{R}^m}^{*} \mathcal{H}^{n-m}(f^{-1}\{y\} \setminus D)\, dy \leq \frac{\alpha_m \alpha_{n-m}}{\alpha_n} \operatorname{Lip}(f)^m \, \mathcal{H}^n(E \setminus D) = 0$$

by Theorem 9.1 (coarea inequality). This gives the result. □

Proposition 10.4 (countably rectifiable sets in $\mathbb{R}^n$) *A set $E \subset \mathbb{R}^n$ is countably $\mathcal{H}^m$-rectifiable if and only if there exists a sequence of m-dimensional C^1 submanifolds M_k of $\mathbb{R}^n$ such that*

$$\mathcal{H}^m\big(E \setminus \textstyle\bigcup_k M_k\big) = 0.$$

See [Fed1969, Theorem 3.2.29], [Sim2014, Ch. 3].

Proof: Suppose that $\mathcal{H}^m(E \setminus \bigcup_i f_i(\mathbb{R}^m)) = 0$ for a sequence of Lipschitz maps $f_i: \mathbb{R}^m \to \mathbb{R}^n$. By Theorem 5.2, we can assume without loss of generality that the f_i are C^1. Let $U_i \subset \mathbb{R}^m$ be the set of all $x \in \mathbb{R}^m$ where df_x has rank m. By the area formula, $\mathcal{H}^m(f_i(\mathbb{R}^m \setminus U_i)) = 0$. Hence, $\mathcal{H}^m(E \setminus \bigcup_i f_i(U_i)) = 0$. Finally, it follows from the inverse function theorem that each $f_i(U_i)$ is a countable union of C^1 submanifolds.

The other implication is clear. □

We now turn to linear approximation properties of countably $\mathcal{H}^m$-rectifiable subsets of $\mathbb{R}^n$. There are different definitions of approximate tangent spaces for such sets in the literature, compare [Fed1969, 3.2.16] and [Mat1995, Ch. 15]. Here we adopt the approach from [Sim2014, Ch. 3].

For $x \in \mathbb{R}^n$ and $\lambda > 0$, define $\eta_{x,\lambda}: \mathbb{R}^n \to \mathbb{R}^n$, $\eta_{x,\lambda}(y) = (y-x)/\lambda$. Note that $\eta_{x,\lambda}$ maps $B(x, \lambda r)$ onto $B(0, r)$.

Definition 10.5 (approximate tangent space) Suppose that $E \subset \mathbb{R}^n$ is an $\mathcal{H}^m$-measurable set with $\mathcal{H}^m(E) < \infty$. Let $x \in \mathbb{R}^n$. An m-dimensional linear subspace $L \subset \mathbb{R}^n$ is called the ($\mathcal{H}^m$-)*approximate tangent space* of E at x if

$$\lim_{\lambda \to 0+} \int_{\eta_{x,\lambda}(E)} \phi \, d\mathcal{H}^m = \int_L \phi \, d\mathcal{H}^m$$

for all $\phi \in C_c(\mathbb{R}^n)$. Then we write $L =: \mathrm{Tan}^m(E,x)$.

Clearly $\mathrm{Tan}^m(E,x)$ is uniquely determined if it exists. For an m-dimensional C^1 submanifold $M \subset \mathbb{R}^n$, $\mathrm{Tan}^m(M,x)$ agrees with the usual tangent space T_xM for every $x \in M$.

Theorem 10.6 (existence of approximate tangent spaces) *Suppose that $E \subset \mathbb{R}^n$ is an $\mathcal{H}^m$-measurable and countably $\mathcal{H}^m$-rectifiable set with $\mathcal{H}^m(E) < \infty$. Then for $\mathcal{H}^m$-almost every $x \in E$,* $\mathrm{Tan}^m(E,x)$ *exists and* $\Theta^m(E,x) = 1$.

Proof: Choose a sequence of m-dimensional C^1 submanifolds M_k of $\mathbb{R}^n$ such that $\mathcal{H}^m(E \setminus \bigcup_k M_k) = 0$, compare Proposition 10.4. The sets $E_k := E \cap M_k$ are $\mathcal{H}^m$-measurable with $\mathcal{H}^m(E_k) < \infty$, so it follows from the second statement in Theorem 7.2 that for $\mathcal{H}^m$-almost every $x \in E_k$,

$$\Theta^m(E,x) = \Theta^m(M_k,x) - \Theta^m(M_k \setminus E_k,x) + \Theta^m(E \setminus E_k,x) = 1 - 0 + 0.$$

Similarly, if $\phi \in C_c(\mathbb{R}^n)$ with $\mathrm{spt}(\phi) \subset B(0,r)$, then

$$\int_{\eta_{x,\lambda}(E)} \phi \, d\mathcal{H}^m = \int_{\eta_{x,\lambda}(M_k)} \phi \, d\mathcal{H}^m - \int_{\eta_{x,\lambda}(M_k \setminus E_k)} \phi \, d\mathcal{H}^m + \int_{\eta_{x,\lambda}(E \setminus E_k)} \phi \, d\mathcal{H}^m$$

for all $\lambda > 0$, and for $\mathcal{H}^m$-almost every $x \in E_k$, the last two integrals tend to 0 as $\lambda \to 0$. For example,

$$\begin{aligned} \left| \int_{\eta_{x,\lambda}(E \setminus E_k)} \phi \, d\mathcal{H}^m \right| &\le \sup|\phi| \cdot \mathcal{H}^m(\eta_{x,\lambda}(E \setminus E_k) \cap B(0,r)) \\ &= \sup|\phi| \cdot \frac{\mathcal{H}^m((E \setminus E_k) \cap B(x,\lambda r))}{\lambda^m}, \end{aligned}$$

which tends to $\sup|\phi|\, r^m \,\Theta^m(E \setminus E_k, x) = 0$. It now follows that

$$\lim_{\lambda \to 0+} \int_{\eta_{x,\lambda}(E)} \phi \, d\mathcal{H}^m = \lim_{\lambda \to 0+} \int_{\eta_{x,\lambda}(M_k)} \phi \, d\mathcal{H}^m = \int_{T_xM_k} \phi \, d\mathcal{H}^m$$

and thus $\mathrm{Tan}^m(E,x) = T_xM_k$ for $\mathcal{H}^m$-almost every $x \in E_k$. □

The following two converses to Theorem 10.6 hold. For the first part, see again [Mat1995, Ch. 15] and [Sim2014, Ch. 3]. The second statement is a deep result of Preiss [Pre1987]; see [Mat1995, Ch. 17] and [DeL2008] for some accounts.

Theorem 10.7 (rectifiability criteria) *Let $E \subset \mathbb{R}^n$ be an $\mathcal{H}^m$-measurable set with $\mathcal{H}^m(E) < \infty$.*

(1) *If* $\mathrm{Tan}^m(E,x)$ *exists for $\mathcal{H}^m$-almost every $x \in E$, then E is countably $\mathcal{H}^m$-rectifiable.*

(2) *If the density* $\Theta^m(E, x)$ *exists for* $\mathcal{H}^m$*-almost every* $x \in E$*, then* E *is countably* $\mathcal{H}^m$*-rectifiable.*

Finally, we state the Besicovitch–Federer projection theorem which played a central role in the development of the theory of rectifiable currents. This deep result was proved in [Bes1939] for $m = 1$ and $n = 2$ and in [Fed1947] for general dimensions. See [Fed1969, Theorem 3.3.13], [Mat1995, Theorem 18.1].

A subset F of a metric space Y is *purely* $\mathcal{H}^m$*-unrectifiable* if $\mathcal{H}^m(F \cap E) = 0$ for every countably $\mathcal{H}^m$-rectifiable set $E \subset Y$. Every set $A \subset Y$ with $\mathcal{H}^m(A) < \infty$ can be written as the disjoint union of a countably $\mathcal{H}^m$-rectifiable set E and a purely $\mathcal{H}^m$-unrectifiable set F (compare [Mat1995, Theorem 15.6]).

Theorem 10.8 (Besicovitch, Federer) *Let* $F \subset \mathbb{R}^n$ *be a purely* $\mathcal{H}^m$*-unrectifiable set with* $\mathcal{H}^m(F) < \infty$*. Then* $\mathcal{H}^m(\pi_L(F)) = 0$ *for* $\gamma_{n,m}$*-almost every* $L \in G(n, m)$ *, where* $\gamma_{n,m}$ *denotes the Haar measure on* $G(n, m)$ *(see below), and* $\pi_L \colon \mathbb{R}^n \to L$ *is orthogonal projection.*

Recall that on every locally compact, Hausdorff topological group G, there exists a left-invariant Radon measure $\mu \not\equiv 0$ that is uniquely determined up to a positive factor. As for the canonical measure $\gamma_{n,m}$ on the Grassmannian manifold $G(n, m)$ of m-dimensional linear subspaces of $\mathbb{R}^n$, one can take the normalized Haar measure μ on the compact group $G = SO(n)$ and define $\gamma_{n,m}$ as the push-forward $\alpha_\# \mu$ under the map $\alpha \colon SO(n) \to G(n, m), g \mapsto gL_0$, for any fixed subspace $L_0 \in G(n, m)$. Thus, for every Borel set $B \subset G(n, m)$,

$$\gamma_{n,m}(B) := \mu\big(\{g \in SO(n) : gL_0 \in B\}\big).$$

For illustration, we describe a simple example of compact purely $\mathcal{H}^1$-unrectifiable subset of $\mathbb{R}^2$ with positive finite $\mathcal{H}^1$-measure.

Example 10.9 (Cantor dust) Start with the unit square $[0, 1]^2$ in $\mathbb{R}^2$, subdivide it into 16 closed squares of equal size, and let C_1 denote the union of the first and third square in the top row and the second and fourth square in the bottom row. Subdivide each of these four squares using the same pattern, and let C_2 denote the corresponding union of the sixteen little squares. Continuing this process, one gets a decreasing sequence of compact sets C_k consisting of 4^k squares of edge length 4^{-k}. By construction, the set $C := \bigcap_{k=1}^\infty C_k$ has the property that $\pi(C) = [0, 1]$ for the projection $\pi \colon (x, y) \mapsto x$, hence $\mathcal{H}^1(C) \geq 1$, and clearly $\mathcal{H}^1(C) < \infty$. One can check that C has no approximate tangents, so it follows from Theorem 10.6 that C is purely $\mathcal{H}^1$-unrectifiable. According to Theorem 10.8, $\mathcal{H}^1(\pi_L(C)) = 0$ for $\gamma_{2,1}$-almost every line $L \in G(2, 1)$.

Using Theorem 10.8 and Example 10.9, one can give a simple construction of a *Besicovitch set*, that is, a set $B \subset \mathbb{R}^2$ of $\mathcal{L}^2$-measure zero containing a line in every direction (compare [Mat1995, Thm. 18.1]).

Example 10.10 (Besicovitch set) The idea is to use the set C from Example 10.9 as a parameter space for a family of lines. Define $f\colon C \times \mathbb{R} \to \mathbb{R}^2$ by $f(a, b, x) := (x, ax + b)$, and put $B := f(C \times \mathbb{R})$. Note that B is σ-compact. For every $a \in [0, 1]$ there exists a point $(a, b) \in C$, so B contains the graph of the function $x \mapsto ax + b$. Thus there are lines of every slope $a \in [0, 1]$ in B, and by taking the union of four suitably rotated copies of B one gets a Borel set containing a line in every direction. It remains to show that $\mathcal{L}^2(B) = 0$. For $t \in \mathbb{R}$, consider the set

$$B \cap (\{t\} \times \mathbb{R}) = \{(t, at + b) : (a, b) \in C\} = \{t\} \times \rho_t(C),$$

where $\rho_t(x, y) = tx + y$. Write $t = \cot(\phi)$ for $\phi \in (0, \pi)$, and let $\pi_\phi\colon \mathbb{R}^2 \to L_\phi$ denote the orthogonal projection to $L_\phi = \mathbb{R}(\cos(\phi), \sin(\phi))$. Then $\pi_\phi(x, y) = r(\cos(\phi), \sin(\phi))$ for

$$r = \cos(\phi)x + \sin(\phi)y = \rho_t(x, y)\sin(\phi).$$

It follows that $\mathcal{L}^1(\rho_t(C)) = 0$ for almost every $t \in \mathbb{R}$ if and only if $\mathcal{H}^1(\pi_\phi(C)) = 0$ for almost every $\phi \in (0, \pi)$, which holds by Theorem 10.8. Hence, by Fubini's Theorem, $\mathcal{L}^2(B) = 0$.

References

[Alm1962] F. J. Almgren, Jr., The homotopy groups of the integral cycle groups, Topology 1 (1962), 257–299.

[AmbK2000a] L. Ambrosio, B. Kirchheim, Rectifiable sets in metric and Banach spaces, Math. Ann. 318 (2000), 527–555.

[AmbK2000b] L. Ambrosio, B. Kirchheim, Currents in metric spaces, Acta Math. 185 (2000), 1–80.

[AmbT2004] L. Ambrosio, P. Tilli, Topics on analysis in metric spaces, Oxford Lecture Series in Math. and its Appl. 25, Oxford Univ. Press 2004.

[BenL2000] Y. Benyamini, J. Lindenstrauss, Geometric Nonlinear Functional Analysis, Volume 1, Amer. Math. Soc. Colloquium Publ., Vol. 48, Amer Math. Soc. 2000.

[Bes1939] A. S. Besicovitch, On the fundamental geometric properties of linearly measurable plane sets of points III, Math. Ann. 116 (1939), 349–357.

[Bla1916] W. Blaschke, Kreis und Kugel, Veit 1916, de Gruyter 1956.

[BurBI2001] D. Burago, Y. Burago, S. Ivanov, A Course in Metric Geometry, Graduate Studies in Math., Vol. 33, Amer. Math. Soc. 2001.

[BurZ1988] Y. D. Burago, V. A. Zalgaller, Geometric Inequalities, Springer 1988.

[Dav1970] O. Davies, Increasing sequences of sets and Hausdorff measure, Proc. London Math. Soc. 20 (1970), 222–236.

[DeL2008] C. De Lellis, Rectifiable Sets, Densities and Tangent Measures, Zurich Lect. Adv. Math., European Math. Soc. 2008.

[EsmH2021] B. Esmayli, P. Hajlasz, The coarea inequality. Ann. Fenn. Math. 46 (2021), no. 2, 965–991.

[EvaG1992] L. C. Evans, R. E. Gariepy, Measure Theory and Fine Properties of Functions, Boca Raton, CRC Press 1992.

[Fed1947] H. Federer, The (ϕ, k) rectifiable subsets of n space, Trans. Amer. Math. Soc. 62 (1947), 114–192.

[Fed1969] H. Federer, Geometric Measure Theory, Springer 1969, 1996.

[Gro1981] M. Gromov, Groups of polynomial growth and expanding maps, Publ. Math. IHES 53 (1981), 53–73.

[Gro1999] M. Gromov, with Appendices by M. Katz, P. Pansu, and S. Semmes, Metric Structures for Riemannian and Non-Riemannian Spaces, Birkhäuser 1999.

[Hei2001] J. Heinonen, Lectures on Analysis on Metric Spaces, Springer 2001.

[Hei2003] J. Heinonen, Geometric embeddings of metric spaces, Rep. Univ. Jyväskylä Dept. Math. Stat., no. 90, 2003.

[JohLS1986] W. B. Johnson, J. Lindenstrauss, G. Schechtman, Extensions of Lipschitz maps into Banach spaces, Israel J. Math. 54 (1986), 129–138.

[Kir1994] B. Kirchheim, Rectifiable metric spaces: local structure and regularity of the Hausdorff measure, Proc. Amer. Math. Soc. 121 (1994), 113–123.

[Kirs1934] M. D. Kirszbraun, Über die zusammenziehende und Lipschitzsche Transformationen, Fundamenta Math. 22 (1934), 77–108.

[KorS1993] N. J. Korevaar, R. M. Schoen, Sobolev spaces and harmonic maps for metric space targets, Comm. Anal. Geom. 1 (1993), 561–659.

[LanS1997] U. Lang, V. Schroeder, Kirszbraun's theorem and metric spaces of

bounded curvature, Geom. Funct. Anal. 7 (1997), 535–560.
[Lan1999] U. Lang, Extendability of large-scale Lipschitz maps, Trans. Amer. Math. Soc. 351 (1999), 3975–3988.
[Mal1999] J. Maly, A simple proof of the Stepanov theorem on differentiability almost everywhere, Exposition. Math. 17 (1999), 59–61.
[Mat1995] P. Mattila, Geometry of Sets and Measures in Euclidean Spaces, Fractals and Rectifiability, Cambridge Studies in Advanced Math., Vol. 44, Cambridge U. Press 1995.
[McS1934] E. J. McShane, Extension of range of functions, Bull. Amer. Math. Soc. 40 (1934), 837–842.
[Rad1919] H. Rademacher, Über partielle und totale Differenzierbarkeit I, Math. Ann. 79 (1919), 340–359.
[Rei2009] L. P. Reichel, The Coarea Formula for Metric Space Valued Maps, Diss. ETH No. 18540, 2009.
[Pre1987] D. Preiss, Geometry of measures in R^n: distribution, rectifiability, and densities, Ann. of Math. 125 (1987), 537–643.
[Sim2014] L. Simon, Introduction to Geometric Measure Theory, Tsinghua Lecture Notes 2014.
[Ste1970] E. M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton Univ. Press, New Jersey 1970.
[Step1923] W. Stepanoff, Über totale Differenzierbarkeit, Math. Ann. 90 (1923), 318–320.
[Val1945] F. A. Valentine, A Lipschitz condition preserving extension for a vector function, Amer. J. Math. 67 (1945), 83–93.
[Wen2008] S. Wenger, Characterizations of metric trees and Gromov hyperbolic metric spaces, Math. Res. Lett. 15 (2008), 1017–1026.
[Whit1934] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, Trans. Amer. Math. Soc. 36 (1934), 63–89.
[Whit1935] H. Whitney, A function not constant on a connected set of critical points, Duke Math. J. 1 (1935), 514–517.

Urs Lang
lang@math.ethz.ch
Department of Mathematics
ETH Zurich
8092 Zurich
Switzerland