Notes on Rectifiability

Urs Lang

2007

Abstract

These are the notes to a first part of a lecture course on Geometric Measure Theory I taught at ETH Zurich in Fall 2006. They partially overlap with [Lan1].

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 X, $l^\infty(X) = (l^\infty(X), \|\cdot\|_\infty)$ denotes the Banach space of all functions $s\colon X \to \mathbb{R}$ with $\|s\|_\infty := \sup_{x\in X} |s(x)| < \infty$. Similarly, $l^\infty := l^\infty(\mathbb{N}) = \{(s_k)_{k\in\mathbb{N}} : \|(s_k)\|_\infty := \sup_{k\in\mathbb{N}} |s_k| < \infty\}$.

1.1 Proposition (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: (1) Fix a basepoint $z \in X$ and define $f\colon X \to l^\infty(X)$,

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

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

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

and equality occurs for $y = x'$.

(2) Choose a basepoint $z \in X$ and a countable dense set $D = \{x_k : k \in \mathbb{N}\}$ in X. Define $f\colon X \to l^\infty$,

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

By (1), $f|_D\colon D \to l^\infty = l^\infty(D)$ is an isometric embedding. Since D is dense and f is continuous, f is an isometric embedding. □

Note that if X is bounded, we do not need to subtract the term $d(y,z)$ or $d(x_k,z)$, 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 [Hei2].

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.

1.2 Definition (uniformly precompact family)
A family $(X_j)_{j\in J}$ 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_j can be covered by n closed balls of radius ϵ. The family $(X_j)_{j\in J}$ is* uniformly bounded *if $\sup_{j\in J} \operatorname{diam} X_j < \infty$.*

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

We follow essentially the original proof from [Gro1].

Proof: For $i \in \mathbb{N}$, let $\epsilon_i := 2^{-i}$ and pick $n_i \in \mathbb{N}$ such that each X_j 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 for each $i \in \mathbb{N}$,

$$\pi^{-1}(N_i) = N_{i+1}, \quad \#\pi^{-1}\{k\} = n_{i+1} \quad \forall k \in N_i.$$

In each X_j, we construct a sequence $(x_k^j)_{k\in\mathbb{N}}$ according to the following inductive scheme. For $i = 1$, the points x_k^j with $k \in N_i = N_1$ are chosen such that the n_1 balls $\mathrm{B}(x_k^j, \epsilon_1)$ cover X_j. For $i \geq 1$, if the $n_1 \dots n_i$ centers x_k^j with $k \in N_i$ are selected, the $n_1 \dots n_i n_{i+1}$ points x_l^j with $l \in N_{i+1}$ are chosen such that for each $k \in N_i$, the ball $\mathrm{B}(x_k^j, \epsilon_i)$ is covered by the n_{i+1} balls

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

with $l \in \pi^{-1}\{k\}$. This way we obtain for every $j \in J$ a dense sequence $(x_k^j)_{k\in\mathbb{N}}$ in X_j which gives rise to an isometric embedding $f_j\colon X_j \to l^\infty$, mapping x to $(d(x, x_k^j))_{k\in\mathbb{N}}$. Whenever $i \in \mathbb{N}$, $k \in N_i$, and $l \in \pi^{-1}\{k\}$, then

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

Hence, each $f_j(X_j)$ lies in the set Z of all sequences $(s_k)_{k\in\mathbb{N}}$ with $0 \leq s_k \leq \sup_j \operatorname{diam} X_j$ for all $k \in \mathbb{N}$ and

$$|s_k - s_l| \leq 2\epsilon_i \quad \forall i \in \mathbb{N},\, k \in N_i,\, l \in \pi^{-1}\{k\}.$$

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

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

the closed δ-neighborhood of A and by

$$d_H(A, B) = \inf\{\delta \geq 0 : A \subset \mathrm{N}_\delta(B),\, B \subset \mathrm{N}_\delta(A)\}$$

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

2.1 Theorem (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 [Bla] for compact convex bodies in $\mathbb{R}^3$ to settle the existence question in the isoperimetric problem.

Proof: (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

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

Let $\epsilon > 0$. Choose i_0 such that $d_H(C_i, C_j) < \epsilon/2$ 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 $d(x, C_i) \leq d(x, C_j) + d_H(C_i, C_j) < \epsilon$ for all $i \geq i_0$. This shows that $C \subset \mathrm{N}_\epsilon(C_i)$ for $i \geq i_0$.

Now suppose $x \in C_i$ for some $i \geq i_0$. Pick a sequence $i = i_1 < i_2 < \dots$ such that $d_H(C_m, C_n) < \epsilon/2^k$ 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 $C_{i_k} \cup C_{i_{k+1}} \cup \dots$ for all k. Thus $y \in C$ and $d(x, C) < \epsilon$. This shows that $C_i \subset \mathrm{N}_\epsilon(C)$ whenever $i \geq i_0$.

(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 $\mathrm{N}_\epsilon(Z) = X$. We show that every $C \in \mathcal{C}$ is at Hausdorff distance at most ϵ of some subset of Z, namely $Z_C := Z \cap \mathrm{N}_\epsilon(C)$. 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 $C \subset \mathrm{N}_\epsilon(Z_C)$. Since also $Z_C \subset \mathrm{N}_\epsilon(C)$, we have $d_H(C, Z_C) \leq \epsilon$. As there are only finitely many distinct subsets of Z, we conclude that $\mathcal{C}$ is precompact.
□

2.2 Definition (Gromov–Hausdorff distance)
The Gromov–Hausdorff distance *of two metric spaces X, Y is the number*

$$d_{GH}(X, Y) = \inf d_H^Z(X', Y'),$$

where the infimum is taken over all triples (Z, X', Y') such that $Z = (Z, d^Z)$ is a metric space, $X' \subset Z$ is an isometric copy of X, and $Y' \subset Z$ is an isometric copy of Y.

Cf. [Gro1], [Gro2]. 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

$$d_{GH}(X, Y) = \inf \bar{d}_H(X, Y)$$

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

For instance, suppose that $\operatorname{diam}(X), \operatorname{diam}(Y) \leq D < \infty$. 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 $d_{GH}(X, Y) \leq D/2$.

2.3 Proposition

(1) *d_{GH} satisfies the triangle inequality, i.e. $d_{GH}(X, Z) \leq d_{GH}(X, Y) + d_{GH}(Y, Z)$ for all metric spaces X, Y, Z.*

(2) *d_{GH} defines a metric on the set of isometry classes of compact metric spaces.*

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

2.4 Theorem (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. $\lim_{j \to \infty} d_{GH}(X_{i_j}, Z) = 0$.

This was proved in [Gro1].

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

$$d(f(x), f(x')) \leq \lambda d(x, x') \quad \forall x, x' \in X;$$

f is *Lipschitz* if

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

We say that $f \colon X \to Y$ is *bi-Lipschitz* if f is λ-*bi-Lipschitz* for some $\lambda \in [1, \infty)$, i.e.

$$\lambda^{-1} d(x, x') \leq d(f(x), f(x')) \leq \lambda d(x, x') \quad \forall x, x' \in X.$$

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

3.1 Proposition (McShane, Whitney)
Suppose 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 J be an arbitrary set. Every λ-Lipschitz map $f\colon A \to l^\infty(J)$ possesses a λ-Lipschitz extension $\bar{f}\colon X \to l^\infty(J)$.*

Proof: (1) For $n = 1$, put

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

For $n \geq 2$, $f = (f_1, \dots, f_n)$, extend each f_i separately.

(2) For $f = (f_j)_{j \in J}$, extend each f_j separately. □

In (1), the factor $\sqrt{n}$ cannot be replaced by a constant $< n^{1/4}$, cf. [JohLS] and [Lan]. 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:

3.2 Theorem (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 [Kirs], [Val], or [Fed, Theorem 2.10.43]. A generalization to metric spaces with curvature bounds was given in [LanS].

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 9.1). We call a metric space Y *Lipschitz m-connected* if there is a constant $c \geq 1$ such that for $k \in \{0, \dots, m\}$, every λ-Lipschitz map $f\colon \mathrm{S}^k \to Y$ admits a $c\lambda$-Lipschitz extension $\bar{f}\colon \mathrm{B}^{k+1} \to Y$; 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 $m \geq 0$. The sphere S^n is Lipschitz $(n-1)$-connected.

3.3 Theorem (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 $f\colon A \to Y$, $A \subset \mathbb{R}^m$, has a $c\lambda$-Lipschitz extension $\bar{f}\colon \mathbb{R}^m \to Y$.*

The idea of the proof goes back to Whitney [Whit]. Compare [Alm1, Theorem (1.2)] and [JohLS].

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

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

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 otherwise the next bigger dyadic cube C' containing C would still fulfil

$$\operatorname{diam} C' = 2 \operatorname{diam} C \le 2(d(A, C) - \operatorname{diam} C) \le 2\, d(A, C').$$

Denote by $\Sigma_k \subset \mathbb{R}^m$ the k-skeleton of this cubical decomposition. Extend f to a Lipschitz map $f_0: A \cup \Sigma_0 \to Y$ by precomposing f with a nearest point retraction $A \cup \Sigma_0 \to A$. Then, for $k = 0, \dots, m-1$, successively extend f_k to $f_{k+1}: A \cup \Sigma_{k+1} \to Y$ by means of the Lipschitz $(m-1)$-connectedness of Y. As $A \cup \Sigma_m = \mathbb{R}^m$, $\bar{f} := f_m$ is the desired extension of f. □

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

This is taken from [Hei1, Theorem 6.8].

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

$$f_j(x) := \inf\{f(a) + jd(a, x) : a \in X\}.$$

Then $f_j(x) \ge \inf f > -\infty$, and f_j is j-Lipschitz (compare the proof of Proposition 3.1). Since $f_j(x) \le f(x)$, we have

$$f_j(x) \le f(a) + f(x) - f(a) \le f(a) + 2 \sup |f|$$

for all $a, x \in X$. Setting $\delta_j := 2 \sup |f|/j$, we conclude that

$$f_j(x) = \inf\{f(a) + jd(a, x) : a \in X,\ d(a, x) \le \delta_j\}.$$

Hence,

$$0 \le f(x) - f_j(x) \le \sup\{f(x) - f(a) : a \in X,\ d(a, x) \le \delta_j\} \le \omega(\delta_j)$$

for all $x \in X$, i.e. $f_j \to f$ $(j \to \infty)$ uniformly on X. □

4 Differentiability of Lipschitz maps

Recall the following definitions.

4.1 Definition (Gâteaux and Fréchet differential)
Suppose 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

$$L(v) = D_v f(x) \quad \forall v \in X.$$

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

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

whenever $|t| \leq \delta$ and $u \in \mathrm{S}(0, 1) \subset X$.

4.2 Lemma (differentiable Lipschitz maps)
Suppose 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 $u \in \mathrm{S}^{m-1}$ there is a $u' \in D'$ with $|u - u'| \leq \epsilon$. Then there is a $\delta > 0$ such that

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

whenever $|t| \le \delta$ and $u' \in D'$. Given $u \in \mathrm{S}^{m-1}$, 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 $|t| \le \delta$. □

4.3 Theorem (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 [Rad].

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 $m \ge 2$. For $u \in \mathrm{S}^{m-1}$, denote by B_u the set of all $x \in \mathbb{R}^m$ where $D_u f(x)$ exists and by H_u 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

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

Note that B_u is a Borel set, in particular, the characteristic function of $\mathbb{R}^m \setminus B_u$ is Lebesgue measurable. Fubini's theorem then implies that $\mathcal{L}^m(\mathbb{R}^m \setminus B_u) = 0$. Now choose a dense countable subset D of S^{m-1}, and put $B := \bigcap_{u \in D} B_u$. Then

$$\mathcal{L}^m(\mathbb{R}^m \setminus B) = 0,$$

and for all $x \in B$, $D_u f(x)$ and $D_{e_i} f(x)$ exist for all $u \in D$ and $i = 1, \dots, m$; in particular, the formal gradient

$$\nabla f(x) := (D_{e_i} f(x), \dots, D_{e_m} f(x))$$

exists. We show that for $\mathcal{L}^m$-almost all $x \in B$ we have, in addition, the usual relation

$$D_u f(x) = \langle \nabla f(x), u \rangle \quad \forall u \in D.$$

The theorem then follows from Lemma 4.2. Let $\varphi \in C_c^\infty(\mathbb{R}^m)$. By Lebesgue's bounded convergence theorem,

$$\begin{aligned}\lim_{t \to 0} \int \frac{f(x+tu) - f(x)}{t} \varphi(x)\,dx &= \int D_u f(x) \varphi(x)\,dx, \\ \lim_{t \to 0} \int f(x) \frac{\varphi(x-tu) - \varphi(x)}{t}\,dx &= -\int f(x) D_u \varphi(x)\,dx.\end{aligned}$$

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

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

and similarly

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

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

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

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

This generalization of Rademacher's theorem was proved in [Step]. The following elegant argument is due to Malý [Mal].

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

$$A_i := \{x \in U_i : \text{both } a_i \text{ and } b_i \text{ are differentiable at } x\}.$$

By Rademacher's theorem, $Z := \bigcup_{i=1}^\infty U_i \setminus A_i$ has measure zero. Let $x \in L(f) \setminus Z$. We show that there exists an index i such that $x \in A_i$ and $a_i(x) = b_i(x)$; 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 $y \in \mathrm{B}(x, r)$ and for some $\lambda \ge 0$ independent of y. Choose i such that $i \ge \lambda$ and $x \in U_i \subset \mathrm{B}(x, r)$. Since $x \notin Z$, $x \in A_i$. By the definition of a_i and b_i,

$$f(x) - i|y - x| \le a_i(y) \le f(y) \le b_i(y) \le f(x) + i|y - x|$$

for all $y \in U_i$. Hence, $a_i(x) = b_i(x)$. □

Generalizations of these results to maps between Banach spaces or even more general classes of metric spaces are a topic of current research.

5 Extension of smooth functions

We state Whitney's extension theorem for C^1 functions and an application, cf. [Whit], [Fed, Theorem 3.1.14] and [Sim, Theorem 5.3], [Fed, Theorem 3.1.16].

5.1 Theorem (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 $\nabla \bar{f}|_A = g$.

Note 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.

5.2 Theorem (C^1 approximation of Lipschitz functions)
If $f : \mathbb{R}^m \to \mathbb{R}$ 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: By Rademacher's theorem, f is almost everywhere differentiable, and $g := \nabla f$ is a measurable function. According to Lusin's theorem, there is a closed set $C \subset \mathbb{R}^m$ with $\mathcal{L}^m(\mathbb{R}^m \setminus C) < \epsilon/2$ such that $g|_C$ is continuous. For $x \in C$ and $i \in \mathbb{N}$, let

$$r_i(x) := \sup \frac{|f(y) - f(x) - \langle g(x), y - x\rangle|}{|y - x|},$$

the supremum taken over all $y \in C$ with $0 < |y - x| \le 1/i$. We know that $r_i \to 0$ pointwise on C as $i \to \infty$. By Egorov's theorem, there is a closed set $A \subset C$ with $\mathcal{L}^m(C \setminus A) < \epsilon/2$ such that $r_i \to 0$ uniformly on compact subsets of A. Now extend $f|_A$ to $\mathbb{R}^m$ by means of Theorem 5.1. $\square$

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

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

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

6.1 Theorem (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

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

Compare [Kir, Proposition 1], [BurBI, Theorem 2.7.6], and [AmbT, Theorem 4.1.1]. The proof given below is taken from [AmbT].

Proof: Choose a dense sequence $(y_j)_{j\in\mathbb{N}}$ in $\gamma([a,b])$ and define $r_j\colon [a,b] \to \mathbb{R}$, $r_j(t) = d(\gamma(t), y_j)$. Note that $\mathrm{Lip}(r_j) \le \mathrm{Lip}(\gamma)$ since

$$|r_j(s) - r_j(t)| \le d(\gamma(s),\gamma(t)).$$

Hence, for almost every $t \in [a,b]$, the derivative $\dot r_j(t)$ exists for all j, and

$$\liminf_{h\to 0} \frac{d(\gamma(t+h),\gamma(t))}{|h|} \ge \sup_j \lim_{h\to 0} \frac{|r_j(t+h) - r_j(t)|}{|h|} = \sup_j |\dot r_j(t)|.$$

On the other hand, whenever $a \le t < t+h \le b$, then

$$d(\gamma(t+h),\gamma(t)) = \sup_j |r_j(t+h) - r_j(t)| \le \int_t^{t+h} \sup_j |\dot r_j(\tau)|\,d\tau;$$

for the first step, compare Proposition 1.1. Hence

$$\limsup_{h\to 0} \frac{d(\gamma(t+h),\gamma(t))}{|h|} \le \limsup_{h\to 0} \frac{1}{|h|} \int_t^{t+h} \sup_j |\dot r_j(\tau)|\,d\tau.$$

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

The above argument also shows that

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

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 $t_k := a + kh$, $k = 0,1,\dots,N$. Then

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

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 $\int_a^b |\dot{\gamma}|(t)\,dt \le L(\gamma)$. $\square$

6.2 Definition (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* $\mathrm{md}\, f_x$.

6.3 Theorem (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 [Kir] and [KorS]. For the proof we need the following technical proposition.

6.4 Proposition
Let $f\colon U \to Y$ be given as in Theorem 6.3.
(1) *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 $v \in \mathbb{R}^m$. Then $\mathcal{L}^m(U \setminus B) = 0$. For every $x \in B$, $\sigma_x\colon \mathbb{R}^m \to \mathbb{R}$ is $\mathrm{Lip}(f)$*-Lipschitz, and*

$$\sigma_x(sv) = |s|\sigma_x(v) \quad \forall v \in \mathbb{R}^m,\ s \in \mathbb{R}.$$

(2) *There exists a sequence of compact sets $K_1, K_2, \ldots \subset B$ with $\mathcal{L}^m(B \setminus \bigcup_{j=1}^\infty K_j) = 0$ and with 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 $x \in K_j$, $v, w \in \mathbb{R}^m$, $|v|, |w| \le \delta$, and $x + w \in K_j$.

The proof of (2) given below is taken from [Wen].

Proof: (1): Choose a countable dense set $D \subset \mathrm{S}^{m-1}$. Using Theorem 6.1, we conclude similarly as in the first part of the proof of Rademacher's Theorem 4.3 that there exists a measurable set $B \subset U$ with $\mathcal{L}^m(U \setminus B) = 0$ such that the limit

$$\sigma_x(u) = \lim_{t \to 0} \frac{d(f(x+tu), f(x))}{|t|}$$

exists whenever $x \in B$ and $u \in D$.

Let $x \in B$. For a fixed $t \neq 0$, the map $\mathbb{R}^m \ni v \mapsto d(f(x+tv), f(x))/|t|$ is $\mathrm{Lip}(f)$-Lipschitz. It follows that the limit $\sigma_x(u)$ exists uniformly for all $u \in \mathrm{S}^{m-1}$. (Compare Lemma 4.2.) The existence of $\sigma_x(u)$ also implies the existence of $\sigma_x(ru)$ for all $r \neq 0$, and

$$\sigma_x(sv) = |s|\sigma(v)$$

for all $v \in \mathbb{R}^m$ and $s \in \mathbb{R}$. Moreover, $\sigma_x \colon \mathbb{R}^m \to \mathbb{R}$ is $\mathrm{Lip}(f)$-Lipschitz.

(2): Now consider the map $\sigma \colon B \to C(\mathrm{S}^{m-1})$, $x \mapsto \sigma_x$, where $C(\mathrm{S}^{m-1})$ denotes the space of continuous real-valued functions on S^{m-1}, endowed with the supremum norm. 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

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

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

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

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 $|v|, |w| \le \delta$, $v \neq w$, 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

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

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.,

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

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

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

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 $v, w \in \mathbb{R}^m$ we have

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

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

7 Hausdorff measures

For $m \in \mathbb{N}$, denote by

$$\alpha_m := \mathcal{L}^m(\mathrm{B}(0,1)) = \frac{\pi^{m/2}}{\Gamma(\frac{m}{2} + 1)}$$

the Lebesgue measure of the unit ball in $\mathbb{R}^m$, and put $\alpha_0 := 1$. Now let X be a metric space. For $m \ge 0$, $0 < \delta \le \infty$, and $A \subset X$, define

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

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 $(\operatorname{diam}\emptyset)^m = 0$, $0^0 = 1$ are used.) Then

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

is the *m-dimensional Hausdorff measure* of the set A. For every m, $\mathcal{H}^m$ is a Borel regular metric outer measure on X. 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

$$\mathcal{H}^m(f(A)) \le \operatorname{Lip}(f)^m \mathcal{H}^m(A).$$

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

$$\Theta^{*m}(A,x) = \limsup_{r\to 0+} \frac{\mathcal{H}^m(A \cap \mathrm{B}(x,r))}{\alpha_m r^m},$$
$$\Theta^m_*(A,x) = \liminf_{r\to 0+} \frac{\mathcal{H}^m(A \cap \mathrm{B}(x,r))}{\alpha_m r^m},$$

respectively. If the two coincide, then the common value $\Theta^m(A,x)$ is the *density* of A at x. If $A, B \subset X$ are two $\mathcal{H}^m$-measurable sets with $A \subset B$ and $\mathcal{H}^m(B) < \infty$, then

$$2^{-m} \le \Theta^{*m}(B,x) \le 1$$

for $\mathcal{H}^m$-almost all $x \in B$,

$$\Theta^m(B,x) = 0$$

for $\mathcal{H}^m$-almost all $x \in X \setminus B$, and

$$\Theta^{*m}(A,x) = \Theta^{*m}(B,x), \quad \Theta^m_*(A,x) = \Theta^m_*(B,x)$$

for $\mathcal{H}^m$-almost all $x \in A$. (See e.g. [Mat, Theorem 6.2 and Corollary 6.3].)

We remark that if $\|\cdot\|$ is a norm on $\mathbb{R}^m$ with norm ball $B_{\|\cdot\|} := B_{\|\cdot\|}(0,1)$, then

$$\mathcal{H}^m_{\|\cdot\|}(B_{\|\cdot\|}) = \alpha_m$$

(cf. [Kir, Lemma 6]). The inequality $\mathcal{H}^m_{\|\cdot\|}(B_{\|\cdot\|}) \le \alpha_m$ follows from the fact that the quotient $\mathcal{H}^m_{\|\cdot\|}(B_{\|\cdot\|}(x,r))/r^m$ is constant for all $x \in \mathbb{R}^m$ and $r > 0$ and, hence, is less than or equal to α_m since $\Theta^{*m}(\mathbb{R}^m_{\|\cdot\|}, x) \le 1$. The reverse inequality is a consequence of the isodiametric inequality in $\mathbb{R}^m_{\|\cdot\|}$, cf. [BurZ, Theorem 11.2.1], which states that

$$\mathcal{L}^m(C) \le \mathcal{L}^m(B_{\|\cdot\|})\big(\tfrac{1}{2}\operatorname{diam}_{\|\cdot\|} C\big)^m \quad \forall C \subset \mathbb{R}^m.$$

It implies that for every covering $(C_i)_{i\in\mathbb{N}}$ of $B_{\|\cdot\|}$, $\sum_{i=1}^\infty (\frac{1}{2}\operatorname{diam}_{\|\cdot\|} C_i)^m \ge 1$, thus $\mathcal{H}^m_{\|\cdot\|}(B_{\|\cdot\|}) \ge \alpha_m$.

8 Area formula

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

8.1 Lemma (Borel partition)
Suppose 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 $\operatorname{md} f_x$ *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 $\|\cdot\|_i$ on $\mathbb{R}^m$ such that*

$$\lambda^{-1}\|x-x'\|_i \le d(f(x),f(x')) \le \lambda\|x-x'\|_i,$$
$$\lambda^{-1}\|v\|_i \le \operatorname{md} f_x(v) \le \lambda\|v\|_i$$

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 $\|\cdot\|_i$ may be chosen to be euclidean, i.e., induced by an inner product, cf. [Fed, Lemma 3.2.2], [EvaG, p. 94].

Proof: Choose a sequence of norms $\|\cdot\|_i$ on $\mathbb{R}^m$ such that for every norm $\|\cdot\|$ on $\mathbb{R}^m$ and for every $\mu > 1$ there is an $i \in \mathbb{N}$ with

$$\mu^{-1}\|v\|_i \le \|v\| \le \mu\|v\|_i \quad \forall v \in \mathbb{R}^m.$$

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 $x \in B$ with

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

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

To see that the $B_{i,k}$ cover B, let $x \in B$, choose $i \in \mathbb{N}$ such that (i) holds, let $c_i > 0$ be such that $|v| \le c_i\|v\|_i$ for all $v \in \mathbb{R}^m$, and pick $k \in \mathbb{N}$ such that (ii) holds with $\delta\|v\|_i$ replaced by $(\delta/c_i)|v|$; then $x \in B_{i,k}$. Now if $C \subset B_{i,k}$ is a set with $\operatorname{diam} C \le 1/k$, then

$$d(f(x+v),f(x)) \le \operatorname{md} f_x(v) + \delta\|v\|_i \le \lambda\|v\|_i,$$
$$d(f(x+v),f(x)) \ge \operatorname{md} f_x(v) - \delta\|v\|_i \ge \lambda^{-1}\|v\|_i$$

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

8.2 Definition (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 $\mathbf{J}(L)$ *of L is the number satisfying*

$$\mathcal{H}^m(L(A)) = \mathbf{J}(L)\mathcal{H}^m(A) \quad \forall A \subset X.$$

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

$$\mathcal{H}^m_\sigma(A) = \mathbf{J}(\sigma)\mathcal{L}^m(A) \quad \forall A \subset \mathbb{R}^m$$

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 $f(A)$.

8.3 Theorem (area formula)
Suppose 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*

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

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

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

See [Kir]. Note that $\# = \mathcal{H}^0$. In the case $Y = \mathbb{R}^n$ we obtain the classical area formula as $\mathbf{J}(\operatorname{md} f_x)$ coincides with $\mathbf{J}(Df_x)$ for $\mathcal{L}^m$-almost every $x \in \mathbb{R}^m$. Cf. [Fed, Theorem 3.2.3], [EvaG, 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: (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

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

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 $\operatorname{md} f_x$ is a norm, i.e. $\mathbf{J}(\operatorname{md} f_x) > 0$.

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 $\|\cdot\|_i$ on $\mathbb{R}^m$ such that $f|_{A_i}$ is injective,

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

and $\lambda^{-1}\|\cdot\|_i \le \operatorname{md} f_x \le \lambda\|\cdot\|_i$ for all $x \in A_i$. This last assertion yields

$$\lambda^{-m}\mathbf{J}(\|\cdot\|_i) \le \mathbf{J}(\operatorname{md} f_x) \le \lambda^m \mathbf{J}(\|\cdot\|_i).$$

Since $f|_{A_i}$ is injective, $y \mapsto \#(f^{-1}\{y\} \cap A_i)$ is the characteristic function of the set $f(A_i)$, which is $\mathcal{H}^m$-measurable. We conclude that

$$\begin{aligned}\int_Y \#(f^{-1}\{y\} \cap A_i)\, d\mathcal{H}^m(y) &= \mathcal{H}^m(f(A_i)) \\ &\le \lambda^m \mathcal{H}^m_{\|\cdot\|_i}(A_i) \\ &= \lambda^m \mathbf{J}(\|\cdot\|_i)\mathcal{L}^m(A_i) \\ &\le \lambda^{2m} \int_{A_i} \mathbf{J}(\mathrm{md}\, f_x)\, dx.\end{aligned}$$

Similarly, we obtain the estimate

$$\int_Y \#(f^{-1}\{y\} \cap A_i)\, d\mathcal{H}^m(y) \ge \lambda^{-2m} \int_{A_i} \mathbf{J}(\mathrm{md}\, f_x)\, dx.$$

As $\lambda > 1$ was arbitrary, the two integrals are equal. This shows that (1) holds for each A_i and hence 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 $\epsilon > 0$, and define

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

Equip $Y \times \mathbb{R}^m$ with the l^1 product metric $d_1((y,z),(y',z')) = d(y,y') + |z - z'|$. Clearly, g is metrically differentiable at x whenever f is, and

$$\mathrm{md}\, g_x(v) = \mathrm{md}\, f_x(v) + \epsilon|v| \quad \forall v \in \mathbb{R}^m.$$

In particular, $\mathrm{md}\, g_x$ is a norm on $\mathbb{R}^m$ for every $x \in A''$. Fix $x \in A''$ for the moment, and put $\|\cdot\| := \mathrm{md}\, g_x$. Consider the norm ball $B := \mathrm{B}_{\|\cdot\|}(0,1) \subset \mathbb{R}^m$. As remarked earlier, $\mathcal{H}^m_{\|\cdot\|}(B) = \alpha_m$. Since $\mathrm{md}\, f_x$ is not a norm, there is a $v_0 \in \mathbb{R}^m$ with $\|v_0\| = 1$ and $\mathrm{md}\, f_x(v_0) = 0$, hence $|v_0| = 1/\epsilon$. Moreover, for every v with $\|v\| = 1$, we have $1 = \mathrm{md}\, f_x(v) + \epsilon|v| \le (\mathrm{Lip}(f) + \epsilon)|v|$, thus $|v| \ge 1/(\mathrm{Lip}(f) + \epsilon)$. In other words, the convex set B contains $\{v_0, -v_0\} \cup \mathrm{B}(0, 1/(\mathrm{Lip}(f) + \epsilon))$, where $|v_0| = 1/\epsilon$. It follows that

$$\mathcal{L}^m(B) \ge \frac{c_m}{\epsilon(\mathrm{Lip}(f) + \epsilon)^{m-1}}$$

for some constant c_m depending only on m, hence

$$\mathbf{J}(\mathrm{md}\, g_x) = \mathbf{J}(\|\cdot\|) = \frac{\mathcal{H}^m_{\|\cdot\|}(B)}{\mathcal{L}^m(B)} \le \frac{\alpha_m \epsilon(\mathrm{Lip}(f) + \epsilon)^{m-1}}{c_m}.$$

Applying the above result for (f, A') to (g, A''), we get

$$\mathcal{H}^m(g(A'')) = \int_{A''} \mathbf{J}(\mathrm{md}\, g_x)\, dx \le \frac{\alpha_m \epsilon(\mathrm{Lip}(f) + \epsilon)^{m-1}}{c_m} \mathcal{L}^m(A'').$$

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

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

9 Rectifiable sets

The following notion is fundamental in geometric measure theory.

9.1 Definition (countably rectifiable set)
Let Y be a metric space, and let $E \subset Y$.
(1) *The set E is* countably m-rectifiable *if there is a countable family of Lipschitz maps $f_i \colon A_i \to Y$, $A_i \subset \mathbb{R}^m$, such that $E \subset \bigcup_i f_i(A_i)$.*
(2) *The set E is* countably $\mathcal{H}^m$-rectifiable *if there is a countably m-rectifiable set $E' \subset Y$ such that $\mathcal{H}^m(E \setminus E') = 0$.*

In (1), 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).

9.2 Proposition (bi-Lipschitz parametrization)
Suppose Y is a locally complete 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 $f_i \colon C_i \to f_i(C_i) \subset E$, with $C_i \subset \mathbb{R}^m$ compact, such that the $f_i(C_i)$ are pairwise disjoint and

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

Compare [Fed, Lemma 3.2.18] and [AmbK2, Lemma 4.1]. In case $Y = \mathbb{R}^n$ it is possible to choose all f_i to be λ-bi-Lipschitz, for any given $\lambda > 1$.

Proof: First we assume that $E \subset Y$ is a Borel set contained in the image of a single Lipschitz map $h \colon A \to Y$, where $A \subset \mathbb{R}^m$ is compact. Consider an isometric embedding $Y \subset \bar{Y} := l^\infty(Y)$, and extend h to a Lipschitz map $\bar{h} \colon \mathbb{R}^m \to \bar{Y}$. Since $h(A)$ is closed in $\bar{Y}$, $\bar{h}^{-1}(E)$ is a Borel set. Using Lemma 8.1 (Borel partition), Theorem 8.3 (area formula), and the inner regularity of $\mathcal{L}^m$, we find a sequence of bi-Lipschitz maps $g_k \colon D_k \to g_k(D_k) \subset E$, with $D_k \subset \mathbb{R}^m$ compact, such that $\mathcal{H}^m(E \setminus \bigcup_k g_k(D_k)) = 0$. Consider the Borel sets

$$D_k' := D_k \setminus g_k^{-1}\big(\textstyle\bigcup_{j=1}^{k-1} g_j(D_j)\big).$$

Then $\bigcup_k g_k(D_k') = \bigcup_k g_k(D_k)$, and the $g_k(D_k')$ are pairwise disjoint. For every k, choose a sequence of pairwise disjoint compact sets $C_{k,l} \subset D_k'$ such that $\mathcal{L}^m(D_k' \setminus \bigcup_l C_{k,l}) = 0$. It follows that $\mathcal{H}^m(E \setminus \bigcup_{k,l} g_k(C_{k,l})) = 0$, and the $g_k(C_{k,l})$ are pairwise disjoint.

To prove the general result, partition $\mathcal{H}^m$-almost all of E into a sequence of $\mathcal{H}^m$-measurable sets E_j with $E_j \subset h_j(A_j)$ for some Lipschitz map $h_j\colon A_j \to Y$, where $A_j \subset \mathbb{R}^m$ is compact. This is possible since Y is locally complete, thus every Lipschitz map $h\colon A \to Y$ extends locally to the closure of $A \subset \mathbb{R}^m$. Then $\mathcal{H}^m(E_j) < \infty$, and by the regularity of $\mathcal{H}^m$ there is an F_σ set $F_j \subset E_j$ with $\mathcal{H}^m(E_j \setminus F_j) = 0$. Now apply the above result to each F_j. □

9.3 Proposition (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

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

See [Fed, Theorem 3.2.29], [Sim, Lemma 11.1].

Proof: Suppose that $\mathcal{H}^m(E \setminus \bigcup_i f_i(\mathbb{R}^m)) = 0$ for a sequence of Lipschitz maps $f_i\colon \mathbb{R}^m \to \mathbb{R}^n$. By Theorem 5.2, we assume w.l.o.g. 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. □

For $x \in \mathbb{R}^n$ and $r > 0$, define $T_{x,r}\colon \mathbb{R}^n \to \mathbb{R}^n$, $T_{x,r}(y) = (y - x)/r$. Note that $T_{x,r}$ maps $\mathrm{B}(x, r)$ onto $\mathrm{B}(0, 1)$.

9.4 Definition (approximate tangent space)
*Suppose $E \subset \mathbb{R}^n$ 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*

$$\lim_{r \to 0+} \int_{T_{x,r}(E)} \varphi \, d\mathcal{H}^m = \int_L \varphi \, d\mathcal{H}^m$$

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

Clearly $\mathrm{Tan}^m(E, x)$ is uniquely determined if it exists. There are various definitions of approximate tangent spaces in the literature, compare [Sim, Definition 11.2], [Fed, 3.2.16], and [Mat, Definition 15.17].

9.5 Theorem (existence of approximate tangent spaces)
Suppose $E \subset \mathbb{R}^n$ 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 $\Theta^m(E, x) = 1$.

For this result and Theorem 9.6(1) below, see [Fed, Theorem 3.2.19], [Sim, Theorem 11.6], and [Mat, Theorem 15.19].

Proof: Choose a sequence of m-dimensional C^1 submanifolds M_k of $\mathbb{R}^n$ such that $\mathcal{H}^m(E \setminus \bigcup_k M_k) = 0$, cf. Proposition 9.3. Put $E_k := E \cap M_k$; then $\mathcal{H}^m(E \setminus \bigcup_k E_k) = 0$. Since M_k is C^1, it follows that for $\mathcal{H}^m$-almost every $x \in E_k$, we have $\Theta^m(E_k, x) = 1$ and $\mathrm{Tan}^m(E_k, x) = T_x M_k$. Moreover, for $\mathcal{H}^m$-almost every $x \in E_k$, $\Theta^m(E \setminus E_k, x) = 0$. Combining these two properties we conclude that for $\mathcal{H}^m$-almost every $x \in E_k$, $\Theta^m(E, x) = 1$ and $\mathrm{Tan}^m(E, x) = T_x M_k$. □

The following two converses to Theorem 9.5 hold. The second is a deep result of Preiss [Pre]; an account of the theorem and its history is given in [Mat, Sect. 17].

9.6 Theorem (rectifiability criteria)
Suppose $E \subset \mathbb{R}^n$ is 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 very important role in the development of the theory of currents. This deep result was proved in [Bes] for $m = 1$ and $n = 2$ and in [Fed0] for general dimensions. See [Fed, Theorem 3.3.13] and [Mat, Theorem 18.1]. A set $F \subset \mathbb{R}^n$ is *purely $\mathcal{H}^m$-unrectifiable* if $\mathcal{H}^m(F \cap f(\mathbb{R}^m)) = 0$ for every Lipschitz map $f \colon \mathbb{R}^m \to \mathbb{R}^n$. Every set $A \subset \mathbb{R}^n$ 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 (cf. [Mat, Theorem 15.6]).

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

10 Coarea formula

For the proof of the coarea formula, Theorem 10.4, we need the following general coarea inequality, which is also of independent interest.

10.1 Theorem (coarea inequality)
Suppose that X, Y are metric spaces, $f \colon X \to Y$ is Lipschitz, $A \subset X$, and $m, k \geq 0$. Then

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

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 (i.e. closed bounded subsets of X are compact) and A is $\mathcal{H}^{k+m}$-measurable with $\mathcal{H}^{k+m}(A) < \infty$, 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, cf. [Fed, 2.10.26]. Theorem 10.1 is stated with some additional assumptions in [Fed, Theorem 2.10.25]. As remarked in [Dav, p. 236], these are superfluous. We prove the result in the case Y is an m-dimensional normed space.

Proof: For every $j \in \mathbb{N}$, choose a covering $(C_n^j)_{n \in \mathbb{N}}$ of A such that $\operatorname{diam} C_n^j \le 1/j$ for all n and

$$\sum_{n=1}^{\infty} \alpha_{k+m} \big(\tfrac{1}{2} \operatorname{diam} C_n^j\big)^{k+m} \le \mathcal{H}^{k+m}_{1/j}(A) + 1/j.$$

Denote by D_n^j the closure of $f(C_n^j)$, and by g_n^j the characteristic function of D_n^j multiplied by $\alpha_k(\frac{1}{2} \operatorname{diam} C_n^j)^k$. For all $y \in Y$ and $i \le j$ we have

$$\mathcal{H}^k_{1/i}\big(f^{-1}\{y\} \cap A\big) \le \sum_{n=1}^{\infty} g_n^j(y).$$

Using the isodiametric inequality in Y (see p. 16) we obtain

$$\begin{aligned}
\int_Y^* \mathcal{H}^k_{1/i}\big(f^{-1}\{y\} \cap A\big)\, d\mathcal{H}^m(y) &\le \int_Y \sum_{n=1}^{\infty} g_n^j(y)\, d\mathcal{H}^m(y) \\
&= \sum_{n=1}^{\infty} \int_Y g_n^j(y)\, d\mathcal{H}^m(y) \\
&= \sum_{n=1}^{\infty} \alpha_k \big(\tfrac{1}{2} \operatorname{diam} C_n^j\big)^k \mathcal{H}^m(D_n^j) \\
&\le \sum_{n=1}^{\infty} \alpha_k \big(\tfrac{1}{2} \operatorname{diam} C_n^j\big)^k \alpha_m \big(\tfrac{1}{2} \operatorname{diam} D_n^j\big)^m \\
&\le \sum_{n=1}^{\infty} \alpha_k \alpha_m \operatorname{Lip}(f)^m \big(\tfrac{1}{2} \operatorname{diam} C_n^j\big)^{k+m} \\
&\le \frac{\alpha_k \alpha_m}{\alpha_{k+m}} \operatorname{Lip}(f)^m \big(\mathcal{H}^{k+m}_{1/j}(A) + 1/j\big).
\end{aligned}$$

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

10.2 Lemma (factorization)
Suppose $f \colon \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, $n \ge m$, 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 a family $(A_i)_{i \in \mathbb{N}}$ 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 are assuming 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 $(A_i)_{i \in \mathbb{N}}$ of A such that each $u|_{A_i}$ 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 $(u|_{A_i})^{-1}$. There is a measurable set $B_i \subset A_i$ with $\mathcal{L}^n(A_i \backslash B_i) = 0$ such that for all $x \in B_i$, v_i is differentiable at $u(x)$ and the differential is the inverse of Du_x. Let $\lambda > 1$. Applying Lemma 8.1 to v_i, we find a measurable partition $(C_{i,k})_{k \in \mathbb{N}}$ of $u(B_i)$ and a sequence of linear isomorphisms $T_{i,k} \colon \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^n$ such that

$$\begin{aligned} \lambda^{-1}|T_{i,k}(y,z) - T_{i,k}(y',z')| &\le |v_i(y,z) - v_i(y',z')| \\ &\le \lambda |T_{i,k}(y,z) - T_{i,k}(y',z')| \end{aligned}$$

and $\lambda^{-1}|T_{i,k}(\cdot)| \le |D(v_i)_{(y,z)}(\cdot)| \le \lambda |T_{i,k}(\cdot)|$ for all $(y,z),(y',z') \in C_{i,k}$. It follows that the restriction of

$$h_{i,k} := T_{i,k} \circ u$$

to $v_i(C_{i,k})$ is λ-bi-Lipschitz, and $D(h_{i,k})_x$ is λ-bi-Lipschitz for all $x \in v_i(C_{i,k})$ as well. Finally, we define $q \colon \mathbb{R}^m \times \mathbb{R}^{n-m} \to \mathbb{R}^m$ by $q(y,z) = y$ and put

$$L_{i,k} := q \circ T_{i,k}^{-1}.$$

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

10.3 Definition (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 $\mathcal{H}^n(A) < \infty$.

See [AmbK1, Sect. 9]. Note that the right side is invariant under translations of A and, by Theorem 10.1 (coarea inequality), less than or equal to $(\alpha_{n-m}\alpha_m/\alpha_n)\,\mathrm{Lip}(L)^m\mathcal{H}^n(A)$. 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 $Z \subset X$ complementary to the kernel $\ker L$ and a set A of the form $A = B + C$ for $B \subset \ker L$ and $C \subset Z$, 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 $\mathcal{H}^m(L(C)) = \mathbf{J}(L|_Z)\mathcal{H}^m(C)$, this yields the identity

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

In the case X is a euclidean space, it follows that

$$\mathbf{C}_m(L) = \mathbf{J}\big(L|_{(\ker L)^\perp}\big) \geq \mathbf{J}(L|_Z),$$

where $(\ker L)^\perp$ 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 $Z := H^{-1}((\ker L)^\perp)$. Then

$$\mathbf{C}_m(L \circ H) \geq \mathbf{J}(L \circ H|_Z) = \mathbf{J}(H|_Z)\mathbf{J}\big(L|_{(\ker L)^\perp}\big) \geq \lambda^{-m}\mathbf{C}_m(L),$$

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.

10.4 Theorem (coarea formula)
Suppose $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, $n \geq m \geq 1$.
(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: (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 10.1 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'. Using Lemma 10.2 we choose a family $(A_i)_{i\in\mathbb{N}}$ 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 $\mathcal{L}^n(A' \setminus \bigcup A_i) = 0$, $h_i|_{A_i}$ is λ-bi-Lipschitz, $f = L_i \circ h_i$, and $D(h_i)_x$ exists and is λ-bi-Lipschitz for all $x \in A_i$. We have $Df_x = L_i \circ D(h_i)_x$ and hence

$$\lambda^{-m}\mathbf{C}_m(L_i) \le \mathbf{C}_m(Df_x) \le \lambda^m \mathbf{C}_m(L_i)$$

for all $x \in A_i$. We conclude that

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

Similarly, we obtain the estimate

$$\int_{\mathbb{R}^m} \mathcal{H}^{n-m}(f^{-1}\{y\} \cap A_i)\, dy \ge \lambda^{-2n} \int_{A_i} \mathbf{C}_m(Df_x)\, dx.$$

As $\lambda > 1$ was arbitrary, the two integrals are equal. This shows that (1) holds for each 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

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

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

$$Dg_{(x,z)}(v,w) = Df_x(v) + \epsilon w \quad \forall (v,w) \in \mathbb{R}^n \times \mathbb{R}^m.$$

In particular, $Dg_{(x,z)}$ has rank m, and $\dim(\ker Dg_{(x,z)}) = n$. We have

$$\dim\big(\ker Dg_{(x,z)} \cap (\mathbb{R}^n \times \{0\})\big) = \dim(\ker Df_x) \ge n - (m-1),$$

hence $\dim W \le m-1$ for $W := p(\ker Dg_{(x,z)})$ and $\dim W^\perp \ge 1$ for the orthogonal complement of W in $\mathbb{R}^m$. Since $\{0\}\times W^\perp \subset Z := (\ker Dg_{(x,z)})^\perp$, it follows that

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

Let $C := [0,1]^m \subset \mathbb{R}^m$. Using Fubini's theorem and Theorem 10.1 (coarea inequality), we obtain

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

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

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

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

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

10.5 Theorem (rectifiable level sets)
Suppose X is a metric space, $n \ge k \ge 1$, $E \subset X$ is countably $\mathcal{H}^n$-rectifiable, and $f\colon E \to \mathbb{R}^k$ is Lipschitz. Then for $\mathcal{L}^k$-almost every $y \in \mathbb{R}^k$, $f^{-1}\{y\}$ is countably $\mathcal{H}^{n-k}$-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 k. Choose a Borel partition $(B_i)_{i\in\mathbb{N}}$ of B and coordinate projections $p_i\colon \mathbb{R}^n \to \mathbb{R}^{n-k}$ such that each $u_i|_{B_i}$ is bi-Lipschitz, where $u_i = (f, p_i)$, cf. the first part of the proof of Lemma 10.2. For all $y \in \mathbb{R}^k$,

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

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

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

This shows that for $\mathcal{L}^k$-almost every $y \in \mathbb{R}^k$, $f^{-1}\{y\}$ is countably $\mathcal{H}^{n-k}$-rectifiable.

Now consider the general case. Choose Lipschitz maps $h_i : A_i \to E$, $A_i \subset \mathbb{R}^n$, such that $\mathcal{H}^n(E \setminus \bigcup_{i=1}^\infty h_i(A_i)) = 0$. For each i, put $f_i := f \circ h_i : A_i \to \mathbb{R}^k$ and pick a Lipschitz extension $\bar{f}_i : \mathbb{R}^n \to \mathbb{R}^k$. For $\mathcal{L}^k$-almost every $y \in \mathbb{R}^k$, we know that $\bar{f}_i^{-1}\{y\}$ is countably $\mathcal{H}^{n-k}$-rectifiable, thus

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

is countably $\mathcal{H}^{n-k}$-rectifiable. Moreover, for $\mathcal{L}^k$-almost every $y \in \mathbb{R}^k$, $\mathcal{H}^{n-k}(f^{-1}\{y\} \setminus \bigcup_{i=1}^\infty h_i(A_i)) = 0$ since

$$\begin{aligned} &\int_{\mathbb{R}^k}^* \mathcal{H}^{n-k}(f^{-1}\{y\} \setminus \textstyle\bigcup_{i=1}^\infty h_i(A_i))\, dy \\ &\qquad \le \frac{\alpha_{n-k}\alpha_k}{\alpha_n} \operatorname{Lip}(f)^k \mathcal{H}^n(E \setminus \textstyle\bigcup_{i=1}^\infty h_i(A_i)) = 0 \end{aligned}$$

by Theorem 10.1 (coarea inequality). □

References

[Alm1] F. J. Almgren, Jr., The homotopy groups of the integral cycle groups, Topology 1 (1962), 257–299.

[AmbK1] L. Ambrosio, B. Kirchheim, Rectifiable sets in metric and Banach spaces, Math. Ann. 318 (2000), 527–555.

[AmbK2] L. Ambrosio, B. Kirchheim, Currents in metric spaces, Acta Math. 185 (2000), 1–80.

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

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

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

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

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

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

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

[Fed0] H. Federer, The (φ, k) rectifiable subsets of n space, Trans. Amer. Math. Soc. 62 (1947), 114–192.

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

[Gro1] M. Gromov, Groups of polynomial growth and expanding maps, Publ. Math. IHES 53 (1981), 53–73.

[Gro2] M. Gromov, with Appendices by M. Katz, P. Pansu, and S. Semmes, Metric Structures for Riemannian and Non-Riemannian Spaces, Birkhäuser 1999.

[Hei1] J. Heinonen, Lectures on Analysis on Metric Spaces, Springer 2001.

[Hei2] J. Heinonen, Geometric embeddings of metric spaces, Rep. Univ. Jyväskylä Dept. Math. Stat., no. 90, 2003.

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

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

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

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

[LanS] U. Lang, V. Schroeder, Kirszbraun's theorem and metric spaces of bounded curvature, Geom. Funct. Anal. 7 (1997), 535–560.

[Lan] U. Lang, Extendability of large-scale Lipschitz maps, Trans. Amer. Math. Soc. 351 (1999), 3975–3988.

[Lan1] U. Lang, Introduction to Geometric Measure Theory, Lecture Notes, 2005, http://www.math.ethz.ch/∼lang/diab05_notes.pdf.

[Mal] J. Maly, A simple proof of the Stepanov theorem on differentiability almost everywhere, Exposition. Math. 17 (1999), 59–61.

[Mat] P. Mattila, Geometry of Sets and Measures in Euclidean Spaces, Fractals and Rectifiability, Cambridge Studies in Advanced Math., Vol. 44, Cambridge U. Press 1995.

[McS] E. J. McShane, Extension of range of functions, Bull. Amer. Math. Soc. 40 (1934), 837–842.

[Rad] H. Rademacher, Über partielle und totale Differenzierbarkeit I, Math. Ann. 79 (1919), 340–359.

[Pre] D. Preiss, Geometry of measures in R^n: distribution, rectifiability, and densities, Ann. of Math. 125 (1987), 537–643.

[Sim] L. Simon, Lectures on Geometric Measure Theory, Proc. Cent. Math. Anal., Vol. 3, Austr. Nat. U., Canberra 1983.

[Ste] E. M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton Univ. Press, New Jersey 1970.

[Step] W. Stepanoff, Über totale Differenzierbarkeit, Math. Ann. 90 (1923), 318–320.

[Val] F. A. Valentine, A Lipschitz condition preserving extension for a vector function, Amer. J. Math. 67 (1945), 83–93.

[Wen] S. Wenger, Characterizations of metric trees and Gromov hyperbolic metric spaces, Math. Res. Lett. 15 (2008), 1017–1026.

[Whit] H. Whitney, Analytic extensions of differentiable functions defined in closed sets, Trans. Amer. Math. Soc. 36 (1934), 63–89.

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