LECTURES ON LIPSCHITZ ANALYSIS

JUHA HEINONEN

1. Introduction

A function $f : A \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, is said to be L-Lipschitz, $L \geq 0$, if

$$|f(a) - f(b)| \leq L\,|a - b| \tag{1.1}$$

for every pair of points $a, b \in A$. We also say that a function is *Lipschitz* if it is L-Lipschitz for some L.

The Lipschitz condition as given in (1.1) is a purely metric condition; it makes sense for functions from one metric space to another. In these lectures, we concentrate on the theory of Lipschitz functions in Euclidean spaces. In Section 2, we study extension problems and Lipschitz retracts. In Section 3, we prove the classical differentiability theorems of Rademacher and Stepanov. In Section 4, we briefly discuss Sobolev spaces and Lipschitz behavior; another proof of Rademacher's theorem is given there based on the Sobolev embedding. Section 5 is the most substantial. Therein we carefully develop the basic theory of flat differential forms of Whitney. In particular, we give a proof of the fundamental duality between flat chains and flat forms. The Lipschitz invariance of flat forms is also discussed. In the last section, Section 6, we discuss some recent developments in geometric analysis, where flat forms are used in the search for Lipschitz variants of the measurable Riemann mapping theorem.

Despite the Euclidean framework, the material in these lectures should be of interest to students of general metric geometry. Many basic results about Lipschitz functions defined on subsets of $\mathbb{R}^n$ are valid in great generality, with similar proofs. Moreover, fluency in the classical theory is imperative in analysis and geometry at large.

Lipschitz functions appear nearly everywhere in mathematics. Typically, the Lipschitz condition is first encountered in the elementary theory of ordinary differential equations, where it is used in existence theorems. In the basic courses on real analysis, Lipschitz functions appear as examples of functions of bounded variation, and it is proved

Lectures at the 14th Jyväskylä Summer School in August 2004.
Supported by NSF grant DMS 0353549 and DMS 0244421.

that a real-valued Lipschitz function on an open interval is almost everywhere differentiable. Among more advanced topics, Lipschitz analysis is extensively used in geometric measure theory, in partial differential equations, and in nonlinear functional analysis. The Lipschitz condition is one of the central concepts of metric geometry, both finite and infinite dimensional. There are also striking applications to topology. Namely, every topological manifold outside dimension four admits a unique Lipschitz structure, while such a manifold may have no smooth or piecewise linear structures or it may have many such. On a more practical side, questions about Lipschitz functions arise in image processing and in the study of internet search engines, for example. Finally, even when one considers rougher objects, such as functions in various Sobolev spaces or quasiconformal mappings, vestiges of Lipschitz behavior are commonly found in them, and the theory is applicable.

In many ways, the Lipschitz condition is more natural, and more ubiquitous, than say the condition of infinite smoothness. For example, families of Lipschitz functions are often (pre-)compact, so that Arzelà-Ascoli type arguments can be applied. Compactness in the smooth context is typically more complicated.

Some of the preceding issues will be studied in these lectures in more detail, while others will only briefly be alluded to. Many important topics are not covered at all.

References to the topics advertized in this introduction include [18], [50], [17], [80], [20], [59], [16], [25], [63], [5], [67], [64], [47], [73], [62], [43].

1.1. **Notation.** Our notation is standard. Euclidean n-space $\mathbb{R}^n$, $n \geq 1$, is equipped with the distance

$$|x-y| := (\sum_{i=1}^{n}(x_i-y_i)^2)^{1/2}$$

unless otherwise stipulated. The Lebesgue n-measure of a measurable set $E \subset \mathbb{R}^n$ is denoted by $|E|$, and integration against Lebesgue measure by

$$\int_E f(x)\,dx\,.$$

Open and closed balls in $\mathbb{R}^n$ are denoted by $B(x,r)$ and $\overline{B}(x,r)$, respectively; here $x \in \mathbb{R}^n$ and $r > 0$. If we need to emphasize the dimension of the underlying space, we write $B^n(x,r)$. We also write $\mathbb{B}^n := B^n(0,1)$ and $\mathbb{S}^{n-1} := \partial\mathbb{B}^n$. The closure of a set $E \subset \mathbb{R}^n$ is $\overline{E}$, and the complement $E^c := \mathbb{R}^n \setminus E$.

Other standard or self-explanatory notation will appear.

1.2. **Acknowledgements.** I thank the organizers of the 14th Jyväskylä Summer School, especially Professors Tero Kilpeläinen and Raimo Näkki, for inviting me to give these lectures. I am grateful to Eero Saksman for many illuminating conversations about the Whitney theory. I also thank Ole Jacob Broch, Bruce Kleiner, and Peter Lindqvist for some useful information, and Bruce Hanson, Leonid Kovalev, Seppo Rickman and Jussi Väisälä for carefully reading the manuscript and for their comments.

2. Extension

Every Lipschitz function $f : A \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, can be extended to a Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}^m$. This means that F is Lipschitz and $F|A = f$. In this section, we offer three proofs of this fundamental result, and discuss the related problem of Lipschitz retracts. The development of this section reveals the great flexibility afforded by Lipschitz functions; they can be glued, pasted, and truncated without impairing the Lipschitz property.

We begin with some preliminaries.

2.1. **Distance functions and quasiconvexity.** Distance functions are simple but important examples of Lipschitz functions. The distance can be taken either to a point $x_0 \in \mathbb{R}^n$,

$$x \mapsto \operatorname{dist}(x, x_0) := |x - x_0| \,, \tag{2.1}$$

or more generally to a set $E \subset \mathbb{R}^n$,

$$x \mapsto \operatorname{dist}(x, E) := \inf\{|x - a| : a \in E\} \,. \tag{2.2}$$

That $\operatorname{dist}(\cdot, x_0)$ is 1-Lipschitz is a direct consequence of the triangle inequality. It is similarly straightforward to check from the definitions that the function $\operatorname{dist}(\cdot, E)$ in (2.2) is 1-Lipschitz, but it is worthwhile to record the following general fact.

Lemma 2.1. *Let $\{f_i : i \in I\}$ be a collection of L-Lipschitz functions $f_i : A \to \mathbb{R}$, $A \subset \mathbb{R}^n$. Then the functions*

$$x \mapsto \inf_{i \in I} f_i(x) \,, \qquad x \in A \,,$$

and

$$x \mapsto \sup_{i \in I} f_i(x) \,, \qquad x \in A \,,$$

are L-Lipschitz on A, if finite at one point.

The proof of Lemma 2.1 is an easy exercise.

Note that the set E in (2.2) is not assumed to be closed. On the other hand, we have that $\operatorname{dist}(x, E) = \operatorname{dist}(x, \overline{E})$. Therefore, one typically considers closed sets E in this connection. More generally, every L-Lipschitz function $f : A \to \mathbb{R}^m$ extends to an L-Lipschitz function defined on the closure $\overline{A}$, simply by uniform continuity.

Lipschitz condition (1.1) is global; it requires control over each pair of points a, b in A. Sometimes we only have local information. There is a simple but useful lemma which shows that under special circumstances local information can be turned into global.

A set $A \subset \mathbb{R}^n$ is said to be C-*quasiconvex*, $C \geq 1$, if every pair of points $a, b \in A$ can be joined by a curve γ in A such that

$$\operatorname{length}(\gamma) \leq C\,|a - b|\,. \tag{2.3}$$

We also say that A is *quasiconvex* if it is C-quasiconvex for some $C \geq 1$. By the *length* of a curve γ we mean as usual the quantity,

$$\operatorname{length}(\gamma) := \sup \sum_{i=0}^{N-1} |\gamma(t_{i+1}) - \gamma(t_i)|\,,$$

where the supremum is taken over all partitions $0 = t_0 < t_1 < \cdots < t_N = 1$ for a curve $\gamma : [0,1] \to \mathbb{R}^n$.

A function $f : A \to \mathbb{R}^m$ is called *locally L-Lipschitz* if every point in A has a neighborhood on which f is L-Lipschitz.

Lemma 2.2. *If $A \subset \mathbb{R}^n$ is C-quasiconvex and $f : A \to \mathbb{R}^m$ is locally L-Lipschitz, then f is CL-Lipschitz.*

We leave the straightforward proof of this lemma to the reader. Now consider the "slit plane",

$$A := \{(r, \theta) : 0 < r < \infty, -\pi < \theta < \pi\} \subset \mathbb{R}^2\,,$$

in polar coordinates. The function

$$(r, \theta) \mapsto (r, \theta/2)\,, \quad A \to \mathbb{R}^2\,,$$

is locally 1-Lipschitz, but not globally Lipschitz. This example shows the relevance of quasiconvexity in the situation of Lemma 2.2.

The distance function in (2.1) can be defined by using the intrinsic metric of a set. Let $A \subset \mathbb{R}^n$ be a set such that every pair of points in A can be joined by a curve of finite length in A. The *intrinsic metric* δ_A in A is defined as

$$\delta_A(a, b) := \inf \operatorname{length}(\gamma)\,, \tag{2.4}$$

where the infimum is taken over all curves γ joining a and b in A. Expression (2.4) indeed defines a metric in A, and A is quasiconvex if and only if the identity mapping between the two metrics is bi-Lipschitz. We recall here that a homeomorphism between metric spaces is *bi-Lipschitz* if it is Lipschitz and has a Lipschitz inverse.

The function

$$x \mapsto \operatorname{dist}_A(x, x_0) := \delta_A(x, x_0) \tag{2.5}$$

is 1-Lipschitz with respect to the intrinsic metric; it is Lipschitz if A is quasiconvex. We will return to quasiconvexity in connection with Lipschitz retracts later in this section.

Finally, we say that a curve γ in a set A, joining two points a and b, is an *intrinsic geodesic* if $\operatorname{length}(\gamma) = \delta_A(a, b)$.

2.2. **Extension theorems.** We prove the important extension theorems of McShane-Whitney and Kirszbraun.

Theorem 2.3 (McShane-Whitney extension theorem)**.** *Let $f : A \to \mathbb{R}$, $A \subset \mathbb{R}^n$, be an L-Lipschitz function. Then there exists an L-Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}$ such that $F|A = f$.*

Proof. Because the functions

$$f_a(x) := f(a) + L|x - a|\,, \quad a \in A\,,$$

are L-Lipschitz on $\mathbb{R}^n$, the function

$$F(x) := \inf_{a \in A} f_a(x)\,, \qquad F : \mathbb{R}^n \to \mathbb{R}\,,$$

is L-Lipschitz by Lemma 2.1. It is obvious that $F(a) = f(a)$ whenever $a \in A$. □

The extension F in Theorem 2.3 is the largest L-Lipschitz extension of f in the sense that if $G : \mathbb{R}^n \to \mathbb{R}$ is L-Lipschitz and $G|A = f$, then $G \le F$. One can also find the smallest L-Lipschitz extension of f, by setting

$$F(x) := \sup_{a \in A} f(a) - L|x - a|\,, \quad x \in \mathbb{R}^n\,.$$

Corollary 2.4. *Let $f : A \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, be an L-Lipschitz function. Then there exists an $\sqrt{m}L$-Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}^m$ such that $F|A = f$.*

Corollary 2.4 follows by applying Theorem 2.3 to the coordinate functions of f. The multiplicative constant $\sqrt{m}$ in the corollary is in fact redundant, but this is harder to prove.

Theorem 2.5 (Kirszbraun's theorem)**.** *Let $f : A \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, be an L-Lipschitz function. Then there exists an L-Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}^m$ such that $F|A = f$.*

Proof. By dividing the function f by L, we may assume that $f : A \to \mathbb{R}^m$ is 1-Lipschitz.

To prove the theorem, the following is a key lemma.

Lemma 2.6. *If f is an $\mathbb{R}^m$-valued 1-Lipschitz function on a finite set $F \subset \mathbb{R}^n$, and if $x \in \mathbb{R}^n$, then there is an extension of f to an $\mathbb{R}^m$-valued 1-Lipschitz function on $F \cup \{x\}$.*

To prove Lemma 2.6, we consider in turn the following assertion.

Lemma 2.7. *Let $\{x_1, \ldots, x_k\}$ be a finite collection of points in $\mathbb{R}^n$, and let $\{y_1, \ldots, y_k\}$ be a collection of points in $\mathbb{R}^m$ such that*

$$|y_i - y_j| \leq |x_i - x_j| \tag{2.6}$$

for all $i, j \in \{1, \ldots, k\}$. If $r_1, \ldots, r_k$ are positive numbers such that

$$\bigcap_{i=1}^{k} \overline{B}(x_i, r_i) \neq \emptyset\,,$$

then

$$\bigcap_{i=1}^{k} \overline{B}(y_i, r_i) \neq \emptyset\,.$$

Let us first prove Lemma 2.6 by the aid of Lemma 2.7. Indeed, let $F = \{x_1, \ldots, x_k\} \subset \mathbb{R}^n$, let $f : F \to \mathbb{R}^m$ be a 1-Lipschitz map, and let $x \in \mathbb{R}^n$. Set $r_i := |x - x_i|$ and $y_i := f(x_i)$. By Lemma 2.7, there exists a point $y \in \mathbb{R}^m$ such that $|y - f(x_i)| \leq |x - x_i|$ for each i. The desired extension is accomplished by setting $f(x) = y$. This proves Lemma 2.6 assuming Lemma 2.7.

Now we turn to the proof of Lemma 2.7. Put

$$G(y) := \max_{i=1,\ldots,k} \frac{|y - y_i|}{r_i}\,, \quad y \in \mathbb{R}^m\,.$$

Then $G : \mathbb{R}^m \to \mathbb{R}$ is a continuous function (Lipschitz, in fact) with $G(y) \to \infty$ as $|y| \to \infty$. It follows that G achieves its minimum at a point $w \in \mathbb{R}^m$, and we need to show that $G(w) \leq 1$.

Towards a contradiction, assume that $G(w) =: \lambda > 1$. Let J denote those indices $j \in \{1, \ldots, k\}$ for which $|w - y_j| = r_j \lambda$. Pick a point

$$x \in \bigcap_{j \in J} \overline{B}(x_j, r_j)\,,$$

and consider the following two sets of directions,

$$D := \{\frac{x_j - x}{|x_j - x|} : j \in J\} \subset \mathbb{S}^{n-1}\,, \quad D' := \{\frac{y_j - w}{|y_j - w|} : j \in J\} \subset \mathbb{S}^{m-1}\,.$$

It is easy to see from the definitions, and from the contrapositive assumption, that the natural map $D \to D'$ strictly decreases distances. We therefore require the following additional lemma.

Lemma 2.8. *Let $g : K \to \mathbb{S}^{m-1}$ be an L-Lipschitz map, $L < 1$, where $K \subset \mathbb{S}^{n-1}$ is compact. Then $g(K)$ is contained in an open hemisphere.*

Before we prove Lemma 2.8, let us point out how Lemma 2.7 follows from it. Indeed, the map between directions, $D \to D'$, strictly decreases the distances, and so is L-Lipschitz for some $L < 1$ because the sets in question are finite. It follows that D' is contained in an open hemisphere; say $D' \subset \mathbb{S}^{m-1} \cap \{x_m > 0\}$. But then by moving w slightly in the direction of the mth basis vector e_m, the value of the function G decreases, contradicting the fact that G assumes its minimum at w.

It therefore suffices to prove Lemma 2.8. To do so, let C be the convex hull of $g(K)$ in $\overline{\mathbb{B}}^m$. We need to show that C does not contain the origin. Thus, assume

$$\lambda_1 g(v_1) + \cdots + \lambda_k g(v_k) = 0$$

for some vectors $v_i \in K$, and for some real numbers $\lambda_i \in [0, 1]$ such that $\sum_{i=1}^k \lambda_i = 1$. Because g is L-Lipschitz with $L < 1$, we have that

$$\langle g(v_i), g(v_j)\rangle > \langle v_i, v_j\rangle$$

for every $i \neq j$. Thus, writing $b_i := \lambda_i v_i$, we find

$$\sum_{i=1}^k \langle b_j, b_i\rangle < 0$$

for each j. But this implies

$$\langle (b_1 + \cdots + b_k), (b_1 + \cdots + b_k)\rangle = \sum_{i,j=1}^k \langle b_j, b_i\rangle < 0\,,$$

which is absurd.

This completes our proof of Lemma 2.8, and hence that of Lemma 2.6. It remains to indicate how Kirszbraun's theorem 2.5 follows from Lemma 2.6.

We use a standard Arzelà-Ascoli type argument. Choose countable dense sets $\{a_1, a_2, \dots\}$ and $\{b_1, b_2, \dots\}$ in A and in $\mathbb{R}^n \setminus A$, respectively. We may assume that both of these sets are infinite. (If $\mathbb{R}^n \setminus A$ is finite, the extension is automatic; if A is finite, the ensuing argument requires only minor notational modifications.) For each $k = 1, 2, \dots,$ we can use Lemma 2.6 repeatedly so as to obtain a 1-Lipschitz map

$$f_k : \{a_1, \dots, a_k, b_1, \dots, b_k\} \to \mathbb{R}^m$$

such that $f_k(a_i) = f(a_i)$ for every $i = 1, \dots, k$. The sequence $(f_k(b_1)) \subset \mathbb{R}^m$ is bounded, and hence has a convergent subsequence, say $(f_{k_j^1}(b_1))$. Similarly, from the mappings corresponding to this subsequence we can subtract another subsequence, say $(f_{k_j^2})$, such that the sequence $(f_{k_j^2}(b_2)) \subset \mathbb{R}^m$ converges. Continuing this way, and finally passing to the diagonal sequence (g_j), $g_j := f_{k_j^j}$, we find that the limit

$$g(c) := \lim_{j\to\infty} g_j(c) \in \mathbb{R}^m$$

exists for every $c \in C := \{a_1, a_2, \dots\} \cup \{b_1, b_2, \dots\}$. Moreover, $g : C \to \mathbb{R}^m$ is 1-Lipschitz, and $g(a_i) = f(a_i)$ for each $i = 1, 2, \dots$. Because C is dense in $\mathbb{R}^n$, and because $\{a_1, a_2, \dots\}$ is dense in A, we have that g extends to a 1-Lipschitz map $\mathbb{R}^n \to \mathbb{R}^m$ as required.

This completes the proof of Kirszbraun's theorem 2.5. $\square$

Remark 2.9. (a) The crucial lemma in the preceding proof of Kirszbraun's theorem was Lemma 2.7. Gromov has asserted [19] an interesting volume monotonicity property that also can be used to derive Lemma 2.7. Namely, assume that

$$\overline{B}(x_1, r_1), \dots, \overline{B}(x_k, r_k) \quad \text{and} \quad \overline{B}(y_1, r_1), \dots, \overline{B}(y_k, r_k)$$

are closed balls in $\mathbb{R}^n$, $k \le n+1$, such that $|y_i - y_j| \le |x_i - x_j|$ for each $i, j \in \{1, \dots, k\}$. Then

$$|\bigcap_{i=1}^{k} \overline{B}(x_i, r_i)| \le |\bigcap_{i=1}^{k} \overline{B}(y_i, r_i)|\,. \tag{2.7}$$

It is easy to see that Lemma 2.7 follows from this assertion, thus providing another route to Kirszbraun's theorem.

(b) The preceding proof of Kirszbraun's theorem 2.5 works the same when one replaces $\mathbb{R}^n$ by an arbitrary separable Hilbert space, and $\mathbb{R}^m$ by an arbitrary finite dimensional Hilbert space. Standard proofs of Kirszbraun's theorem typically use Zorn's lemma (in conjunction with Lemma 2.7 or a similar auxiliary result). The preceding Arzelà-Ascoli argument does not work for infinite-dimensional targets.

2.3. **Exercises.** (a) Let $\{\overline{B}_i : i \in I\}$ be an arbitrary collection of closed balls in a Hilbert space with the property that

$$\bigcap_{i\in F} \overline{B}_i \ne \emptyset$$

for every finite subcollection $F \subset I$. Prove that

$$\bigcap_{i\in I} \overline{B}_i \ne \emptyset\,.$$

(Remember that bounded closed convex sets are compact in the weak topology of a Hilbert space.)

(b) Prove Kirszbraun's theorem in arbitrary Hilbert spaces.

It is a problem of considerable current research interest to determine for which metric spaces Kirszbraun's theorem remains valid. There are various variants on this theme. One can consider special classes of source spaces and target spaces, or even special classes of subspaces from where the extension is desired. Moreover, the Lipschitz constant may be allowed to change in a controllable manner. It would take us too far afield to discuss such general developments (references are given in the Notes to this section), but let us examine a bit further the case of subsets of Euclidean spaces.

2.4. **Lipschitz retracts.** A set $Y \subset \mathbb{R}^m$ is said to have the *Lipschitz extension property with respect to Euclidean spaces*, or the *Lipschitz extension property*, for short, if for every Lipschitz function $f : A \to Y$, $A \subset \mathbb{R}^n$, extends to a Lipschitz function $F : \mathbb{R}^n \to Y$. Note that we are asking for the mildest form of extension, with no control of the constants. In applications, a more quantitative requirement is often necessary. Sets with the Lipschitz extension property can be characterized as Lipschitz retracts of Euclidean spaces.

A set $Y \subset \mathbb{R}^m$ is said to be a (*Euclidean*) *Lipschitz retract* if there is a Lipschitz function $\rho : \mathbb{R}^m \to Y$ such that $\rho(y) = y$ for all $y \in Y$. Such a function ρ is called a *Lipschitz retraction* (*onto* Y). We also say that Y is a *Lipschitz retract of* $\mathbb{R}^m$ in this case. Note that if Y is a Lipschitz retract of some $\mathbb{R}^m$, then it is a Lipschitz retract of every $\mathbb{R}^M$ containing Y. Thus the term "Euclidean Lipschitz retract" is appropriate.

A Lipschitz retract is necessarily closed, as it is the preimage of zero under the continuous map $y \mapsto \rho(y) - y$. Therefore it is no loss of generality to consider only closed sets in the ensuing discussion.

Proposition 2.10. *A closed set $Y \subset \mathbb{R}^m$ has the Lipschitz extension property if and only if Y is a Lipschitz retract of $\mathbb{R}^m$.*

Proof. If Y has the Lipschitz extension property, then Y is a Lipschitz retract of $\mathbb{R}^m$, for the identity function $Y \to Y$ must have a Lipschitz extension to $\mathbb{R}^m$. On the other hand, if $\rho : \mathbb{R}^m \to Y$ is a Lipschitz retraction and if $f : A \to Y$, $A \subset \mathbb{R}^n$, is a Lipschitz function, then $\rho \circ F : \mathbb{R}^n \to Y$ provides a Lipschitz extension of f, where $F : \mathbb{R}^n \to \mathbb{R}^m$ is an extension guaranteed by the McShane-Whitney extension theorem. $\square$

Every Euclidean Lipschitz retract must be contractible. Indeed, if $\rho : \mathbb{R}^m \to Y$ is a retraction, and $h : \mathbb{R}^m \times [0,1] \to \mathbb{R}^m$ is a continuous deformation of $\mathbb{R}^m$ to a point $y_0 \in Y$ (that is, $h(x,0) = x$ and $h(x,1) = y_0$ for all $x \in \mathbb{R}^m$), then

$$H : Y \times [0,1] \to Y\,, \quad H(y,t) := \rho \circ h(y,t)\,,$$

provides a required homotopy.

Another basic observation is that every Euclidean Lipschitz retract Y must be quasiconvex. Indeed, if $\rho : \mathbb{R}^m \to Y$ is an L-Lipschitz retraction and if $[a,b]$ is the line segment in $\mathbb{R}^m$ connecting two points $a, b \in Y$, then $\rho([a,b])$ is a curve in Y joining a and b of length at most $L|a-b|$.

It is remarkable that the preceding two obvious necessary conditions for a retract are also sufficient in dimension $m = 2$.

Theorem 2.11. *A closed set Y in $\mathbb{R}^2$ is a Euclidean Lipschitz retract if and only if Y is contractible and quasiconvex. The statement is quantitative in the sense that the quasiconvexity constant of the retract and the Lipschitz constant of the retraction depend only on each other.*

I learned this result from Jason Miller, who discovered a proof while working on an REU-project at the University of Michigan during the summer of 2004.[1] We soon found out that Theorem 2.11 follows from a more general result of Lang and Schroeder [42]. Namely, every contractible planar continuum is a CAT(0)-space in its intrinsic metric, and [42, Theorem A] implies that Kirszbraun's theorem holds for such target spaces; the identity map $Y \to Y$ extends to a map $\mathbb{R}^2 \to Y$ that is Lipschitz with respect to the intrinsic distance. The presumed quasiconvexity guarantees that a Lipschitz condition holds also with respect to the Euclidean metric. See [42] and [10] for the definition of CAT(0) spaces, and [10, p. 310] for the fact cited here. Further important extensions of Kirszbraun's theorem in terms of curvature conditions can be found in [41].

We will not prove Theorem 2.11 in these notes.

In higher dimensions there is a lack of good geometric criteria for a set to be a Euclidean Lipschitz retract. The following result due to Hohti [32] provides an implicit characterization.

Theorem 2.12. *Let $Y \subset \mathbb{R}^m$ be a closed set. Then Y is a Lipschitz retract of $\mathbb{R}^m$ if and only if Y is quasiconvex and there exists a Lipschitz*

[1]REU is a U.S. National Science Foundation funded program *Research Experience for Undergraduates.*

map

$$\sigma : Y \times Y \times Y \to Y \tag{2.8}$$

satisfying

$$\sigma(a, a, b) = \sigma(a, b, a) = \sigma(b, a, a) = a \tag{2.9}$$

whenever $a, b \in Y$.

In (2.8), the Lipschitz condition for σ is understood with respect to the metric of $\mathbb{R}^{3m} \supset Y \times Y \times Y$. Condition (2.9) stipulates the existence of a “center of mass”, that varies in a Lipschitz manner.

Proof of Theorem 2.12. First we prove the necessity. Thus, assume that Y is a Lipschitz retract. It was observed earlier that Y is quasi-convex. Moreover, by using the retraction map, it is enough to show that there is a map σ as required for $Y = \mathbb{R}^m$. If $m = 1$, then we choose σ to be the “mid-point map”, i.e., from each given triple of real numbers σ picks the one that lies in the middle (with respect to the natural order of $\mathbb{R}$). In higher dimensions, we apply the mid-point map to the coordinates. It is easy to check that σ thus defined is indeed Lipschitz; condition (2.9) is automatic.

The sufficiency of trickier to prove. We take this opportunity to introduce a Whitney decomposition of an open set in $\mathbb{R}^n$.

2.5. **Whitney decomposition.** A system of *dyadic cubes* in $\mathbb{R}^n$ is the collection $\mathcal{D}$ of cubes consisting of all (closed) cubes Q in $\mathbb{R}^n$ that have sides parallel to the coordinate axes, side length 2^k and vertices in the set $2^k\mathbb{Z}^n$, $k \in \mathbb{Z}$. Thus, $\mathcal{D}$ is divided into *generations*, each consisting of (essentially) disjoint cubes with side length 2^k for a fixed k.

Now suppose that Ω is an open set in $\mathbb{R}^n$. It is possible to single out a collection $\mathcal{W}_\Omega$ of dyadic cubes Q contained in Ω with the following properties:

$$\mathcal{W}_\Omega \text{ consists of pairwise essentially disjoint cubes;} \tag{2.10}$$

$$\bigcup_{Q \in \mathcal{W}_\Omega} Q = \Omega; \tag{2.11}$$

$$\ell(Q) \le \operatorname{dist}(Q, \Omega^c) \le 4\sqrt{n}\,\ell(Q); \tag{2.12}$$

$$\frac{6}{5}Q \subset \Omega, \qquad c(n)\, Q \cap \Omega^c \ne \emptyset. \tag{2.13}$$

Here $\ell(Q)$ denotes the side length of Q, and λQ, $\lambda > 0$, is the cube with same center as Q and with sides parallel to the coordinate axes such that $\ell(\lambda Q) = \lambda\ell(Q)$.

Note that (2.13) follows from (2.12); one can choose $c(n) = 1+8\sqrt{n}$.

The collection $\mathcal{W}_\Omega$ as above is called the *Whitney decomposition* of Ω. There are various ways to construct a Whitney decomposition for an open set; the precise choice is immaterial in applications. What really matters is that each cube in the decomposition is roughly the size of its distance to the complement of Ω, and that only a fixed number of cubes overlap, even when slightly expanded.

The reader is invited to construct a Whitney decomposition with the listed properties as an exercise. Alternatively, one can consult [65, pp. 167-168].

We use the Whitney decomposition to record a characterization for Euclidean Lipschitz retracts.

2.6. **Lipschitz contractibility.** A set $Y \subset \mathbb{R}^m$ is said to have the property LC(k), $k \geq 1$, if for every $L > 0$ there exists $L' > 0$ such that every L-Lipschitz map $f : \partial Q \to Y$ extends to an L'-Lipschitz map $F : Q \to Y$, whenever $Q \subset \mathbb{R}^k$ is a k-dimensional cube.

Note that Y has the property LC(1) if and only if Y is quasiconvex.

It is easy to see (by using the bi-Lipschitz equivalence between cubes and balls) that Y has the property LC(k) if and only if every Lipschitz map from a $(k-1)$-dimensional sphere $\partial B^k(x,r)$ to Y can be extended to a Lipschitz map from the ball $B^k(x,r)$ to Y; the Lipschitz constant of the extension can only depend on the constant of the boundary map, and the dimension.

The letters LC stand for *Lipschitz contractibility.*

Proposition 2.13. *Let $Y \subset \mathbb{R}^m$ be a closed set. Then Y is a Lipschitz retract of $\mathbb{R}^m$ if and only if it has the property* LC(k) *for every $1 \leq k \leq m$.*

Proof. The necessity is immediate. Indeed, if $f : \partial Q \to Y$ is a Lipschitz map, then f extends to a Lipschitz map $F : Q \to \mathbb{R}^m$ by the McShane-Whitney extension theorem. Thus $\rho \circ F : Q \to Y$ is the required extension, where $\rho : \mathbb{R}^m \to Y$ is a Lipschitz retraction onto Y.

To prove the sufficiency, let $\Omega := \mathbb{R}^m \setminus Y$, and fix a Whitney decomposition $\mathcal{W}_\Omega$ of Ω. Let S_0 denote the set of all vertices of all the cubes in $\mathcal{W}_\Omega$. Then define a Lipschitz map

$$f_0 : S_0 \to Y$$

by choosing for each point $v \in S_0$ a point in Y that is closest to v. It follows from the properties of the Whitney decomposition that f_0 is

Lipschitz with constant L_0 that depends only on m. Also note that f_0 extends continuously to Y by setting $f_0(y) = y$ for $y \in Y$.

Next, let S_1 denote the *1-skeleton* of the Whitney decomposition; that is, S_1 is the union of all the line segments that are the edges of the cubes in $\mathcal{W}_\Omega$. We extend the Lipschitz map f_0 to S_1 by using the LC(1) property (i.e. quasiconvexity). This extension is easily seen to be L_1-Lipschitz in S_1 with L_1 depending only on m and the data in the LC(1) hypothesis (cf. Lemma 2.2). By continuing in this manner, we get a sequence of maps

$$f_k : S_k \to Y\,, \qquad 1 \le k \le m\,,$$

from the k-skeleton S_k of the Whitney decomposition (the definition for S_k should be clear) that are L_k-Lipschitz with L_k depending only on m and the data in the LC(k)-hypothesis. Moreover, each f_k extends continuously to Y by setting $f_k(y) = y$ for $y \in Y$.

It follows that $f_m : \mathbb{R}^m \to Y$ is a Lipschitz retraction to Y as required. This proves Proposition 2.13. □

Remark 2.14. Jason Miller [51] has recently proved that every Lipschitz map from an n-dimensional sphere $\mathbb{S}^n$, $n \ge 2$, into the plane extends to a Lipschitz map from $\overline{\mathbb{B}}^{n+1}$ to $\mathbb{R}^2$ with the same image. Moreover, the Lipschitz constant of the extension depends only on the Lipschitz constant of the boundary map. This result shows that *every set in* $\mathbb{R}^2$ *has the property* LC(k)*, for every* $k \ge 3$.

Miller's result is a Lipschitz version of the fact that *every planar set is aspherical*, meaning that all the homotopy groups beyond the first two are trivial for such a set. See [12].

We now continue the proof for the sufficiency part of Theorem 2.12. Thus, let $Y \subset \mathbb{R}^m$ be a closed set, and let σ be a map as in (2.8) and (2.9). We will show, by using σ, that Y has the property LC(k) for each $k \ge 2$. This suffices by Proposition 2.13, as LC(1) is part of the assumptions.

To this end, we use (as we may) balls rather than cubes. Thus, let $S = \partial B$ be the boundary of a k-dimensional ball $B = B(x, r)$, $k \ge 2$, and let $f : S \to Y$ be a Lipschitz map. In what follows, we consider various Lipschitz maps without specifying their Lipschitz constants; none of these constants depend on the ball B.

We may assume that $x = 0$. Write $A := \overline{B} \setminus \frac{1}{2}B$ for the (closed) annulus, and let $f_1 : A \to S$ be the standard radial Lipschitz retraction, i.e., f_1 maps every line segment $[\frac{1}{2}s, s]$ to $s \in S$. Next, denote by S^+ and S^- the (closed) upper and lower hemispheres of S. Both hemispheres are bi-Lipschitz homeomorphic to a $(k-1)$-dimensional cube, and hence are Lipschitz retracts. In particular, there exist Lipschitz maps

$$f_2 : \overline{B} \to S^+\,, \qquad f_3 : \overline{B} \to S^-\,,$$

such that

$$f_2(\frac{1}{2}B) = f_3(\frac{1}{2}B) = e := (r, 0, \ldots, 0) \in S^+ \cap S^-\,,$$

and that $f_2|S^+ =$identity and $f_3|S^- =$identity.

The preceding understood, we define a map $F : B \to Y$ as follows. First, we have a map

$$G : A \to Y \times Y \times Y\,, \quad G(x) := (f \circ f_1(x), f \circ f_2(x), f \circ f_3(x))\,,$$

which may not have a continuous extension to $\frac{1}{2}B$. But upon defining

$$F := \sigma \circ G : A \to Y\,,$$

we obtain a map that has a continuous extension to all of B by setting

$$F(\frac{1}{2}B) \equiv \sigma(f(e))\,.$$

This last assertion readily follows from the properties of σ. Because $F|S = f$, and because all the maps in question are Lipschitz (with constants independent from B), we have established the required LC(k) property.

This completes the proof of Theorem 2.12.

□

Remark 2.15. It is necessary to assume a priori that Y is quasiconvex in the sufficiency part of Theorem 2.12. For example, every so called *snowflake arc* in $\mathbb{R}^2$ admits a Lipschitz map as in (2.8) and (2.9), but contains no rectifiable curves. More precisely, let $f : [0,1] \to \mathbb{R}^2$ be an embedding that satisfies

$$C^{-1}\,|x-y|^\alpha \;\le\; |f(x)-f(y)| \;\le\; C\,|x-y|^\alpha\,, \tag{2.14}$$

for all $x, y \in [0,1]$, for some $\frac{1}{2} < \alpha < 1$ and $C \ge 1$. (It is not difficult to construct such embeddings, but see for example [72].) Now let $Y := f([0,1])$, and define

$$\sigma : Y \times Y \times Y \to Y$$

by

$$\sigma(a,b,c) := f \circ \sigma_1(f^{-1}(a), f^{-1}(b), f^{-1}(c))\,,$$

where $\sigma_1 : [0,1]^3 \to [0,1]$ is the mid point map mentioned in the beginning of the proof of Theorem 2.12. It is easy to see that σ is Lipschitz and satisfies (2.9).

2.7. **Example.** Let us define a *tree* to be a connected and contractible subset T of $\mathbb{R}^m$ that can be written as a union of countably many line segments,

$$T = \bigcup_{i \geq 1} I_i \,,$$

such that any two segments meet at most at one common end point, that no point in T belong to more than finitely many of the segments, and that every pair of points in T can be joined by a finite union of line segments from the collection $\{I_1, I_2, \ldots\}$. Note that under this definition, T need not be a closed subset of $\mathbb{R}^m$.

Every tree T has its intrinsic metric as defined earlier in Section 2.1, and every pair of points in T can be joined by a unique (intrinsic) geodesic. We will now look for a map $\sigma : T \times T \times T \to T$ that is Lipschitz in the intrinsic metric and such that (2.9) holds. Indeed, if $a, b, c \in T$, then the union of the three intrinsic geodesics between the three points is homeomorphic either to a line segment or to a union of three segments meeting at a point. In the first case, one of the three points a, b, c lies in between the other two, and we let $\sigma(a, b, c)$ to be that point. In the second case, we let $\sigma(a, b, c)$ to be the unique meeting point.

It is easy to check that σ thus defined is Lipschitz in the intrinsic metric; it is locally uniformly Lipschitz and then a variant of Lemma 2.2 can be used. In particular, if T is quasiconvex as a subset of $\mathbb{R}^m$, then σ is Lipschitz with respect to the underlying Euclidean metric. Because the closure of every quasiconvex set in $\mathbb{R}^m$ remains quasiconvex, we conclude that the closure of a quasiconvex tree T in $\mathbb{R}^m$ possesses a Lipschitz map $\sigma : \overline{T} \times \overline{T} \times \overline{T} \to \overline{T}$ as in Theorem 2.12. In particular, we have the following result.

Theorem 2.16. *The closure of a quasiconvex tree in $\mathbb{R}^m$ is a Euclidean Lipschitz retract.*

It is not difficult to construct quasiconvex trees in $\mathbb{R}^m$, $m \geq 2$, whose closure has Hausdorff dimension larger than $m - \epsilon$, for any prescribed $\epsilon > 0$. Note that although every tree T as defined above has Hausdorff dimension one always, the closure $\overline{T}$ may have much larger Hausdorff dimension. We leave the details as an exercise for the reader.

2.8. **Exercises.** (a) Construct, for a given $m \geq 2$ and $1 \leq \alpha \leq m$, a compact Lipschitz retract $Y \subset \mathbb{R}^m$ such that the Hausdorff dimension of Y is α.

(b) Show that every Lipschitz retract $Y \subset \mathbb{R}^m$ with empty interior is porous with constant that depends only on m and on the Lipschitz constant of a retraction $\rho : \mathbb{R}^m \to Y$.

(A subset Y of $\mathbb{R}^m$ is said to be *porous* if there exists a constant $c \in (0,1)$ such that for every $y \in Y$ and every $r > 0$ there exists a point $z \in B(y,r)$ such that $B(z,cr) \cap Y = \emptyset$.)

Then conclude that a Lipschitz retract in $\mathbb{R}^m$ either has no interior, or has Hausdorff dimension strictly less than m (only depending on m and the Lipschitz constant of the retraction).

(Hint: Use the local degree theory as explained, for example, in [56, Section I. 4].)

For the definition and properties of Hausdorff measure and dimension, see [50], and for facts about porosity, see [50, p. 156].

Remark 2.17. That closed quasiconvex trees in $\mathbb{R}^m$, as explained in the preceding example, are Lipschitz retracts also follows from a more general result of Lang and Schroeder [42, Theorem B]. This result asserts that an arbitrary complete and geodesic metric space, where all triangles are κ-*thin* for every $\kappa \in \mathbb{R}$, is an absolute Lipschitz retract, i.e. it satisfies the conclusion of the McShane-Whitney extension theorem 2.3 as a target space with respect to arbitrary metric source spaces. The proof in [42], albeit still elementary, is more involved than the one given here for the special case.

2.9. **Open problem.** Let $2 \leq n < m$ and let $Y \subset \mathbb{R}^m$ be a set that is homeomorphic to $\mathbb{R}^n$. Assume moreover that Y satisfies the conditions of n-*Ahlfors regularity* and *linear local contractibility*. The first condition means that Y has Hausdorff dimension n and that the Hausdorff n-measure $\mathcal{H}_n$ on Y satisfies the following: there exists a constant $C \geq 1$ such that

$$C^{-1}\, r^n \leq \mathcal{H}_n(B(y,r) \cap Y) \,\leq\, C\, r^n \tag{2.15}$$

for every $y \in Y$ and $r > 0$. The second condition means that there exists a constant $C \geq 1$ such that every set $B(y,r) \cap Y$ can be contracted to a point inside $B(y,Cr) \cap Y$, for every $y \in Y$ and $r > 0$.

Is Y then a Euclidean Lipschitz retract?

It is known that under the said assumptions Y need not be bi-Lipschitz equivalent to $\mathbb{R}^n$ if $n \geq 3$. Every known example to this effect is nontrivial [61]. It is not known whether, for $n = 2$, every such set Y is bi-Lipschitz equivalent to $\mathbb{R}^2$.[2]

[2]Added in August 2005. Chris Bishop has recently shown that also for $n = 2$ such a set Y need not be bi-Lipschitz equivalent to $\mathbb{R}^2$.

See [61], [60], [26], [7] for more information about metric parametrization problems.

2.10. **Another proof for the extension.** In this subsection, we outline yet another proof for the fact that every L-Lipschitz map $f : A \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, can be extended to an L'-Lipschitz map $F : \mathbb{R}^n \to \mathbb{R}^m$. By Kirszbraun's theorem, we know that one can choose $L' = L$. By the simple McShane-Whitney argument, one obtains $L' = \sqrt{m}L$. Here we give a proof where L' depends only on L and n. In fact, this proof generalizes for arbitrary Banach space targets.

Thus, let $A \subset \mathbb{R}^n$ and let $f : A \to \mathbb{R}^m$ be L-Lipschitz. We may assume that A is closed. Then consider a Whitney decomposition $\mathcal{W}_\Omega$ of the complement $\Omega := \mathbb{R}^n \setminus A$. There is associated with the Whitney decomposition a *Lipschitz partition of unity* $\{\varphi_Q : Q \in \mathcal{W}_\Omega\}$ with the following properties. By letting $C > 1$ denote an appropriate dimensional constant, not necessarily the same at each occurrence, we have

(a) $0 \le \varphi_Q \le 1$ and $\varphi_Q|Q \ge C^{-1}$;
(b) φ_Q is supported in $CQ \subset \Omega$;
(c) for every $x \in \Omega$ we have $\varphi_Q(x) \ne 0$ for only at most C cubes Q;
(d) φ_Q is Lipschitz with constant $C/\ell(Q)$;
(e) $\sum_{Q\in\mathcal{W}_\Omega} \varphi_Q \equiv 1$.

Now let z_Q denote the center of a Whitney cube Q, and pick a point $y_Q \in A$ that is closest to z_Q in A. Note in particular that $|z_Q - y_Q| \approx \ell(Q)$. Then define

$$F(z) := \sum_{Q\in\mathcal{W}_\Omega} \varphi_Q(z)\, f(y_Q)\,, \qquad z \in \Omega\,. \tag{2.16}$$

It is easy to see that $F : \Omega \to \mathbb{R}^m$ is Lipschitz and admits continuous extension to $\mathbb{R}^n$ with $F|A = f$. We leave the details as an exercise.

2.11. **Exercise.** Show that a partition of unity as in (a)–(e) above exists. Then show that the function F given in (2.16) extends continuously to A with $F|A = f$; morever, this extension is Lipschitz with constant L', where L' depends only on L and n, and

2.12. **Notes to Section 2.** The beginning material of the section is standard and can be found in many texts, e.g. in [17], [18], [50]. Our proof for Kirszbraun's theorem 2.5 is somewhat different from the usual sources; typical proofs, even in finite dimensions, seem to use Zorn's lemma (as Peter Lindqvist pointed out to me). I learned the proof of Lemma 2.7 from Bruce Kleiner; similar arguments can be found in [38, Section 4], [42], [41], [20, p. 21]. Recently, there has been much interest in finding sharper versions of the McShane-Whitney-Kirszbraun results, where the extension is required to have minimal possible Lipschitz constant in all subregions; see [2].

Euclidean Lipschitz retracts have been considered in geometric measure theory, e.g. in [18], but there are only a few papers that study them as such. In particular, as mentioned in the text, very few sufficient criteria for a set to be a Euclidean Lipschitz retract are known. Proposition 2.13 goes back to Almgren [1]. I learned Theorem 2.12 from Hohti [32] who builds on an earlier work [75]. The example in 2.14 was known to Hohti [32], who also discusses Lipschitz retracts in more general contexts. Papers [41] and [43] contain more information about Lipschitz extension problem in general settings. See also [5, Chapter 1]. The proof in subsection 2.10 is due to Johnson, Lindenstrauss, and Schechtman [35].

3. Differentiability

This section is devoted to the proof of the following fundamental result.

Theorem 3.1 (Rademacher's theorem)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, and let $f : \Omega \to \mathbb{R}^m$ be Lipschitz. Then f is differentiable at almost every point in Ω.*

Recall that a function $f : \Omega \to \mathbb{R}^m$, where $\Omega \subset \mathbb{R}^n$ is open, is *differentiable at* $a \in \Omega$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{x\to a} \frac{|f(x) - f(a) - L(x-a)|}{|x-a|} = 0\,. \tag{3.1}$$

If such a linear map L exists, it is unique, called the *derivative* of f at a, and denoted by $Df(a)$. We also note that $f = (f_1, \ldots, f_m)$ is differentiable at a if and only if each of the coordinate functions f_i are differentiable at a.

To analyze condition (3.1) more carefully, suppose that $f : \Omega \to \mathbb{R}$ is a real-valued function, differentiable at a point $a \in \Omega$. For $t \in \mathbb{R}$, $t \neq 0$, consider the functions

$$f_t(x) := \frac{f(a+tx) - f(a)}{t}\,, \tag{3.2}$$

that are defined for all t small enough. Then

$$\lim_{t\to 0} |f_t(x) - L(x)| = 0 \tag{3.3}$$

uniformly in $x \in \mathbb{B}^n$. This procedure can be reversed and we conclude that a function f is differentiable at a point a if and only if the sequence of renormalized functions (f_t) as in (3.2) converges uniformly in $\mathbb{B}^n$ to a linear function as $t \to 0$.

Assume now that $f : \Omega \to \mathbb{R}$ is L-Lipschitz, and that $a \in \Omega$. Then the family (f_t) consists of uniformly bounded L-Lipschitz functions on $\mathbb{B}^n$ (for small enough t). The Arzelà-Ascoli theorem (see e.g. [31, p. 44]) therefore guarantees that there is a subsequence of the sequence (f_t) that converges uniformly to an L-Lipschitz function on $\mathbb{B}^n$. What Rademacher's theorem claims, in effect, is that for almost all points a in Ω this limit is independent of the subsequence, and that the limit function is linear. It is important to notice the two separate assertions.

To prove Rademacher's theorem 3.1, we may assume that $m = 1$. The proof is based on a reduction to the case where also $n = 1$. It is therefore appropriate to recall this special case in some detail.

Theorem 3.2 (Lebesgue)**.** *Let $f : (a, b) \to \mathbb{R}$ be Lipschitz. Then f is differentiable at almost every point in (a, b).*

Proof. The well known theorem of Lebesgue asserts in fact that every function of bounded variation on an interval is almost everywhere differentiable. Lipschitz functions are examples of functions of bounded variation. The key point in establishing this result is that every function $f : (a, b) \to \mathbb{R}$ of bounded variation can be written as a difference of two increasing functions, namely

$$f(x) = V_f(x) - (V_f(x) - f(x)), \tag{3.4}$$

with

$$V_f(x) := \sup \sum_{i=1}^{N} |f(x_{i+1}) - f(x_i)|, \tag{3.5}$$

where the supremum is taken over all finite sequences $a < x_1 < \cdots < x_{N+1} < x$.

Thus, assume that $f : [a, b] \to \mathbb{R}$ is continuous and increasing (we may clearly assume that f is defined and continuous on $[a, b]$). For $x \in (a, b)$, set

$$D^+ f(x) := \limsup_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

and

$$D^- f(x) := \liminf_{h \to 0} \frac{f(x + h) - f(x)}{h}.$$

We need to show that $D^+ f(x) = D^- f(x) \in \mathbb{R}$ for almost every x.

To do so, we use the important *Vitali covering theorem*. Let $\mathcal{I}$ be a collection of closed intervals in (a,b) with the property that if a point x belongs to some interval from $\mathcal{I}$, then for every $\epsilon > 0$ there is an interval from $\mathcal{I}$ containing x and having length less than ϵ. Under this assumption, the Vitali covering theorem asserts that there is a pairwise disjoint subcollection $\mathcal{I}_1 = \{I_1, I_2, \dots\} \subset \mathcal{I}$ such that

$$|\bigcup_{\mathcal{I}_1} I_i| = |\bigcup_{\mathcal{I}} I|\,.$$

By using this covering theorem, and the definitions for D^+f and D^-f, it is easy to check the following inequalities:

$$q\,|E_q| \;\le\; |f(E_q)|\,, \tag{3.6}$$

if $D^+f(x) > q$ at every $x \in E_q$, and

$$|f(E_p)| \;\le\; p\,|E_p|\,,$$

if $D^-f(x) < p$ at every $x \in E_p$. Similarly, we find for the set

$$E_{pq} := \{x : D^-f(x) < p < q < D^+f(x)\}$$

that

$$q\,|E_{pq}| \;\le\; |f(E_{pq})| \;\le\; p\,|E_{pq}|\,,$$

which implies that $|E_{pq}| = 0$. It follows that $D^+f(x) = D^-f(x)$ for almost every $x \in (a,b)$. Because (3.6) implies that

$$|\{x : D^+f(x) = \infty\}| = 0$$

(we must have that $|f(E_q)| \le |f(b) - f(a)| < \infty$ for every q), the proof is complete. □

Note that the Lipschitz condition was completely erased from the preceding proof. Indeed, the result is really a theorem about differentiation of measures; the (one-dimensional) argument with the variation function enables us to use measure theoretic and covering arguments.

In the proof of the general Rademacher theorem, we need another important one variable fact; namely, that the fundamental theorem of calculus holds for Lipschitz functions.

Theorem 3.3. *Let $f : [a,b] \to \mathbb{R}$ be Lipschitz. Then*

$$f(b) - f(a) = \int_a^b f'(t)\,dt\,. \tag{3.7}$$

Theorem 3.3 again is a special case of a more general fact; namely, (3.7) holds for all *absolutely continuous functions*. For the proof of Theorem 3.3, we refer to any of the standard texts in real analysis. Recall however that (3.7) does not hold for all functions of bounded variation, and that not every absolutely continuous function is Lipschitz.

Armed with Theorems 3.2 and 3.3, we will proceed with the proof of Rademacher's theorem.

Proof of Theorem 3.1. By using the extension theorems, we may assume for simplicity and without loss of generality that $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz. The proof splits into three parts. First the one-dimensional result is used to conclude that the partial derivatives $(\frac{\partial f}{\partial x_i})$ of f exists almost everywhere. This gives us a candidate for the total derivative, namely the (almost everywhere defined) formal gradient

$$\nabla f(x) := (\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}) . \tag{3.8}$$

Next, it is shown that all directional derivatives exist almost everywhere, and can be given in terms of the gradient. Finally, by using the fact that there are only "countably many directions" in $\mathbb{R}^n$, the total derivative is shown to exist; it is only in this last step that the Lipschitz condition is seriously used.

We will now carry out these steps. The first claim is a direct consequence of Theorem 3.2. Indeed, for every $x, v \in \mathbb{R}^n$, the function

$$f_{x,v}(t) := f(x + tv) , \quad t \in \mathbb{R} ,$$

is Lipschitz as a function of one real variable, and hence differentiable at almost every $t \in \mathbb{R}$. Keeping now $v \in \mathbb{R}^n$ fixed, we conclude from Fubini's theorem and the preceding remark that

$$D_v f(x) := \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} \tag{3.9}$$

exists for almost every $x \in \mathbb{R}^n$. (To be precise here, in order to use Fubini's theorem, one has to first show that the set of those points x for which the limit in (3.9) exists is measurable.) In particular, as

$$\frac{\partial f}{\partial x_i} = D_{e_i} f$$

for each $i = 1, \dots, n$, where e_i is the ith standard basis vector in $\mathbb{R}^n$, the formal gradient $\nabla f(x)$ as given above in (3.8) exists at almost every $x \in \mathbb{R}^n$.

As a second step, we show that for every $v \in \mathbb{R}^n$ we have

$$D_v f(x) = v \cdot \nabla f(x) \tag{3.10}$$

for almost every $x \in \mathbb{R}^n$. To do so, fix $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$. Then fix a test function $\eta \in C_0^\infty(\mathbb{R}^n)$. We have that

$$\begin{aligned}
\int_{\mathbb{R}^n} D_v f(x)\,\eta(x)\,dx &= \int_{\mathbb{R}^n} \lim_{t\to 0} \frac{f(x+tv)-f(x)}{t}\,\eta(x)\,dx \\
&= \lim_{t\to 0} \int_{\mathbb{R}^n} \frac{f(x+tv)-f(x)}{t}\,\eta(x)\,dx \\
&= \lim_{t\to 0} \int_{\mathbb{R}^n} -f(x) \frac{\eta(x)-\eta(x-tv)}{t}\,dx \\
&= -\int_{\mathbb{R}^n} f(x) \lim_{t\to 0} \frac{\eta(x)-\eta(x-tv)}{t}\,dx \\
&= -\int_{\mathbb{R}^n} f(x)\,D_v\eta(x)\,dx \\
&= -\sum_{i=1}^n v_i \int_{\mathbb{R}^n} f(x)\,\frac{\partial \eta}{\partial x_i}(x)\,dx \\
&= \sum_{i=1}^n v_i \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i}(x)\,\eta(x)\,dx \\
&= \int_{\mathbb{R}^n} v \cdot \nabla f(x)\,\eta(x)\,dx\,.
\end{aligned}$$

Because η was arbitrary, equality (3.10) holds for almost every $x \in \mathbb{R}^n$.

In the above string of equalities, the second and the fourth are valid by the dominated convergence theorem (by using the Lipschitz condition in the first case), and the third is valid by change of variables. The penultimate equality in turn is valid by using the integration by parts on almost every line parallel to the coordinate axes, which is possible by Theorem 3.3.

The last step is to prove the original claim. To this end, fix a countable dense set of directions in $\mathbb{R}^n$; that is, fix a countable dense set of vectors (v_i) in $\partial\mathbb{B}^n$. By the first two steps, we infer that there is a set $A \subset \mathbb{R}^n$ such that $|\mathbb{R}^n \setminus A| = 0$ and that

$$D_{v_i} f(a) = v_i \cdot \nabla f(a) \tag{3.11}$$

for every v_i and for every $a \in A$, where we also understand that both sides of (3.11) exist (the gradient $\nabla f(a)$ is still understood formally as in (3.8)). Now fix $a \in A$. For $v \in \partial\mathbb{B}^n$ and $t \in \mathbb{R}$, $t \neq 0$, set

$$D(v,t) := \frac{f(a+tv)-f(a)}{t} - v \cdot \nabla f(a)\,.$$

To prove the differentiability of f at a, we need to show that $D(v,t) \to 0$ as $t \to 0$ independently of v. To do this, fix $\varepsilon > 0$. Then choose an ε-dense set of vectors $v_1, \dots, v_N$ from the chosen dense collection of directions; i.e., for each $v \in \partial\mathbb{B}^n$ we have that $|v - v_i| < \varepsilon$ for some $i = 1, \dots, N$. We then find that

$$\begin{aligned} |D(v,t) - D(v_i,t)| \;&\leq\; \left|\frac{f(a+tv) - f(a+tv_i)}{t}\right| + |(v - v_i)\cdot \nabla f(a)| \\ &\leq\; C \cdot |v - v_i| \;<\; C \cdot \varepsilon\,, \end{aligned}$$

where C is a constant independent of v, by the Lipschitz assumption. Because $\lim_{t\to 0} D(v_i,t) = 0$ for each v_i, we can choose $\delta > 0$ such that

$$D(v_i,t) < \varepsilon$$

for $|t| < \delta$, for each $i = 1, \dots, N$. By combining the preceding inequalitites, we obtain that

$$|D(v,t)| < \; C \cdot \varepsilon$$

whenever $|t| < \delta$, where C is independent of v, as required.

This completes the proof of Rademacher's theorem 3.1. □

We will give a different proof of Rademacher's theorem in the next section (Theorem 4.9).

3.1. **Stepanov's theorem.** There is a generalization of Rademacher's theorem due to Stepanov. The *pointwise Lipschitz constant* of a function $f : A \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, is

$$\text{Lip} f(x) := \limsup_{y\to x, y\in A} \frac{|f(x) - f(y)|}{|x - y|}\,. \tag{3.12}$$

Theorem 3.4 (Stepanov's theorem)**.** *Let $\Omega \subset \mathbb{R}^n$ be open, and let $f : \Omega \to \mathbb{R}^m$ be a function. Then f is differentiable almost everywhere in the set*

$$L(f) := \{x \in \Omega : \text{Lip} f(x) < \infty\}\,.$$

The following elegant proof is due to Malý [48].

Proof. We may assume that $m = 1$. Let $\{B_1, B_2, \dots\}$ be the countable collection of all balls contained in Ω such that each B_i has rational center and radius, and that $f|B_i$ is bounded. In particular, this collection covers $L(f)$. Define

$$u_i(x) := \inf\{u(x) : u \text{ is } i\text{-Lipschitz with } u \geq f \text{ on } B_i\}$$

and

$$v_i(x) := \sup\{v(x) : v \text{ is } i\text{-Lipschitz with } v \leq f \text{ on } B_i\}\,.$$

Then (see Lemma 2.1) functions $u_i, v_i : B_i \to \mathbb{R}$ are i-Lipschitz for each i, and $v_i \leq f|B_i \leq u_i$. It is clear that f is differentiable at every point a, where, for some i, both u_i and v_i are differentiable with $v_i(a) = u_i(a)$. We claim that almost every point in $L(f)$ is such a point.

By Rademacher's theorem, the set

$$Z := \bigcup_{i=1}^{\infty} \{x \in B_i : \text{either } u_i \text{ or } v_i \text{ is not differentiable at } x\}$$

has measure zero. If $a \in L(f) \setminus Z$, then there is a radius $r > 0$ such that

$$|f(a) - f(x)| \leq M|a - x|$$

for all $x \in B(a, r)$ and for some M independent of x. Clearly there is an index $i > M$ such that $a \in B_i \subset B(a, r)$. In particular,

$$f(a) - i|a - x| \leq v_i(x) \leq u_i(x) \leq f(a) + i|a - x|$$

for $x \in B_i$, which gives the claim. The theorem follows. $\square$

3.2. **Differentiability of quasiconformal mappings.** The mild hypotheses of Stepanov's theorem makes the theorem valuable in practice. We next give an example to this effect, by showing that quasiconformal mappings are almost everywhere differentiable.

Recall that a homeomorphism $f : \Omega \to \Omega'$ between two domains in $\mathbb{R}^n$, $n \geq 2$, is *quasiconformal* if

$$\sup_{a \in \Omega} H(a, f) < \infty \,, \tag{3.13}$$

where

$$H(a, f) := \limsup_{r \to 0} \sup_{|a-x|=|a-y|=r} \frac{|f(a) - f(x)|}{|f(a) - f(y)|} \,. \tag{3.14}$$

Quasiconformal mappings need not be Lipschitz (in any nonempty open set), nevertheless they satisfy the hypotheses of Stepanov's theorem in the sense that $\mathrm{Lip}(f) < \infty$ almost everywhere.

Theorem 3.5. *Quasiconformal mappings are differentiable almost everywhere.*

Proof. Denote by H the supremum in (3.13). Then, for $a \in \Omega$, we have that

$$\begin{aligned}\limsup_{x\to a} \frac{|f(a)-f(x)|}{|a-x|} &\leq \limsup_{x\to a} C(n)\, H\, \frac{|f(B(a,|a-x|))|^{1/n}}{|a-x|} \\ &\leq \limsup_{x\to a} C(n)\, H\, \frac{|f(B(a,|a-x|))|^{1/n}}{|B(a,|a-x|)|^{1/n}} \\ &\leq C(n)\, H \mu_f(a)^{1/n}\,,\end{aligned}$$

where $C(n)$ is a dimensional constant and $\mu_f(a)$ denotes the almost everywhere finite Radon-Nikodym derivative between the Lebesgue measure and its pullback under f. (Note that quasiconformality was used in the first inequality only.) The assertion now follows from Stepanov's theorem. $\square$

3.3. **Notes to Section 3.** Except perhaps the proof of the Stepanov theorem 3.4, the material in this section is standard. The proof of the Rademacher theorem here essentially follows the presentation in [17]. For a more complete discussion of the classical case of one real variable, see e.g. [57, Chapter 8] or [9]. For the theory of quasiconformal mappings, see [74].

There has been recently a great deal of activity in extending classical differential analysis to certain (finite dimensional) metric measure spaces. In particular, a version of Rademacher's theorem in such a context was given by Cheeger [14]. See also [36]. For a similar extension of Stepanov's theorem, see [4]. Differentiability of Lipschitz functions between infinite dimensional Banach spaces has been a topic of extensive research for a long time; see e.g. [5].

4. Sobolev spaces

In this section, we discuss Sobolev functions. In general, Sobolev functions possess less regularity than Lipschitz functions. Nevertheless, Lipschitz analysis is useful in this context as well. Here we assume as known the basic definitions and facts in the theory of distributions and Sobolev spaces.

Let $\Omega \subset \mathbb{R}^n$ be open and let $1 \leq p \leq \infty$. The *Sobolev space* $W^{1,p}(\Omega)$ consists of all real-valued functions $u \in L^p(\Omega)$ whose distributional partial derivatives $\frac{\partial u}{\partial x_i}$, $i = 1, \ldots, n$, are also in $L^p(\Omega)$.

According to the well known Sobolev embedding theorem, functions in $W^{1,p}(\Omega)$, $p > n$, have continuous representatives. We also have continuity for functions in $W^{1,1}(\Omega)$ for $n = 1$, and, in all cases, for $p < \infty$, these continuous representatives need not be Lipschitz. For

$1 < p \le n$, there are nowhere continuous functions in $W^{1,p}(\Omega)$. For example, it is an easy exercise to show that the function

$$u(x) := \sum_i 2^{-i}|x - q_i|^{-\alpha}\,, \quad \alpha > 0\,, \tag{4.1}$$

is in $W^{1,p}(\mathbb{B}^n)$ for $p < n/(\alpha+1)$ whenever (q_i) is a countable set in $\mathbb{B}^n$, and that u is nowhere continuous if (q_i) is dense in $\mathbb{B}^n$. Similarly, the function

$$u(x) := \sum_i \log\log\frac{1}{|x - q_i|} \tag{4.2}$$

is in $W^{1,n}(B^n(0,e^{-e}))$ whenever (q_i) is a countable set in $B^n(0,e^{-e})$.

It is easy to see that the spaces $W^{1,p}(\Omega)$ are Banach spaces with norm

$$||u||_{1,p} := ||u||_p + ||\nabla u||_p\,,$$

where $\nabla u = (\frac{\partial u}{\partial x_1},\ldots,\frac{\partial u}{\partial x_n})$ is the distributional gradient of u.

We recall the standard approximation procedure. If $\eta \in C_0^\infty(\mathbb{R}^n)$ is a function with

$$\int_{\mathbb{R}^n} \eta(x)\,dx = 1\,, \tag{4.3}$$

then the *convolution*

$$u_\varepsilon(x) := u * \eta_\varepsilon(x) = \int_{\mathbb{R}^n} u(y)\,\eta_\varepsilon(x-y)\,dy\,, \tag{4.4}$$

where

$$\eta_\varepsilon(x) := \varepsilon^{-n}\eta(x/\varepsilon)\,, \tag{4.5}$$

is C^∞-smooth, and $u_\varepsilon \to u$ in $L^p(\Omega)$, if $u \in L^p(\Omega)$ and $1 \le p < \infty$. (To integrate over $\mathbb{R}^n$ in (4.4), we set u to be zero outside Ω.)

We also have that $u_\varepsilon \to u$ locally uniformly, if u is continuous. Moreover,

$$\partial_i u_\varepsilon = u * \partial_i\eta_\varepsilon = \partial_i u * \eta_\varepsilon\,,$$

if $u \in W^{1,p}(\Omega)$. It follows that smooth functions are dense in the Sobolev space $W^{1,p}(\Omega)$ for $1 \le p < \infty$.

Essentially, $W^{1,\infty}(\Omega)$ consists of Lipschitz functions.

Theorem 4.1. *The space $W^{1,\infty}(\Omega)$ consists of those bounded functions on Ω that are locally L-Lipschitz (for some L depending on the function). In particular, if Ω is quasiconvex, then $W^{1,\infty}(\Omega)$ consists of all bounded Lipschitz functions on Ω.*

Proof. Note that the second claim follows from the first and Lemma 2.2. To prove the first claim, assume first that $u : \Omega \to \mathbb{R}$ is locally L-Lipschitz for some L. Then u is Lipschitz on each line that is parallel to a coordinate axis. By using integration by parts on such a line (see Theorem 3.3), and then Fubini's theorem, we find that

$$\int_\Omega \frac{\partial u}{\partial x_i}(x)\,\eta(x)\,dx = -\int_\Omega u(x)\,\frac{\partial \eta}{\partial x_i}(x)\,dx$$

for each test function $\eta \in C_0^\infty(\Omega)$, and for each $i = 1, \ldots, n$. This proves that the almost everywhere existing classical gradient of u is the distributional gradient as well. Moreover, $||\nabla u||_\infty \le L$ by (3.10).

Next, assume that $u \in W^{1,\infty}(\Omega)$. Fix a ball B with compact closure in Ω. The convolutions u_ε converge to u pointwise almost everywhere in B. Moreover, we have that

$$||\nabla u_\varepsilon||_{\infty,B} \;\le\; ||\nabla u||_\infty < \infty$$

for all small enough ε. On the other hand, the functions u_ε are smooth, so that

$$u_\varepsilon(a) - u_\varepsilon(b) = \int_0^1 \nabla u_\varepsilon(b + t(a-b)) \cdot (a-b)\,dt\,,$$

and, consequently,

$$|u_\varepsilon(a) - u_\varepsilon(b)| \;\le\; ||\nabla u||_\infty |a-b|\,,$$

whenever $a, b \in B$. By letting $\varepsilon \to 0$, we find that

$$|u(a) - u(b)| \;\le\; ||\nabla u||_\infty |a-b| \tag{4.6}$$

for a, b outside a set of measure zero in B. Obviously, then, u has a continuous representative for which (4.6) holds everywhere in B. The theorem follows. □

Remark 4.2. The proof of Theorem 4.1 gives the following: If $u \in W^{1,\infty}(\Omega)$, then u is locally $||\nabla u||_\infty$-Lipschitz. Conversely, if u is bounded and locally L-Lipschitz, then $u \in W^{1,\infty}(\Omega)$, the distributional gradient of u agrees almost everywhere with the classical gradient, and $||\nabla u||_\infty \le L$.

Although Sobolev functions can exhibit rather singular behavior, as witnessed by the examples in (4.1) and (4.2), there is some regularity beneath the rough surface. We will next prove the following result.

Theorem 4.3. *Let* $u \in W^{1,p}(\Omega)$, $1 \le p \le \infty$. *Then*

$$\Omega = \bigcup_{i=1}^\infty E_i \cup Z\,,$$

where E_i are measurable sets such that $u|E_i$ is i-Lipschitz, and Z has measure zero.

Theorem 4.3 is an immediate consequence of the following proposition.

Proposition 4.4. *Let $u \in W^{1,p}(B)$, $1 \le p \le \infty$, where $B \subset \mathbb{R}^n$ is a ball. Then there exist a measurable function $g : B \to \mathbb{R}$ and a set of measure zero $Z \subset B$ such that*

$$|u(x) - u(y)| \le |x - y| \, (g(x) + g(y)) \tag{4.7}$$

whenever $x, y \in B \setminus Z$.

We will show that essentially one can take for g in (4.7) the maximal function of the gradient of u, cf. Remark 4.6.

Recall that the *maximal function* of a locally integrable function $f \in L^1_{loc}(\mathbb{R}^n)$ is defined as

$$Mf(x) := \sup_{r>0} \fint_{B(x,r)} |f(y)| \, dy \,, \tag{4.8}$$

where the barred integral sign denotes the mean value over the integration domain. The well known *maximal function theorem* of Hardy-Littlewood-Wiener is one of the fundamental results in analysis. It states that M maps $L^1(\mathbb{R}^n)$ to weak-$L^1(\mathbb{R}^n)$, and $L^p(\mathbb{R}^n)$ to $L^p(\mathbb{R}^n)$ if $1 < p \le \infty$. More precisely, we have that

$$|\{x : Mf(x) > t\}| \le C(n) \frac{||f||_1}{t} \,, \quad t > 0 \,, \tag{4.9}$$

and that

$$\int_{\mathbb{R}^n} |Mf(x)|^p \, dx \le C(n,p) \int_{\mathbb{R}^n} |f(x)|^p \, dx \,, \quad 1 < p \le \infty \,. \tag{4.10}$$

We will not prove (4.9) and (4.10) here. The standard proof can be found in many texts, e.g. in [65], [80], [50], [25].

The *Riesz potential* (of order 1) of a function f is

$$I_1(f)(x) := \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-1}} \, dy \,. \tag{4.11}$$

The mapping properties of the Riesz potential are important in proving the Sobolev embedding theorems, cf. Theorem 4.8. The starting point is the following pointwise estimate.

Proposition 4.5. *Let $u \in C^\infty(B)$ for some ball $B \subset \mathbb{R}^n$, and let $x, y \in B$. Then*

$$|u(x) - u(y)| \le C(n) \left(I_1(|\nabla u| \cdot \chi_B)(x) + I_1(|\nabla u| \cdot \chi_B)(y) \right) . \tag{4.12}$$

Proof. We have by the fundamental theorem of calculus that

$$\begin{aligned} u(x) - u(y) &= -\int_0^{|y-x|} \frac{d}{dt} u(x + t\omega)\, dt \\ &= -\int_0^{|y-x|} \nabla u(x + t\omega) \cdot \omega\, dt\,, \end{aligned}$$

where $\omega = \frac{y-x}{|y-x|} \in \mathbb{S}^{n-1}$. Integrating over y then gives

$$|B|(u(x) - u_B) \;=\; -\int_B \int_0^{|y-x|} \nabla u(x + t\omega) \cdot \omega\, dt\, dy\,,$$

where u_B denotes the mean value of u over B. Next, we extend $|\nabla u|$ to an integrable function on all of $\mathbb{R}^n$ by setting it equal to zero outside of B, and obtain from the preceding that

$$\begin{aligned} |B|\,|u(x) - u_B| \;&\leq\; \int_{\mathbb{R}^n} \int_0^{\operatorname{diam}(B)} |\nabla u(x + t\omega)|\, dt\, dy \\ &= \int_0^{\operatorname{diam}(B)} \int_{\mathbb{R}^n} |\nabla u(x + t\omega)|\, dy\, dt \\ &= \int_0^{\operatorname{diam}(B)} \int_{\mathbb{S}^{n-1}} \int_0^{\operatorname{diam}(B)} |\nabla u(x + t\omega)|\, r^{n-1}\, dr\, d\omega\, dt \\ &= \frac{(\operatorname{diam}(B))^n}{n} \int_0^{\operatorname{diam}(B)} \int_{\mathbb{S}^{n-1}} |\nabla u(x + t\omega)|\, d\omega\, dt \\ &= C(n)\,|B| \int_{\mathbb{S}^{n-1}} \int_0^{\operatorname{diam}(B)} |\nabla u(x + t\omega)|\, t^{1-n} t^{n-1} dt\, d\omega \\ &= C(n)\,|B| \int_B \frac{|\nabla u(y)|}{|x - y|^{n-1}}\, dy\,. \end{aligned}$$

In conclusion,

$$|u(x) - u_B| \;\leq\; C(n) \int_B \frac{|\nabla u(y)|}{|x - y|^{n-1}}\, dy\,,$$

from which (4.12) follows by the triangle inequality. The proposition is proved. $\square$

Proof of Proposition 4.4. Assume first that $u \in C^\infty(B)$. We use the definition for the maximal function together with the pointwise estimate (4.12) to achieve (4.7). Thus, fix $x, y \in B$, and assume that x and y lie in a ball $B' \subset B$ whose diameter does not exceed $2|x - y|$. (If this is not the case, one has to perform a standard "chaining argument". In any case, for the purposes of Theorem 4.3, a weaker version of Proposition 4.4 would suffice, where one considers only points $x, y \in \frac{1}{2}B$. We leave the details here to the reader.) We apply Proposition 4.5 to the ball B', and get

$$\begin{aligned} I_1(|\nabla u| \cdot \chi_{B'})(x) &= \sum_{i=0}^{\infty} \int_{B(x,2^{-i}\mathrm{diam}(B'))\setminus B(x,2^{-i-1}\mathrm{diam}(B'))} \frac{|\nabla u(z)|}{|x-z|^{n-1}}\, dz \\ &\leq C(n) \sum_{i=1}^{\infty} 2^{-i}\mathrm{diam}(B') \fint_{B(x,2^{-i}\mathrm{diam}(B'))} |\nabla u(z)|\, dz \\ &\leq C(n)|x-y|\, M(|\nabla u|)(x)\,, \end{aligned}$$

as required.

To finish the proof, we use a routine approximation argument together with (4.9) and (4.10). We leave the details to the reader, and conclude the proof of Proposition 4.4. □

Remark 4.6. The proof of Proposition 4.4 shows that in (4.7) one can choose

$$g(x) := C(n,p)M(|\nabla u|)(x)\,,$$

for some appropriate constant $C(n,p)$ depending only on n and p.

4.1. **Approximate differentiability.** A function $f : A \to \mathbb{R}^m$, $A \subset \mathbb{R}^n$, is said to be *approximately differentiable* at $a \in A$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that the differential quotient

$$\frac{|f(x) - f(a) - L(x-a)|}{|x-a|}\,, \quad x \in A \setminus \{a\}\,,$$

has approximate limit zero at a. Recall that a function $g : A \to \mathbb{R}$, $A \subset \mathbb{R}^n$, has *approximate limit* $b \in \mathbb{R}$ at $a \in A$ if

$$\lim_{r\to 0} \frac{|B(a,r) \cap (\mathbb{R}^n \setminus g^{-1}(B(a,\varepsilon)))|}{|B(a,r)|} = 0$$

for every $\varepsilon > 0$. The linear map L is called the *approximate derivative* of f at a, and denoted by $\mathrm{ap}Df(a)$. It is easy to see that an approximate derivative, if it exists, is unique, so the terminology is justified. The definitions imply that if f is approximately differentiable at $a \in A$, then a is a point of Lebesgue density of A.

Theorem 4.3 has the following interesting corollary.

Theorem 4.7. *Let $u \in W^{1,p}(\Omega)$, $1 \leq p \leq \infty$. Then u is approximately differentiable almost everywhere.*

Proof. Consider $u_i := u|E_i$, where E_i is one of the sets provided by Theorem 4.3. Then u_i can be extended to be a Lipschitz function in all of $\mathbb{R}^n$. By Rademacher's theorem, this extension is differentiable almost everywhere in E_i. Because almost every point of E_i is a Lebesgue density point, it is easy to see that the derivative (in the sense of (3.1)) of the extension of u_i is an approximate derivative of u almost everywhere in E_i. This proves the theorem. □

4.2. **Proof of Rademacher's theorem via Sobolev embedding.** Let us begin with the following Sobolev embedding theorem, cf. Exercise 4.3 (b).

Theorem 4.8. *Let $u \in W^{1,p}(B)$, where $B \subset \mathbb{R}^n$ is a ball, and let $p > n$. Then u agrees almost everywhere with a continuous function such that*

$$|u(x) - u(y)| \leq C(n,p)\, |x-y|^{1-n/p} ||\nabla u||_{p,B}\,. \tag{4.13}$$

Proof. We use the pointwise estimate (4.12) together with Hölder's inequality. Indeed,

$$I_1(|\nabla u| \cdot \chi_B)(x) \leq ||\nabla u||_{p,B} ||\,|\cdot - x|^{1-n}||_{p/(p-1),B}$$

and a computation gives for the second norm the bound

$$C(n,p)\, (\mathrm{diam}(B))^{1-n/p}\,.$$

Thus, the claim follows for $u \in C^\infty(\Omega)$, and the general case follows by using the convolution approximations. □

4.3. **Exercises.** (a) Fill in the details in the proofs for Proposition 4.4 and Theorem 4.8.

(b) Prove that I_1 maps $L^p(\mathbb{R}^n)$ to $L^{np/(n-p)}(\mathbb{R}^n)$ for $1 < p < n$, and $L^1(\mathbb{R}^n)$ to weak-$L^{n/(n-1)}(\mathbb{R}^n)$, boundedly. (Hint: Use the maximal function theorem as in [80, Section 2.8] or [25, Chapter 3], for example.)

The following extension of Rademacher's theorem is due to Cesari [13] and Calderón [11].

Theorem 4.9. *Let $u \in W^{1,p}(\Omega)$ for $p > n$. Then u is differentiable almost everywhere in Ω.*

Proof. We will show that u is differentiable at every L^p-Lebesgue point of the gradient ∇u. (The definition for such a point becomes clear from the ensuing argument.) By basic real analysis, almost every point is such a point. Indeed, let a be an L^p-Lebesgue point of ∇u, and let

$$f(x) := u(x) - \nabla u(a) \cdot x\,.$$

Then $f \in W^{1,p}(B(a,r))$ for all small enough $r > 0$, and it follows from (4.13) that

$$\begin{aligned}
&|u(x) - u(a) - \nabla u(a) \cdot (x-a)| \; = \; |f(a) - f(x)| \\
&\quad \le C(n,p)\, |x-a|^{1-n/p} \Big(\int_{B(a,|a-x|)} |\nabla u(x) - \nabla u(a)|^p \, dx \Big)^{1/p} \\
&\quad \le \; C(n,p)\, |a-x| \Big(\fint_{B(a,|a-x|)} |\nabla u(x) - \nabla u(a)|^p \, dx \Big)^{1/p} .
\end{aligned}$$

By the Lebesgue point assumption, the last integral average tends to 0 as $|a-x| \to 0$. This proves the differentiability. $\square$

4.4. **Notes to Section 4.** The material in this section is standard. For more information about Sobolev spaces, see e.g. [65], [17], [80]. In [21], Hajłasz took the conclusion (4.7) in Proposition 4.4 as a definition for Sobolev functions, starting an extensive development of the Sobolev space theory in general metric measure spaces. See [23], [25], [39], [22], and the references there. For careful studies on approximate derivatives, see [18] and [17]. Approximate derivatives for Sobolev functions in general spaces have been studied by Keith [37].

5. Whitney flat forms

The *differential* of a Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$,

$$df(x) := \frac{\partial f}{\partial x_1}(x)\, dx_1 + \cdots + \frac{\partial f}{\partial x_n}(x)\, dx_n \,, \tag{5.1}$$

is a differential 1-form in $\mathbb{R}^n$ with bounded measurable coefficients,

$$\frac{\partial f}{\partial x_i} \in L^\infty(\mathbb{R}^n) \,.$$

This is a plain observation in view of the results in previous sections.

There is a deeper connection between Lipschitz functions and differential forms with bounded measurable coefficients, which we will explore in this section. This connection was first realized by Whitney, who initiated a geometric integration theory in the 1940s [77]. The theory, which is fully presented in Whitney's monograph [78], is based upon geometric objects called flat chains, and their dual objects called flat cochains. The latter turn out to be identifiable as bounded measurable differential forms with bounded exterior derivative, according to a result of J. H. Wolfe. Such differential forms are called flat forms; they are Lipschitz invariant.

In this section, we present the basic theory of flat forms, including a proof of Wolfe's theorem. We begin by reviewing some basic exterior algebra.

5.1. **Exterior algebra.** Let V be a real vector space of finite dimension n. The *exterior algebra of* V is a graded anticommutative algebra

$$\wedge_* V = \wedge_0 V \oplus \wedge_1 V \oplus \cdots \oplus \wedge_n V \oplus 0 \oplus \dots ,$$

where we have the *exterior multiplication*

$$v \wedge w \in \wedge_{k+l} V \tag{5.2}$$

satisfying

$$v \wedge w = (-1)^{kl} w \wedge v \,, \tag{5.3}$$

whenever $v \in \wedge_k V$ and $w \in \wedge_l V$. (Here property (5.2) explains the term *graded* and property (5.3) the term *anticommutative*.) In addition, we have that

$$\wedge_0 V = \mathbb{R}\,, \qquad \wedge_1 V = V\,.$$

It follows that if $\{e_1, \dots, e_n\}$ is a basis of V, then

$$\{e_{i_1} \wedge \dots \wedge e_{i_k} : 1 \le i_1 < \cdots < i_k \le n\}$$

is a basis of $\wedge_k V$. In particular,

$$\dim \wedge_k V = \binom{n}{k}\,.$$

Elements in $\wedge_k V$ are called *k-vectors of* V.

If V^* is the dual space of V, we write

$$\wedge^k V := \wedge_k V^*\,, \qquad \wedge^* V := \wedge_* V^*\,.$$

Thus, if $\{e_1^*, \dots, e_n^*\}$ is a basis of V^*, dual to $\{e_1, \dots, e_n\}$, then

$$\{e_{i_1}^* \wedge \dots \wedge e_{i_k}^* : 1 \le i_1 < \cdots < i_k \le n\}$$

is a basis of $\wedge^k V$. In particular,

$$\wedge^k V = (\wedge_k V)^*\,.$$

An *orientation* of a vector space V is an equivalence class of ordered bases, where two ordered bases are equivalent if they can be transformed to each other by a linear transformation with positive determinant. Alternatively, an orientation of an n-dimensional real vector space V is a choice of one of the two components of the complement of 0 in the one-dimensional space $\wedge_n V$.

Every inner product in V determines an inner product in $\wedge_k V$, and hence in $\wedge^k V$. We simply declare

$$\{e_{i_1} \wedge \dots \wedge e_{i_k} : 1 \le i_1 < \cdots < i_k \le n\}$$

to be an orthonormal basis of $\wedge_k V$, if $\{e_1, \ldots, e_n\}$ is an orthonormal basis of V. If V is an inner product space, there is canonical isomorphism between $\wedge_k V$ and $\wedge^k V$.

The following particular case will be important to us. If V is a k-dimensional vector subspace of $\mathbb{R}^n$, then it inherits the standard inner product from $\mathbb{R}^n$. An orientation of V can be signified by a unit k-vector

$$v_1 \wedge \ldots \wedge v_k \in \wedge_k V\,.$$

Such a k-vector is called a *k-direction of V*. If $\{v_1, \ldots, v_k\}$ is any collection of linearly independent vectors in an oriented k-dimensional subspace V of $\mathbb{R}^n$, such that $v_1 \wedge \ldots \wedge v_k$ falls in the chosen component of $\wedge_k V$, then

$$\frac{v_1 \wedge \ldots \wedge v_k}{|v_1 \wedge \ldots \wedge v_k|} \tag{5.4}$$

is a k-direction of V. (Here $|\cdot|$ denotes the norm determined by the inner product.)

More generally, if V is a k-dimensional affine subspace of $\mathbb{R}^n$, then its orientation is a choice of equivalent k-vectors of the form $v_1 \wedge \ldots \wedge v_k$, where $\{v_1, \ldots, v_k\} \subset V - a$, $a \in V$, is a linearly independent set. Similarly, a k-direction of V is a unit k-vector $v_1 \wedge \ldots \wedge v_k \in V - a$.

For a more detailed discussion of exterior algebra we refer to [18, Chapter 1] and [78, Chapter I].

5.2. **Mass and comass.** If V has an inner product, then, as mentioned earlier, there is an associated inner product in each of the spaces $\wedge_k V$ and $\wedge^k V$. We denote the inner product by $\langle \cdot, \cdot \rangle$ and the induced norm by $|\cdot|$. (Compare (5.4).) Naturally, these norms in $\wedge_k V$ and $\wedge^k V$ are dual norms. In geometric integration theory, a different pair of dual norms is needed.

We call a k-vector $\varphi \in \wedge_k V$ *simple* if it can be written as an exterior product of vectors in V, that is, if

$$\varphi = v_1 \wedge \ldots \wedge v_k$$

for some collection of vectors $\{v_1, \ldots, v_k\} \subset V$. There is a natural one-to-one correspondence between unit simple k-vectors and oriented k-dimensional subspaces of V. We define simple k-covectors similarly. See [18, 1.6.1] or [78, I. 9] for more about the geometry of simple vectors.

The *comass* of a k-covector $\omega \in \wedge^k V$ is defined as

$$||\omega|| := \sup\{\langle \omega, \varphi \rangle : \varphi \in \wedge_k V \text{ is simple and } |\varphi| \leq 1\}. \tag{5.5}$$

Next, the *mass* of a k-vector $\varphi \in \wedge_k V$ is defined as

$$||\varphi|| := \sup\{\langle \omega, \varphi \rangle : \omega \in \wedge^k V \text{ and } ||\omega|| \leq 1\}. \tag{5.6}$$

We have that $|\omega| = ||\omega||$ for $\omega \in \wedge^k V$ if and only if ω is simple, and similarly for k-vectors. Both mass and comass are norms, dual to each other. They are comparable to the norms coming from the inner product; see [18, 1.8] for more precise statements.

5.3. **Differential forms.** We denote by $\{e_1, \dots, e_n\}$ the standard basis of $\mathbb{R}^n$, and by $\{dx_1, \dots, dx_n\}$ the dual basis. A *k-vectorfield* in an open set $\Omega \subset \mathbb{R}^n$ is a map

$$\Omega \to \wedge_k \mathbb{R}^n \,,$$

and a (*differential*) *k-form* in Ω is a map

$$\Omega \to \wedge^k \mathbb{R}^n \,.$$

Notice that 0-vectorfields and 0-forms are simply real-valued functions. As the target space for vectorfields and forms is a finite dimensional vector space, we have natural notions of smooth, measurable, etc. vectorfields and forms. In this article, differential forms are more important than vectorfields although the latter will briefly appear. In the ensuing discussion, we concentrate on forms although much of the terminology goes over to vectorfields as well.

Thus, a differential k-form in Ω is an expression of the form

$$\omega(x) = \sum a_{i_1 \dots i_k}(x)\, dx_{i_1} \dots dx_{i_k} \,, \tag{5.7}$$

where the sum is taken over all increasing sequences $i_1 < \cdots < i_k$ of numbers from $\{1, \dots, n\}$. The functions $a_{i_1 \dots i_k}$ are the *coefficients* of ω.

Differential k-forms in Ω, with coefficients in a fixed linear function space, form a vector space in an obvious manner. Moreover, the collection of all differential forms with coefficients in a fixed linear function space has the structure of a graded anticommutative algebra, obviously inherited from $\wedge^* \mathbb{R}^n$. The multiplication between generating 1-forms obeys the rule

$$dx_i dx_j = -dx_j dx_i \,,$$

where, as customary, we abbreviate $dx_i dx_j := dx_i \wedge dx_j$. (For arbitrary forms, we typically write $\omega \wedge \eta$ rather than $\omega\, \eta$.)

When one multiplies a k-form with an $(n-k)$-form, the resulting n-form,

$$\omega(x) = a(x)\, dx_1 \dots dx_n \,,$$

often called a *volume form*,[3] can be integrated, provided the coefficient function $a \in L^1(\Omega)$. We write

$$\int_\Omega \omega = \int_\Omega a(x)\, dx_1 \dots dx_n = \int_\Omega a(x)\, dx\,.$$

We denote by

$$\wedge^k(\Omega; F)$$

the k-forms in Ω with coefficients in a function space F. Note that

$$\wedge^0(\Omega; F) = F\,.$$

We also abuse notation and write $\omega \in F$, instead of $\omega \in \wedge^k(\Omega; F)$, if there is no danger of confusion.

For forms with smooth coefficients we have the *exterior differential*

$$d : \wedge^k(\Omega; C^\infty(\Omega)) \to \wedge^{k+1}(\Omega; C^\infty(\Omega))$$

defined by

$$df := \sum_{i=1}^n \frac{\partial f}{\partial x_i}\, dx_i\,,$$

for $k = 0$, and

$$d\omega(x) = \sum da_{i_1 \dots i_k}(x)\, dx_{i_1} \dots dx_{i_k}\,,$$

if ω is as in (5.7).

The exterior differential can be extended to forms with locally integrable coefficients as in the theory of distributions. Thus, let $\omega \in \wedge^k(\Omega; L^1_{\mathrm{loc}}(\Omega))$. We say that a $(k+1)$-form $\alpha \in \wedge^{k+1}(\Omega; L^1_{\mathrm{loc}}(\Omega))$ is the *distributional exterior differential* of ω if

$$\int_\Omega \alpha \wedge \eta = (-1)^{k+1} \int_\Omega \omega \wedge d\eta$$

for every $\eta \in \wedge^{n-k-1}(\Omega; C_0^\infty(\Omega))$. It is easy to see that if a distributional exterior differential exists, it is unique; therefore we write $\alpha =: d\omega$.

The following proposition is a direct consequence of the definition, and of the fact that $dd\eta = 0$ for smooth forms η.

Proposition 5.1. *We have that*

$$dd\omega = 0$$

in the sense of distributions.

---

[3]Sometimes the term "volume form" is reserved for n-forms with nonzero or positive coefficient.

5.4. **Forms acting on oriented polyhedra.** There is fundamental duality between differential forms and oriented polyhedra of correct dimension, which is an integrated version of the pointwise duality between forms and vectorfields. For simplicity, let us assume in this section that the vectorfields and forms are defined in all of $\mathbb{R}^n$.

Consider first the case $k = 1$. Let ω be a smooth 1-form, and let $[a, b]$ be an oriented line segment in $\mathbb{R}^n$. Here by an *oriented line segment* we mean an orientation in the affine line determined by the two points $a, b \in \mathbb{R}^n$ as explained in the end of Section 5.1. The notation $[a, b]$ moreover signifies that the orientation is given by the unit vector

$$v_{a,b} := \frac{b-a}{|b-a|}\,.$$

Now we can let ω "act" on $[a, b]$ by integration:

$$\langle \omega, [a, b]\rangle := \int_{[a,b]} \omega := \int_0^{|b-a|} \langle \omega(a + tv_{a,b}), v_{a,b}\rangle \, dt\,. \tag{5.8}$$

Notice the obvious sign change,

$$\langle \omega, [a, b]\rangle = -\langle \omega, [b, a]\rangle\,,$$

as well as the fundamental theorem of calculus, or *Stokes's theorem*,

$$\langle df, [a, b]\rangle = f(b) - f(a)\,, \quad f \in C^\infty(\mathbb{R}^n)\,.$$

Naturally, we can interpret (5.8) as $[a, b]$ acting on smooth 1-forms. In this interpretation, the action is clearly linear. We will later see how the action of ω on oriented line segments can be thought of as a linear action as well.

Consider now the case $k = 2$. Let ω be a smooth 2-form, and let v_1 and v_2 be two linearly independent vectors in $\mathbb{R}^n$. Then v_1 and v_2, in this order, determine an oriented two-dimensional subspace V of $\mathbb{R}^n$, and we have the corresponding k-direction

$$\frac{v_1 \wedge v_2}{|v_1 \wedge v_2|}\,.$$

If P is any 2-simplex such that its translate $P - p \subset V$ for some (all) $p \in P$, then we can think of P being oriented as V is, and define the "action" of ω on P by

$$\langle \omega, P\rangle := \int_P \omega := \int_P \langle \omega(x), \frac{v_1 \wedge v_2}{|v_1 \wedge v_2|}\rangle \, d\mathcal{H}_2(x)\,, \tag{5.9}$$

where $\mathcal{H}_2$ is the Hausdorff 2-measure (area measure) on P.

Keeping P fixed, we have a linear action on forms in (5.9). The linearity of the action on P will be studied in the next subsection.

It is now clear how to continue for forms and simplexes of higher degree and dimension. We leave the details to the reader.

Remark 5.2. (a) The action of forms as described above can be extended to more general *oriented rectifiable sets*, in particular to smooth manifolds. For example, consider a bounded subset M of $\mathbb{R}^n$ of finite Hausdorff k-measure that possesses an *approximate tangent plane* at $\mathcal{H}_k$-almost every point. We can measurably orient these tangent planes by choosing a k-direction, then pair each choice with a smooth form, and finally integrate the outcome over M by using the Hausdorff measure.

This procedure, as before, can also be thought of as M acting linearly on smooth forms. In fact, such an M is an example of a *rectifiable current*. We will discuss currents later in these lectures. See [18, Chapter 4] for a thorough exposition of these ideas.

(b) Initially, in his book, Whitney defines the integration of a continuous form over an oriented simplex somewhat differently by using essentially Riemann integration [78, Chapter III]. Later, in [78, Chapter IX, Section 5], when only measurable forms are discussed, an approach based on the Lebesgue theory is taken.

5.5. **Flat forms.** We can equip differential forms with various norms, or topologies, depending on our goals. Standard choices are various L^p and Sobolev norms, or locally convex topologies as in the distribution theory. In geometric measure theory, the flat norm is a pivotal concept. To avoid certain technical issues, we only consider globally defined flat forms.

A k-form ω in $\mathbb{R}^n$ is called *flat* if $\omega \in L^\infty(\mathbb{R}^n)$ and if also $d\omega \in L^\infty(\mathbb{R}^n)$ in the sense of distributions. The vector space of flat k-forms in $\mathbb{R}^n$ is denoted by $\mathcal{F}^k(\mathbb{R}^n)$. It is a Banach space under the *flat norm*

$$||\omega||_\flat := \max\{||\omega||_\infty, ||d\omega||_\infty\}\,. \tag{5.10}$$

Here and later the L^∞-norm $||\alpha||_\infty$ for a form α stands for the L^∞-norm of the *pointwise comass*,

$$||\alpha|| := \operatorname{ess\,sup}_{x\in\mathbb{R}^n} ||\alpha(x)||\,, \qquad \alpha(x) \in \wedge^k(\mathbb{R}^n)\,. \tag{5.11}$$

Remark 5.3. (a) We could have used any of the standard equivalent norms in the finite dimensional space $\wedge^k(\mathbb{R}^n)$ when defining flat forms. From an analytic point of view this makes little difference. The comass, as defined in (5.5), is the most suitable in the present geometric context as we will see.

(b) Obviously, one can consider the Banach space of flat forms defined in a given open set $\Omega \subset \mathbb{R}^n$.

Later we will give a fundamental description of the analytically described space of flat k-forms as the dual Banach space of a geometrically described space of flat k-chains.

One can also show that Lipschitz maps pull back flat forms to flat forms. Let us next see why such an assertion is not trivial. Thus, let $f : \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz and let $\omega = a_1 dx_1 + \cdots + a_m dx_m$ be a flat 1-form in $\mathbb{R}^m$. If ω is smooth, its pullback is well defined,

$$f^*\omega(x) := a_1(f(x))df_1 + \cdots + a_m(f(x))df_m\,, \tag{5.12}$$

and is obviously a flat 1-form in $\mathbb{R}^n$. But if ω only has bounded measurable coefficients, the expression in (5.12) is easily meaningless a priori. Namely, f can map a set of positive measure to a point, where the values of the coefficients a_i are not well defined. To counter this example, one may argue that at such points the differential df_i must vanish, and we can set $f^*\omega(x) = 0$. In a sense, this is true, but in order to make everything precise, we need to understand the dual space nature of flat forms.

Before we discuss flat chains, I want to present another hallmark of flat forms. If $[a, b]$ is an oriented line segment in $\mathbb{R}^n$, then a line integration of a sort,

$$\int_{[a,b]} \omega\,, \tag{5.13}$$

can be defined for all flat 1-forms ω. This is surprising because lines in $\mathbb{R}^n$, $n \geq 2$, have measure zero, and a priori there is no well defined trace of a flat 1-form on a given line segment. As in the case of Lipschitz pullback, the integration (5.13) has to be understood through appropriate duality.

5.6. **Flat chains.** We describe the space of flat k-chains; this space will be shown to be a predual of the space of flat k-forms. We continue to consider global objects only, and begin with chains of small degree.

A *polyhedral 0-chain* in $\mathbb{R}^n$ is a finite real linear combination of points. The *mass* of a 0-chain

$$\sigma = \sum_{i=1}^{N} \lambda_i a_i\,, \qquad \lambda_i \in \mathbb{R}\,,\ a_i \in \mathbb{R}^n\,, \tag{5.14}$$

is

$$|\sigma| := \sum_{i=1}^{N} |\lambda_i|\,.$$

Here it is assumed that there are no multiple appearances among the points a_i in (5.14). Two 0-chains can be added in a natural manner, and so we obtain a vector space.

The description of polyhedral 1-chains is slightly more complicated. Consider a formal linear combination

$$\sigma = \sum_{i=1}^{N} \lambda_i [a_i, b_i]\,, \qquad \lambda_i \in \mathbb{R}\,, \tag{5.15}$$

of oriented line segments $[a_i, b_i] \subset \mathbb{R}^n$. First we break each of the line segments $[a_i, b_i]$ into line segments $[a_{i_1}, b_{i_1}], \dots, [a_{i_k}, b_{i_k}]$, keeping with the orientation, such that the total collection of new line segments $\{[a_{i_j}, b_{i_j}]\}$ has the property that any two segments from the collection either coincide as sets or meet at most at one point. Then we replace each of the summand in (5.15) by

$$\sum_{j=1}^{k} \lambda_i [a_{i_j}, b_{i_j}]$$

to get another formal linear combination σ'. In this combination, we add up the coefficients in front of each two similarly oriented coinciding line segments. Moreover, we stipulate that

$$\lambda[a, b] = -\lambda[b, a] \tag{5.16}$$

so that any two oppositely oriented coinciding line segments with same coefficient will cancel each other. At the end, we can assume that any two line segments from the expression making up σ' meet at most at one point. Such a new combination σ' is called a *refinement* of σ.

The preceding understood, a *polyhedral 1-chain* is an equivalence class of expressions σ as in (5.15), with two expressions identified should they have a common refinement.

Polyhedral 1-chains form a real vector space in an obvious manner.

We define *polyhedral 2-chains* similarly. These are equivalence classes of formal real linear combinations of oriented 2-simplexes in $\mathbb{R}^n$, where we identify two combinations if they can be refined so as to agree, keeping with the rule that a change in the orientation of a simplex changes the sign of the associated coefficient. We leave it to the reader to make this definition more rigorous.

Finally, *polyhedral k-chains* in $\mathbb{R}^n$ for each $k \leq n$ can be defined in a similar manner.

An oriented 2-simplex in $\mathbb{R}^n$ can be signified by $[a, b, c]$ for three points $a, b, c \in \mathbb{R}^n$ not lying on a line. As a set it is the convex hull of the three points, and the orientation is determined by the given order.

Analogous notation can be used for an oriented k-simplex, $k \geq 1$. There is no orientation for 0-chains.

As is customary, we will speak of a polyhedral chain in connection with expressions like in (5.15), without referring to the full equivalence class. The *mass* of a general polyhedral k-chain $\sigma = \sum_i \lambda_i \sigma_i$ is defined as

$$|\sigma| := \sum_i |\lambda_i||\sigma_i|\,, \tag{5.17}$$

where $|\sigma_i|$ stands for the k-dimensional area (Lebesgue measure) of a k-simplex σ_i. In addition, it is understood in (5.17) that in the expression for σ the k-simplexes σ_i meet only along lower dimensional parts; according to the rules of equivalence, this can be assumed.

The *boundary* of a polyhedral k-chain is a polyhedral $(k-1)$-chain, defined in the usual way. For example,

$$\partial[a,b] = b - a\,,$$

and

$$\partial[a,b,c] = [a,b] + [b,c] + [c,a] = [a,b] - [a,c] + [b,c]\,.$$

We have that $\partial\partial\sigma = 0$.

The *flat norm* of a polyhedral k-chain σ is defined as

$$|\sigma|_\flat := \inf\{|\sigma - \partial\tau| + |\tau|\}\,, \tag{5.18}$$

where the infimum is taken over all polyhedral $(k+1)$-chains τ in $\mathbb{R}^n$.

We observe that for every $(k+1)$-chain τ,

$$|\partial\sigma|_\flat \leq |\partial\sigma - \partial(\sigma - \partial\tau)| + |\sigma - \partial\tau| \leq |\sigma - \partial\tau| + |\tau|\,,$$

which gives that

$$|\partial\sigma|_\flat \leq |\sigma|_\flat\,. \tag{5.19}$$

It is instructive to study the flat norm in low degrees. Let σ be a polyhedral 0-chain. If σ is presented just by one point (with a weight), then the flat norm agrees with the mass. This is true also if $\sigma = a + b$ for $a, b \in \mathbb{R}^n$ (cf. Exercise 5.7 (b)). But if $\sigma = a - b$ for $a, b \in \mathbb{R}^n$, then

$$|\sigma|_\flat \leq |a - b - \partial[a,b]| + |a - b| = |a - b|\,,$$

which is less than the mass if $|a - b| < 2$. It is in fact easy to see that

$$|\sigma|_\flat = \min\{|a-b|, 2\} \tag{5.20}$$

if $\sigma = a - b$.

Next, consider a polyhedral 1-chain σ. If σ is presented by just a single line segment (with a weight), then the flat norm and the mass agree. If there are two line segments involved, then their mutual location becomes relevant. For example, the flat norm of the 2-chain

$$\sigma := [0, e_1] + [e_1 + \varepsilon e_2, \varepsilon e_2]\,, \quad \varepsilon > 0\,,$$

in $\mathbb{R}^2$ is at most 3ε.

5.7. **Exercise.** (a) Let $\sigma = [e, 0] + [0, e']$ for two unit vectors e, e'. Show that $|\sigma|_\flat = |\sigma| = 2$ if and only if $e = -e'$.

(b) Suppose that $\sigma_1, \ldots, \sigma_N$ are similarly oriented disjoint k-simplexes in a k-dimensional affine subspace of $\mathbb{R}^n$. Prove that

$$|\sigma| = |\sigma|_\flat$$

if

$$\sigma = \sum_{i=1}^{N} \lambda_i \sigma_i\,, \quad \lambda_i \geq 0\,.$$

The precise value of the flat norm of a general 1-chain σ as in the preceding exercise 5.7 (a) seems difficult to determine. See the comment in [78, Example 5 (a), p. 158]. I do not know if anyone has studied the general question of determining flat norms for various standard polyhedral chains.

Nevertheless, the flat norm is always positive for a nontrivial chain.

Proposition 5.4. *The flat norm is a norm.*

To prove the proposition, it suffices to show that $|\sigma|_\flat > 0$ whenever σ is a nontrivial polyhedral k-chain; the rest is routine. To do this, however, is not an entirely trivial matter. We will prove Proposition 5.4 later after we have discussed currents in the next subsection.

The vector space of polyhedral k-chains equipped with the flat norm is called the space of *polyhedral flat k-chains* in $\mathbb{R}^n$, and denoted by $\mathcal{P}_k(\mathbb{R}^n)$. The completion of this normed space is the Banach space of *flat k-chains* in $\mathbb{R}^n$, denoted by $\mathcal{F}_k(\mathbb{R}^n)$.

Despite the relative simplicity of its definition, the members in $\mathcal{F}_k(\mathbb{R}^n)$ do not lend themselves to easy description. There are moreover some surprising examples.

5.8. **Exercise.** Show that every Jordan curve in $\mathbb{R}^2$ can be viewed as a flat 1-chain. (Hint: Every Jordan region can be exhausted by Jordan regions with polygonal boundary. Here by a *Jordan curve* we mean a topological circle, and a *Jordan region* in $\mathbb{R}^2$ is the bounded component of the complement of a Jordan curve.)

The following fundamental duality was proved by J. H. Wolfe in his Harvard thesis in 1948 (see [78, p. viii]).

Theorem 5.5. *The space $\mathcal{F}^k(\mathbb{R}^n)$ of flat k-forms is the Banach space dual of the space $\mathcal{F}_k(\mathbb{R}^n)$ flat k-chains.*

We emphasize that Theorem 5.5 asserts that the identification of $\mathcal{F}^k(\mathbb{R}^n)$ as the dual space of $\mathcal{F}_k(\mathbb{R}^n)$ is isometric; the flat norm agrees with the dual norm. Also recall that the comass is used in the definition for the flat norm in $\mathcal{F}^k(\mathbb{R}^n)$, cf. (5.11).

Remark 5.6. (a) Whitney's presentation of Wolfe's theorem in [78] does not involve the language of distributional derivatives. Presumably, the same is true for Wolfe's proof. (Note that L. Schwartz's treatise [58] appeared two years after Wolfe finished his thesis.) Flat forms were defined somewhat differently, but equivalently, in [78, Chapter IX].

(b) We could have considered flat chains in an arbitrary open subset Ω of $\mathbb{R}^n$. In this case, the definition requires some technical modifications, which I do not want to go into here. See [78, Chapter VIII].

We will prove the duality by viewing the flat chains as currents.

5.9. **Flat chains as currents.** The theory of currents is an extension of the theory of distributions; currents act on forms of any given degree. More precisely, denote by

$$\mathcal{D}^k(\mathbb{R}^n)$$

the vector space of smooth compactly supported k-forms in $\mathbb{R}^n$ (the *test forms*). A *k-dimensional current* is a linear map

$$T : \mathcal{D}^k(\mathbb{R}^n) \to \mathbb{R}$$

satisfying a continuity condition analogous to that of distributions: for every compact set $K \subset \mathbb{R}^n$ there exist a constant $C \geq 0$ and an integer $N \geq 0$ such that

$$|T(\omega)| \leq C \max_{|\alpha| \leq N} ||\partial^\alpha \omega||_\infty \tag{5.21}$$

for every $\omega \in \mathcal{D}^k(\mathbb{R}^n)$ with support in K. The maximum in (5.21) is taken over all partial derivatives of the components of ω up to order N. When $k = 0$, we have that

$$\mathcal{D}^0(\mathbb{R}^n) = C_0^\infty(\mathbb{R}^n)\,,$$

so that 0-dimensional currents are precisely the Schwartz distributions. The smallest integer N that works for every compact set in the definition (5.21) is called the *order* of a current; the order is infinite if no such integer exists.

We denote the vector space of k-dimensional currents on $\mathbb{R}^n$ by

$$\mathcal{D}_k(\mathbb{R}^n)\,.$$

Note that neither the space $\mathcal{D}^k(\mathbb{R}^n)$ of test forms nor the space $\mathcal{D}_k(\mathbb{R}^n)$ of currents has a canonical norm. (They have a natural structure of a topological vector space, but we will not go into this.) However, both spaces can be normed in various ways depending on what applications one has in mind. We will consider one such norming, pertaining to the theory of flat forms.

Thus, we equip $\mathcal{D}^k(\mathbb{R}^n)$ with the flat norm (5.10), and denote the resulting normed space by $\mathbf{F}^k(\mathbb{R}^n)$. The dual space of this normed space is a Banach space when normed by the dual norm, also called a *flat norm*. (Recall that the dual of every normed space is complete under the dual norm.) We denote the dual space by

$$\mathbf{F}_k(\mathbb{R}^n) := \mathbf{F}^k(\mathbb{R}^n)^*\,.$$

Thus,

$$||T||_\flat := \sup\{|T(\omega)| : ||\omega||_\flat \le 1\} \tag{5.22}$$

for $T \in \mathbf{F}_k(\mathbb{R}^n)$. It is clear that every element in the dual space $\mathbf{F}_k(\mathbb{R}^n)$ is a current (of order one); the members of $\mathbf{F}_k(\mathbb{R}^n)$ are called *flat k-currents.*

Note here that the space $\mathbf{F}^k(\mathbb{R}^n)$ of smooth forms equipped with the flat norm is not a Banach space; it is not complete in the flat norm. We do not bother to understand the completion of $\mathbf{F}^k(\mathbb{R}^n)$, but view this space as an auxiliary tool. Its dual space $\mathbf{F}_k(\mathbb{R}^n)$ is a huge Banach space that provides us a stage for our play.

The key fact is that the space of flat k-chains sits isometrically in $\mathbf{F}_k(\mathbb{R}^n)$.

Proposition 5.7. *We have a canonical isometric embedding*

$$\mathcal{F}_k(\mathbb{R}^n) \subset \mathbf{F}_k(\mathbb{R}^n)\,. \tag{5.23}$$

The meaning of the term "canonical" in the preceding statement requires an explanation. Indeed, each polyhedral chain has a natural interpretation as a current, and it is this action that determines the isometric embedding for the dense set of polyhedral chains. A polyhedral k-chain

$$\sigma = \sum_{i=1}^N \lambda_i \sigma_i$$

acts on test forms $\omega \in \mathcal{D}^k(\mathbb{R}^n)$ by integration; each simplex σ_i in a representation of σ is oriented, so that integration is well defined, and independent of the representation. This was explained in subsection 5.4. We write

$$\langle \omega, \sigma \rangle := \sum_{i=1}^{N} \lambda_i \int_{\sigma_i} \omega \,,$$

and note that the Stokes theorem holds in this context; namely,

$$\langle \omega, \partial\tau \rangle = \langle d\omega, \tau \rangle \tag{5.24}$$

for all polyhedral $(k+1)$-chains τ and for all test forms $\omega \in \mathcal{D}^k(\mathbb{R}^n)$.

5.10. **Exercise.** Prove (5.24).

The preceding understood, we first prove one half of the statement in Proposition 5.7.

Lemma 5.8. *Let σ be a polyhedral k-chain in $\mathbb{R}^n$ and let $\omega \in \mathbf{F}^k(\mathbb{R}^n)$. Then*

$$|\langle \omega, \sigma \rangle| \;\leq\; |\sigma|_\flat \cdot ||\omega||_\flat \,. \tag{5.25}$$

In particular, the dual norm of σ as an element of $\mathbf{F}_k(\mathbb{R}^n)$ does not exceed its flat norm.

Proof. Let τ be a polyhedral $(k+1)$-chain in $\mathbb{R}^n$. Then it follows from the definitions (5.10) and (5.11), and from the Stokes theorem (5.24), that

$$\begin{aligned}
\langle \omega, \sigma \rangle &= \langle \omega, \sigma - \partial\tau \rangle + \langle \omega, \partial\tau \rangle \\
&\leq\; |\sigma - \partial\tau| \cdot ||\omega||_\infty + |\tau| \cdot ||d\omega||_\infty \\
&\leq\; (|\sigma - \partial\tau| + |\tau|) \cdot ||\omega||_\flat \,.
\end{aligned}$$

By taking the infimum over all τ, we obtain (5.25) and the proposition follows. ☐

We need some further results in order to prove that the dual norm of a polyhedral chain bounds its flat norm. Notice that this then suffices for the isometric inclusion (5.23), because polyhedral chains are dense in $\mathcal{F}_k(\mathbb{R}^n)$ by definition, and because $\mathbf{F}_k(\mathbb{R}^n)$ is a Banach space.

However, armed with Lemma 5.8, we can prove Proposition 5.4.

Proof of Proposition 5.4. Assume that a polyhedral k-chain

$$\sigma = \sum_{i=1}^{N} \lambda_i \sigma_i$$

is represented by a sum of oriented k-simplexes σ_i that meet (at most) along lower dimensional sides. By (5.25), it suffices to exhibit a smooth compactly supported form ω in $\mathbb{R}^n$ such that $\langle \omega, \sigma \rangle \neq 0$. To this end, fix ε positive and small, say much smaller than the diameter of any of the sides of the σ_i's. Then we write

$$\sigma = \sigma_\varepsilon + \hat{\sigma}_\varepsilon \, ,$$

where $\hat{\sigma}_\varepsilon$ denotes an ε-neighborhood of the boundary of σ, within σ, and $\sigma_\varepsilon := \sigma - \hat{\sigma}_\varepsilon$. Note that

$$\sigma_\varepsilon = \sum_{i=1}^{N} \lambda_i \sigma_{i,\varepsilon}$$

is a k-chain made up of simplexes that have pairwise positive distance. We can therefore find a smooth compactly supported k-form ω_ε such that ω_ε takes on a constant value on each $\sigma_{\varepsilon,i}$ such that

$$\langle \omega_\varepsilon, \lambda_i \sigma_{\varepsilon,i} \rangle = |\lambda_i| \cdot |\sigma_{i,\varepsilon}| \, .$$

This can be done so that $||\omega||_\infty \leq 1$. Thus,

$$\langle \omega_\varepsilon, \sigma_\varepsilon \rangle = |\sigma_\varepsilon| \geq |\sigma| - C(\sigma) \cdot \varepsilon \, ,$$

where $C(\sigma) > 0$ depends only on σ. Because also

$$|\langle \omega_\varepsilon, \hat{\sigma}_\varepsilon \rangle| \leq C(\sigma) \cdot \varepsilon \, ,$$

we find that

$$\begin{aligned} \langle \omega_\varepsilon, \sigma \rangle &= \langle \omega_\varepsilon, \sigma_\varepsilon \rangle + \langle \omega_\varepsilon, \hat{\sigma}_\varepsilon \rangle \\ &\geq |\sigma| - C(\sigma) \cdot \varepsilon + \langle \omega_\varepsilon, \hat{\sigma}_\varepsilon \rangle \\ &\geq |\sigma| - C(\sigma) \cdot \varepsilon - C(\sigma) \cdot \varepsilon \, . \end{aligned}$$

This finishes the task, provided $\varepsilon > 0$ is chosen small enough, depending on σ only. The proposition follows. $\square$

5.11. **Discussion.** Before we continue, it is worthwhile to pause and review the situation. Recall that our main goal is to prove Theorem 5.5. In symbols, this is

$$\mathcal{F}_k(\mathbb{R}^n)^* = \mathcal{F}^k(\mathbb{R}^n) \, , \tag{5.26}$$

which by the basic Banach space theory implies that

$$\mathcal{F}_k(\mathbb{R}^n) \subset \mathcal{F}^k(\mathbb{R}^n)^* \tag{5.27}$$

isometrically. Now the space $\mathbf{F}^k(\mathbb{R}^n)$ of smooth compactly supported k-forms with the flat norm is clearly a subspace of $\mathcal{F}^k(\mathbb{R}^n)$, so has a bigger dual. Therefore, by (5.27), we should have

$$\mathcal{F}_k(\mathbb{R}^n) \subset \mathbf{F}^k(\mathbb{R}^n)^* = \mathbf{F}_k(\mathbb{R}^n) \tag{5.28}$$

continuously embedded. En route to our main goal, we will first Proposition 5.7 which asserts that (5.28) holds isometrically.

5.12. **Flat cochains.** We are in pursuit of a concrete description of the dual space of polyhedral flat chains. To that end, it is necessary to study this space first as an abstract entity.

Following Whitney [78], we call the elements in the dual space $\mathcal{P}_k(\mathbb{R}^n)^*$ *flat k-cochains*, and denote them by capital letters $X, Y, \ldots$. Cochains come equipped with the dual norm,

$$|X|_\flat := \sup_{|\sigma|_\flat \le 1} \langle X, \sigma\rangle \,. \tag{5.29}$$

We also have the *comass* of a cochain defined by

$$|X| := \sup_{|\sigma| \le 1} \langle X, \sigma\rangle \,. \tag{5.30}$$

Because $|\sigma|_\flat \le |\sigma|$, we have that

$$|X| \le |X|_\flat \,.$$

The *coboundary* of a k-cochain X is a $(k+1)$-cochain dX defined by

$$\langle dX, \sigma\rangle := \langle X, \partial\sigma\rangle \,, \quad \sigma \in \mathcal{P}_{k+1}(\mathbb{R}^n)\,.$$

The coboundary is indeed a cochain, because

$$|\langle dX, \sigma\rangle| \le |X|_\flat \cdot |\partial\sigma|_\flat \le |X|_\flat \cdot |\sigma|_\flat$$

by (5.19). It follows that

$$|dX| \le |dX|_\flat \le |X|_\flat \,.$$

Proposition 5.9. *For a flat cochain X we have that*

$$|X|_\flat = \max\{|X|, |dX|\}\,. \tag{5.31}$$

Proof. By the discussion preceding the proposition, we only have to prove that the left hand side of (5.31) does not exceed the right hand side. For this, fix a polyhedral k-chain σ, and then a polyhedral $(k+1)$-chain τ. We get

$$\begin{aligned} |\langle X, \sigma\rangle| &\le |\langle X, \sigma - \partial\tau\rangle| + |\langle X, \partial\tau\rangle| \\ &\le |X| \cdot |\sigma - \partial\tau| + |dX| \cdot |\tau| \\ &\le \max\{|X|, |dX|\} \cdot (|\sigma - \partial\tau| + |\tau|)\,, \end{aligned}$$

and by taking the infimum over all chains τ, we obtain the desired inequality. The proposition follows. □

We next show that the mass for cochains can be localized.

Proposition 5.10. *For every k-cochain X we have that*

$$|X| = \sup\langle X, \sigma\rangle\,, \tag{5.32}$$

where the supremum is taken over all (oriented) k-simplexes σ with $|\sigma| \le 1$.

Proof. We only have to show that the right hand side of (5.32) is at least $|X|$. To this end, fix $\varepsilon > 0$, and suppose that the right hand side in (5.32) is less than $|X| - \varepsilon$. Let $\sigma = \sum_{i=1}^N \lambda_i\sigma_i$ be a k-chain such that $|\sigma| = 1$ and that

$$|X| - \varepsilon < \langle X, \sigma\rangle\,.$$

Then

$$\begin{aligned} |X| - \varepsilon < |\langle X, \sigma\rangle| &\le \sum_{i=1}^N |\lambda_i| \cdot |\langle X, \sigma_i\rangle| \\ &\le \sum_{i=1}^N |\lambda_i| \cdot |\sigma_i| \cdot (|X| - \varepsilon) = |X| - \varepsilon\,, \end{aligned}$$

which is absurd. This proves the proposition. □

From the preceding proof we obtain the following corollary.

Corollary 5.11. *The mass of a k-cochain X is the supremum of the values*

$$\frac{\langle X, \sigma\rangle}{|\sigma|}\,,$$

where σ runs through k-simplexes of diameter less than any prescribed number.

The corollary together with Proposition 5.9 expresses the important fact that the flat norm of a cochain is locally determined, thus anticipating the identification of every cochain as a differential form.

5.13. **Smooth cochains.** Every smooth differential k-form ω with bounded flat norm determines a k-cochain in a natural way, as explained in subsection 5.4. We call such cochains *smooth.*

The discussion in 5.4 is equally valid for forms with continuous coefficients. With a slight underuse of terminology, we call a cochain X *continuous* if the action of both X and dX on polyhedral chains is given by continuous differential forms. The proof of Lemma 5.8 gives that *the dual norm $|\omega|_\flat$ of a continuous cochain ω does not exceed its flat norm $||\omega||_\flat$.* We will prove later in Lemma 5.21 that in fact $|\omega|_\flat = ||\omega||_\flat$ for continuous cochains.

5.14. **The end point cases.** We prove Theorem 5.5 for the values $k = 0$ and $k = n$.

Theorem 5.12. *The space of flat* 0*-cochains can be identified with the Banach space of bounded real-valued Lipschitz functions on* $\mathbb{R}^n$ *equipped with the norm*

$$||f||_L := \max\{||f||_\infty, \mathrm{Lip}(f)\}\,. \tag{5.33}$$

Here

$$\mathrm{Lip}(f) := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$$

denotes the best Lipschitz constant of f.

Note that the content of Theorem 5.12 agrees with Theorem 5.5, by the results in previous sections. See, in particular, Remark 4.2, and observe that the comass and the Euclidean norm agree for 1-forms.

Proof. Let X be a 0-cochain; obviously X can be thought of as a pointwise defined function on $\mathbb{R}^n$. By Proposition 5.10, we have that

$$|X| = \sup_{a \in \mathbb{R}^n} |\langle X, a\rangle| = ||X||_\infty$$

and that

$$|dX| = \sup_{|a-b| \leq 1} |\langle dX, [a, b]\rangle| = \sup_{|a-b| \leq 1} |\langle X, b\rangle - \langle X, a\rangle| = \mathrm{Lip}(X)\,.$$

The preceding reasoning can be inverted, and we have that every bounded Lipschitz function determines a 0-cochain. The correspondence is bijective and isometric. (Note that the convexity of $\mathbb{R}^n$ has been used here, cf. Lemma 2.2.) The theorem follows. $\square$

Remark 5.13. Bounded Lipschitz functions on a metric space equipped with a norm as in (5.33) form a Banach algebra, that is also a dual Banach space. Such Lipschitz algebras play an interesting role in recent developments in analysis on metric spaces. See [76].

Theorem 5.14. *The space of flat* n*-cochains can be identified with the Banach space* $L^\infty(\mathbb{R}^n)$*.*

Note, again, that Theorem 5.14 is in agreement with Theorem 5.5.

Proof. A polyhedral n-chain is a nothing but a function that is supported in a finite collection of n-simplexes; the flat norm agrees with the L^1-norm of the function. Therefore, the dual space is $L^\infty(\mathbb{R}^n)$ by basic real analysis. $\square$

5.15. L^1**-flat chains.** Polyhedral chains are rather discrete objects; they can be thought of as finite valued functions supported on simplexes, with orientation regarded. Next we describe another Banach space, whose members are more diffused chains. This space ultimately will be shown to agree with the space $\mathcal{F}_k(\mathbb{R}^n)$ of flat k-chains.

Let

$$\mathbf{G}_k(\mathbb{R}^n) := \wedge_k(\mathbb{R}^n; L^1(\mathbb{R}^n)) \oplus \wedge_{k+1}(\mathbb{R}^n; L^1(\mathbb{R}^n))\,,$$

equipped with the L^1-norm,

$$||(\varphi, \psi)||_1 := ||\varphi||_{L^1(\mathbb{R}^n)} + ||\psi||_{L^1(\mathbb{R}^n)}\,,$$

where the L^1-norm is taken with respect to the pointwise mass norm for each of the multivectors as defined in (5.6).

We define the *divergence* $\operatorname{div}\psi$ of a locally integrable $(k+1)$-vectorfield ψ to be a locally integrable k-vectorfield that satisfies

$$\int_{\mathbb{R}^n} \langle \omega, \operatorname{div}\psi\rangle\, dx \;=-\int_{\mathbb{R}^n} \langle d\omega, \psi\rangle\, dx \tag{5.34}$$

for every smooth compactly supported k-form ω. It is easy to see that such a vector field $\operatorname{div}\psi$, if exists, is unique. It is also easy to see, by using convolution approximations as defined in (4.4), that (5.34) holds in the following two more general instances: ω is a flat form of compact support, or ω is an arbitrary flat form and both ψ and $\operatorname{div}\psi$ are integrable .

Now consider the space

$$\mathbf{E_k}(\mathbb{R}^n) := \{(\varphi, \psi) \in \mathbf{G}_k(\mathbb{R}^n) : \varphi = \operatorname{div}\psi\}\,.$$

It is evident that $\mathbf{E}_k(\mathbb{R}^n)$ is a closed subspace of $\mathbf{G}_k(\mathbb{R}^n)$. We call the quotient Banach space

$$\mathcal{G}_k(\mathbb{R}^n) := \mathbf{G}_k(\mathbb{R}^n)/\mathbf{E}_k(\mathbb{R}^n)$$

the space of *integrable k-chains*. The terminology will become clear later.

The members of $\mathbf{G}_k(\mathbb{R}^n)$ naturally act on flat k-forms via

$$\langle(\varphi, \psi), \omega\rangle := \int_{\mathbb{R}^n} \langle \omega, \varphi\rangle\, dx + \int_{\mathbb{R}^n} \langle d\omega, \psi\rangle\, dx\,. \tag{5.35}$$

By the duality of mass and comass (see 5.2), we have that

$$|\langle(\varphi, \psi), \omega\rangle| \le ||(\varphi, \psi)||_1 \cdot ||\omega||_\flat\,, \tag{5.36}$$

which implies that $\mathbf{G}_k(\mathbb{R}^n)$ embeds continuously to the dual of flat k-forms. In particular,

$$\mathbf{G}_k(\mathbb{R}^n) \subset \mathbf{F}_k(\mathbb{R}^n) = \mathbf{F}^k(\mathbb{R}^n)^*\,,$$

where we recall $\mathbf{F}^k(\mathbb{R}^n)$ is the space of smooth compactly supported flat k-forms; its dual $\mathbf{F}_k(\mathbb{R}^n)$ is the space of flat k-currents.

The next proposition identifies the space of flat k-forms $\mathcal{F}^k(\mathbb{R}^n)$ as the dual of the quotient space $\mathcal{G}_k(\mathbb{R}^n)$, and provides an analog of Proposition 5.7 integrable chains.

Proposition 5.15. *We have a canonical isometry*

$$\mathcal{G}_k(\mathbb{R}^n)^* = \mathcal{F}^k(\mathbb{R}^n)\,. \tag{5.37}$$

Moreover,

$$\mathcal{G}_k(\mathbb{R}^n) \subset \mathbf{F}_k(\mathbb{R}^n) \tag{5.38}$$

isometrically.

For the proof, we recall the following simple functional analytic fact.

Lemma 5.16. *Let W be a closed subspace of a Banach space V and let*

$$W^\perp := \{v^* \in V^* : \langle v^*, w\rangle = 0 \text{ for all } w \in W\}\,,$$

where V^* denotes the dual Banach space and $\langle v^*, w\rangle$ the dual action. Then we have

$$(V/W)^* = W^\perp \tag{5.39}$$

canonically and isometrically.

The proof of Lemma 5.16 is left as an exercise to the reader.

Proof of Proposition 5.15. Fix k and denote for short $V := \mathbf{G}_k(\mathbb{R}^n)$. Then by basic real analysis, and by the duality of mass and comass,

$$V^* = \wedge^k(\mathbb{R}^n; L^\infty(\mathbb{R}^n)) \oplus \wedge^{k+1}(\mathbb{R}^n; L^\infty(\mathbb{R}^n))\,,$$

isometrically when the latter is equipped with the norm

$$||(\omega,\eta)||_\infty := \max\{||\omega||_\infty, ||\eta||_\infty\}\,.$$

Equality (5.37) now follows from Lemma 5.16, provided we can show that

$$W^\perp = \mathcal{F}^k(\mathbb{R}^n)\,, \tag{5.40}$$

where $W := \mathbf{E}_k(\mathbb{R}^n)$. To see this, let $(\omega,\eta) \in V^*$ be such that

$$\int_{\mathbb{R}^n} \langle \omega, \operatorname{div}\psi\rangle\, dx + \int_{\mathbb{R}^n} \langle \eta, \psi\rangle\, dx = 0$$

for every $(\operatorname{div}\psi, \psi) \in W$. By unraveling the various definitions, we obtain that $\eta = d\omega$ in the sense of distributions. On the other hand, if $\omega \in \mathcal{F}^k(\mathbb{R}^n) \subset V^*$, then by the remarks made after (5.34) we have that

$$\int_{\mathbb{R}^n} \langle \omega, \operatorname{div} \psi \rangle \, dx + \int_{\mathbb{R}^n} \langle d\omega, \psi \rangle \, dx = 0$$

for every $(\operatorname{div} \psi, \psi) \in \mathbf{E}_k(\mathbb{R}^n)$. Thus (5.40) holds.

Finally, to prove (5.38), we first observe that

$$\mathcal{G}_k(\mathbb{R}^n) \subset \mathcal{G}_k(\mathbb{R}^n)^{**} = \mathcal{F}^k(\mathbb{R}^n)^* ,$$

where the first inclusion is canonical and isometric by the basic Banach space theory, and the second equality follows from (5.37). Because $\mathbf{F}^k(\mathbb{R}^n) \subset \mathcal{F}^k(\mathbb{R}^n)$, we have, therefore, an inclusion

$$\mathcal{G}_k(\mathbb{R}^n) \subset \mathbf{F}^k(\mathbb{R}^n)^* = \mathbf{F}_k(\mathbb{R}^n) .$$

We claim that this inclusion is isometric as well. Indeed, given $T \in \mathcal{G}_k(\mathbb{R}^n)$ and $\omega \in \mathcal{F}^k(\mathbb{R}^n)$, we use standard approximation arguments, also alluded to in the first part of this proof, and find a sequence $(\omega_i) \subset \mathbf{F}^k(\mathbb{R}^n)$ such that

$$\lim_{i\to\infty} \langle T, \omega_i \rangle = \langle T, \omega \rangle .$$

In other words, $\mathbf{F}^k(\mathbb{R}^n)$ is dense in $\mathcal{F}^k(\mathbb{R}^n)$ in the weak topology determined by $\mathcal{G}_k(\mathbb{R}^n) \subset \mathcal{F}^k(\mathbb{R}^n)^*$. We obtain (5.38) from these remarks, by standard functional analytic arguments. The proof of the proposition is thereby complete. □

Remember that our goal is to identify the space $\mathcal{F}_k(\mathbb{R}^n)$ of flat k-chains as a predual of the space $\mathcal{F}^k(\mathbb{R}^n)$ of flat k-forms. Proposition 5.15 identifies $\mathcal{G}_k(\mathbb{R}^n)$ as a predual of $\mathcal{F}^k(\mathbb{R}^n)$. What we will do next, is to show that the spaces $\mathcal{F}_k(\mathbb{R}^n)$ and $\mathcal{G}_k(\mathbb{R}^n)$ are identical as Banach spaces, when considered as subspaces of $\mathbf{F}_k(\mathbb{R}^n) = \mathbf{F}^k(\mathbb{R}^n)^*$. Note that the (canonical and isometric) inclusions,

$$\mathcal{F}_k(\mathbb{R}^n) \subset \mathbf{F}^k(\mathbb{R}^n)^* , \qquad \mathcal{G}_k(\mathbb{R}^n) \subset \mathbf{F}^k(\mathbb{R}^n)^* , \tag{5.41}$$

follow from Propositions 5.7 and 5.15, respectively. (We still have not verified Proposition 5.7, but this is done momentarily.)

Proposition 5.17.

$$\mathcal{F}_k(\mathbb{R}^n) = \mathcal{G}_k(\mathbb{R}^n) .$$

Notice that Theorem 5.5 follows from Proposition 5.17, in view of the preceding remarks.

We break the proof of Proposition 5.17 into two separate propositions. After the first proposition, we will prove Proposition 5.7.

Proposition 5.18. *Given any $\xi \in \mathcal{G}_k(\mathbb{R}^n)$, and given any $\varepsilon > 0$, there exists a polyhedral k-chain σ such that*

$$\sup \frac{|\langle \omega, \xi - \sigma \rangle|}{||\omega||_\flat} < \varepsilon \,, \tag{5.42}$$

where the supremum is taken over all forms $\omega \in \mathbf{F}^k(\mathbb{R}^n)$. In particular, flat k-chains are dense in $\mathcal{G}_k(\mathbb{R}^n) \subset \mathbf{F}_k(\mathbb{R}^n)$ with respect to the dual norm.

Proof. Fix $\xi \in \mathcal{G}_k(\mathbb{R}^n)$. By the density of smooth compactly supported vectorfields in $\mathbf{G}_k(\mathbb{R}^n)$, we can assume that ξ can be represented by a pair (φ, ψ) of smooth vectorfields of compact support. By further subtracting $(\operatorname{div} \psi, \psi)$, we can assume that $\xi = (\varphi, 0)$ consists of a single smooth k-vectorfield of compact support, with the action

$$\langle \varphi, \omega \rangle = \int_{\mathbb{R}^n} \langle \varphi, \omega \rangle \, dx$$

for $\omega \in \mathbf{F}^k(\mathbb{R}^n)$. With a slight abuse of notation, we identify φ with ξ, and show that φ can be approximated in $\mathbf{F}^k(\mathbb{R}^n)^*$ by polyhedral k-chains.

To this end, fix $\varepsilon > 0$. We choose a dyadic decomposition of $\mathbb{R}^n$ into small enough cubes such that φ is essentially constant in each of the cubes; in particular, we assume that for each such dyadic cube Q,

$$||\varphi|Q - \varphi_Q||_\infty < \varepsilon \tag{5.43}$$

for some k-vector field φ_Q that is constant in Q and zero outside of Q. In (5.43), $\varphi|Q$ is the restriction of φ to Q, and we use the sup-norm of the pointwise mass for the k-vectors as defined in (5.6). It is clearly no loss of generality to assume that the side length $\ell(Q)$ for cubes in the chosen dyadic decomposition is less than ε. We fix such a cube, and assume that

$$\varphi_Q \,=\, e_{i_1} \wedge \cdots \wedge e_{i_k} \cdot \chi_Q \,.$$

In general, φ_Q is a linear combination of such basic vectors, but the proof will show that this simplifying assumption is of no consequence.

Let Q^k be the k-dimensional face of Q that is parallel to the subspace determined by $e_{i_1} \wedge \cdots \wedge e_{i_k}$, and oriented accordingly, and let Q^{n-k} be the complementary face, so that

$$Q = Q^k \times Q^{n-k} \,.$$

Next, for $q \in Q^{n-k}$ let $P(q) = Q^k \times q$ be an oriented k-cube that is parallel to Q^k, with the same orientation. Define

$$\sigma := |Q^{n-k}| \, P(q_0) \,, \tag{5.44}$$

where q_0 is the center of Q^{n-k}.

We compare the action of σ on smooth k-forms to that of φ. Thus, for $\omega \in \mathbf{F}^k(\mathbb{R}^n)$, we have that

$$|\langle \omega, \sigma - \varphi \rangle| \leq |\langle \omega, \sigma - \varphi_Q \rangle| + |\langle \omega, \varphi_Q - \varphi \rangle| \,. \tag{5.45}$$

The second term in (5.45) admits an estimate

$$|\langle \omega, \varphi_Q - \varphi \rangle| \,\leq\, \varepsilon \cdot |Q| \cdot ||\omega||_\flat + |\int_{\mathbb{R}^n \setminus Q} \langle \omega, \varphi \rangle \, dx| \,. \tag{5.46}$$

To estimate the first term in (5.45), we find first that

$$\langle \omega, \sigma \rangle = \int_{Q^{n-k}} \langle \omega, P(q_0) \rangle \, dq \,,$$

where we think of $P(q_0)$ as a constant function from Q^{n-k} to k-chains. It follows that

$$|\langle \omega, \sigma - \varphi_Q \rangle| = |\int_{Q^{n-k}} \langle \omega, P(q_0) - P(q) \rangle \, dq \,| \,. \tag{5.47}$$

We have from estimate (5.50) below that

$$|P(q_0) - P(q)|_\flat \,\leq\, (\sqrt{n}\, \ell(Q)) \cdot (|Q| + |\partial Q|) \,,$$

so that (5.47) and Lemma 5.8 give

$$|\langle \omega, \sigma - \varphi_Q \rangle| \leq\, (\sqrt{n}\, \ell(Q)) \cdot\, (|Q| + |\partial Q|) \cdot ||\omega||_\flat \cdot |Q^{n-k}| \,.$$

Combining this with (5.45) and (5.46), and with the fact that $\ell(Q) < \varepsilon$, we obtain

$$\begin{aligned} |\langle \omega, \sigma - \varphi \rangle| \leq\, & \varepsilon \cdot (\sqrt{n} \cdot (|Q| + |\partial Q|) \cdot |Q^{n-k}| + |Q|) \cdot ||\omega||_\flat \\ & + |\int_{\mathbb{R}^n \setminus Q} \langle \omega, \varphi \rangle \, dx| \,. \end{aligned}$$

By performing a similar approximation in each cube Q, and by forming a sum of all the k-chains as in (5.44), we obtain a k-chain σ with the property that

$$|\langle \omega, \sigma - \varphi \rangle| \leq\, \varepsilon \cdot C \cdot ||\omega||_\flat \,,$$

where $C > 0$ depends only on φ. This completes the proof of Proposition 5.18. $\square$

Proposition 5.18 allows us, finally, to prove Proposition 5.7. There is one more ingredient that is required. This is discussed in the next subsection.

5.16. **Smoothening of cochains.** It is a standard fact in the theory of distributions that every distribution can be approximated by smooth functions in the weak topology. This goes over to currents; for every k-dimensional current T there exists a sequence of smooth k-forms ω_ε, $\varepsilon > 0$, such that

$$T(\omega) = \lim_{\varepsilon \to 0} \langle \omega_\varepsilon, \omega \rangle \,,$$

where

$$\langle \omega_\varepsilon, \omega \rangle := \int_{\mathbb{R}^n} \langle \omega_\varepsilon, \omega \rangle \, dx \,.$$

A similar approximation procedure holds for cochains. Recall the terminology from 5.13.

Proposition 5.19. *Let $X \in \mathcal{F}^k(\mathbb{R}^n)$ be a flat k-cochain. Then there exists a sequence of smooth k-cochains ω_ε, $\varepsilon > 0$, such that*

$$\lim_{\varepsilon \to 0} \langle \omega_\varepsilon, \sigma \rangle = \langle X, \sigma \rangle \tag{5.48}$$

for every polyhedral k-chain σ. Moreover,

$$|\omega_\varepsilon|_\flat \le |X|_\flat \,. \tag{5.49}$$

We require a translation operator

$$T_v : \mathcal{P}_k(\mathbb{R}^n) \to \mathcal{P}_k(\mathbb{R}^n) \,, \qquad v \in \mathbb{R}^n \,,$$

defined as follows: for an oriented k-simplex σ, $T_v\sigma$ is a similarly oriented simplex that is the translation of σ by the vector v; for a general polyhedral chain T_v is defined by linearity. Obviously, T_v is a linear isometry of polyhedral chains. Moreover,

$$|T_v\sigma - \sigma|_\flat \;\le\; |v| \cdot (|\sigma| + |\partial\sigma|) \tag{5.50}$$

for all polyhedral chains σ.

5.17. **Exercise.** Prove estimate (5.50).

We now discuss the proof for Proposition 5.19. Let η_ε be a bump function as in (4.5). For a given k-cochain X, we define a new k-cochain X_ε by

$$\langle X_\varepsilon, \sigma \rangle := \int_{\mathbb{R}^n} \langle X, T_y\sigma \rangle \, \eta_\varepsilon(y) \, dy \tag{5.51}$$

for a polyhedral k-chain σ. It is clear from the properties of η_ε (see (4.3)) that X_ε is a cochain with

$$|X_\varepsilon|_\flat \;\le\; |X|_\flat \,. \tag{5.52}$$

Moreover, because T_v commutes with the boundary operator, we also have that

$$dX_\varepsilon = (dX)_\varepsilon \,. \tag{5.53}$$

Next, by (5.50),

$$\begin{aligned} |\langle X_\varepsilon - X, \sigma\rangle| &= |\int_{\mathbb{R}^n} \langle X, T_y\sigma - \sigma\rangle \, \eta_\varepsilon(y)\, dy| \\ &\le \, |X|_\flat \, (|\sigma| + |\partial\sigma|) \int_{\mathbb{R}^n} |y| \, \eta_\varepsilon(y)\, dy \\ &\le \, |X|_\flat \, (|\sigma| + |\partial\sigma|) \, \cdot C \cdot \varepsilon \,, \end{aligned}$$

where $C > 0$ only depends on the size of the support of η. It follows that

$$\lim_{\varepsilon \to 0} \langle X_\varepsilon, \sigma\rangle = \langle X, \sigma\rangle$$

for all polyhedral k-chains σ.

The preceding understood, it only suffices to show that X_ε is a smooth cochain. We outline a proof of this, and refer to Whitney's book [78] for the details. The concept of a *sharp cochain* is used here.

Following [78, V. 7], we define the *Lipschitz constant* of a k-cochain X by

$$\mathcal{L}(X) := \sup \frac{\langle X, T_v\sigma - \sigma\rangle}{|\sigma| \, |v|} \,, \tag{5.54}$$

where the supremum is taken over all polyhedral k-chains σ and their translations by a nonzero vector v. (By an argument similar to that in Proposition 5.10, one can show that the supremum in (5.54) is attained over oriented simplexes and their translations [78, p. 161].) The *sharp norm* of a k-cochain X is

$$|X|_\sharp := \max\{|X|_\flat, (k+1)\mathcal{L}(X)\} \,, \tag{5.55}$$

and a cochain is called *sharp* if it has finite sharp norm. (See Remark 5.20 below.) Thus, sharp cochains are flat, but the converse is not true in general. Whitney shows in [78, V. Theorem 10A] that to each sharp k-cochain X there corresponds a unique differential k-form ω_X with Lipschitz continuous coefficients such that the action of X on polyhedral chains correspond to integral action of ω_X as explained in 5.4. The proof of this claim is elementary, albeit a bit technical, using the definitions and standard Riemann sum type approximation for integrals. See [78, pp. 167–170] for the details.

Now we can easily show that the cochains X_ε as defined in (5.51) are sharp. Indeed,

$$\begin{aligned}|\langle X_\varepsilon, T_v\sigma - \sigma\rangle| &= |\int_{\mathbb{R}^n} \langle X, T_{y+v}\sigma - T_y\sigma\rangle\, \eta_\varepsilon(y)\, dy| \\ &= |\int_{\mathbb{R}^n} \langle X, T_y\sigma\rangle\, (\eta_\varepsilon(y-v) - \eta_\varepsilon(y))\, dy| \\ &\le |X|\,|\sigma| \int_{\mathbb{R}^n} |\eta_\varepsilon(y-v) - \eta_\varepsilon(y)|\, dy \\ &\le C\,|X|\,|\sigma|\,|v|\end{aligned}$$

where $C > 0$ depends only on η_ε.

Accepting the fact that each sharp cochain corresponds to a continuous differential form, we obtain from the preceding that X_ε is a continuous cochain as defined in 5.13. (Here we need (5.53) as well.)

One can show that X_ε is smooth for every cochain X, but essentially this requires Wolfe's theorem. Indeed, there is a representation

$$\omega_{X_\varepsilon}(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y)\, \omega_X(y)\, dy\,, \tag{5.56}$$

where ω_X is the bounded form associated with a cochain X, and ω_{X_ε} is the continuous form associated with the continuous cochain X_ε. Formula (5.56) is proved in [78, p. 261].

To avoid this circular argument, we can prove Proposition 5.19 by approximating each continuous form ω_{X_ε} by convolutions. These convolutions are smooth and converge to ω_{X_ε} locally uniformly.

This discussion completes the proof of Proposition 5.19.

Remark 5.20. Whitney defines sharp cochains as continuous functionals on the space of polyhedral chains with respect to a *sharp norm* on chains; the sharp norm is a weaker norm than the flat norm. Thus, in effect, the expression in (5.55) should be regarded as a theorem, akin to Proposition 5.9, rather than a definition. We will not discuss the sharp norm on chains here.

Next we verify the following lemma. Recall the terminology from 5.13.

Lemma 5.21. *Let ω be a continuous k-cochain. Then*

$$|\omega|_\flat = ||\omega||_\flat\,, \tag{5.57}$$

where we have the dual norm on the left and the flat norm on the right.

Proof. The inequality $|\omega|_\flat \le ||\omega||_\flat$ was pointed out in 5.13. To prove the reverse in equality, let $\varepsilon > 0$. By the duality $L^\infty = (L^1)^*$ (with respect to mass and comass), and by the density of smooth functions in L^1, we can pick a smooth compactly supported k-vectorfield φ in $\mathbb{R}^n$ such that $||\varphi||_1 \leq 1$ and that

$$||\omega||_\flat \leq \int_{\mathbb{R}^n} \langle \omega, \varphi \rangle \, dx + \varepsilon \, .$$

On the other hand, by Proposition 5.18, we can find a polyhedral k-chain σ such that the dual norm of σ in $\mathbf{F}_k(\mathbb{R}^n)$ does not exceed $1 + \varepsilon$ and such that

$$|\int_{\mathbb{R}^n} \langle \omega, \varphi \rangle \, dx - \langle \omega, \sigma \rangle| < \varepsilon \, ||\omega||_\flat \, .$$

Combining the last two inequalities, we obtain

$$||\omega||_\flat \, \leq \, \varepsilon + \varepsilon \, ||\omega||_\flat + |\omega|_\flat |\sigma|_\flat \, \leq \, \varepsilon + \varepsilon \, ||\omega||_\flat + |\omega|_\flat \, (1 + \varepsilon) \, .$$

The claim follows by letting $\varepsilon \to 0$. □

Proof of Proposition 5.7. By Lemma 5.8, it suffices to prove that the flat norm of a polyhedral k-chain σ does not exceed its dual norm. To this end, we first observe that

$$|\sigma|_\flat = \sup_{|X|_\flat \leq 1} |\langle X, \sigma \rangle| \, ,$$

by standard functional analysis. By Proposition 5.19, the preceding supremum is achieved over a sequence of smooth cochains. Obviously, because σ is fixed, we can assume these smooth cochains are compactly supported. We have $|\omega|_\flat = ||\omega||_\flat$ for such cochains by Lemma 5.21, so that

$$|\sigma|_\flat \, = \, \sup_{||\omega||_\flat \leq 1} |\langle \omega, \sigma \rangle| \, .$$

The right hand side of the last inequality is precisely the dual norm of σ in $\mathbf{F}_k(\mathbb{R}^n) = (\mathbf{F}^k(\mathbb{R}^n))^*$, and the proof of Proposition 5.7 is thereby complete. □

We turn back to the proof of Proposition 5.17. The following is a converse to Proposition 5.18.

Proposition 5.22. *Given any polyhedral k-chain σ in $\mathbb{R}^n$, and given any $\varepsilon > 0$, there exists a k-vectorfield $\varphi \in \wedge^k(\mathbb{R}^n; L^1(\mathbb{R}^n))$ such that*

$$|\sigma - \varphi|_\flat < \varepsilon \, .$$

In particular,

$$\mathcal{F}_k(\mathbb{R}^n) \subset \mathcal{G}_k(\mathbb{R}^n) \, . \tag{5.58}$$

Proof. It is enough to consider the case where σ is a single oriented k-simplex. Now such a chain is a limit in the flat norm of a linear combination of oriented k-cubes, so we can in fact assume (for notational simplicity) that σ consists of a single oriented k-cube Q^k. As a further reduction, we assume that Q^k is parallel to the first k-coordinate axes, and oriented by $e_1 \wedge \cdots \wedge e_k$.

The preceding understood, fix $\varepsilon > 0$. Let $q_0 \in Q^k$ be the center of the cube, and let Q^{n-k} be an $(n-k)$-cube that is orthogonal to Q^k, also centered at q_0, and of side length ε. For each $q \in Q^{n-k}$ let $P(q)$ denote the k-cube that is parallel to Q^k, similarly oriented, and meets Q^{n-k} orthogonally at q. By (5.50), we have that

$$|P(q) - \sigma|_\flat \leq \varepsilon \left(|Q^k| + |\partial Q^k|\right) . \tag{5.59}$$

Define a k-vectorfield φ by

$$\varphi := \frac{1}{|Q^{n-k}|} \cdot e_1 \wedge \cdots \wedge e_k \cdot \chi_{Q^k \times Q^{n-k}} .$$

Then, for every smooth compactly supported k-form ω, we have that

$$\begin{aligned} |\langle \varphi - \sigma, \omega\rangle| &= \frac{1}{|Q^{n-k}|} \Big| \int_{Q^{n-k}} \langle P(q) - \sigma, \omega\rangle \, dq \Big| \\ &\leq |P(q) - \sigma|_\flat \cdot ||\omega||_\flat \leq C\, \varepsilon\, ||\omega||_\flat , \end{aligned}$$

where, by (5.59), $C > 0$ depends only on σ. (In the preceding equality, we interpret σ as a constant covector on Q^{n-k}, similarly to the discussion before (5.47).)

This completes the proof of Proposition 5.22. $\square$

Proposition 5.17 now follows from Propositions 5.18 and 5.22. As remarked after the statement of Proposition 5.17, this also accomplishes the proof of our main result, Theorem 5.5.

Remark 5.23. We have now shown that every k-cochain X corresponds to a unique bounded measurable form ω_X with bounded exterior derivative $d\omega_X$ (in the sense of distributions). Conversely, given any such flat form ω, there is a corresponding cochain X_ω. The concrete action of a flat form ω on polyhedral chains can be given by using smoothing. Indeed, if $\omega_\varepsilon = \omega * \eta_\varepsilon$ denotes the convolution of a flat k-form ω, then

$$\langle \omega, \sigma\rangle := \langle X_\omega, \sigma\rangle = \lim_{\varepsilon \to 0} \langle \omega_\varepsilon, \sigma\rangle , \tag{5.60}$$

where the action $\langle \omega_\varepsilon, \sigma\rangle$ makes sense as integration as explained in 5.4. Equality (5.60) follows from the discussion in this section by standard arguments.

5.18. **Lipschitz invariance of flat chains.** We finish this long section by discussing the important Lipschitz invariance of flat chains. For simplicity, we consider globally defined flat forms and mappings. As should be clear by now, this entails no essential loss of generality.

Theorem 5.24. *Let ω be a flat k-form in $\mathbb{R}^m$ and let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a Lipschitz map. Then $f^*\omega$ is a flat k-form in $\mathbb{R}^n$.*

The statement of the theorem requires an explanation. If

$$\omega(y) = \sum_I \omega_I(y) dy_I$$

is a flat k-form as in the theorem, then the natural definition for $f^*\omega$ would be

$$f^*\omega(x) := \sum_I \omega(f(x)) df_I(x)\,, \tag{5.61}$$

where

$$df_I := df_{i_1} \wedge \ldots \wedge df_{i_k}\,, \qquad I = (i_1, \ldots, i_k)\,.$$

The problem with (5.61) is that the coefficients ω_I of ω are only functions in L^∞, so that $\omega_I(f(x))$ may be undefined for x in a set of positive measure in $\mathbb{R}^n$.

In trying to understand $f^*\omega$ by way of duality, the natural definition for $f^*\omega$ would be as a linear functional,

$$\langle f^*\omega, \sigma\rangle := \langle \omega, f_*\sigma\rangle\,. \tag{5.62}$$

The problem with this definition in turn is that we need to show that a natural pushforward $f_*\sigma$ exists as a flat k-chain for every polyhedral k-chain σ in $\mathbb{R}^n$.

Both of the above problems can be overcome, so that we have two approaches to Theorem 5.24.

In the first case, one first shows that ω has a well defined action on all k-directions at every point. Such a result can be viewed as a sharpening of Theorem 5.5. It follows that if df has rank at least k at a point x,then $f^*\omega(x)$ can be defined as in (5.61). On the other hand, if the rank is less than k, then $df_I = 0$ for every $I = (i_1, \ldots, i_k)$, and we set $f^*\omega(x) = 0$ in this case.

In the second case, one shows that $f_*\sigma$ is a flat k-chain by invoking Rademacher's theorem: σ can be thought of as a region in $\mathbb{R}^k$, so that f is almost everywhere differentiable on σ with respect to the Hausdorff k-measure. The differentials of f can be used to find polyhedral approximations in the flat norm to the image of σ under f.

It would take us too far afield to fully discuss the outlined two approaches. We refer to [78, Chapter X], or [18, Chapter 4] for the details.

Remark 5.25. An important consequence of Theorem 5.24 is that flat forms can be defined on Lipschitz manifolds. We obtain a differential complex (see Lemma 5.1) whose cohomology agrees with singular cohomology of the manifold akin to the de Rham theory. Whitney hints at such a result in his works on geometric integration theory, but it seems that it is nowhere explicitely stated in his book. In [78, p. viii], Whitney promises to return to "Lipschitz spaces" in a "separate memoir", but apparently this never happened. In any event, that the de Rham theorem holds in the Lipschitz context follows by standard sheaf theoretic arguments by using the local Poincaré lemma. See e.g. [71].

5.19. **Notes to Section 5.** The proof of Wolfe's theorem 5.5 in this section essentially follows the argument in Whitney's book. Some streamlining was achieved by using the functional analytic framework and the language of distributions and currents. A similar approach was taken by Federer in his book, except that the argument there is embedded in a more general discussion of various types of currents. See [18, 4.1.19].

Whitney's theory has been used in topology, for example in explaining Novikov's theorem about topological invariance of rational Pontryagin classes. See [70].

Recently, Harrison has presented an interesting modification of Whitney's theory. See [24].

6. Locally standard Lipschitz structures

This last section surveys some recent results and open problems in geometric analysis related to Lipschitz functions, where also flat forms play an important role. The format of this section is somewhat different from the previous sections, as we will not prove much, and will take many mathematical concepts as known. The discussion is in large part inspired by Dennis Sullivan's talk at a conference held in memory of Lars Ahlfors at Stanford University in 1997 [66].

6.1. **Locally standard smooth structures.** A smooth manifold M can be equipped with various geometric structures. One such structure is a *conformal structure*, which can be defined as an understanding of what is meant by a round ball in each tangent space. More precisely, a conformal structure on a manifold is a smoothly varying assignment of equivalence classes of inner products attached with each tangent space T_pM, $p \in M$, where two inner products are equivalent if one is a real multiple of the other. Thus, in a conformal structure we can say what a ball in each tangent space is, but cannot specify its size.

A *Riemannian structure* is obtained from a conformal structure when we fix a representative in each equivalence class of inner products, in a smoothly varying way. In a Riemannian manifold, we can speak about the size of objects, and informally one can say that a Riemannian structure is a conformal structure plus *volume*.

Every diffeomorphism from a region in M^n into $\mathbb{R}^n$ induces a conformal or a Riemannian structure in the region by pulling back the standard structure from $\mathbb{R}^n$. With respect to this pullback structure such a diffeomorphism is a *conformal map*, or an *isometry*, as the case may be.

Conversely, if a conformal or a Riemannian structure on M is given, one can ask whether it arises locally as a pullback of the standard structure of $\mathbb{R}^n$ by some diffeomorphism, where by *locally* we mean that every point in M has a neighborhood where the structure agrees with a pullback structure. If this is the case, we say that such a structure on M is *locally standard*. In particular, with respect to locally standard conformal or Riemannian structures, M is locally conformally or isometrically equivalent to a patch in $\mathbb{R}^n$.

Consider as an example first the case when the dimension of M is one, and M has a Riemannian structure. Now M is nothing but a smooth curve locally, and thus there exists a smooth arc length parametrization that gives a local isometry between M and an interval in $\mathbb{R}$. Thus, *every Riemannian structure on a 1-manifold is locally standard*.

The preceding assertion is not true in dimensions above one. There is an obstruction *qua curvature* of the structure. Indeed, for a Riemannian n-manifold to be locally isometric to a patch in $\mathbb{R}^n$, its *Riemannian curvature tensor must vanish identically*, and in every dimension at least two there are Riemannian manifolds with non-zero curvature. On the other hand, by a fundamental theorem in Riemannian geometry, the curvature tensor provides the only obstruction in this case. Thus, *a Riemannian manifold is locally standard if and only if the condition* $R \equiv 0$ *is fulfilled*, where R is the associated Riemannian curvature tensor.

For conformal structures there is another fundamental theorem, going back to cartography and Gauss (see Remark 6.1), stating that *every conformal structure on a smooth 2-manifold is locally standard*. In particular, every smooth Riemannian 2-manifold is locally conformally equivalent to a patch in $\mathbb{R}^2$, although typically it is not isometrically so. Thus, as is the case for Riemannian structures on 1-manifolds, *there is no obstruction in dimension two for a conformal structure to be locally standard*.

See [79, Theorems 2.4.11 and 2.5.14] for the proofs of the above cited facts from geometry.

We cannot expect conformal structures to be locally standard beyond dimension two; several obstructions emerge. For example, in dimensions above three a necessary and sufficient condition for a Riemannian manifold to be locally conformally equivalent to $\mathbb{R}^n$ is that the so called *Weyl component*, or the *Weyl-Schouten tensor*, of the curvature tensor vanishes. See, for example, Lafontaine's lectures in [40].

6.2. **Measurable conformal structures.** Both conformal and Riemannian structures allow for a weaker formulation; it is not necessary to have them smooth. The distribution of inner products in the definition of these structures could be asked to be only measurable, for example. Measurability, or some such requirement, makes sense by interpreting a distribution of inner products as a section of the vector bundle of symmetric bilinear 2-forms on the manifold; this is a map between two smooth manifolds. Our problems are eventually local, so that M could just as well be an open set in $\mathbb{R}^n$, in which case the phrases "measurable distribution of inner products" and "almost everywhere" have an obvious meaning.

A measurable conformal structure on M can be said to be *locally standard* if every point in M has a neighborhood together with an almost everywhere differentiable homeomorphism from the neighborhood onto an open set in $\mathbb{R}^n$ such that, at almost every point, the tangent map is conformal from the given inner product to the standard Euclidean inner product. In a more picturesque language, a measurable conformal structure is locally standard if locally there exists a homeomorphism into $\mathbb{R}^n$ that takes, at almost every point, infinitesimal balls as determined by the given structure to infinitesimal round balls in $\mathbb{R}^n$. In practice, it is necessary to require more regularity from the homeomorphism in the preceding definition, e.g. a membership in some Sobolev class.

The celebrated *measurable Riemann mapping theorem*, first proved by Morrey in 1938 [53], gives that *every measurable conformal structure on a 2-manifold is locally standard, provided the defining circles on tangent spaces have uniformly bounded eccentricity when measured against some fixed background smooth Riemannian metric.*

The condition on eccentricity in Morrey's theorem is clearly independent of the chosen metric. To explain this theorem in analytic terms, let us assume, as we may by the smooth result mentioned earlier, that $M = \Omega$ is an open subset of $\mathbb{R}^2$. Then a measurable conformal structure on Ω can be viewed as a measurable field of ellipsoids on Ω. The eccentricity of the ellipsoid that is attached with a point $z \in \Omega$ is by definition the ratio of the lengths of its major and minor axes. In this way, we arrive at the following partial differential equation, known as the *Beltrami equation*,

$$\overline{\partial} f(z) = \mu(z) \partial f(z)\,, \tag{6.1}$$

where

$$\partial := \frac{1}{2}\Big(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\Big)\,, \qquad \overline{\partial} := \frac{1}{2}\Big(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\Big)$$

for a complex valued function $f : \Omega \to \mathbb{C}$. The function $\mu(z)$ is a measurable complex-valued function, measuring the eccentricity and the direction of the axes of an ellipse at a point z. We stipulate the condition

$$||\mu||_\infty < 1 \tag{6.2}$$

for the bounded eccentricity. For a nonsmooth μ equation (6.1) has to be understood in the sense of distributions.

Morrey proved that homeomorphic solutions to equation (6.1) exist in the Sobolev space $W^{1,2}_{loc}$, and that they are unique up to postcomposing by conformal mapping. One can show that every homeomorphic solution f to (6.1) is almost everywhere differentiable, and by working out the analytic details, as already Gauss did for smooth mappings, we find that f pulls back the standard conformal structure to the given measurable structure in the sense discussed earlier.

Equation (6.1) expresses more than just the conformality of the solution between the two structures. One can also prescribe the infinitesimal rotation of the mapping. Moreover, as was proved by Bojarski in 1955 [8], every homeomorphic solution to (6.1) belongs to $W^{1,2+\varepsilon}_{\mathrm{loc}}$ for some $\varepsilon > 0$ depending only on the L^∞-norm of μ. This extra degree of smoothness is crucial for many applications of the Beltrami equation.

Homeomorphic solutions to the Beltrami equation (6.1) are called *quasiconformal mappings*. The theory of quasiconformal mappings, extensively developed during the last fifty years, has applications that extend far beyond what can be discussed here. We refer to [46] for a complete account of the early years, and to [45] and [34] for recent developments.

Remark 6.1. (a) Morrey's theorem in the smooth context is often credited to Gauss, but this seems an exaggeration as far as a rigorous proof is concerned. It is probably true that Gauss was the first to speak of *isothermal coordinates*, and that he also was the first to realize that such coordinates can be found for every smooth surface.

(b) One can relax condition (6.2) and still obtain the existence of homeomorphisms that pull back the standard conformal structure to that determined by μ. Such results were considered by Lehto already in the 1960s [44]. A paper by David [15] triggered an extensive development in this direction. See [34] and the references there.

There is no direct analog of the measurable Riemann mapping theorem in dimensions higher than two. That is, one cannot at will distribute a measurable ellipsoid field with uniformly bounded eccentricity in a region in $\mathbb{R}^n$, $n \geq 3$, and expect this to be a pullback distribution under a homeomorphism. Analytically one sees this from the associated partial differential equation which becomes overdetermined in dimensions higher than two. Geometrically, the lack of such a theorem reflects the lack of nontrivial conformal mappings in space. Recall that according to the *Liouville theorem* every conformal mapping in a domain in $\mathbb{R}^n$, $n \geq 3$, is a restriction of a Möbius transformation. See [34] for more discussion on this topic.

It is however unknown if an extra *integrability condition* attached with a measurable conformal structure would yield a positive existence result. The term "integrability condition" in this connection was used by Sullivan in his lecture [66]. The vanishing of the Riemannian curvature tensor, $R \equiv 0$, is an example of such a condition; it guarantees that a structure is locally standard. Similar remark holds for condition (6.2) in dimension two. Sullivan also viewed the *Darboux theorem* in symplectic geometry in this light. Recall that a *symplectic manifold* is a smooth even dimensional manifold together with a nondegenerate closed 2-form on the manifold. The Darboux theorem asserts that locally every such form is a pullback of the standard symplectic form $\omega_0 = \sum_{i=1}^n dx_i \wedge dy_i$ of $\mathbb{R}^{2n}$ under a diffeomorphism. Now one could consider more generally pairs (M^{2n}, ω), where ω is a nondegenerate 2-form on M, and conclude that there is an integrability condition, namely $d\omega = 0$, which implies that (M, ω) is locally equivalent to $(\mathbb{R}^{2n}, \omega_0)$. Sullivan raised the interesting question whether there is a measurable formulation of this integrability condition and Darboux's theorem, by using flat forms for example. For an introduction to symplectic geometry, see e.g. [3].

It is not clear what kind of integrability conditions one should be looking for in the case of measurable conformal structures in dimensions higher than two. For example, it is not known if there are measurable analogs of the Weyl-Schouten tensor mentioned earlier at the end of subsection 6.1. In the next subsection, we will discuss some possible integrability conditions for measurable Riemannian metrics.

6.3. **Measurable Riemannian structures.** A measurable Riemannian structure on a smooth manifold makes sense, as described in the beginning of the previous subsection. To say that such a structure is locally standard is to say that locally there exists an almost everywhere differentiable homeomorphism (preferrably in some appropriate Sobolev class) onto an open set in $\mathbb{R}^n$ such that the tangent map of this homeomorphism induces almost everywhere an isometric isomorphism between the given inner product and the standard inner product in $\mathbb{R}^n$.

To see what kind of integrability conditions one might have in this case, we consider a measurable Riemannian structure in an open subset Ω of $\mathbb{R}^n$. Such a structure is determined by a measurable map from Ω into the Lie group $GL(n, \mathbb{R})$ of linearly independent frames, each frame representing an orthonormal basis in the inner product. By using the canonical isomorphism between $\mathbb{R}^n = T_x\mathbb{R}^n$ and $T_x^*\mathbb{R}^n$, we can think of such a measurable structure as a map

$$\rho : \Omega \to \wedge^1(\mathbb{R}^n) \times \cdots \times \wedge^1(\mathbb{R}^n)\,,$$

where on the right we have an n-fold product of 1-forms. Thus,

$$\rho = (\rho_1, \ldots, \rho_n)$$

is an n-tuple of 1-forms. It is natural to stipulate that *each form* ρ_i *is a flat 1-form in* Ω. The measurable Riemannian volume form in such a situation would be the bounded flat n-form

$$\rho_1 \wedge \ldots \wedge \rho_n\,.$$

It is further natural to stipulate that this "volume form" is nondegenerate and has one sign; that is, we require that the form $\rho_1 \wedge \ldots \wedge \rho_n$ lies in a fixed component of $\wedge^n(\mathbb{R}^n) \setminus \{0\}$ almost everywhere. We next strengthen this requirement to a uniform distance from the origin.

Definition 6.2. A *Cartan-Whitney presentation* in Ω is an n-tuple $\rho = (\rho_1, \ldots, \rho_n)$ of flat 1-forms such that

$$\operatorname{essinf} * \,(\rho_1 \wedge \ldots \wedge \rho_n) \,>\, 0\,, \tag{6.3}$$

where the *Hodge star* $*$ renders the canonical isomorphism between $\wedge^n(\mathbb{R}^n)$ and $\mathbb{R}$.

Thus, a Cartan-Whitney presentation can be viewed as a measurable (Whitney flat) coframe at almost every point, such that the associated volume form is uniformly bounded and bounded away from zero.

When is a Cartan-Whitney presentation locally standard? The assumption that the forms be flat, together with (6.3), suggests that we should be looking for bi-Lipschitz homeomorphisms, and being locally standard means that

$$\rho_i = f^*(dx_i) = df_i$$

for all $i = 1, \ldots, n$, for some bi-Lipschitz map f. Note that if ρ is locally standard, then necessarily

$$d\rho = ddf = 0\,.$$

We call a Cartan-Whitney presentation *closed* if $d\rho = 0$. It turns out that the preceding obvious necessary condition is almost sufficient, as we will soon see (Theorem 6.3).

Before this, let us discuss an interesting idea that was put forward by Sullivan in his 1997 lecture. Namely, suppose that a Cartan-Whitney presentation ρ in Ω is given. It is a purely algebraic fact that there exists a skew-symmetric matrix θ of 1-forms such that the equality

$$d\rho = \theta \wedge \rho \tag{6.4}$$

holds. More precisely, at almost every point x there is a skew-symmetric matrix $\theta(x)$ with entries 1-forms such that $d\rho(x) = \theta(x) \wedge \rho(x)$. (See [52, p. 302].) If the Cartan-Whitney presentation ρ is smooth, the matrix θ is the connection matrix associated with the Levi-Civita connection of the metric. Further, in the smooth case, the Riemannian curvature tensor R is a matrix of 2-forms,

$$R = d\theta - \theta \wedge \theta\,. \tag{6.5}$$

Now we could stipulate that, given a Cartan-Whitney presentation ρ, the unique matrix θ in (6.4) is also made of flat forms. In particular, an exterior differential $d\theta$ can be formed in the sense of distributions. This stipulation understood, we can define the *curvature* R of the Cartan-Whitney presentation ρ by the formula (6.5).

In subsection 6.1, we discussed how there is just one obstruction for a smooth Riemannian structure to be locally standard, the vanishing of the Riemannian curvature tensor. Is there an analog of this result in the measurable context by using formula (6.5)? As Sullivan observed, the curvature associated with every closed Cartan-Whitney presentation vanishes. Moreover, he proved the following result in [68].

Theorem 6.3. *Every closed Cartan-Whitney presentation in an open set in $\mathbb{R}^n$, $n \geq 2$, is locally standard outside a closed exceptional set of measure zero and of topological dimension at most $n-2$. The exceptional set can really occur.*

The proof of Theorem 6.3 relies on two major mathematical ideas. First, there is the theory of flat forms and duality, as developed in Section 5. Second, there is a theory of quasiregular mappings initiated by

Reshetnyak in the 1960s. We will discuss the latter theory momentarily. Before that, let us consider an example showing that an exceptional set of topological codimension two, as allowed in the theorem, can really appear.

The *k-times winding map*, $k \geq 2$,

$$w : \mathbb{R}^n \to \mathbb{R}^n\,, \qquad w(r, \theta, z) := (r, k\theta, z)\,, \tag{6.6}$$

in cylindrical coordinates, is a Lipschitz map with the property that

$$\det Dw \;=\; k \tag{6.7}$$

almost everywhere. The mapping w is a local homeomorphism outside the $(n-2)$-dimensional subspace $B_w := \{r = 0\}$ in $\mathbb{R}^n$. Near every point of B_w we have a k-to-one map. The set B_w is called the *branch set* of w. The pullback forms

$$\rho_i := w^*(dx_i) = dw_i\,, \qquad i = 1, \ldots n\,, \tag{6.8}$$

constitute a Cartan-Whitney presentation in $\mathbb{R}^n$, because of (6.7). This presentation cannot be locally standard at points in B_w.

The curvature R associated with the presentation $\rho = (\rho_1, \ldots, \rho_n)$ in (6.8) is zero, as explained earlier. In this example, there is (to quote Sullivan) "new kind of curvature", resting on the branch set of the winding map.

The winding map is a particular example of a mapping of bounded length distortion.

6.4. **BLD-mappings.** A Lipschitz mapping $f : \Omega \to \mathbb{R}^n$, where $\Omega \subset \mathbb{R}^n$, $n \geq 2$, is open, is said to be a *BLD-mapping*, or a *mapping of bounded length distortion*, if

$$\operatorname{essinf} \det Df > 0\,. \tag{6.9}$$

BLD-mappings form a subclass of *quasiregular mappings*[4] that were alluded to earlier. As such, BLD-mappings were first studied by Martio and Väisälä [49], who gave several equivalent definitions for this class of mappings. In particular, they showed that the analytic condition (6.9), together with the Lipschitz condition, imply that BLD-maps preserve the lengths of paths up to a multiplicative error. Conversely, such length preserving property characterizes BLD-maps among discrete, open, and sensepreserving maps.[5]

[4]We forgo the definition for general quasiregular mappings here as it is not needed; see [55], [56].

[5]Instead of a global Lipschitz requirement, Martio and Väisälä used a uniform local Lipschitz condition. For local considerations, as in these notes, the difference is immaterial. See also [29].

Recall that a continuous mapping between topological spaces is *open* if it maps open sets to open sets, and *discrete* if the preimage of every point consists of isolated points.

The following fundamental theorem is due to Reshetnyak [54], [55].

Theorem 6.4. *Every BLD-mapping is an open and discrete mapping that is locally bi-Lipschitz outside a closed set of measure zero and of topological dimension at most $n-2$.*

Proof of Theorem 6.3. Let ρ be a closed Cartan-Whitney presentation in Ω. Pick a point $p \in \Omega$ and $r > 0$ such that $B(p,r) \subset \Omega$. Define

$$F_p(x) := \langle \rho, [p,x] \rangle \tag{6.10}$$

for $x \in B(p,r)$, where the right hand side denotes the duality between flat 1-forms and oriented polyhedral 1-chains, proved in Section 5.[6] Moreover, we apply the duality to the components of $\rho = (\rho_1, \dots, \rho_n)$.

We claim that the function $F_p : B(p,r) \to \mathbb{R}^n$ is Lipschitz. Indeed,

$$\begin{aligned} |F_p(x) - F_p(y)| &= |\langle \rho, [p,x] - [p,y] \rangle| \\ &\leq |\langle \rho, [p,x] + [x,y] - [p,y] \rangle| + |\langle \rho, [x,y] \rangle| \\ &= |\langle d\rho, [p,x,y] \rangle| + |\langle \rho, [x,y] \rangle| \\ &\leq 0 + ||\rho||_\infty \cdot |x-y| \,, \end{aligned}$$

so that F_p is L-Lipschitz in $B(p,r)$ for $L = ||\rho||_\infty$, where

$$||\rho||_\infty := \max_{i=1,\dots,n} ||\rho_i||_\infty \,.$$

Next we show that condition (6.9) holds. For this, observe that

$$\begin{aligned} \langle dF_p - \rho, [x,y] \rangle &= \langle \rho, [p,y] - [p,x] - [x,y] \rangle \\ &= \langle \rho, [p,y] + [y,x] + [x,p] \rangle \\ &= \langle d\rho, [p,y,x] \rangle = 0 \end{aligned}$$

whenever $[x,y] \subset B(p,r)$ is a line segment. This implies that $dF_p = \rho$, by the theory of flat chains and forms as given in Section 5. Hence (6.9) is automatically satisfied.

It follows that $F_p : B(p,r) \to \mathbb{R}^n$ is a BLD-mapping. Because F_p is a locally bi-Lipschitz map outside a closed set B_{F_p} of measure zero and topological dimension at most $n-2$ (Theorem 6.4), we have that ρ is locally standard in the complement of B_{F_p}. The theorem follows. $\square$

[6] The fact that the forms here are not globally defined as required by the theory in Section 5 is of no consequence, for one can always multiply by a smooth cut-off function, for example.

There is an analog of Theorem 6.3 for general, not necessarily closed Cartan-Whitney presentations. In a sense, this result is more satisfactory than Theorem 6.3, as it may be difficult to verify whether a given flat form is closed. But the conclusion has to be weaker, as locally standard forms have to be closed. The following theorem is again due to Sullivan [68] (with a simpler proof in [30]).

Theorem 6.5. *Let ρ be a Cartan-Whitney presentation in an open set Ω in $\mathbb{R}^n$, $n \geq 2$. Then for every point $p \in \Omega$ there exists $r_0 > 0$ and a BLD-mapping $F_p : B(p, r_0) \to \mathbb{R}^n$ such that*

$$||dF_p - \rho||_{\infty,B(p,r)} \leq \, ||d\rho||_{\infty,B(p,r)} \cdot r \tag{6.11}$$

for all $r < r_0$. The BLD-data of F_p depends only on n and the flat norm of ρ.

Proof. Fix $p \in \Omega$, assume that $B(p, r_0) \subset \Omega$, and define F_p as in (6.10) for $x \in B(p, r_0)$. A computation as in the proof of Theorem 6.3, using the elementary estimate

$$|[p, x, y]| \, \leq \, \max\{|p - x|, |p - y|\} \cdot |x - y| \tag{6.12}$$

for the area of the triangle $[p, x, y]$, yields that F_p is L-Lipschitz for $L = \max\{||\rho||_\infty, r_0 \cdot ||d\rho||_\infty\}$. Moreover,

$$\langle dF_p - \rho, [x, y]\rangle = \langle d\rho, [p, y, x]\rangle \, . \tag{6.13}$$

For $x \neq p$ and $|y - x| << |x - p|$, we use again estimate (6.12) and obtain from (6.13) that

$$||dF_p - \rho||_{\infty,B(p,r)} \leq \, ||d\rho||_{\infty,B(p,r)} \cdot r \, , \qquad r < r_0 \, . \tag{6.14}$$

(Here we also use the fact that the flat norm is determined locally, cf. Corollary 5.11.)

It is now clear from (6.14) that upon choosing r_0 small enough, we obtain that (6.9) holds for F_p, so that F_p is a BLD-map. with constants depending only on n and the constants associated with ρ. Moreover, (6.11) holds. The proof is complete. □

It follows from Theorem 6.5 that every Cartan-Whitney presentation is *asymptotically*, near every point p, a pullback of the standard presentation under a BLD-map. Such a BLD-map can branch at p, preventing us from concluding that ρ is "asymptotically locally standard". In the next subsection, we discuss an additional hypothesis that guarantees there is no branching at p. Interestingly, it is still true that an arbitrary Cartan-Whitney presentation is asymptotically locally standard outside an exceptional set as in Theorem 6.3. The following result is due to Keith and myself (it is proved in a more general context in [27]).

Theorem 6.6. *Let ρ be a Cartan-Whitney presentation in an open set Ω in $\mathbb{R}^n$, $n \geq 2$. Then there exists a closed set $E \subset \Omega$ of measure zero and of topological dimension at most $n-2$ such that the mappings F_p as in* (6.10) *are bi-Lipschitz near every $p \in \Omega \setminus E$.*

Proof. Fix $p \in \Omega$, and let r_0 and F_p be as in the proof of Theorem 6.5. In particular, F_p is a BLD-map. Fix $0 < \varepsilon < r_0$ and $q \in B(p, \varepsilon)$ such that F_p is a local homeomorphism near q; that is, we pick q outside the branch set of F_p. (Here we do not exclude the possibility that $q = p$.) We consider two families of rescaled maps,

$$(F_p)_{q,r}(x) := \frac{F_p(q + rx) - F_p(q)}{r},$$
$$(F_q)_{q,r}(x) := \frac{F_q(q + rx) - F_q(q)}{r} = \frac{F_q(q + rx)}{r},$$

for $x \in \mathbb{R}^n$, $|x| < 1$, and $r > 0$ small enough so that the maps are defined. Then, for all small enough $r > 0$, the maps $(F_p)_{q,r}$, $(F_q)_{q,r}$ are uniformly BLD-maps in the unit ball $\mathbb{B}^n$ of $\mathbb{R}^n$, taking the origin to itself. By standard Arzelà-Ascoli type arguments, we can choose a sequence of numbers (r_i) converging to zero such that both $(F_p)_{q,r_i}$ and $(F_q)_{q,r_i}$ converge uniformly to BLD-maps G_p and G_q, respectively, defined in $\mathbb{B}^n$. Because F_p is a local homeomorphism near q, the map G_p is bi-Lipschitz. On the other hand, one easily obtains from the definitions for the maps F_p and F_q, given in (6.10), and from the Stokes theorem as in the proof of Theorem 6.5, that

$$|(F_q)_{q,r_i}(x) - (F_p)_{q,r_i}(x)| \leq 2\,||\rho||_\flat\,|x|\,|p - q|\,, \tag{6.15}$$

for all $r_i > 0$ small enough.

Every BLD-map f, in addition to being Lipschitz, satisfies the following uniform lower bound for local squeezing,

$$\liminf_{y \to x} \frac{|f(x) - f(y)|}{|x - y|} \geq c > 0\,, \tag{6.16}$$

for every x in the domain of f, where $c > 0$ depends only on the BLD-data of f [49, Corollary 2.13]. By combining (6.15) and (6.16), we obtain that the maps G_p and G_q have a common local degree at 0, provided that $\varepsilon > 0$ is small enough, depending only on n and the flat norm of ρ. It follows that G_q is a local homeomorphism at q. Because $(F_q)_{q,r_i}$ converges to G_q locally uniformly, we must have (by the basic degree theory) that also $(F_q)_{q,r_i}$, and hence F_q is a local homeomorphism at q.

The assertion of the theorem now follows from the basic properties of BLD-mappings (as in Theorem 6.4), by covering Ω by balls of the form $B(p, \varepsilon)$ as in the preceding, and observing that the property "F_q is a local homeomorphism at q" is an open condition. The proof is complete. □

6.5. **A Sobolev condition that removes branching.** It is an interesting open problem to describe a bi-Lipschitz invariant hypothesis on a Cartan-Whitney presentation that would imply that the maps F_p as defined in the proofs of Theorems 6.3 and 6.5 are local homeomorphisms at p.

The following sufficient condition was given in [27].

Theorem 6.7. *Let ρ be a Cartan-Whitney presentation in an open set Ω in $\mathbb{R}^n$, $n \geq 2$. Assume that $\rho \in W^{1,2}_{loc}(\Omega)$. Then for every point $p \in \Omega$ there exists $r_0 > 0$ and a bi-Lipschitz mapping $F_p : B(p, r_0) \to \mathbb{R}^n$ such that*

$$||dF_p - \rho||_{\infty, B(p,r)} \leq ||d\rho||_{\infty, B(p,r)} \cdot r \tag{6.17}$$

for all $r < r_0$. The bi-Lipschitz constant of F_p depends only on n and the flat norm of ρ.

We will not prove Theorem 6.7 in these notes, but refer to [27]. For an earlier result, where ρ was assumed to be closed, see [28].

By the hypothesis $\rho \in W^{1,2}_{loc}(\Omega)$ in Theorem 6.7, we understand that the components of ρ as vector-valued functions belong to the local Sobolev space.

One can rephrase Theorem 6.7 by saying that *Cartan-Whitney presentations in the Sobolev space $W^{1,2}$ are asymptotically locally standard.* The Sobolev condition provides, therefore, a sought after integrability condition that guarantees that a given Cartan-Whitney presentation is locally standard, at least asymptotically; if the forms in question are in addition closed, then they are locally standard as required by our earlier discussion.

Interestingly, the imposed Sobolev condition in Theorem 6.7 is sharp. A straightforward computation shows that the pullback presentation under the winding map (6.6) is in the Sobolev space $W^{1,2-\varepsilon}$ for every $\varepsilon > 0$ near points on the branch set B_w. Unfortunately, the Sobolev condition in Theorem 6.7 is not bi-Lipschitz invariant.

Remark 6.8. Theorems 6.5 and 6.7 were proved in [30], [27] in a more general context than that of Euclidean spaces. In this way, there are applications to the problem of finding bi-Lipschitz parametrizations of metric spaces by Euclidean spaces, and to the smoothability of Lipschitz manifolds. For more discussion of these problems, and for references, see [68], [29], [30], [26], [27].

6.6. **Notes to Section 6.** The idea of looking for locally standard measurable structures in a Riemannian context, as discussed in this section, is due to Sullivan [66]. See also [68], [69]. Measurable Riemannian metrics in the context of quasiconformal geometry have been studied in [33], [6]. Cartan-Whitney presentations have been studied in [30], [29], [27]. The text in this section contains references to the related classical literature.

References

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

[2] Aronsson, G., Crandall, M. G., and Juutinen, P. A tour of the theory of absolutely minimizing functions. *Bull. Amer. Math. Soc. (N.S.) 41*, 4 (2004), 439–505 (electronic).

[3] Audin, M., and Lafontaine, J., Eds. *Holomorphic curves in symplectic geometry*, vol. 117 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1994.

[4] Balogh, Z. M., Rogovin, K., and Zürcher, T. The Stepanov differentiability theorem in metric measure spaces. *J. Geom. Anal. 14*, 3 (2004), 405–422.

[5] Benyamini, Y., and Lindenstrauss, J. *Geometric Nonlinear Functional Analysis, Volume I*, vol. 48 of *Colloquium Publications*. Amer. Math. Soc., 2000.

[6] Bojarski, B., Iwaniec, T., and Kopiecki, R. Riemannian manifolds with nonsmooth metric tensors and quasiconformal mappings. In *Monge-Ampère equations and related topics (Florence, 1980)*. Ist. Naz. Alta Mat. Francesco Severi, Rome, 1982, pp. 123–167.

[7] Bonk, M., Heinonen, J., and Saksman, E. The quasiconformal Jacobian problem. In *In the tradition of Ahlfors and Bers, III*, vol. 355 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 2004, pp. 77–96.

[8] Boyarskiĭ, B. V. Homeomorphic solutions of Beltrami systems. *Dokl. Akad. Nauk SSSR (N.S.) 102* (1955), 661–664.

[9] Bruckner, A., Bruckner, J., and Thomson, B. *Real Analysis*. Prentice-Hall, N.J., 1997.

[10] Burago, D., Burago, Y., and Ivanov, S. *A course in metric geometry*, vol. 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.

[11] Calderón, A. P. On differentiability of absolute continuous functions. *Riv. Mat. Univ. Parma 2* (1951), 203–213.

[12] Cannon, J. W., Conner, G. R., and Zastrow, A. One-dimensional sets and planar sets are aspherical. *Topology Appl. 120*, 1-2 (2002), 23–45. In memory of T. Benny Rushing.

[13] Cesari, L. Sulle funzioni assolutamente continue in due variabili. *Ann. Scuola Norm. Super. Pisa (2) 10* (1941), 91–101.

[14] Cheeger, J. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal. 9* (1999), 428–517.

[15] David, G. Solutions de l'équation de Beltrami avec $\|\mu\|_\infty = 1$. *Ann. Acad. Sci. Fenn. Ser. A I Math. 13*, 1 (1988), 25–70.

[16] David, G., and Semmes, S. *Fractured fractals and broken dreams: self-similar geometry through metric and measure*, vol. 7 of *Oxford Lecture Series in Mathematics and its Applications*. Clarendon Press Oxford University Press, 1997.

[17] Evans, L. C., and Gariepy, R. F. *Measure Theory and Fine Properties of Functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, Florida, 1992.

[18] Federer, H. *Geometric Measure Theory*, vol. 153 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York, 1969.

[19] Gromov, M. Monotonicity of the volume of intersection of balls. In *Geometrical aspects of functional analysis (1985/86)*, vol. 1267 of *Lecture Notes in Math.* Springer, Berlin, 1987, pp. 1–4.

[20] Gromov, M. *Metric structures for Riemannian and non-Riemannian spaces*, vol. 152 of *Progress in Mathematics*. Birkhäuser Boston Inc., Boston, MA, 1999. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[21] Hajłasz, P. Sobolev spaces on an arbitrary metric space. *Potential Anal. 5* (1996), 403–415.

[22] Hajłasz, P. Sobolev spaces on metric-measure spaces. In *Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002)*, vol. 338 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 2003, pp. 173–218.

[23] Hajłasz, P., and Koskela, P. Sobolev met Poincaré. *Memoirs Amer. Math. Soc. 145*, 688 (2000).

[24] Harrison, J. Geometric representations of currents and distributions. In *Fractal geometry and stochastics III*, vol. 57 of *Progr. Probab.* Birkhäuser, Basel, 2004, pp. 193–204.

[25] Heinonen, J. *Lectures on analysis on metric spaces*. Springer-Verlag, New York, 2001.

[26] Heinonen, J. The branch set of a quasiregular mapping. In *Proceedings of the International Congress of Mathematicians, Vol. II (Beijing, 2002)* (Beijing, 2002), Higher Ed. Press, pp. 691–700.

[27] Heinonen, J., and Keith, S. Work in preparation.

[28] Heinonen, J., and Kilpeläinen, T. BLD-mappings in $W^{2,2}$ are locally invertible. *Math. Ann. 318*, 2 (2000), 391–396.

[29] Heinonen, J., and Rickman, S. Geometric branched covers between generalized manifolds. *Duke Math. J. 113*, 3 (2002), 465–529.

[30] Heinonen, J., and Sullivan, D. On the locally branched Euclidean metric gauge. *Duke Math. J. 114*, 1 (2002), 15–41.

[31] Hirsch, F., and Lacombe, G. *Elements of functional analysis*, vol. 192 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1999. Translated from the 1997 French original by Silvio Levy.

[32] Hohti, A. On absolute Lipschitz neighbourhood retracts, mixers, and quasiconvexity. *Topology Proc. 18* (1993), 89–106.

[33] Iwaniec, T. Regularity theorems for solutions of partial differential equations for quasiconformal mappings in several dimensions. *Dissertationes Math. (Rozprawy Mat.) 198* (1982), 45.

[34] Iwaniec, T., and Martin, G. *Geometric function theory and non-linear analysis*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2001.
[35] Johnson, W. B., Lindenstrauss, J., and Schechtman, G. Extensions of Lipschitz maps into Banach spaces. *Israel J. Math. 54*, 2 (1986), 129–138.
[36] Keith, S. A differentiable structure for metric measure spaces. *Adv. Math. 183*, 2 (2004), 271–315.
[37] Keith, S. Measurable differentiable structures and the Poincaré inequality. *Indiana Univ. Math. J. 53*, 4 (2004), 1127–1150.
[38] Kleiner, B. The local structure of length spaces with curvature bounded above. *Math. Z. 231*, 3 (1999), 409–456.
[39] Koskela, P. Upper gradients and Poincaré inequalities. In *Lecture notes on analysis in metric spaces (Trento, 1999)*, Appunti Corsi Tenuti Docenti Sc. Scuola Norm. Sup., Pisa, 2000, pp. 55–69.
[40] Kulkarni, R. S., and Pinkall, U., Eds. *Conformal geometry*. Aspects of Mathematics, E12. Friedr. Vieweg & Sohn, Braunschweig, 1988. Papers from the seminar held in Bonn, 1985–1986.
[41] Lang, U., Pavlović, B., and Schroeder, V. Extensions of Lipschitz maps into Hadamard spaces. *Geom. Funct. Anal. 10*, 6 (2000), 1527–1553.
[42] Lang, U., and Schroeder, V. Kirszbraun's theorem and metric spaces of bounded curvature. *Geom. Funct. Anal. 7*, 3 (1997), 535–560.
[43] Lee, J. R., and Naor, A. Extending Lipschitz functions via random metric partitions. *Invent. Math. 160*, 1 (2005), 59–95.
[44] Lehto, O. Homeomorphic solutions of a Beltrami differential equation. In *Festband 70. Geburtstag R. Nevanlinna*. Springer, Berlin, 1966, pp. 58–65.
[45] Lehto, O. *Univalent functions and Teichmüller spaces*. Springer-Verlag, New York-Heidelberg, 1987.
[46] Lehto, O., and Virtanen, K. I. *Quasiconformal mappings in the plane*, second ed. Springer-Verlag, New York, 1973. Translated from the German by K. W. Lucas, Die Grundlehren der mathematischen Wissenschaften, Band 126.
[47] Luukkainen, J., and Väisälä, J. Elements of Lipschitz topology. *Ann. Acad. Sci. Fenn. Ser. A I Math. 3* (1977), 85–122.
[48] Maly, J. A simple proof of the Stepanov theorem on differentiability almost everywhere. *Exposition. Math. 17*, 1 (1999), 59–61.
[49] Martio, O., and Väisälä, J. Elliptic equations and maps of bounded length distortion. *Math. Ann. 282*, 3 (1988), 423–443.
[50] Mattila, P. *Geometry of sets and measures in Euclidean spaces*, vol. 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.
[51] Miller, J. Spherical Lipschitz extensions. Preprint, 2005.
[52] Milnor, J. W., and Stasheff, J. D. *Characteristic classes*. Princeton University Press, Princeton, N. J., 1974. Annals of Mathematics Studies, No. 76.
[53] Morrey, Jr., C. B. On the solutions of quasi-linear elliptic partial differential equations. *Trans. Amer. Math. Soc. 43*, 1 (1938), 126–166.
[54] Reshetnyak, Y. G. Space mappings with bounded distortion. *Sibirsk. Mat. Z. 8* (1967), 629–659.

[55] Reshetnyak, Y. G. *Space mappings with bounded distortion*, vol. 73 of *Translations of Mathematical Monographs*. American Mathematical Society, 1989. Translated from the Russian by H. H. McFaden.

[56] Rickman, S. *Quasiregular Mappings*. Springer-Verlag, Berlin, 1993.

[57] Rudin, W. *Real and Complex Analysis*, third ed. McGraw-Hill, New York, 1987.

[58] Schwartz, L. *Théorie des distributions. Tome I*. Actualités Sci. Ind., no. 1091 = Publ. Inst. Math. Univ. Strasbourg 9. Hermann & Cie., Paris, 1950.

[59] Semmes, S. Finding structure in sets with little smoothness. In *Proc. of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)* (Basel, 1995), Birkhäuser, pp. 875–885.

[60] Semmes, S. Good metric spaces without good parameterizations. *Rev. Mat. Iberoamericana 12* (1996), 187–275.

[61] Semmes, S. On the nonexistence of bi-Lipschitz parameterizations and geometric problems about A_∞-weights. *Rev. Mat. Iberoamericana 12* (1996), 337–410.

[62] Semmes, S. Mappings and spaces. In *Quasiconformal mappings and analysis (Ann Arbor, MI, 1995)*. Springer, New York, 1998, pp. 347–368.

[63] Semmes, S. *Some novel types of fractal geometry*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2001.

[64] Siebenmann, L., and Sullivan, D. On complexes that are Lipschitz manifolds. In *Geometric topology (Proc. Georgia Topology Conf., Athens, Ga., 1977)*. Academic Press, New York, 1979, pp. 503–525.

[65] Stein, E. M. *Singular integrals and differentiability properties of functions*. Princeton University Press, Princeton, N.J., 1970. Princeton Mathematical Series, No. 30.

[66] Sullivan, D. The Ahlfors-Bers measurable Riemann mapping theorem for higher dimensions. Lecture at the Ahlfors celebration, Stanford University, September 1997, http://www.msri.org/publications/ln/hosted/ahlfors/1997 /sullivan/1/index.html.

[67] Sullivan, D. Hyperbolic geometry and homeomorphisms. In *Geometric topology (Proc. Georgia Topology Conf., Athens, Ga., 1977)*. Academic Press, New York, 1979, pp. 543–555.

[68] Sullivan, D. Exterior d, the local degree, and smoothability. In *Prospects in topology (Princeton, NJ, 1994)*, vol. 138 of *Ann. of Math. Stud.* Princeton Univ. Press, Princeton, NJ, 1995, pp. 328–338.

[69] Sullivan, D. On the foundation of geometry, analysis, and the differentiable structure for manifolds. In *Topics in low-dimensional topology (University Park, PA, 1996)*. World Sci. Publishing, River Edge, NJ, 1999, pp. 89–92.

[70] Sullivan, D., and Teleman, N. An analytic proof of Novikov's theorem on rational Pontrjagin classes. *Inst. Hautes Études Sci. Publ. Math.*, 58 (1983), 79–81 (1984).

[71] Teleman, N. The index of signature operators on Lipschitz manifolds. *Inst. Hautes Études Sci. Publ. Math.*, 58 (1983), 39–78 (1984).

[72] Tukia, P. A quasiconformal group not isomorphic to a Möbius group. *Ann. Acad. Sci. Fenn. Ser. A I Math. 6* (1981), 149–160.

[73] Tukia, P., and Väisälä, J. Lipschitz and quasiconformal approximation and extension. *Ann. Acad. Sci. Fenn. Ser. A I Math. 6*, 2 (1981), 303–342 (1982).

[74] Väisälä, J. *Lectures on n-dimensional quasiconformal mappings.* No. 229 in Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1971.

[75] van Mill, J., and van de Vel, M. On an internal property of absolute retracts. In *The Proceedings of the 1979 Topology Conference (Ohio Univ., Athens, Ohio, 1979)* (1979), vol. 4, pp. 193–200 (1980).

[76] Weaver, N. Lipschitz algebras and derivations. II. Exterior differentiation. *J. Funct. Anal. 178*, 1 (2000), 64–112.

[77] Whitney, H. Algebraic topology and integration theory. *Proc. Nat. Acad. Sci. U. S. A. 33* (1947), 1–6.

[78] Whitney, H. *Geometric integration theory.* Princeton University Press, Princeton, N. J., 1957.

[79] Wolf, J. A. *Spaces of constant curvature*, fifth ed. Publish or Perish Inc., Houston, TX, 1984.

[80] Ziemer, W. P. *Weakly Differentiable Functions*, vol. 120 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, 1989.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA

E-mail address: juha@umich.edu