Geometric Measure Theory

Ernst Kuwert, University of Freiburg

Course, Summer Term 2023

Contents

Chapter 1

Introduction

The rough idea of Geometric Measure Theory (GMT) is to model surfaces in Euclidean spaces by measures. From the Analysis course or from Differential Geometry, we already know two ways to describe surfaces:

- a parametrized surface of dimension n in $\mathbb{R}^{n+k}$ is an immersion $f \in C^1(U, \mathbb{R}^{n+k})$ where $U \subset \mathbb{R}^n$ is an open parameter domain.
- a C^1 submanifold of dimension n in $\mathbb{R}^{n+k}$ is a subset M with the following property: for any $p \in M$ there is an open neighborhood W and a diffeomorphism $\phi : W \to \phi(W)$, such that $\phi(M \cap W) = (\mathbb{R}^n \times \{0\}) \cap \phi(W)$.

These concepts are definitively very useful. However, in Geometric Calculus of Variations we are dealing with sequences Σ_i of surfaces. For instance, in the classical Plateau problem we want to minimize the area among surfaces Σ having a given boundary $\partial\Sigma = \Gamma$. The approach is to choose a minimizing sequence Σ_i, i.e. the areas of the Σ_i converge to the infimum. The goal is obtain the minimizing surface Σ as the limit of the Σ_i. In situations of this type, the above concepts are of limited use.

For submanifolds $M_i \subset \mathbb{R}^{n+k}$ the data, i.e. the covering by open sets W and the diffeomorphisms ϕ, can degenerate in many ways along the sequence. For instance, the open sets W may shrink or the diffeomorphsms may lose their rank etc. This concept is really inappropriate when dealing with sequences.

The concept of parametrized surfaces is somewhat better, at least if we consider immersions $f_i : U \to \mathbb{R}^{n+k}$ from a fixed parameter domain, since then we can ask about the convergence of the functions f_i. However this also leads to serious difficulties:

- If the immersions f_i were uniformly Lipschitz, then we could apply the Arzela-Ascoli theorem to get a Lipschitz limit. But in problems of interest, like for example in the Plateau problem, we only have a bound for the total area, and this does not give us any pointwise control. In fact a minimizing sequence for the Plateau problem could get really wild, forming many thin tentacles and so on.
- By invariance of the area under reparametrization, there is the possibility of degeneration due to bad parametrizations, even if all f_i just describe one nice fixed surface. There are concepts to handle this by choosing a preferred parametrization, e.g. a parametrization

by arclength for curves or a conformal parametrization for two-dimensional surfaces. Also in higher dimensions there are situations where a particular parametrization plays a role, for example if the surfaces are graphs. However in general there is no special parametrization and this poses a problem for the concept.

By contrast the space of measures has a simple and general compactness property. Let μ_i be a sequence of Radon measures on $\mathbb{R}^{n+k}$ such that $\sup_{i\in\mathbb{N}} \mu_i(U) < \infty$ for all $U \subset\subset \mathbb{R}^{n+k}$. Then a subsequence converges to a Radon measure μ, in the sense that

$$\int_{\mathbb{R}^{n+k}} \varphi\, d\mu = \lim_{i\to\infty} \int_{\mathbb{R}^{n+k}} \varphi\, d\mu_i \quad \text{ for all } \varphi \in C^0_{\mathrm{c}}(\mathbb{R}^{n+k}).$$

This suggests to model surfaces in $\mathbb{R}^{n+k}$ just by Radon measures, but this notion is too general. We want to have concepts which still capture some geometry. The above interprets Radon measures as nonnegative linear functionals on $C^0_{\mathrm{c}}(\mathbb{R}^{n+k})$. We now list three main concepts of GMT, using in each case a description as linear functionals:

Varifolds (Almgren, Allard)
Let $G(n,k)$ be the set of n-dimensional subspaces of $\mathbb{R}^{n+k}$. For any n-dimensional, properly embedded surface Σ one has a functional V_Σ acting on functions $\phi : \mathbb{R}^{n+k} \times G(n,k) \to \mathbb{R}$ by

$$V_\Sigma(\phi) = \int_\Sigma \phi(x, T_x\Sigma)\, d\mu_\Sigma(x).$$

This motivates the definition of an n-varifold as a nonnegative continuous linear functional V on $C^0_{\mathrm{c}}(\mathbb{R}^{n+k} \times G(n,k))$. Equivalently, V is a Radon measure on $\mathbb{R}^{n+k} \times G(n,k)$. This concept allows to define the first variation and a weak notion of mean curvature, in particular one introduces the class of stationary varifolds generalizing classical minimal surfaces.

Currents (Federer, Fleming)
To any n-dimensional, oriented, properly embedded surface $\Sigma \subset \mathbb{R}^{n+k}$ one associates a functional T_Σ acting on differential n-forms with compact support by integration, that is

$$T_\Sigma(\omega) = \int_\Sigma \omega \quad \text{ for all } \omega \in C^\infty_{\mathrm{c}}\big(\mathbb{R}^{n+k}, \Lambda^n(\mathbb{R}^{n+k})\big).$$

The space of n-dimensional currents is then defined as the set of all continuous linear functionals on that space of differential forms. It is easy to define the boundary of an n-current, one just puts $\partial T(\eta) = T(d\eta)$ for all $(n-1)$-forms. For regular surfaces Σ, this is consistent by the theorem of Stokes. A big success of the concept is that it allows to formulate and solve Plateau's problem.

Caccioppoli sets (Caccioppoli, Di Giorgi)
A Borel set $E \subset \mathbb{R}^{n+1}$ is a Caccioppoli set if its characteristic function χ_E has locally bounded variation. This means that for any bounded open set $U \subset \mathbb{R}^n$ one has

$$|D\chi_E|(U) := \sup\Big\{ \int_E \operatorname{div} g(x)\, dx : \operatorname{spt} g \subset U,\ |g| \leq 1 \Big\} < \infty.$$

The definition is motivated by the case when E is a domain of class C^1, with inward unit normal ν_E and boundary measure $\mu_{\partial E}$. Namely then the theorem of Gauß implies for g as

above
$$\int_E \operatorname{div} g(x)\, dx = -\int_{\partial E} \langle g(x), \nu_E(x)\rangle \, d\mu_{\partial E}(x) \leq \mu_{\partial E}(U) < \infty.$$
The concept of Caccioppoli sets is of great importance in phase transitions, where two materials are separated by a common phase boundary. A classical example is the Stefan problem modelling the melting of ice in water.

All three concepts lead to a beautiful theory, however we will not be able to cover them for reasons of time. We decided to focus on varifolds since they are rather general, have many applications and relate to my own recent work.

Chapter 2

Measure theory in metric spaces

Definition 2.1. *A measure on X is a function $\mu : 2^X \to [0,\infty]$ with $\mu(\emptyset) = 0$, such that*

$$\mu(A) \leq \sum_{i=1}^{\infty} \mu(A_i) \quad \text{whenever } A \subset \bigcup_{i=1}^{\infty} A_i. \tag{2.1}$$

It follows that μ is monotone, that is $\mu(A) \leq \mu(B)$ for $A \subset B$, and countably subadditive:

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

Reversely, these two properties imply (2.1). We should mention that μ as above is often called an *outer measure*, while by a measure one means a countably additive function on a σ-algebra which is given a priori. However, in geometry it is convenient to have μ defined on all sets.

Definition 2.2. *A set $A \subset X$ is μ-measurable if the following holds:*

$$\mu(S) \geq \mu(S \cap A) + \mu(S \backslash A) \quad \text{for all } S \subset X. \tag{2.2}$$

As S is the union of $S \cap A$ and $S \backslash A$, we always have $\mu(S) \leq \mu(S \cap A) + \mu(S \backslash A)$, and hence equality if A is measurable. Null sets are measurable: if $\mu(N) = 0$ then

$$\mu(S \cap N) = 0, \quad \text{and} \quad \mu(S) \geq \mu(S \backslash N) = \mu(S \cap N) + \mu(S \backslash N).$$

We will see that the system $\mathcal{M}$ of measurable subsets of X is closed under set operations and limits. To formulate this one introduces the following basic concept.

Definition 2.3 (σ-Algebra)**.** *A system $\mathcal{A} \subset 2^X$ is a σ-algebra, if the following hold:*

(i) $X \in \mathcal{A}$

(ii) $A \in \mathcal{A} \quad \Rightarrow \quad X \backslash A \in \mathcal{A}$

(iii) $A_i \in \mathcal{A}$ *for* $i = 1, 2, \ldots \quad \Rightarrow \quad \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

A σ-algebra $\mathcal{A}$ is also closed under countable intersections, this follows by writing

$$\bigcap_{i=1}^{\infty} A_i = X \backslash \left(\bigcup_{i=1}^{\infty} X \backslash A_i \right).$$

Moreover for $A, B \in \mathcal{A}$ we also have $A \backslash B = A \cap (X \backslash B) \in \mathcal{A}$.

Lemma 2.4. *Let $A_1, A_2, \dots, A_k \subset X$ be pairwise disjoint, μ-measurable sets. Then*

$$\mu\big(S \cap \bigcup_{i=1}^{k} A_i\big) = \sum_{i=1}^{k} \mu(S \cap A_i) \quad \textit{for all } S \subset X.$$

Proof. This is trivial for $k = 1$, and for $k \geq 2$ we get by induction, as A_k is measurable,

$$\begin{aligned}
\mu\big(S \cap \bigcup_{i=1}^{k} A_i\big) &= \mu\big((S \cap \bigcup_{i=1}^{k} A_i) \cap A_k\big) + \mu\big((S \cap \bigcup_{i=1}^{k} A_i) \backslash A_k\big) \\
&= \mu(S \cap A_k) + \mu\big(S \cap \bigcup_{i=1}^{k-1} A_i\big) \\
&= \sum_{i=1}^{k} \mu(S \cap A_i).
\end{aligned}$$

□

Theorem 2.5. *The system $\mathcal{M}$ of μ-measurable subsets of X is a σ-algebra. Moreover*

$$A_i \in \mathcal{M},\ i \in \mathbb{N},\ \textit{pairwise disjoint} \quad \Rightarrow \quad \mu\big(\bigcup_{i=1}^{\infty} A_i\big) = \sum_{i=1}^{\infty} \mu(A_i). \tag{2.3}$$

Proof. We have $X \in \mathcal{M}$ since $\mu(S \cap X) = \mu(S)$ and $\mu(S \backslash X) = \mu(\emptyset) = 0$. For $A \in \mathcal{M}$ we also have $X \backslash A \in \mathcal{M}$, because

$$\mu\big(S \cap (X \backslash A)\big) = \mu(S \backslash A) \quad \text{and} \quad \mu\big(S \backslash (X \backslash A)\big) = \mu(S \cap A).$$

Next we show that $A \cup B \in \mathcal{M}$ for $A, B \in \mathcal{M}$, in fact we have for any $S \subset X$

$$\begin{aligned}
\mu\big(S \cap (A \cup B)\big) + \mu\big(S \backslash (A \cup B)\big) &\leq \mu(S \cap A) + \mu\big((S \backslash A) \cap B\big) + \mu\big((S \backslash A) \backslash B\big) \\
&\leq \mu(S \cap A) + \mu(S \backslash A) \quad (\text{using } B \in \mathcal{M}) \\
&\leq \mu(S) \quad (\text{using } A \in \mathcal{M}).
\end{aligned}$$

This yields further $A \cap B = X \backslash ((X \backslash A) \cup (X \backslash B)) \in \mathcal{M}$ and $A \backslash B = A \cap (X \backslash B) \in \mathcal{M}$. Induction shows that $\mathcal{M}$ is closed under finite unions and intersections. We now prove

$$A_i \in \mathcal{M} \quad \text{for } i = 1, 2, \dots \quad \Rightarrow \quad A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{M}.$$

We can assume that $A_i \cap A_j = \emptyset$ for $i \neq j$, otherwise we consider $\tilde{A}_i = A_i \backslash (A_1 \cup \ldots \cup A_{i-1})$. We conclude for any S, using $\bigcup_{i=1}^{k} A_i \in \mathcal{M}$,

$$\mu(S) = \mu\big(S \cap \bigcup_{i=1}^{k} A_i\big) + \mu\big(S \backslash \bigcup_{i=1}^{k} A_i\big) \geq \sum_{i=1}^{k} \mu(S \cap A_i) + \mu(S \backslash A).$$

The inequality uses Lemma 2.4 and monotonicity of μ. Letting $k \to \infty$ we conclude

$$\mu(S) \geq \sum_{i=1}^{\infty} \mu(S \cap A_i) + \mu(S \backslash A) \geq \mu\big(\bigcup_{i=1}^{\infty} (S \cap A_i)\big) + \mu(S \backslash A) = \mu(S \cap A) + \mu(S \backslash A).$$

Thus $A = \bigcup_{i=1}^{\infty} A_i$ is measurable. Putting $S = X$ in Lemma 2.4 we finally have

$$\lim_{k\to\infty} \mu\big(\bigcup_{i=1}^{k} A_i\big) = \sum_{i=1}^{\infty} \mu(A_i) \geq \mu\big(\bigcup_{i=1}^{\infty} A_i\big) \geq \lim_{k\to\infty} \mu\big(\bigcup_{i=1}^{k} A_i\big).$$

$\square$

Theorem 2.6 (Continuity of measure)**.** *Let $A_1, A_2, \ldots$ be measurable sets. Then*

(i) *if $A_1 \subset A_2 \subset \ldots$ then $\mu\big(\bigcup_{i=1}^{\infty} A_i\big) = \lim_{k\to\infty} \mu(A_k)$,*

(ii) *if $A_1 \supset A_2 \supset \ldots$ and $\mu(A_1) < \infty$, then $\mu\big(\bigcap_{i=1}^{\infty} A_i\big) = \lim_{k\to\infty} \mu(A_k)$.*

Proof. For (i) let $\tilde{A}_k = A_k \backslash \bigcup_{i=1}^{k-1} A_i$ and compute using (2.3)

$$\mu\big(\bigcup_{i=1}^{\infty} A_i\big) = \mu\big(\bigcup_{i=1}^{\infty} \tilde{A}_i\big) = \sum_{i=1}^{\infty} \mu(\tilde{A}_i) = \lim_{k\to\infty} \mu\big(\bigcup_{i=1}^{k} \tilde{A}_i\big) = \lim_{k\to\infty} \mu(A_k).$$

For (ii) consider the increasing sequence $A'_k = A_1 \backslash A_k$. We have

$$\mu(A_1) = \mu(A_1 \cap A_k) + \mu(A_1 \backslash A_k) = \mu(A_k) + \mu(A'_k).$$

We conclude using statement (i)

$$\mu(A_1) - \lim_{k\to\infty} \mu(A_k) = \lim_{k\to\infty} \mu(A'_k) = \mu\big(\bigcup_{i=1}^{\infty} A'_i\big) = \mu(A_1 \backslash \bigcap_{i=1}^{\infty} A_i) = \mu(A_1) - \mu\big(\bigcap_{i=1}^{\infty} A_i\big).$$

$\square$

Example 2.7. *The condition $\mu(A_1) < \infty$ in (ii) cannot be dropped, e.g. consider the counting measure and $A_k = \{k, k+1, \ldots\} \subset \mathbb{N}$.*

A useful construction is the following mapping of a measure.

Lemma 2.8. *For a measure μ on X and $f : X \to Y$, the pushforward on Y is defined by*

$$f(\mu) : 2^Y \to [0, \infty],\ f(\mu)(B) = \mu(f^{-1}(B)).$$

$f(\mu)$ is an outer measure. If $f^{-1}(B)$ is μ-measurable, then B is $f(\mu)$-measurable.

Proof. It is easy to see that $f(\mu)$ is an outer measure. By definition, $f^{-1}B$ is μ-measurable if

$$\mu(S) = \mu(S \cap f^{-1}B) + \mu(S \backslash f^{-1}B) \quad \text{for all } S \subset X.$$

On the other hand, the $f(\mu)$-measurability of B means that

$$\mu(f^{-1}T) = \mu(f^{-1}T \cap f^{-1}B) + \mu(f^{-1}T \backslash f^{-1}B) \quad \text{for all } T \subset Y.$$

Clearly, the first implies the second. $\square$

Up to now we considered measures on an arbitrary set X, but from now on we assume that X carries a metric d. In particular we have the Borel algebra, which is by definition the smallest σ-algebra that contains all open sets (or equivalently, all closed sets). It is then a natural requirement for μ to have these sets measurable. This can be guaranteed via the following nice result.

Theorem 2.9 (Caratheodory's criterion). *Let μ be a measure on (X, d), and assume*

$$\mu(A \cup B) \geq \mu(A) + \mu(B) \quad \text{whenever} \quad \operatorname{dist}(A, B) = \inf_{a \in A, b \in B} d(a, b) > 0. \tag{2.4}$$

Then all Borel sets are μ-measurable.

Proof. We show that any closed set $C \subset X$ is measurable. Let $S \subset X$ be arbitrary, without loss of generality $\mu(S) < \infty$. Consider the parallel sets

$$C_j = \{x \in X : \operatorname{dist}(x, C) \leq \frac{1}{j}\}.$$

We have $\operatorname{dist}(S \backslash C_j, S \cup C) \geq \frac{1}{j} > 0$, hence the assumption implies

$$\mu(S \backslash C_j) + \mu(S \cap C) \leq \mu\big((S \backslash C_j) \cup (S \cap C)\big) \leq \mu(S).$$

The theorem follows if we can show

$$\mu(S \backslash C) \leq \lim_{j \to \infty} \mu(S \backslash C_j). \tag{2.5}$$

Now consider for $k \in \mathbb{N}$ the parallel strips

$$S_k = \big\{x \in S : \frac{1}{k+1} < \operatorname{dist}(x, C) \leq \frac{1}{k}\big\} = S \cap (C_k \backslash C_{k+1}).$$

We have for any $j \in \mathbb{N}$

$$S \backslash C = \big(S \backslash C_j\big) \cup \bigcup_{k=j}^{\infty} S_k.$$

Taking the limit $j \to \infty$ we can write, using the monotonicity,

$$\mu(S \backslash C) \leq \lim_{j \to \infty} \mu(S \backslash C_j) + \lim_{j \to \infty} \mu\Big(\bigcup_{k=j}^{\infty} S_k\Big). \tag{2.6}$$

Now for $k \geq j+2$ we have $\operatorname{dist}(S_j, S_k) \geq \frac{1}{j+1} - \frac{1}{k} > 0$. Therefore by assumption

$$\begin{aligned} \sum_{i=1}^{N} \mu(S_{2i}) &= \mu\Big(\bigcup_{i=1}^{N} S_{2i}\Big) \leq \mu(S) < \infty, \\ \sum_{i=1}^{N} \mu(S_{2i-1}) &= \mu\Big(\bigcup_{i=1}^{N} S_{2i-1}\Big) \leq \mu(S) < \infty. \end{aligned}$$

From the Cauchy criterion for series, we conclude

$$\lim_{j \to \infty} \mu\Big(\bigcup_{k=j}^{\infty} S_k\Big) \leq \lim_{j \to \infty} \sum_{k=j}^{\infty} \mu(S_k) = 0.$$

Recalling (2.6) we obtain (2.5), and the theorem is proved. □

Definition 2.10 (Borel regularity). *A masure μ on a metric space (X,d) is called Borel regular, if it has the following two properties:*

(1) *All Borel sets are μ-measurable.*

(2) *Any $S \subset X$ has a Borel hull: there exists a Borel set $B \supset S$ with $\mu(B) = \mu(S)$.*

We want to discuss the Borel regularity in connection with the operation of restricting a measure, which is defined as follows.

Lemma 2.11 (restriction measure). *Let μ be a measure on X and $E \subset X$. Then the restriction of μ to E is the measure given by*

$$(\mu \llcorner E)(A) = \mu(A \cap E) \quad \text{for all } A \subset X.$$

If $A \subset X$ is μ-measurable, then A is also $\mu \llcorner E$-measurable.

Proof. For any $S \subset X$ we have, if A is μ-measurable,

$$\begin{aligned}(\mu \llcorner E)(S) &= \mu(S \cap E) \\ &\geq \mu\big((S \cap E) \cap A\big) + \mu\big((S \cap E) \backslash A\big) \\ &= \mu\big((S \cap A) \cap E\big) + \mu\big((S \backslash A) \cap E\big) \\ &= (\mu \llcorner E)(S \cap A) + (\mu \llcorner E)(S \backslash A).\end{aligned}$$

□

Note that in Lemma 2.11 the set E need not be μ-measurable.

Theorem 2.12 (Borel regularity of $\mu \llcorner E$). *Let μ be a Borel regular measure on X. Then for $E \subset X$ the measure $\mu \llcorner E$ is also Borel regular, if one of the following conditions hold:*

(1) *E is a Borel set.*

(2) *E is μ-measurable, and E is a countable union of sets with finite μ-measure.*

Proof. Lemma 2.11 implies that Borel sets are $\mu \llcorner E$-measurable, thus we only have to construct a Borel hull for a given set $S \subset X$. We first give the argument for condition (1). As μ is Borel regular, there exists a Borel set B such that

$$(S \cap E) \subset B \quad \text{and} \quad \mu(B) = \mu(S \cap E).$$

Let $\tilde{B} = B \cap (X \backslash E)$. Then $\tilde{B}$ is Borel with $\tilde{B} \supset S$, and we compute

$$(\mu \llcorner E)(\tilde{B}) = \mu(B \cap E) \leq \mu(B) = \mu(S \cap E) = (\mu \llcorner E)(S).$$

This proves our claim. We now show the result for condition (2), first in the case when $\mu(E) < \infty$. Choose $B \supset E$ Borel with $\mu(B) = \mu(E)$, and compute for any $S \subset X$

$$\begin{aligned}(\mu \llcorner B)(S) &= \mu(S \cap B) \\ &\leq \mu(S \cap E) + \mu\big(S \cap (B \backslash E)\big) \\ &\leq \mu(S \cap E) + \mu(B \backslash E) \\ &= (\mu \llcorner E)(S) + \underbrace{\mu(B) - \mu(E)}_{=0} \leq (\mu \llcorner B)(S).\end{aligned}$$

This shows $\mu \llcorner E = \mu \llcorner B$, and the claim follows by case (1). Finally, let E be μ-measurable with $E = \bigcup_{j=1}^{\infty} E_j$ where $\mu(E_j) < \infty$. We can assume:

- E_j is μ-measurable, otherwise choose Borel sets $\tilde{E}_j \supset E_j$ with $\mu(\tilde{E}_j) = \mu(E_j)$, and consider $\tilde{E}_j \cap E$ instead of E_j.
- $E_1 \subset E_2 \subset \ldots$, otherwise we pass to $\bigcup_{i=1}^{j} E_i$ instead of E_j.

Now let $S \subset X$. As shown just before, we can choose $B_j \supset S$ Borel with $(\mu \llcorner E_j)(B_j) = (\mu \llcorner E_j)(S)$. For $B = \bigcap_{j=1}^{\infty} B_j$ we have $B \supset S$ and, by continuity of the measure,

$$\begin{aligned}(\mu \llcorner E)(B) &= \mu(B \cap E)\\ &= \lim_{j\to\infty} \mu(B \cap E_j)\\ &\leq \limsup_{j\to\infty} \mu(B_j \cap E_j)\\ &= \lim_{j\to\infty} \mu(S \cap E_j)\\ &\leq (\mu \llcorner E)(S).\end{aligned}$$

This settles the remaining case. $\square$

Theorem 2.13 (Approximation)**.** *Let μ be a Borel measure on (X, d). Then for any Borel set $A \subset X$ the following hold:*

(1) $\mu(A) = \inf\{\mu(U) : U$ *open,* $A \subset U\}$*, if the right hand side is finite.*

(2) $\mu(A) = \sup\{\mu(C) : C$ *closed,* $C \subset A\}$*, if* $\mu(A) < \infty$*.*

Proof. We first assume that $\mu(X) < \infty$, and consider

$$\mathcal{A} = \{A \subset X : A \text{ is Borel and satisfies (1)}\}.$$

We claim that $\mathcal{A}$ is closed under countable unions and intersections. Let $A_j \in \mathcal{A}$, $j \in \mathbb{N}$, hence there exist open $U_j \supset A_j$ with $\mu(U_j) \leq \mu(A_j) + 2^{-j}\varepsilon$. Since A_j is measurable we have

$$\mu(U_j \backslash A_j) \leq \mu(U_j) - \mu(A_j) < 2^{-j}\varepsilon.$$

We conclude

$$\begin{aligned}\mu\Big(\bigcup_{j=1}^{\infty} U_j\Big) &\leq \mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) + \mu\Big(\bigcup_{j=1}^{\infty} U_j \backslash \bigcup_{j=1}^{\infty} A_j\Big)\\ &\leq \mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) + \mu\Big(\bigcup_{j=1}^{\infty} U_j \backslash A_j\Big)\\ &\leq \mu\Big(\bigcup_{j=1}^{\infty} A_j\Big) + \varepsilon.\end{aligned}$$

For intersections we argue similarly

$$\begin{aligned}\mu\Big(\bigcap_{j=1}^{\infty} U_j\Big) &\leq \mu\Big(\bigcap_{j=1}^{\infty} A_j\Big) + \mu\Big(\bigcap_{j=1}^{\infty} U_j \backslash \bigcap_{j=1}^{\infty} A_j\Big)\\ &\leq \mu\Big(\bigcap_{j=1}^{\infty} A_j\Big) + \mu\Big(\bigcup_{j=1}^{\infty} U_j \backslash A_j\Big)\\ &\leq \mu\Big(\bigcap_{j=1}^{\infty} A_j\Big) + \varepsilon.\end{aligned}$$

Of course the countable intersection of the U_j need not be open, however

$$\bigcap_{j=1}^{\infty} A_j \subset \bigcap_{j=1}^{N} A_j \subset \bigcap_{j=1}^{N} U_j.$$

By continuity of the measure, and recalling $\mu(X) < \infty$,

$$\lim_{N\to\infty} \mu\Big(\bigcap_{j=1}^{N} U_j\Big) = \mu\Big(\bigcap_{j=1}^{\infty} U_j\Big) \leq \mu\Big(\bigcap_{j=1}^{\infty} A_j\Big) + \varepsilon.$$

Now $\mathcal{A}$ trivially contains the open sets. But any closed set C is the intersection of the open sets $\{x \in X : \operatorname{dist}(x, C) < \frac{1}{j}\}$, hence $\mathcal{A}$ also contains the closed sets. Let

$$\tilde{\mathcal{A}} = \{A \in \mathcal{A} : X \backslash A \in \mathcal{A}\}.$$

Clearly $\emptyset, X \in \tilde{\mathcal{A}}$, and $A \in \tilde{\mathcal{A}}$ implies $X \backslash A \in \tilde{\mathcal{A}}$. Moreover for $A_j \in \tilde{\mathcal{A}}$ we have

$$\bigcup_{j=1}^{\infty} A_j \in \mathcal{A} \quad \text{and} \quad X \backslash \bigcup_{j=1}^{\infty} A_j = \bigcap_{j=1}^{\infty} X \backslash A_j \in \mathcal{A}.$$

Thus $\tilde{\mathcal{A}}$ is a σ-algebra which contains the open sets. It follows that $\tilde{\mathcal{A}}$ and hence also $\mathcal{A}$ are equal to the Borel algebra. For claim (2), still in the case $\mu(X) < \infty$, we argue that

$$\begin{aligned} \mu(X \backslash A) &= \inf\{\mu(U) : X \backslash A \subset U \text{ open}\} \\ &= \inf\{\mu(X \backslash C) : A \supset C \text{ closed}\} \\ &= \mu(X) - \sup\{\mu(C) : A \supset C \text{ closed}\}. \end{aligned}$$

For $\mu(X) = \infty$ statement (2) follows from the finite case by considering $\mu \llcorner A$, which is a finite Borel measure by Lemma 2.11. For (1) we consider $\mu \llcorner U_0$ where $U_0 \supset A$ is open with $\mu(U_0) < \infty$. □

Definition 2.14 (Radon measure)**.** *Let (X, d) be a locally compact, separable metric space. A Borel regular measure μ on X is called a Radon measure if*

$$\mu(K) < \infty \quad \textit{for all compact } K \subset X. \tag{2.7}$$

A metric space is locally compact if for any $x \in X$ there is an $r > 0$ such that $\overline{B_r(x)}$ is compact. For instance, a Banach space is locally compact if and only if it is finite-dimensional. A metric space is separable if it contains a countable dense subset. The space $\mathbb{R}^n$ has both properties, as does any submanifold $M \subset \mathbb{R}^n$ (Analysis 3).

Lemma 2.15 (σ-compactness)**.** *Let (X, d) be locally compact and separable. Then there exists an exhaustion $X = \bigcup_{i=1}^{\infty} U_i$, where the U_i are open with $\overline{U_i}$ compact.*

Proof. For any $x \in X$ we have by assumption

$$r(x) = \sup\{r > 0 : \overline{B_r(x)} \text{ is compact}\} > 0.$$

For any $r < r(x)$, the closed ball $\overline{B_r(x)}$ is contained in a compact set, hence these balls are all compact. We now show that

$$\liminf_{y\to x} r(y) \geq r(x). \tag{2.8}$$

For $r < r(x)$ take $\varrho \in (r, r(x))$. Then $\overline{B_r(y)} \subset \overline{B_\varrho(x)}$ for y sufficiently close to x. This shows $\liminf_{y\to x} r(y) \geq r$, and the claim follows by letting $r \nearrow r(x)$. Now if $r(x) = \infty$ for some $x \in X$, then we can take $U_i = B_i(x)$. Otherwise let x_i, $i \in \mathbb{N}$, be a dense subset, and put

$$U_i = B_{\varrho_i}(x_i) \quad \text{ where } \varrho_i = \frac{r(x_i)}{2}..$$

Using (2.8) one verifies that any $x \in X$ is contained in some U_i, and the lemma follows. □

Corollary 2.16. *Let μ be a Radon measure on (X, d). For $A \subset X$ one has*

(1) $\mu(A) = \inf\{\mu(U) : U$ *open*, $A \subset U\}$,

(2) $\mu(A) = \sup\{\mu(K) : K$ *compact*, $K \subset A\}$, *if A is μ-measurable.*

Proof. To prove (1) we may assume $\mu(A) < \infty$ and also A Borel, otherwise consider the Borel hull. Let $U_1 \subset\subset U_2 \subset\subset \ldots$ be the exhaustion from Lemma 2.15. By Theorem 2.13 there exist closed sets $C_i \subset U_i \backslash A$ such that

$$\mu\big(U_i\backslash(A \cap C_i)\big) = \mu\big((U_i\backslash A)\backslash C_i\big) < 2^{-i}\varepsilon.$$

Then $V_i = U_i\backslash C_i$ is open, contains $U_i \cap A$, and for $V = \bigcup_{i=1}^\infty V_i$ we have

$$\mu(V\backslash A) = \mu\Big(\bigcup_{i=1}^\infty V_i\backslash A\Big) \leq \sum_{i=1}^\infty \mu\big(U_i\backslash(A\cap C_i)\big) < \varepsilon.$$

Thus statement (1) is proved. For A μ-measurable we have $\mu(A) = \lim_{i\to\infty} \mu(A\cap U_i)$, therefore it suffices to prove (2) when A is relatively compact. Then $\mu \llcorner A$ is a Radon measure by Theorem 2.12. As proved, there exists $U \supset X\backslash A$ open with

$$\mu\llcorner A(U) < \mu\llcorner A(X\backslash A) + \varepsilon = \varepsilon.$$

Then $C := X\backslash U$ is closed with $C \subset A$, in particular C is compact, and

$$\mu(A) \leq \mu(A\backslash U) + \mu(A\cap U) < \mu(C) + \varepsilon.$$

The corollary is proved. □

At the end of the section we observe some properties of the measure pushforward.

Theorem 2.17 (Borel regularity of $f(\mu)$)**.** *Let X, Y be metric spaces and $f \in C^0(X, Y)$.*

(1) *If μ is a Borel measure, then $f(\mu)$ is also a Borel measure.*

(2) *Let X, Y be locally compact and separable, and assume that $f : X \to Y$ is proper. If μ is a Radon measure, then $f(\mu)$ is also a Radon measure.*

Proof. For (1) we note that the sets $B \subset Y$ for which $f^{-1}(B)$ is μ-measurable form a σ-algebra. The claim follows since $f^{-1}(U)$ is open for U open.

For (2) we have by assumption $f(\mu)(K) = \mu(f^{-1}(K)) < \infty$ for $K \subset Y$ compact. The construction of a Borel hull for given $T \subset Y$ is more tricky, however. Of course we may assume $f(\mu)(T) < \infty$, and it suffices to construct for any $\varepsilon > 0$ an open set $W \supset T$ such that $f(\mu)(W) \leq f(\mu)(T) + \varepsilon$. We first consider the case $T \subset V$ where V is open and relatively compact. Putting $S = f^{-1}(T)$ we choose an open set $U \supset S$ with $\mu(U) < \mu(S) + \varepsilon$. As f is proper, the set $f^{-1}(\overline{V} \backslash U)$ is compact, and hence $W := V \backslash f(f^{-1}\overline{V} \backslash U)$ is open. For any $y \in V$ we have the implication

$$\begin{aligned} y \notin W \quad &\Rightarrow \quad y \in f(f^{-1}\overline{V} \backslash U) \\ &\Rightarrow \quad y \in f(f^{-1}\overline{V} \backslash S) = f(f^{-1}(\overline{V} \backslash T)) \\ &\Rightarrow \quad y \in \overline{V} \backslash T. \end{aligned}$$

This shows $T \subset W$. By definition of W we have further

$$x \in f^{-1}\overline{V} \backslash U \quad \Rightarrow \quad f(x) \notin W, \quad \text{thus } f^{-1}(W) \subset U.$$

We conclude

$$f(\mu)(W) = \mu(f^{-1}W) \leq \mu(U) < \mu(S) + \varepsilon = f(\mu)(T) + \varepsilon.$$

For T arbitrary we choose an exhaustion $Y = \bigcup_{j=1}^{\infty} V_j$ by open sets V_j such that $V_j \subset\subset V_{j+1}$. Put $T_1 = T \cap V_1$ and $T_j = T \cap (V_j \backslash V_{j-1})$ for $j \geq 2$, and choose open sets $W_j \supset T_j$ with $f(\mu)(W_j) \leq f(\mu)(T_j) + 2^{-j}\varepsilon$. Then $W = \bigcup_{j=1}^{\infty} W_j$ contains T, and we have

$$\begin{aligned} f(\mu)(W) \quad &\leq \quad \sum_{j=1}^{\infty} f(\mu)(W_j) \\ &\leq \quad \sum_{j=1}^{\infty} \mu(f^{-1}(T_j)) + \varepsilon \\ &= \quad \sum_{j=1}^{\infty} (\mu \llcorner f^{-1}(T))(f^{-1}(V_j) \backslash f^{-1}(V_{j-1})) \; + \varepsilon \quad (\text{where } V_0 := \emptyset) \\ &= \quad (\mu \llcorner f^{-1}(T))(f^{-1}(Y)) \; + \varepsilon \\ &= \quad f(\mu)(T) + \varepsilon. \end{aligned}$$

We used that open sets are $\mu \llcorner f^{-1}(T)$-measurable.

□

Chapter 3

Hausdorff measure

Let (X,d) be a metric space and $s \in [0,\infty)$. To define the s-dimensional Hausdorff measure of a set $A \subset X$, we consider the set $\mathcal{C}_\delta(A)$ of all countable coverings

$$A \subset \bigcup_{i=1}^{\infty} C_i \quad \text{where } \operatorname{diam} C_i < \delta.$$

We then define an approximating measure at scale $\delta > 0$ by

$$\mathcal{H}^s_\delta(A) = \inf\Big\{ \sum_{i=1}^{\infty} \alpha(s)\Big(\frac{\operatorname{diam} C_i}{2}\Big)^s : (C_i)_{i\in\mathbb{N}} \in \mathcal{C}_\delta(A)\Big\}. \tag{3.1}$$

For $\delta_1 \leq \delta_2$ we have $\mathcal{C}_{\delta_1}(A) \subset \mathcal{C}_{\delta_2}(A)$ and hence $\mathcal{H}^s_{\delta_1}(A) \geq \mathcal{H}^s_{\delta_2}(A)$. We expect that for large $\delta > 0$ the number $\mathcal{H}^s_\delta(A)$ underestimates the true measure. The smaller $\delta > 0$ is chosen, the more the covering has to follow the fine structure of the set A, so that the number $\mathcal{H}^s_\delta(A)$ becomes more accurate. The number $\alpha(s)$ is a normalization constant, in the integer case $s = k \in \mathbb{N}_0$ we take

$$\alpha(k) = \mathcal{L}^k(\{x \in \mathbb{R}^k : |x| < 1\}) = \begin{cases} \frac{\pi^j}{j!} & \text{for } k = 2j, \\ \frac{\pi^j}{(j+\frac{1}{2})(j-\frac{1}{2})\cdot\ldots\cdot\frac{1}{2}} & \text{for } k = 2j+1. \end{cases} \tag{3.2}$$

With this choice, it follows that

$$\alpha(k)\Big(\frac{\operatorname{diam} C_i}{2}\Big)^k = \mathcal{L}^k(B_r) \quad \text{where } r = \frac{\operatorname{diam} C_i}{2}.$$

In particular if C_i is a ball then the number equals the volume of that ball. For $s \notin \mathbb{N}_0$ there is no natural normalization. To be consistent in the integer case we take

$$\alpha(s) = \frac{\pi^{s/2}}{\Gamma(\frac{s}{2}+1)} \quad \text{where } \Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\,dx.$$

To define $\mathcal{H}^s_\delta$ we could restrict to coverings by closed sets since $\operatorname{diam} C = \operatorname{diam} \overline{C}$. It is straightforward to check that the $\mathcal{H}^s_\delta$ are measures. However, not all Borel sets are $\mathcal{H}^s_\delta$ measurable. To give an example we use the following fact, to be proved in the exercises:

$$\operatorname{diam} A = \delta > 0 \quad \Rightarrow \quad \mathcal{H}^1_\delta(A) = \delta.$$

Taking $A = \{x \in \mathbb{R}^n : x_n \geq 0\}$ and $S = B_{\delta/2}(0)$ we obtain

$$\mathcal{H}^1_\delta(S \cap A) + \mathcal{H}^1_\delta(S \backslash A) = 2\delta > \delta = \mathcal{H}^1_\delta(S).$$

Definition 3.1 (Hausdorff measure)**.** *Let (X,d) be a metric space and $s \in [0,\infty)$. The s-dimensional Hausdorff measure of a set $A \subset X$ is defined by*

$$\mathcal{H}^s(A) = \sup_{\delta>0} \mathcal{H}^s_\delta(A) = \lim_{\delta\searrow 0} \mathcal{H}^s_\delta(A).$$

Lemma 3.2. *The Hausdorff measures $\mathcal{H}^s$ on (X,d) are Borel regular.*

Proof. We apply Caratheodory's criterion to show that Borel sets are measurable. Let $A_{1,2} \subset X$ with $\operatorname{dist}(A_1,A_2) = d > 0$. For $\delta < \frac{d}{2}$, let $\mathcal{C}$ be a covering of $A_1 \cup A_2$ by sets of diameter less than δ. Then the families $\mathcal{C}_i = \{C \in \mathcal{C} : C \cap A_i \neq \emptyset\}$ are disjoint, and hence we get

$$\begin{aligned}\sum_{C\in\mathcal{C}} \alpha(s)\Big(\frac{\operatorname{diam} C}{2}\Big)^s &\geq \sum_{C\in\mathcal{C}_1} \alpha(s)\Big(\frac{\operatorname{diam} C}{2}\Big)^s + \sum_{C\in\mathcal{C}_2} \alpha(s)\Big(\frac{\operatorname{diam} C}{2}\Big)^s \\ &\geq \mathcal{H}^s_\delta(A_1) + \mathcal{H}^s_\delta(A_2).\end{aligned}$$

Taking the infimum with respect to $\mathcal{C}$ yields

$$\mathcal{H}^s_\delta(A_1 \cup A_2) \geq \mathcal{H}^s_\delta(A_1) + \mathcal{H}^s_\delta(A_2).$$

Letting $\delta \searrow 0$ proves the criterion. Next let $S \subset X$ and $\delta_i \searrow 0$ be given. For each i we choose a covering C_{ij}, $j \in \mathbb{N}$, of S such that

$$\operatorname{diam} C_{ij} < \delta_i \quad \text{and} \quad \sum_{j=1}^\infty \alpha(s)\Big(\frac{\operatorname{diam} C_{ij}}{2}\Big)^s \leq \mathcal{H}^s_{\delta_i}(S) + \frac{1}{i}.$$

Without loss of generality the C_{ij} are closed, hence we obtain a Borel set by putting

$$B = \bigcap_{i=1}^\infty \Big(\bigcup_{j=1}^\infty C_{ij}\Big).$$

For fixed $\delta > 0$ we estimate, for i large such that $\delta_i \leq \delta$,

$$\mathcal{H}^s_\delta(B) \leq \mathcal{H}^s_\delta\Big(\bigcup_{j=1}^\infty C_{ij}\Big) \leq \sum_{j=1}^\infty \alpha(s)\Big(\frac{\operatorname{diam} C_{ij}}{2}\Big)^s \leq \mathcal{H}^s_{\delta_i}(S) + \frac{1}{i}.$$

Letting $i \to \infty$ yields $\mathcal{H}^s_\delta(B) \leq \mathcal{H}^s(S)$, and $\delta \searrow 0$ proves $\mathcal{H}^s(B) \leq \mathcal{H}^s(S)$ as desired. ☐

Lemma 3.3 (Transformation of $\mathcal{H}^s$ measure)**.** *Under a map $f : X \to Y$ between metric spaces the $\mathcal{H}^s$ measure of a set $A \subset X$ treansforms as follows:*

(1) *For f Lipschitz one has $\mathcal{H}^s(f(A)) \leq L^s\, \mathcal{H}^s(A)$ where $L = \operatorname{Lip}(f)$.*

(2) *For an isometry $f : X \to Y$ one has $\mathcal{H}^s(f(A)) = \mathcal{H}^s(A)$.*

(3) *For $A \subset \mathbb{R}^n$ and $\lambda > 0$ one has $\mathcal{H}^s(\lambda A) = \lambda^s \mathcal{H}^s(A)$.*

Proof. If $L = 0$ in claim (1) then f is constant and the statement is easily checked. Otherwise let $A \subset \bigcup_{i=1}^\infty C_i$ be a covering with $\operatorname{diam} C_i < \delta$. Then $f(A) \subset \bigcup_{i=1}^\infty f(C_i)$ and we have $\operatorname{diam} f(C_i) < L\delta$, hence

$$\mathcal{H}^s_{L\delta}(f(A)) \leq \sum_{i=1}^\infty \alpha(s)\Big(\frac{\operatorname{diam} f(C_i)}{2}\Big)^s \leq L^s \sum_{i=1}^\infty \alpha(s)\Big(\frac{\operatorname{diam} C_i}{2}\Big)^s.$$

Taking the infimum shows $\mathcal{H}^s_{L\delta}(f(A)) \le L^s \mathcal{H}^s_\delta(A)$, and (1) follows by letting $\delta \searrow 0$.

In claim (2) both f and f^{-1} have Lischitz constant one, hence

$$\mathcal{H}^s(A) = \mathcal{H}^s\big(f^{-1}(f(A))\big) \le \mathcal{H}^s(f(A)) \le \mathcal{H}^s(A).$$

Similarly in (3) the map $\tau_\lambda(x) = \lambda x$ has Lipschitz λ, this implies

$$\lambda^s\, \mathcal{H}^s(A) = \lambda^s\, \mathcal{H}^s\Big(\frac{1}{\lambda}\lambda A\Big) \le \mathcal{H}^s(\lambda A) \le \lambda^s\, \mathcal{H}^s(A).$$

□

The measures $\mathcal{H}^s$ and $\mathcal{H}^t$ compare as follows: let $A \subset \bigcup_{i=1}^\infty C_i$ where $\operatorname{diam} C_i < \delta$. Then we have, assuming $s < t$,

$$\mathcal{H}^t_\delta(A) \le \sum_{i=1}^\infty \alpha(t)\Big(\frac{\operatorname{diam} C_i}{2}\Big)^t \le \frac{\alpha(t)}{\alpha(s)}\Big(\frac{\delta}{2}\Big)^{t-s} \sum_{i=1}^\infty \alpha(s)\Big(\frac{\operatorname{diam} C_i}{2}\Big)^s.$$

Taking the infimum with respect to these coverings, we get

$$\mathcal{H}^t_\delta(A) \le \frac{\alpha(t)}{\alpha(s)}\Big(\frac{\delta}{2}\Big)^{t-s} \mathcal{H}^s_\delta(A).$$

In particular we have

$$\begin{aligned} \mathcal{H}^s(A) < \infty \quad &\Rightarrow \quad \mathcal{H}^t(A) = 0 \quad \text{ for } t > s,\\ \mathcal{H}^t(A) > 0 \quad &\Rightarrow \quad \mathcal{H}^s(A) = \infty \quad \text{ for } s < t. \end{aligned}$$

Definition 3.4 (Hausdorff dimension)**.** *The Hausdorff dimension of a set $A \subset (X,d)$ is*

$$\dim_{\mathcal{H}}(A) = \inf\{s \ge 0 : \mathcal{H}^s(A) = 0\}.$$

Example 3.5. *The Cantor set C is the set of all $s \in [0,1]$ having a triadic expansion*

$$s = 0, s_1 s_2 \ldots = \sum_{j=1}^\infty s_j 3^{-j} \qquad \textit{where } s_j \in \{0,2\}.$$

This representation is unique: if $s,t \in C$ satisfy $s_j = t_j$ for $j \le N-1$ but $s_N \ne t_N$ then

$$|s-t| \ge 2\cdot 3^{-N} - \sum_{j=N+1}^\infty 2\cdot 3^{-j} = 2\cdot 3^{-N} - 3^{-N} = 3^{-N} > 0.$$

The estimate also implies that C is complete, hence also closed. Now C is the disjoint union

$$C = C_0 \cup C_2 \qquad \textit{where } C_0 = \{s \in C : s_1 = 0\},\ C_2 = \{s \in C : s_1 = 2\}.$$

Note that $C_0 = \frac{1}{3}C$ and $C_2 = \frac{2}{3} + \frac{1}{3}C$. Therefore Lemma 3.3 implies for any $d \in [0,\infty)$

$$\Big(\frac{1}{3}\Big)^d \mathcal{H}^d(C) = \mathcal{H}^d\Big(\frac{1}{3}C\Big) = \frac{1}{2}\big(\mathcal{H}^d(C_0) + \mathcal{H}^d(C_2)\big) = \frac{1}{2}\mathcal{H}^d(C).$$

Now if d can be chosen with $0 < \mathcal{H}^d(C) < \infty$, then

$$\Big(\frac{1}{3}\Big)^d = \frac{1}{2}, \qquad \textit{that is} \qquad d = \log_3 2 = \frac{\log 2}{\log 3}.$$

The calculation suggests that this number d is the Hausdorff dimension of C.

We now come to the comparison of the measures $\mathcal{H}^n$ and $\mathcal{L}^n$ on $\mathbb{R}^n$. The Lebesgue measure has the following wellknown uniqueness property (see Analysis 3).

Theorem 3.6 (axiomatic characterization of $\mathcal{L}^n$)**.** *If a Radon measure μ on $\mathbb{R}^n$ is translation invariant, then it has the form $\mu = c\,\mathcal{L}^n$ for a constant $c \geq 0$.*

Lemma 3.7. *On $\mathbb{R}^n$ we have $\mathcal{H}^n = c\,\mathcal{L}^n$ where $2^{-n} \leq c \leq 2^{-n}\alpha(n)n^{n/2}$.*

Proof. $\mathcal{H}^n$ is Borel regular by Lemma 3.2. Dividing $[0,1]^n$ into k^n congruent subcubes of diameter $\sqrt{n}/k$, we infer for k large

$$\mathcal{H}^n_\delta([0,1]^n) \leq k^n \alpha(n)\Big(\frac{\sqrt{n}}{2k}\Big)^n = 2^{-n}\alpha(n)n^{n/2}.$$

Letting $\delta \searrow 0$ proves the upper bound, in particular $\mathcal{H}^n$ is a Radon measure. Now consider an arbitrary covering $[0,1]^n \subset \bigcup_{i=1}^\infty C_i$ with $\operatorname{diam} C_i < \delta$. Then

$$1 = \mathcal{L}^n([0,1]^n) \leq \sum_{i=1}^\infty \mathcal{L}^n(C_i) \leq \sum_{i=1}^\infty \alpha(n)\,(\operatorname{diam} C_i)^n = 2^n \sum_{i=1}^\infty \alpha(n)\,\Big(\frac{\operatorname{diam} C_i}{2}\Big)^n.$$

Taking the infimum over all these coverings we conclude

$$2^{-n} \leq \mathcal{H}^n([0,1]^n) \leq 2^{-n}\alpha(n)n^{n/2}.$$

Ths claim follows by Theorem 3.6. $\square$

The equality $\mathcal{H}^n = \mathcal{L}^n$ is more involved. We need the following two facts.

Lemma 3.8. *For open $U \subset \mathbb{R}^n$ and $\delta > 0$ there exist pairwise disjoint, closed balls $B_j \subset U$ with $\operatorname{diam} B_j < \delta$, such that*

$$\mathcal{L}^n\Big(U \backslash \bigcup_{j=1}^\infty B_j\Big) = 0.$$

This will be proved in Chapter 3 using a covering theorem.

Theorem 3.9 (isodiametric inequality)**.**

$$\mathcal{L}^n(A) \leq \alpha(n)\Big(\frac{\operatorname{diam} A}{2}\Big)^n \quad \text{for any } A \subset \mathbb{R}^n.$$

Example 3.10. *The inequality is nontrivial in the sense that not every set A is contained in a ball of radius $\operatorname{diam} A/2$. A standard example is the equilateral triangle: the diameter is equal to the sidelength, and one half the sidelength is shorter than the radius of the perimeter by a factor $\sqrt{3}/2 < 1$. We note however that if A is symmetric with respect to a point p, then it is contained in a ball of radius $\operatorname{diam} A/2$.*

Theorem 3.11. *We have $\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$.*

Proof. For the upper bound choose an exhaustion of $U = (0,1)^n$ by balls B_j with $\operatorname{diam} B_j < \delta$ as in Lemma 3.8. Using that $\mathcal{H}^n_\delta \leq \mathcal{H}^n \leq C\,\mathcal{L}^n$ we obtain

$$\mathcal{H}^n_\delta([0,1]^n) = \mathcal{H}^n_\delta\Big(\bigcup_{j=1}^\infty B_j\Big) \leq \sum_{j=1}^\infty \alpha(n)\Big(\frac{\operatorname{diam} B_j}{2}\Big)^n \leq 1.$$

On the other hand if $[0,1]^n \subset \bigcup_{i=1}^\infty C_i$ then by Theorem 3.9

$$1 \le \sum_{j=1}^\infty \mathcal{L}^n(C_i) \le \sum_{i=1}^\infty \alpha(n)\Big(\frac{\operatorname{diam} C_i}{2}\Big)^n.$$

This yields $\mathcal{H}^n([0,1]^n) \ge 1$. □

To prove the isodiametric inequality we employ a symmetrization. For $v \in \mathbb{S}^{n-1}$ we consider the hyperplane $P_v = \{x \in \mathbb{R}^n : \langle x, v\rangle = 0\}$ and define

$$A_{p,v} = A \cap (p + \mathbb{R}v) \quad \text{ where } p \in P_v.$$

Then the Steiner symmetrization with respect to P_v is defined by

$$S_v(A) = \bigcup_{p\in P_v,\, A_{p,v}\neq\emptyset} \Big\{p + tv : |t| \le \frac{1}{2}\mathcal{L}^1(A_{p,v})\Big\}. \tag{3.3}$$

Lemma 3.12. *For $v \in \mathbb{S}^{n-1}$ and $A \subset \mathbb{R}^n$ we have*

(1) $\operatorname{diam} S_v(A) \le \operatorname{diam} A$,

(2) *If A is $\mathcal{L}^n$-measurable, then so is $S_v(A)$ and $\mathcal{L}^n(S_v(A)) = \mathcal{L}^n(A)$.*

Proof. We may assume $\operatorname{diam} A < \infty$. For $p_i \in P_v$ for $i = 1, 2$ we put

$$a_i = \inf\{t \in \mathbb{R} : p_i + tv \in A\} \quad \text{ and } \quad b_i = \sup\{t \in \mathbb{R} : p_i + tv \in A\}.$$

The points $p_i + a_i v$ and $p_i + b_i v$ are in $\overline{A}$, hence

$$\begin{array}{rcl} \operatorname{diam} A & \ge & |(p_2 + b_2 v) - (p_1 + a_1 v)|,\, |(p_1 + b_1 v) - (p_2 + a_2 v)| \\ & = & \sqrt{|p_2 - p_1|^2 + (b_2 - a_1)^2},\, \sqrt{|p_1 - p_2|^2 + (b_1 - a_2)^2}. \end{array}$$

Now for $p_i + t_i v \in S_v(A)$ we have $|(p_1 + t_1 v) - (p_2 + t_2 v)| = \sqrt{|p_1 - p_2|^2 + |t_1 - t_2|^2}$ where

$$|t_1 - t_2| \le |t_1| + |t_2| \le \frac{1}{2}(b_1 - a_1 + b_2 - a_2) \le \max(|b_2 - a_1|, |b_1 - a_2|).$$

This proves claim (1). By the Cavalieri-Fubini theorem, the sets $A_{p,v}$ are $\mathcal{L}^1$-measurable for almost every $p \in P_v$, the measure $\mathcal{L}^1(A_{p,v})$ is a measurable function on P_v, and

$$\mathcal{L}^n(A) = \int_{P_v} \mathcal{L}^1(A_{p,v})\, d\mathcal{L}^{n-1}(p).$$

Altogether this implies claim (2). □

Proof. (*of the isodiametric inequality*) We assume $\operatorname{diam} A < \infty$ and A closed. Put $A_0 = A$ and define inductively the sets $A_j = S_{e_j}(A_{j-1})$ for $j = 1, \dots, n$. We claim that, for σ_j the reflection at P_{e_j},

$$\sigma_j(A_k) = A_k \quad \text{ for } j = 1, \dots, k. \tag{3.4}$$

By definition of the Steiner symmetrization we have $\sigma_k(A_k) = A_k$, in particular (3.4) holds for $k = 1$. Now let $k \geq 2$ and $1 \leq j \leq k-1$. For $p \in P_{e_k}$ we obtain by induction

$$\begin{aligned}
(A_{k-1})_{\sigma_j(p),e_k} &= A_{k-1} \cap (\sigma_j(p) + \mathbb{R}e_k) \\
&= \sigma_j\big(A_{k-1} \cap (p + \mathbb{R}e_k)\big) \quad \text{(using } \sigma_j(A_{k-1}) = A_{k-1},\ \sigma_j(e_k) = e_k\text{)} \\
&= \sigma_j((A_{k-1})_{p,e_k}).
\end{aligned}$$

In particular $\mathcal{L}^1\big((A_{k-1})_{\sigma_j(p),e_k}\big) = \mathcal{L}^1\big((A_{k-1})_{p,e_k}\big)$. This implies $\sigma_j(A_k) = A_k$ by definition of the Steiner symmetrization. Thus A_n is symmetric with respect to all coordinate hyperplanes, and hence symmetric about the origin. This implies $\operatorname{diam} A_n \geq 2|x|$ for any $x \in A_n$, and we conclude recalling Lemma 3.12

$$\mathcal{L}^n(A) = \mathcal{L}^n(A_n) \leq \alpha(n)\Big(\frac{\operatorname{diam} A_n}{2}\Big)^n \leq \alpha(n)\Big(\frac{\operatorname{diam} A}{2}\Big)^n.$$

$\square$

Chapter 4

Covering theorems

In many arguments in GMT one has a family of balls $B_\varrho(x)$ covering a set A, with certain information on each ball. Typically the radius $\varrho = \varrho(x) > 0$ is small depending on x. In such a situation a covering theorem allows to deduce global information. We discuss two such theorems due to Vitali and Besicovitch. The first produces a disjoint subfamily such that the balls enlarged by a factor 5 are still a covering. This result is particularly useful if we have a measure for which the enlarged balls can be estimated, e.g. the Lebesgue measure. The Besicovitch theorem produces a covering out of the given balls. Of course then the subfamily cannot be disjoint, however the theorem asserts that the overlapping is estimated.

Vitali's theorem applies in any metric space (X, d). For a given closed ball it considers the concentric ball with 5 times the radius, i.e.

$$\hat{B} = \{x \in X : d(x, x_0) \leq 5\varrho\} \quad \text{for } B = \{x \in X : d(x, x_0) \leq \varrho\}. \tag{4.1}$$

Theorem 4.1 (Vitali). *Let $\mathcal{F}$ be a family of closed balls $B \subset X$ with positive diameter and*

$$D := \sup_{B \in \mathcal{F}} \operatorname{diam} B < \infty. \tag{4.2}$$

Then there exists a pairwise disjoint subfamily $\mathcal{G}$ such that

$$\bigcup_{B \in \mathcal{F}} B \subset \bigcup_{B' \in \mathcal{G}} \hat{B}'.$$

In fact, for any $B \in \mathcal{F}$ there exists a $B' \in \mathcal{G}$ such that $B \cap B' \neq \emptyset$ and $B \subset \hat{B}'$.

Proof. We start by ordering the balls by their size. For $k = 1, 2, \ldots$ we put

$$\mathcal{F}_k = \{B \in \mathcal{F} : 2^{-k} D < \operatorname{diam} B \leq 2^{1-k} D\}.$$

Now construct $\mathcal{G}_k \subset \mathcal{F}_k$ inductively as follows:

- let $\mathcal{G}_1$ be a maximal pairwise disjoint subfamily of $\mathcal{F}_1$.
- let $\mathcal{G}_k$, $k \geq 2$, be a maximal pairwise disjoint subfamily of $\mathcal{F}_k'$, by which we mean the set of all $B \in \mathcal{F}_k$ which do not intersect any $B' \in \mathcal{G}_j$ for $j = 1, \ldots, k-1$.

The maximality of $\mathcal{G}_k$ means the following: let B be a ball in $\mathcal{F}_k$. Then either it intersects some ball $B' \in \mathcal{G}_j$ for $j \le k-1$ (for $k=1$ this alternative does not apply), or it intersects a ball $B' \in \mathcal{G}_k$. In fact otherwise B could be added to $\mathcal{G}_k$ contradicting the maximality. For the radii of the balls we have

$$\operatorname{rad} B \le 2^{-k} D \le 2 \operatorname{rad} B'.$$

Thus if x_0 is the center of B' and $x \in B \cap B'$, then for any $y \in B$ we have

$$d(y, x_0) \le d(y,x) + d(x, x_0) \le 4 \operatorname{rad} B' + \operatorname{rad} B' = 5 \operatorname{rad} B'.$$

□

The choice of a maximal subfamily is less obvious as it seems. In $\mathbb{R}^n$ we can argue as follows. Assume that $\mathcal{F}$ is a family of closed balls with lower radius bound $\varrho > 0$. Let $\mathcal{F}_R$ be the subfamily of balls contained in $B_R(0)$. If $B_1, \ldots, B_N \in \mathcal{F}_R$ are pairwise disjoint, then

$$N\alpha(n)\varrho^n \le \sum_{i=1}^N \mathcal{L}^n(B_i) \le \alpha(n) R^n, \quad \text{hence } N \le \Big(\frac{R}{\varrho}\Big)^n.$$

There exists a pairwise disjoint subfamily of $\mathcal{F}_R$ with maximal number of elements. Now choose $R_i = 1, 2, \ldots$ and proceed inductively. In a general metric space (X,d) such an argument is not available. For the existence of a maximal subfamily one then needs Zorn's lemma. We omit the details.

Definition 4.2. *A family $\mathcal{F}$ of sets in (X,d) is a fine covering of A if*

$$\inf\{\operatorname{diam} B : B \in \mathcal{F},\, x \in B\} = 0 \quad \textit{for all } x \in A.$$

Corollary 4.3. *Let $\mathcal{F}$ be a family of closed balls in (X,d) with positive, uniformly bounded diameter. Assume that $\mathcal{F}$ is a fine covering of A. Then there exists a pairwise disjoint subfamily $\mathcal{G}$ with the following property: for any finite collection $B_1, \ldots, B_N$ in $\mathcal{F}$ one has*

$$A \backslash \bigcup_{i=1}^N B_i \quad \subset \bigcup_{B' \in \mathcal{G} \backslash \{B_1, \ldots, B_N\}} \hat{B}'.$$

Proof. Let $\mathcal{G}$ be as in Theorem 4.1, and let $x \in A \backslash \bigcup_{i=1}^N B_i$ be given. As $X \backslash \bigcup_{i=1}^N B_i$ is open and $\mathcal{F}$ is a fine covering, there exists a ball $B \in \mathcal{F}$ with $x \in B$ and $B \cap B_i = \emptyset$ for $i = 1, \ldots, N$. Now by Theorem 4.1 there is a $B' \in \mathcal{G}$ with $B \cap B' \neq \emptyset$ and $B \subset \hat{B}'$. We conclude $x \in \hat{B}'$ and $B' \neq B_i$ for $i = 1, \ldots, N$, thus

$$x \in \bigcup_{B' \in \mathcal{G} \backslash \{B_1, \ldots, B_N\}} \hat{B}'.$$

□

Theorem 4.4. *Let $U \subset \mathbb{R}^n$ be open and $\delta > 0$. There exist pairwise disjoint closed balls $B_j \subset U$ with $\operatorname{diam} B_j < \delta$ such that*

$$\mathcal{L}^n\Big(U \backslash \bigcup_{j=1}^\infty B_j\Big) = 0.$$

Proof. We may assume that U is bounded, otherwise we apply the result to the disjoint open sets $U_k = \{x \in U : k < |x| < k+1\}$. Let $\mathcal{F}_1$ be the set of closed balls $B \subset U$ with $\operatorname{diam} B < \delta$. By Vitali there is a pairwise disjoint subfamily $\mathcal{G}_1 \subset \mathcal{F}_1$ with

$$U \subset \bigcup_{B \in \mathcal{G}_1} \hat{B}.$$

Thus we can estimate

$$\begin{aligned}
\mathcal{L}^n\Big(U \backslash \bigcup_{B \in \mathcal{G}_1} B\Big) &= \mathcal{L}^n(U) - \mathcal{L}^n\Big(\bigcup_{B \in \mathcal{G}_1} B\Big) \\
&= \mathcal{L}^n(U) - \sum_{B \in \mathcal{G}_1} \mathcal{L}^n(B) \\
&= \mathcal{L}^n(U) - 5^{-n} \sum_{B \in \mathcal{G}_1} \mathcal{L}^n(\hat{B}) \\
&\leq (1 - 5^{-n})\, \mathcal{L}^n(U).
\end{aligned}$$

Put $\theta = 1 - \frac{1}{2}5^{-n}$. Then there exists a finite subfamily $\mathcal{G}_1' \subset \mathcal{G}_1$ such that

$$\mathcal{L}^n\Big(U \backslash \bigcup_{B \in \mathcal{G}_1'} B\Big) \leq \theta \mathcal{L}^n(U).$$

Now $U \backslash \bigcup_{B \in \mathcal{G}_1'} B$ is open. Iterating the argument we obtain a decreasing sequence

$$U = U_0 \supset U_1 \supset \ldots, \qquad \text{where } U_k = U_{k-1} \backslash \bigcup_{B \in \mathcal{G}_k'} B.$$

Here $\mathcal{G}_k'$ is a finite, pairwise disjoint collection of closed balls $B \subset U_{k-1}$ with $\operatorname{diam} B < \delta$ and

$$\mathcal{L}^n(U_k) \leq \theta \mathcal{L}^n(U_{k-1}).$$

The family $\mathcal{G}' = \bigcup_{k=1}^{\infty} \mathcal{G}_k'$ is pairwise disjoint, and we have

$$\mathcal{L}^n\Big(U \backslash \bigcup_{B \in \mathcal{G}'} B\Big) = \lim_{k \to \infty} \mathcal{L}^n(U_k) = 0.$$

□

Theorem 4.5 (Besicovitch)**.** *Let $\mathcal{F}$ be a family of closed balls $B_\varrho(a)$ in $\mathbb{R}^n$ with $\varrho > 0$, such that $\varrho^* = \sup\{\varrho : B_\varrho(a) \in \mathcal{F}\} < \infty$. Let A be the set of centers of the balls in $\mathcal{F}$. There exist subfamilies $\mathcal{F}_1, \ldots, \mathcal{F}_N$, each pairwise disjoint, such that*

$$A \subset \bigcup_{j=1}^{N} \bigcup_{B \in \mathcal{F}_j} B \quad \textit{and} \quad N \leq C(n).$$

The theorem asserts the existence of a covering with overlap multiplicity bounded by $C(n)$.

Proof. We first prove the theorem for A bounded. Choose inductively balls $B_1, B_2, \ldots$ in $\mathcal{F}$ as follows: if $B_1, \ldots, B_{j-1}$ are already determined, let

$$\begin{aligned} A_j &= A \setminus \bigcup_{i=1}^{j-1} B_i, \\ \varrho_j^* &= \sup\{\varrho : B_\varrho(a) \in \mathcal{F},\, a \in A_j\}. \end{aligned}$$

Now choose $B_j = B_{\varrho_j}(a_j)$ where $a_j \in A_j$ and $\varrho_j \geq \frac{3}{4}\varrho_j^*$. The start of the process is for $j = 1$ where $A_1 = A$ and $\varrho_1^* = \varrho^*$. By construction we have

$$|a_j - a_i| > \varrho_i \quad \text{for } j > i, \tag{4.3}$$

$$\varrho_j \leq \frac{4}{3}\varrho_i \quad \text{for } j > i. \tag{4.4}$$

For the last inequality we note $A_j \subset A_i$ for $j > i$, this yields

$$\varrho_j \leq \varrho_j^* \leq \varrho_i^* \leq \frac{4}{3}\varrho_i.$$

Combining (4.3) and (4.4) we get for $j > i$

$$|a_j - a_i| > \varrho_i \geq \frac{1}{3}\varrho_i + \frac{2}{3} \cdot \frac{3}{4}\varrho_j \geq \frac{1}{3}\varrho_i + \frac{1}{3}\varrho_j.$$

Thus we have

$$B_{\frac{\varrho_i}{3}}(a_i) \cap B_{\frac{\varrho_j}{3}}(a_j) = \emptyset \quad \text{for } i \neq j. \tag{4.5}$$

We claim that the constructed family of balls covers A. If the process stops at some j, then $A \setminus \bigcup_{i=1}^{j} B_i$ is empty and the claim follows. Otherwise we use that A is bounded and $\varrho^* < \infty$, thus all balls are contained in a fixed large ball. Then (4.5) implies $\varrho_j \to 0$ as $j \to \infty$. Now for each $a \in A$ there is some ball $B_r(a) \in \mathcal{F}$. For j sufficiently large we have

$$r > \frac{4}{3}\varrho_j \geq \varrho_j^* = \sup\{\varrho : B_\varrho(a') \in \mathcal{F}, a' \in A_j\}.$$

We conclude $a \notin A_j$, in other words $a \in B_i$ for some $i < j$.

Now we divide the set of balls into subfamilies. We start by letting $\mathcal{F}_1 = \{B_1\}$. If B_2 is disjoint from B_1, then we add it to $\mathcal{F}_1$. Otherwise we make up a new subfamily $\mathcal{F}_2 = \{B_2\}$. Proceeding by induction, each ball is added to the subfamily with smallest number which keeps disjoint. If there is no subfamily with this property, then we create a new one. Now we claim that there is a constant C_n such that for any $k \in \mathbb{N}$ we have

$$\# J_k \leq C_n \quad \text{where } J_k = \{j < k : B_j \cap B_k \neq \emptyset\}. \tag{4.6}$$

This implies that we need at most $N = C_n + 1$ subfamilies in the above process, and the theorem is proved. We note that the definition of J_k in (4.6) means that

$$|a_j - a_k| \leq \varrho_j + \varrho_k \quad \text{for all } j \in J_k. \tag{4.7}$$

To prove (4.6) we distinguish between balls B_j which are comparable in size to B_k, and other balls which are bigger.

Claim 1. For $J_k' = \{j \in J_k : \varrho_j \leq 3\varrho_k\}$ we have $\# J_k' \leq 20^n$.

We use a packing argument. For $j \in J_k'$ we have $\frac{1}{3}\varrho_j \leq \varrho_k$ and $|a_j - a_k| \leq \varrho_j + \varrho_k \leq 4\varrho_k$. Now we estimate using (4.4) and (4.5)

$$(\# J_k')\,\alpha(n)\Big(\frac{\varrho_k}{4}\Big)^n \leq \sum_{j \in J_k'} \alpha(n)\Big(\frac{\varrho_j}{3}\Big)^n = \mathcal{L}^n\Big(\bigcup_{j \in J_k'} B_{\frac{\varrho_j}{3}}(a_j)\Big) \leq \mathcal{L}^n\big(B_{5\varrho_k}(a_k)\big) = \alpha(n)(5\varrho_k)^n.$$

The claim follows.

Claim 2. Let $J_k'' = \{j \in J_k : \varrho_j > 3\varrho_k\}$. Then for $i, j \in J_k''$, $i \neq j$, we have

$$\sphericalangle(a_i - a_k, a_j - a_k) \geq \delta \quad \text{ for } \delta > 0 \text{ universal.} \tag{4.8}$$

To deduce from (4.8) the estimate for $\# J_k''$ we again use a packing argument, but now on $\mathbb{S}^{n-1}$. For $\omega_j = \frac{a_j - a_k}{|a_j - a_k|}$ we have $|\omega_i - \omega_j| \geq 2d$ where $d = \sin\frac{\delta}{2}$. Let e_n be the north pole, then

$$(\# J_k'')\,\mathcal{H}^{n-1}\big(B_d(e_n) \cap \mathbb{S}^{n-1}\big) = \sum_{j \in J_k''} \mathcal{H}^{n-1}\big(B_d(\omega_j) \cap \mathbb{S}^{n-1}\big) \leq \mathcal{H}^{n-1}(\mathbb{S}^{n-1}).$$

It remains to prove (4.8). Let $i, j \in J_k''$ with $i < j$, and choose coordinates $x = (y, z) \in \mathbb{R}^{n-1} \times \mathbb{R}$ such that $a_k = 0$ and $a_i = |a_i| e_n$. We must find a cone around the z-axis which does not contain a_j, with a universal angle. Consider a point $a \in \mathbb{R}^n$ with $\langle a, e_n\rangle > 2(1-\varepsilon)\varrho_i$, for small $\varepsilon > 0$ to be determined. Using (4.4) and $j \in J_k''$ we estimate

$$|a| > \frac{3}{2}(1-\varepsilon)\varrho_j = \varrho_j + \frac{1}{2}(1-3\varepsilon)\varrho_j \geq \varrho_j + \frac{3}{2}(1-3\varepsilon)\varrho_k = \varrho_j + \varrho_k + \frac{1}{2}(1-9\varepsilon)\varrho_k.$$

Choosing $\varepsilon = \frac{1}{9}$ we obtain, using $|a_j| = |a_j - a_k| \leq \varrho_j + \varrho_k$ by (4.7),

$$\langle a_j, e_n\rangle \leq 2\big(1 - \frac{1}{9}\big)\varrho_i = \frac{16}{9}\varrho_i. \tag{4.9}$$

The plane $\{z = \frac{16}{9}\varrho_i\}$ intersects $B_{\varrho_i}(a_i)$ in an $(n-1)$-disk. Using (4.3), (4.7) for $i \in J_k''$, and (4.4) in the case when $a_k = 0$, we see that

$$\varrho_i \leq |a_i| \leq \varrho_i + \varrho_k \leq \frac{4}{3}\varrho_i,$$

We now estimate the radius R of that disk by

$$R^2 = \varrho_i^2 - \Big(\frac{16}{9}\varrho_i - |a_i|\Big)^2 \geq \varrho_i^2 - \Big(\frac{16}{9}\varrho_i - \varrho_i\Big)^2 = \frac{32}{81}\varrho_i^2.$$

Let C be the cone over that disk, with tip $a_k = 0$. For the angle θ of C we have

$$\tan\theta = \frac{R}{\frac{16}{9}\varrho_i} \geq \frac{1}{\sqrt{8}}. \tag{4.10}$$

We next deal with points close to $a_k = 0$. The spheres $\partial B_{3\varrho_k}(0)$ and $\partial B_{\varrho_i}(a_i)$ intersect in an $(n-2)$-sphere, which is contained in a horizontal plane $\{z = h\}$. We let C' be the cone over the corresponding $(n-1)$-disk. To estimate the angle θ' of C', we apply the cosine law in the triangle with corners $a_k = 0$, a_i and a point on the $(n-2)$-sphere. This yields

$$\cos\theta' = \frac{(3\varrho_k)^2 + |a_i|^2 - \varrho_i^2}{2(3\varrho_k)|a_i|} = \frac{3\varrho_k}{2|a_i|} + \frac{(|a_i| - \varrho_i)(|a_i| + \varrho_i)}{6\varrho_k|a_i|}.$$

Now $\varrho_i \le |a_i| \le \varrho_i + \varrho_k$ by (4.3) and by the definition of J_k in (4.6). The definition of J_k'' further implies $3\varrho_k \le \varrho_i \le |a_i|$. Thus

$$\cos\theta' \le \frac{1}{2} + \frac{1}{3} = \frac{5}{6}.$$

Now $\theta < \theta'$, in fact $\theta = \arctan\frac{1}{\sqrt{8}} \approx 19,47°$ and $\theta' = \arccos\frac{5}{6} \approx 33,56°$. We conclude

$$a_j \notin \{(x,z) \in C : z < h\}. \tag{4.11}$$

Namely, otherwise we had $a_j \in B_{3\varrho_k}(0)$. But $|a_j| = |a_j - a_k| > \varrho_j > 3\varrho_k$ by (4.3) and definition of J_k'', a contradiction. Finally, the part of C with $h \le z \le \frac{16}{9}\varrho_i$ is the convex hull of the two horizontal disks, and is therefore contained in $B_{\varrho_i}(a_i)$. But again by (4.3) we know that $|a_j - a_i| > \varrho_i$, and we conclude $a_j \notin C$. This proves (4.8) with $\delta = \arctan\frac{1}{\sqrt{8}}$. The theorem is proved in the case when A is bounded.

For A unbounded we apply the result to the families

$$\mathcal{F}_k = \{B_\varrho(a) \in \mathcal{F} : 3(k-1)\varrho^* \le |a| < 3k\varrho^*\} \quad \text{where } k \in \mathbb{N}.$$

We obtain subfamilies $\mathcal{F}_{k,1}, \ldots, \mathcal{F}_{k,N}$, each pairwise disjoint, which cover the set of centers of $\mathcal{F}_k$; here $N = N(n)$ is the Besicovitch constant. Now put for $i = 1, \ldots, N$

$$\mathcal{F}_i^0 = \bigcup_{k \text{ even}} \mathcal{F}_{k,i} \quad \text{and} \quad \mathcal{F}_i^1 = \bigcup_{k \text{ odd}} \mathcal{F}_{k,i}.$$

These families are pairwise disjoint, since for $B_\varrho(a) \in \mathcal{F}_k$ and $B_{\varrho'}(a') \in \mathcal{F}_l$ with $|k-l| \ge 2$

$$|a - a'| \ge 3\varrho^* > \varrho + \varrho'.$$

Therefore the theorem holds with constant $2N$. □

Theorem 4.6. *Let μ be a Borel measure on $\mathbb{R}^n$, and let $\mathcal{F}$ be a family of closed balls with positive radius and centers in a set A. Assume that $\mu(A) < \infty$ and*

$$\inf\{\varrho > 0 : B_\varrho(a) \in \mathcal{F}\} = 0 \quad \textit{for each } a \in A.$$

Then for any open U there exists a disjoint subfamily $\mathcal{F}'$ with $B \subset U$ for all $B \in \mathcal{F}'$, such that

$$(\mu \llcorner A)\Big(U \backslash \bigcup_{B \in \mathcal{F}'} B\Big) = 0.$$

Proof. We may assume $\operatorname{diam} B \le 1$ for all $B \in \mathcal{F}$. Putting $U = U_1$, we determine for $j \ge 2$ open sets $U_j = U_{j-1} \backslash \bigcup_{k=1}^{k_j} B_{j,k}$ where $B_{j,k} \in \mathcal{F}$, with the following properties:

- $B_{j,1}, \ldots, B_{j,k_j}$ are contained in U_{j-1} and are pairwise disjoint,

- $(\mu\llcorner A)(U_j) \leq \theta\,(\mu\llcorner A)(U_{j-1})$ where $\theta \in (0,1)$ is a constant.

We then conclude as in Theorem 4.4 that $\mathcal{F}' = \{B_{j,k} : j \geq 2,\, 1 \leq k \leq k_j\}$ has the desired properties. To define U_j we let $\mathcal{F}_j = \{B \in \mathcal{F} : B \subset U_{j-1}\}$, and note

$$\inf\{\varrho > 0 : B_\varrho(a) \in \mathcal{F}_j\} = 0 \quad \text{for each } a \in A \cap U_{j-1}.$$

By Besicovitch there exists a pairwise disjoint subfamily $\mathcal{F}_j'$ such that

$$(\mu\llcorner A)\Big(\bigcup_{B\in\mathcal{F}_j'} B\Big) \geq \frac{1}{N}\,(\mu\llcorner A)(U_{j-1}) \quad \text{where } N = N(n).$$

As the balls $B \in \mathcal{F}_j'$ are disjoint and have positive radius, the family $\mathcal{F}_j'$ is countable. For instance, taking a point with rational coordinates in each B yields an injective map to $\mathbb{Q}^n$. The balls are $\mu\llcorner A$-measurable, see Lemma 2.11. By continuity of the measure, Theorem 2.6, we can choose a finite set $B_{j,1}, \ldots, B_{j,k_j}$ in $\mathcal{F}_j'$ such that

$$(\mu\llcorner A)\Big(\bigcup_{k=1}^{k_j} B_{j,k}\Big) \geq \frac{1}{2N}\,(\mu\llcorner A)(U_{j-1}).$$

Thus we conclude, using again measurability with respect to $\mu\llcorner A$,

$$(\mu\llcorner A)\Big(U_{j-1}\backslash\bigcup_{k=1}^{k_j} B_{j,k}\Big) \leq \Big(1 - \frac{1}{2N}\Big)\,(\mu\llcorner A)(U_{j-1}).$$

The construction is complete. $\square$

Chapter 5

Differentiation of Radon measures

In the chapter we dicuss a version of the Radon-Nikodym theorem for Radon measures on $\mathbb{R}^n$. The proof applies the Besicovitch covering theorem, more precisely Theorem 4.6.

Definition 5.1 (Densities)**.** *Let μ, ν be Radon measures on $\mathbb{R}^n$. The upper/lower density of ν with respect to μ at a point $x \in \operatorname{spt}\mu$ is*

$$\overline{D}_\mu \nu(x) = \limsup_{r \searrow 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} \quad \text{and} \quad \underline{D}_\mu \nu(x) = \liminf_{r \searrow 0} \frac{\nu(B_r(x))}{\mu(B_r(x))}.$$

We write $D_\mu \nu(x)$ when the upper and lower density are equal.

We claim that the densities are Borel measurable functions. To see this we show that the function $x \mapsto \mu(B_\varrho(x))$ is upper semicontinuous, where $B_\varrho(x)$ denotes the closed ball as in the previous chapter. This means that the sets $\{\mu(B_r(x)) < t\}$ are open, and then the claim follows easily. Now if $x_k \to x$ and $r > \varrho$ then $B_r(x) \supset B_\varrho(x_k)$ for k sufficiently large, thus $\mu(B_r(x)) \geq \limsup_{k\to\infty} \mu(B_\varrho(x_k))$. We conclude

$$\mu(B_\varrho(x)) = \lim_{r \searrow \varrho} \mu(B_r(x)) \geq \limsup_{k\to\infty} \mu(B_\varrho(x)).$$

Lemma 5.2. *Let $A \subset \mathbb{R}^n$ and $0 < \alpha < \infty$. Then the following hold:*

(a) $A \subset \{x \in \mathbb{R}^n : \underline{D}_\mu \nu(x) \leq \alpha\} \quad \Rightarrow \quad \nu(A) \leq \alpha\, \mu(A)$,

(b) $A \subset \{x \in \mathbb{R}^n : \overline{D}_\mu \nu(x) \geq \alpha\} \quad \Rightarrow \quad \nu(A) \geq \alpha\, \mu(A)$.

Proof. To prove claim (a) we let U be any open set with $U \supset A$, and define the ball family

$$\mathcal{F} = \{B = B_\varrho(a) \subset U : a \in A,\, \nu(B) \leq (\alpha + \varepsilon)\mu(B)\}.$$

We have $\inf\{\varrho > 0 : B_\varrho(a) \in \mathcal{F}\} = 0$ by assumption. Theorem 4.6 yields a pairwise disjoint subfamily $\mathcal{F}'$ such that

$$\nu\Big(A \backslash \bigcup_{B \in \mathcal{F}'} B\Big) = 0.$$

We conclude

$$\nu(A) \leq \sum_{B \in \mathcal{F}'} \nu(B) \leq (\alpha + \varepsilon) \sum_{B \in \mathcal{F}'} \mu(B) \leq (\alpha + \varepsilon)\, \mu(U).$$

Taking the infimum over all U, see Theorem 2.13, and letting $\varepsilon \searrow 0$ proves (a). For claim (b) we argue that $\overline{D}_\mu \nu(x) \geq \alpha$ implies $\underline{D}_\nu \mu(x) \leq \frac{1}{\alpha}$, and then apply (a). $\square$

Theorem 5.3 (Radon-Nikodym)**.** *For Radon measures μ, ν on $\mathbb{R}^n$ the following holds:*

(1) *The set $Z = \{\overline{D}_\mu \nu = \infty\}$ satisfies $\mu(Z) = 0$, and $D_\mu \nu$ exists μ-almost everywhere.*

(2) *For any μ-measurable set A one has*

$$\nu(A) = \int_A D_\mu \nu \, d\mu + (\nu \llcorner Z)(A).$$

Before entering the proof we mention further terminology:

- ν *is absolutely continuous with respect to* μ, in notation $\nu \ll \mu$, if $\mu(A) = 0$ always implies $\nu(A) = 0$.
- μ *and* ν *are mutually singular*, in notation $\mu \perp \nu$, if there exists a Borel set Z such that $\mu(Z) = 0 = \nu(\mathbb{R}^n \backslash Z)$.

Proof. We first give the proof when $\mu(\mathbb{R}^n)$ and $\nu(\mathbb{R}^n)$ are both finite.

Step 1: Proof of statement (1).
The inequality $\overline{D}_\mu \nu \geq \alpha$ holds on Z for any $\alpha > 0$. Then $\nu(Z) \geq \alpha \mu(Z)$ by Lemma 5.2, and we conclude $\mu(Z) = 0$. Next consider for $0 < a < b < \infty$ the sets

$$R(a,b) = \{\underline{D}_\mu \nu < a < b < \overline{D}_\mu \nu\}.$$

Again by Lemma 5.2 we have

$$b\mu(R(a,b)) \leq \nu(R(a,b)) \leq a\mu(R(a,b)).$$

As $a < b$ this is only possible when $\mu(R(a,b)) = 0$. But now

$$\{\underline{D}_\mu \nu < \overline{D}_\mu \nu\} = \bigcup_{0<a<b<\infty,\, a,b \in \mathbb{Q}} R(a,b).$$

Step 2: Proof of (2) in the case $\nu \ll \mu$.
For $N = \{\underline{D}_\mu \nu = 0\}$ Lemma 5.2 implies $\nu(N) \leq \alpha \mu(N)$ for any $\alpha > 0$, hence $\nu(N) = 0$. Moreover Z and $\{\underline{D}_\mu \nu < \overline{D}_\mu \nu\}$ are null sets for μ, thus also for ν by assumption. Now consider for $t \in (1, \infty)$ and $m \in \mathbb{Z}$ the sets

$$A_m = \{x \in A : t^m \leq D_\mu \nu(x) < t^{m+1}\}.$$

We compute, again by Lemma 5.2,

$$\nu(A) = \sum_{m \in \mathbb{Z}} \nu(A_m) \leq \sum_{m \in \mathbb{Z}} t^{m+1} \mu(A_m) \leq t \sum_{m \in \mathbb{Z}} \int_{A_m} D_\mu \nu \, d\mu = t \int_A D_\mu \, \nu \, d\mu.$$

The lower bound follows in the same way, namely

$$\nu(A) = \sum_{m \in \mathbb{Z}} \nu(A_m) \geq \sum_{m \in \mathbb{Z}} t^m \mu(A_m) \geq \frac{1}{t} \sum_{m \in \mathbb{Z}} \int_{A_m} D_\mu \nu \, d\mu = \frac{1}{t} \int_A D_\mu \, \nu \, d\mu.$$

Letting $t \searrow 1$ finishes Step 2.

Step 3. Proof of (2) for general μ, ν (still finite)
We claim that $\nu_* = \nu \llcorner (\mathbb{R}^n \backslash Z)$ is absolutely continuous with respect to μ. In fact for $\mu(A) = 0$ and any $\alpha > 0$ let $A_\alpha = \{x \in A : \underline{D}_\mu \nu(x) \leq \alpha\}$. Then by Lemma 5.2 we get

$$\nu(A_\alpha) \leq \alpha \mu(A_\alpha) = 0, \quad \text{thus } \nu_*(A) = \nu(A \backslash Z) = \lim_{\alpha \nearrow \infty} \nu(A_\alpha) = 0.$$

Now let $A^\alpha = \{x \in \mathbb{R}^n \backslash Z : \overline{D}_\mu(\nu \llcorner Z)(x) \geq \alpha\}$. Again by Lemma 5.2 we get

$$0 = (\nu \llcorner Z)(A^\alpha) \geq \alpha \mu(A^\alpha),$$

so that $\overline{D}_\mu(\nu \llcorner Z)(x) = 0$ for μ-almost every $x \in \mathbb{R}^n$. This implies $D_\mu \nu_* = D_\mu \nu$ for μ-almost every $x \in \mathbb{R}^n$, and Step 2 implies

$$\nu(A) - (\nu \llcorner Z)(A) = \nu_*(A) = \int_A D_\mu \nu_* \, d\mu = \int_A D_\mu \nu \, d\mu.$$

Step 4. Proof for general μ, ν
Consider the restrictions to $B_R(0)$, and let $R \nearrow \infty$. □

Corollary 5.4 (Lebesgue differentiation theorem)**.** *Let μ be a Radon measure on $\mathbb{R}^n$, and $f \in L^1_{\text{loc}}(\mu)$. Then*

$$\lim_{\varrho \searrow 0} ⨍_{B_\varrho(x)} f(y) \, d\mu(y) = f(x) \quad \textit{for } \mu\textit{-almost every } x \in \mathbb{R}^n.$$

Proof. We may assume $f \geq 0$. There is a unique Radon measure ν on $\mathbb{R}^n$ such that

$$\nu(B) = \int_B f \, d\mu \quad \text{for } B \ \mu\text{-measurable.}$$

ν is absolutely continuous with respect to μ, hence Theorem 5.3 yields the representation

$$\nu(B) = \int_B D_\mu \nu \, d\mu.$$

Taking for B the sets $\{f \geq D_\mu \nu\}$ and $\{f \leq D_\mu \nu\}$ we conclude $f = D_\mu \nu$ μ-almost everywhere. □

We mention a slight improvement which is sometimes useful. Let μ be a Radon measure on $\mathbb{R}^n$ and $f \in L^p_{\text{loc}}(\mu)$ where $1 \leq p < \infty$. A point $x \in \operatorname{spt} \mu$ is an L^p Lebesgue point if

$$\lim_{\varrho \searrow 0} ⨍_{B_\varrho(x)} |f(y) - f(x)|^p \, d\mu(y) = 0. \tag{5.1}$$

We claim that μ-almost all $x \in \operatorname{spt} \mu$ are Lebesgue points. By the Minkowski inequality we have for any $\lambda \in \mathbb{R}$

$$\Big(⨍_{B_\varrho(x)} |f - f(x)|^p \, d\mu \Big)^{\frac{1}{p}} \leq \Big(⨍_{B_\varrho(x)} |f - \lambda|^p \, d\mu \Big)^{\frac{1}{p}} + |f(x) - \lambda|.$$

Applying Corollary 5.4 on the right we obtain for all $x \in \operatorname{spt} \mu$, except a μ-null set $N(\lambda)$,

$$\limsup_{\varrho \searrow 0} \Big(⨍_{B_\varrho(x)} |f - f(x)|^p \, d\mu \Big)^{\frac{1}{p}} \leq 2 \, |f(x) - \lambda|.$$

For $x \notin \bigcup_{\lambda \in \mathbb{Q}} N(\lambda)$ we can take a sequence $\mathbb{Q} \ni \lambda_i \to f(x)$, the claim follows.

For example, consider the characteristic function χ_E of a μ-measurable set $E \subset \mathbb{R}^n$. Then $x \in \operatorname{spt} \mu$ is a Lebsgue point if and only if one of the following holds:

$$\begin{aligned} x \in E \quad &\text{and} \quad \lim_{\varrho \searrow 0} \frac{\mu(E \cap B_\varrho(x))}{\mu(B_\varrho(x))} = 1, \\ x \notin E \quad &\text{and} \quad \lim_{\varrho \searrow 0} \frac{\mu(E \cap B_\varrho(x))}{\mu(B_\varrho(x))} = 0. \end{aligned}$$

The limit (if it exists) is called the μ-density of E in x, in general one considers the upper and lower μ-density by taking the $\limsup$ and $\liminf$. By the above, the μ-density of E is equal to one for μ-almost all $x \in E$, and equal to zero for μ-almost all $x \in \mathbb{R}^n \backslash E$.

Definition 5.5. *Let μ be a Radon measure on $\mathbb{R}^n$, and let $f : \mathbb{R}^n \to \mathbb{R}$. We say that $\lambda \in \mathbb{R}$ is the approximate limit of f with respect to μ at a point $x \in \operatorname{spt} \mu$, if for all $\varepsilon > 0$*

$$\lim_{\varrho \searrow 0} \frac{\mu\big(B_\varrho(x) \cap \{|f - \lambda| \geq \varepsilon\}\big)}{\mu(B_\varrho(x))} = 0.$$

Notation: $\mu\text{-aplim}_{y \to x} f(y) = \lambda$.

The approximate limit is unique: let $|\lambda - \lambda'| = 2\varepsilon > 0$. For any $y \in B_\varrho(x)$ at least one of the inequalities $|f(y) - \lambda| \geq \varepsilon$ or $|f(y) - \lambda'| \geq \varepsilon$ holds. Thus

$$\frac{\mu\big(B_\varrho(x) \cap \{|f - \lambda| \geq \varepsilon\}\big) + \mu\big(B_\varrho(x) \cap \{|f - \lambda'| \geq \varepsilon\}\big)}{\mu(B_\varrho(x))} \geq \frac{\mu(B_\varrho(x))}{\mu(B_\varrho(x))} = 1.$$

Hence $f(y)$ cannot have both λ and λ' as approximate limits. f is approximately continuous at $x \in \operatorname{spt} \mu$, if $\mu\text{-aplim}_{y \to x} f(y) = f(x)$. If $x \in \operatorname{spt} \mu$ is a Lebesgue point of a function $f \in L^1_{\text{loc}}(\mu)$, then f is approximately continuous at x. Namely we have

$$\begin{aligned} \frac{\mu\big(B_\varrho(x) \cap \{|f - f(x)| \geq \varepsilon\}\big)}{\mu(B_\varrho(x))} &= ⨍_{B_\varrho(x)} \chi_{\{f - f(x)| \geq \varepsilon\}} \, d\mu(y) \\ &\leq ⨍_{B_\varrho(x)} \frac{|f(y) - f(x)|}{\varepsilon} \, d\mu(y) \to 0 \quad \text{as } \varrho \searrow 0. \end{aligned}$$

Corollary 5.6 (approximate continuity)**.** *Let μ be A Radon measure on $\mathbb{R}^n$. Then any μ-measurable function $f : \mathbb{R}^n \to \mathbb{R}$ is approximately continuous at μ-almost all points $x \in \mathbb{R}^n$.*

Proof. Consider for $k \in \mathbb{N}$ the truncations

$$f_k(x) = \begin{cases} k & \text{if } f(x) \geq k, \\ f(x) & \text{if } -k \leq f(x) \leq k, \\ -k & \text{if } f(x) \leq -k. \end{cases}$$

Then $f_k \in L^1_{\text{loc}}(\mu)$, and by the above f_k is approximately continuous on $\mathbb{R}^n \backslash N_k$, where N_k is a μ-null set. We claim that f is approximately continuous at all points $x \notin \bigcup_{k=1}^\infty N_k$. For given $\varepsilon > 0$ take $k \in \mathbb{N}$ such that $|f(x)| < k - \varepsilon$, which implies $f_k(x) = f(x)$. Now for $|f(y) - f(x)| \geq \varepsilon$ we have

- if $|f(y)| \le k$ then $|f_k(y) - f_k(x)| = |f(y) - f(x)| \ge \varepsilon$,
- if $|f(y)| > k$ then $|f_k(y) - f_k(x)| \ge \big||f_k(y)| - |f_k(x)|\big| \ge k - (k-\varepsilon) = \varepsilon$.

Thus $|f(y) - f(x)| \ge \varepsilon$ implies $|f_k(y) - f_k(x)| \ge \varepsilon$, and we conclude for $x \notin \bigcup_{k=1}^\infty N_k$

$$\frac{\mu\big(B_\varrho(x) \cap \{|f - f(x)| \ge \varepsilon\}\big)}{\mu(B_\varrho(x))} \le \frac{\mu\big(B_\varrho(x) \cap \{|f_k - f_k(x)| \ge \varepsilon\}\big)}{\mu(B_\varrho(x))} \to 0 \quad \text{as } \varrho \searrow 0.$$

□

We now turn to a another notion of density.

Definition 5.7 (s-dimensonal density)**.** *Let μ be a Borel measure on (X,d) and $A \subset X$. The upper/lower s-dimensional density of A with respect to μ is*

$$\begin{aligned} \overline{\theta}^s(\mu, A, x) &= \limsup_{\varrho \searrow 0} \frac{\mu(A \cap B_\varrho(x))}{\alpha(s)\varrho^s}, \\ \underline{\theta}^s(\mu, A, x) &= \liminf_{\varrho \searrow 0} \frac{\mu(A \cap B_\varrho(x))}{\alpha(s)\varrho^s}. \end{aligned}$$

We write $\theta^s(\mu, A, x)$ in case of equality. The functions $\overline{\theta}^s(\mu, A, \cdot)$ and $\underline{\theta}^s(\mu, A, \cdot)$ are Borel measurable, even when A is not μ measurable. Namely, Borel sets are $\mu \llcorner A$-measurable by Lemma 2.11. The claim then follows from the upper semicontinuity (see definition 5.1)

$$(\mu \llcorner A)(B_\varrho(x)) \ge \limsup_{y \to x} (\mu \llcorner A)(B_\varrho(y)).$$

Unfortunately the two statements in the next theorem are somewhat different in detail. We note that the two sets E and A need not be μ-measurable.

Theorem 5.8 (s-densities)**.** *For a Borel measure μ on (X,d) the following holds:*

(1) *If $\overline{\theta}^s(\mu, E, \cdot) \ge \lambda$ on A, then $\mu(E) \ge \lambda\, \mathcal{H}^s(A)$.*

(2) *If $\overline{\theta}^s(\mu, A, \cdot) \le \lambda$ on A, then $\mu(A) \le 2^s \lambda\, \mathcal{H}^s(A)$ (for μ Borel regular).*

Proof. In (1) we can assume $\mu(E) < \infty$, $\lambda > 0$ and also $\overline{\theta}^s(\mu, E, \cdot) > \lambda$ on A, by eventually letting $\lambda' \nearrow \lambda$. For any $\delta > 0$ we consider the family of balls

$$\mathcal{F} = \Big\{ B_\varrho(a) : a \in A,\ \varrho < \delta,\ \frac{\mu(E \cap B_\varrho(a))}{\alpha(s)\varrho^s} > \lambda \Big\}.$$

By assumption $\mathcal{F}$ is a fine covering of A. Let $\mathcal{G}$ be the disjoint subfamily as in Corollary 4.3. For any finite $\mathcal{G}' \subset \mathcal{G}$ we have

$$\sum_{B \in \mathcal{G}'} \underbrace{(\mu \llcorner E)(B)}_{>0} = (\mu \llcorner E)\Big(\bigcup_{B \in \mathcal{G}'} B \Big) \le \mu(E) < \infty.$$

Hence $\mathcal{G}$ is countable, we write $\mathcal{G} = \{B_j : j \in \mathbb{N}\}$. Now Corollary 4.3 yields that

$$A \subset \bigcup_{j=1}^k B_j \cup \bigcup_{j=k+1}^\infty \hat{B}_j \quad \text{for any } k \in \mathbb{N}.$$

Therefore we can estimate

$$\begin{aligned}
\mathcal{H}^s_{10\delta}(A) &\leq \sum_{j=1}^{k} \alpha(s)\varrho_j^s + \sum_{j=k+1}^{\infty} \alpha(s)(5\varrho_j)^s \\
&\leq \frac{1}{\lambda}\sum_{j=1}^{k}(\mu\llcorner E)(B_j) + \frac{5^s}{\lambda}\sum_{j=k+1}^{\infty}(\mu\llcorner E)(B_j) \\
&\leq \frac{1}{\lambda}\mu(E) + \frac{5^s}{\lambda}\sum_{j=k+1}^{\infty}(\mu\llcorner E)(B_j).
\end{aligned}$$

Letting $k \to \infty$ the series on the right disappears, and the claim then follows by letting $\delta \searrow 0$.

For (2) we can assume $\mathcal{H}^s(A) < \infty$, and also $\overline{\theta}^s(\mu, A, \cdot) < \lambda$ on A (see above). Consider

$$A_k = \Big\{x \in A : \frac{\mu(A \cap B_\varrho(x))}{\alpha(s)\varrho^s} \leq \lambda \ \text{ for all } \ \varrho \in \big(0, \frac{1}{k}\big)\Big\} \quad \text{for } k \in \mathbb{N}.$$

We have $A_1 \subset A_2 \subset \ldots$ and $A = \bigcup_{k=1}^\infty A_k$. Now $\mu(A_k) \to \mu(A)$ as $k \to \infty$ (since μ is Borel regular this holds even when the A_k are not measurable, see Evans-Gariepy, page 5). Therefore it suffices to estimate $\mu(A_k)$. We claim that

$$\mu(A_k) \leq 2^s \lambda\, \mathcal{H}^s_{\frac{1}{k}}(A_k).$$

For this let C_j, $j \in \mathbb{N}$, be any covering of A_k with $\delta_j = \operatorname{diam} C_j < \frac{1}{k}$, and such that $A_k \cap C_j \neq \emptyset$ for all $j \in \mathbb{N}$. Taking points $a_j \in A_k \cap C_j$ we infer

$$A_k \subset \bigcup_{j=1}^{\infty} C_j \subset \bigcup_{j=1}^{\infty} B_{\delta_j}(a_j).$$

This yields the bound

$$\begin{aligned}
\mu(A_k) &\leq \sum_{j=1}^{\infty} \mu\big(A \cap B_{\delta_j}(a_j)\big) \\
&\leq \lambda \sum_{j=1}^{\infty} \alpha(s)\delta_j^s \quad \big(\text{using } a_k \in A_k,\ \delta_j < \frac{1}{k}\big) \\
&\leq 2^s \lambda \sum_{j=1}^{\infty} \alpha(s)\Big(\frac{\operatorname{diam} C_j}{2}\Big)^s.
\end{aligned}$$

Taking the infimum over all coverings of A_k we conclude

$$\mu(A_k) \leq 2^s \lambda\, \mathcal{H}^s_{\frac{1}{k}}(A_k) \leq 2^s \lambda\, \mathcal{H}^s(A).$$

This completes the proof of the theorem. □

Corollary 5.9. *Let μ be a Radon measure on (X, d) and $s \geq 0$. If $\mu(A) = 0$ then*

$$\overline{\theta}^s(\mu, x) = 0 \quad \textit{for } \mathcal{H}^s\textit{-almost every } x \in A.$$

Proof. It suffices to show that for any $\lambda > 0$ the following sets are $\mathcal{H}^s$ null sets:

$$A_\lambda = \{x \in A : \overline{\theta}^s(\mu, x) \geq \lambda\}.$$

Let $U \supset A_\lambda$ be open. Then for $x \in A_\lambda$ we have trivially $\overline{\theta}^s(\mu, U, x) = \overline{\theta}^s(\mu, x) \geq \lambda$. Thus from Theorem 5.8(1) we obtain

$$\lambda\, \mathcal{H}^s(A_\lambda) \leq \mu(U).$$

As $\inf_{U \supset A_\lambda} \mu(U) = \mu(A_\lambda) = 0$ the claim follows. □

Let μ be Borel regular and assume that E is μ-measurable with $\mu(X \backslash E) < \infty$. Then $\mu \llcorner (X \backslash E)$ is a Radon measure for which E is a null set. We conclude

$$\overline{\theta}^s(\mu, X \backslash E, x) = \overline{\theta}^s(\mu \llcorner (X \backslash E), x) = 0 \quad \text{for } \mathcal{H}^s\text{-almost every } x \in E. \tag{5.2}$$

Corollary 5.10. *Let $A \subset (X, d)$ be a Borel set with $\mathcal{H}^s(A) < \infty$. Then*

$$2^{-s} \leq \overline{\theta}^s(\mathcal{H}^s, A, x) \leq 1 \quad \textit{for } \mathcal{H}^s\textit{-almost every } x \in A.$$

Proof. Consider for $\lambda^- < 2^{-s}$ and $\lambda^+ > 1$ the sets

$$\begin{aligned} A^+ &= \{x \in A : \overline{\theta}^s(\mathcal{H}^s, A, x) \geq \lambda^+\}, \\ A^- &= \{x \in A : \overline{\theta}^s(\mathcal{H}^s, A, x) \leq \lambda^-\}. \end{aligned}$$

For open $U \subset A^+$ we apply Theorem 5.8(1) to $\mathcal{H}^s$, $A \cap U$ and A^+, we obtain

$$\lambda^+ \mathcal{H}^s(A^+) \leq \mathcal{H}^s(A \cap U) = (\mathcal{H}^s \llcorner A)(U).$$

Now since A is Borel and $\mathcal{H}^s(A) < \infty$, the measure $\mathcal{H}^s \llcorner A$ is Borel regular by Theorem 2.12. Hence using Theorem 2.13 we get by taking the infimum over all U

$$\lambda^+ \mathcal{H}^s(A^+) \leq (\mathcal{H}^s \llcorner A)(A^+) = \mathcal{H}^s(A^+) \quad \Rightarrow \quad \mathcal{H}^s(A^+) = 0.$$

For the lower bound we note that on A^- we have $\overline{\theta}^s(\mathcal{H}^s, A^-, x) \leq \overline{\theta}^s(\mathcal{H}^s, A, x) \leq \lambda^-$. Applying Theorem 5.8(2) to $\mathcal{H}^s$ and A^- we obtain

$$\mathcal{H}^s(A^-) \leq 2^s \lambda^- \mathcal{H}^s(A^-) \quad \Rightarrow \quad \mathcal{H}^s(A^-) = 0.$$

□

Chapter 6

The Riesz representation theorem

In this chapter (X,d) is always a locally compact, separable metric space, for example $X = \mathbb{R}^n$. By Lemma 2.15 there exists an exhaustion $X = \bigcup_{i=1}^{\infty} U_i$ where the U_i are open with $\overline{U_i}$ compact. Let μ be a Radon measure on X, and assume that $\eta : X \to \mathbb{R}^k$ is μ-measurable and satisfies $|\eta(x)| = 1$ for all $x \in X$. Then we have an induced linear form

$$\phi : C_c^0(X,\mathbb{R}^k) \to \mathbb{R},\ \phi(f) = \int_X \langle f, \eta \rangle \, d\mu. \tag{6.1}$$

Moreover, for any compact set K we have the estimate, with constant $C(K) = \mu(K)$,

$$|\phi(f)| \leq C(K)\, \|f\|_{C^0(X)} \quad \text{for all } f \in C_c^0(X,\mathbb{R}^k) \text{ with } \operatorname{spt} f \subset K. \tag{6.2}$$

In other words, ϕ is a continuous linear functional on the space of $C_c^0(X,\mathbb{R}^k)$ functions with support in K. If X happens to be compact then ϕ is continuous on the whole space $C^0(X,\mathbb{R}^k)$. Any linear form ϕ with (6.2) will be called a linear functional on $C_c^0(X,\mathbb{R}^k)$. The goal of this chapter is to reverse this process: given a linear functional ϕ on $C_c^0(X,\mathbb{R}^k)$, we want to find a Radon measure μ and a μ-measurable function $\eta \in C^0(X,\mathbb{S}^{k-1})$ such that ϕ has the representation in (6.1). Moreover, the data μ and η should be unique.

We start with the construction of the measure μ. To do this we need a partition of unity. The usual statement asserts that that the partition functions χ_i can be chosen subordinate to the covering U_k, in the sense that $\operatorname{spt} \chi_i \subset U_k$ for some k. However, we need below that for $i' \neq i$ one can choose $k' \neq k$, which is not automatic. Therefore we now prove a specific partition lemma.

Lemma 6.1 (simple partition of unity)**.** *Let U_i, $1 \leq i \leq N$, be an open covering of a compact set $K \subset X$. There exist functions $\chi_i \in C_c^0(U_i)$ such that $0 \leq \chi_i \leq 1$ and $\sum_{i=1}^N \chi_i = 1$ on K.*

Proof. We argue by induction. For $N = 1$, that is $K \subset U$, we can take the function

$$\chi(x) = \Big(1 - \frac{\operatorname{dist}(x,K)}{\varepsilon}\Big)^+ \quad \text{for } \varepsilon > 0 \text{ sufficiently small.}$$

Now let $K \subset \bigcup_{i=1}^N U_i$. We construct compact sets K_0, K_N such that $K = K_0 \cup K_N$ and

$$K_0 \subset \bigcup_{i=1}^{N-1} U_i =: U_0 \quad \text{and} \quad K_N \subset U_N.$$

For this we take open sets $V_0 \supset K \backslash U_N$ and $V_N \supset K \backslash U_0$, then we have

$$\begin{aligned} K_0 := K \backslash V_N &\subset K \backslash (K \backslash U_0) \subset U_0, \\ K_N := K \backslash V_0 &\subset K \backslash (K \backslash U_N) \subset U_N. \end{aligned}$$

$K \backslash U_0$ and $K \backslash U_N$ are compact and disjoint by assumption, hence we can choose V_0, V_N disjoint, so that $K = K_0 \cup K_N$ as desired. Using induction we now have functions $\eta_i \in C_c^0(U_i)$ for $i = 1, \ldots, N-1$ and $\eta_N \in C_c^0(U_N)$, all with values in $[0,1]$, such that

$$\sum_{i=1}^{N-1} \eta_i = 1 \text{ on } K_0, \quad \eta_N = 1 \text{ on } K_N.$$

It follows that $\sum_{i=1}^N \eta_i \geq 1$ on K. Thus for $\varepsilon > 0$ small we can finally take

$$\chi_i = \Big(1 - \frac{\operatorname{dist}(x,K)}{\varepsilon}\Big)^+ \frac{\eta_i}{\sum_{i=1}^N \eta_i} \in C_c^0(U_i).$$

$\square$

Definition 6.2 (variation measure)**.** *For any linear functional ϕ on $C_c^0(X, \mathbb{R}^k)$, the variation measure $|\phi| : 2^X \to [0, \infty]$ is defined in two steps as follows:*

(1) $|\phi|(U) = \sup\{\phi(f) : |f| \leq 1, \operatorname{spt} f \subset U\}$ *for U open,*

(2) $|\phi|(E) = \inf\{|\phi|(U) : U \supset E,\ U$ *open*$\}$ *for E general.*

The steps are consistent since in (1) we have $|\phi|(U) \leq |\phi|(V)$ whenever $U \subset V$.

Lemma 6.3 (variation measure)**.** *$|\phi|$ is a Radon measure.*

Proof. We proceed in three steps.

Step 1: *$|\phi|$ is an outer measure.*
For $U = \emptyset$ only the null function is admissible in 6.2(1) and hence $|\phi|(\emptyset) = 0$. Next, let U_i, $i \in \mathbb{N}$, be open and $f \in C_c^0(X)$ such that $|f| \leq 1$ and $\operatorname{spt} f \subset \bigcup_{i=1}^\infty U_i$. By compactness we have $\operatorname{spt} f \subset \bigcup_{i=1}^N U_i$ for some $N \in \mathbb{N}$. Applying Lemma 6.1 we find functions $\chi_i \in C_c^0(U_i)$ with $0 \leq \chi_i \leq 1$ such that

$$\sum_{j=1}^N \chi_i = 1 \text{ on } \operatorname{spt} f.$$

For $f_i = \chi_i f \in C_c^0(U_i)$ we have $|f_i| \leq 1$ and $f = \sum_{i=1}^N f_i$ on $\operatorname{spt} f$. This implies

$$\phi(f) = \sum_{i=1}^N \phi(f_i) \leq \sum_{i=1}^N |\phi|(U_i) \leq \sum_{i=1}^\infty |\phi|(U_i).$$

Taking the supremum over these f we obtain

$$|\phi|\big(\bigcup_{i=1}^\infty U_i\big) \leq \sum_{i=1}^\infty |\phi|(U_i).$$

Now assume $E \subset \bigcup_{j=1}^{\infty} E_j$ where E, E_j are arbitrary. Given $\varepsilon > 0$ we choose open $U_j \supset E_j$ with $|\phi|(U_j) < |\phi|(E_j) + 2^{-j}\varepsilon$, by just using Definition 6.2. Then $E \subset \bigcup_{j=1}^{\infty} U_j$ and hence

$$|\phi|(E) \leq |\phi|(\bigcup_{j=1}^{\infty} U_j) \leq \sum_{j=1}^{\infty} |\phi|(U_j) \leq \sum_{j=1}^{\infty} |\phi|(E_j) + \varepsilon.$$

Letting $\varepsilon \searrow 0$ we conclude that $|\phi|$ is an outer measure.

Step 2: *Borel sets are $|\phi|$-measurable.*
We use Carathedory's criterion, Theorem 2.9. By definition of $|\phi|$, we need to show that whenever $A, B \subset X$ satisfy $\operatorname{dist}(A, B) > 0$ then

$$|\phi|(W) \geq |\phi|(A) + |\phi|(B) \quad \text{for all open } W \supset (A \cup B).$$

For small $\delta > 0$ the sets $U = B_\delta(A) \cap W$ and $V = B_\delta(B) \cap W$ are disjoint. Let $f, g \in C_c^0(X, \mathbb{R}^k)$ with $\operatorname{spt} f \subset U$, $\operatorname{spt} g \subset V$ and $|f|, |g| \leq 1$. Then $\operatorname{spt}(f+g) \subset (\operatorname{spt} f \cup \operatorname{spt} g) \subset W$ and $|f+g| \leq 1$ on all X. We obtain

$$\phi(f) + \phi(g) = \phi(f+g) \leq |\phi|(W).$$

Taking the supremum over all f, g we conclude

$$|\phi|(A) + |\phi|(B) \leq |\phi|(U) + |\phi|(V) \leq |\phi|(W).$$

Step 3: *Construction of Borel hull and finiteness on compact sets.*
Given $E \subset X$ with $|\phi|(E) < \infty$ choose open $U_j \supset E$ with $|\phi|(U_j) \leq |\phi|(E) + \frac{1}{j}$. We may assume that $U_1 \supset U_2 \supset \ldots$. Now $B = \bigcap_{j=1}^{\infty} U_j \supset E$ is Borel and satisfies

$$|\phi|(B) \leq \lim_{j\to\infty} |\phi|(U_j) = |\phi|(E).$$

Hence $|\phi|$ is Borel regular. Finally if K is compact, then by assumption on X there exists a relatively compact, open set $U \supset K$. Assumption (6.2) then yields

$$|\phi|(K) \leq |\phi|(U) \leq C(\overline{U}) < \infty.$$

$\square$

We can now state the central result of the Chapter.

Theorem 6.4 (Riesz). *Let (X, d) be a locally compact, separable metric space. Then for any linear functional ϕ on $C_c^0(X, \mathbb{R}^k)$ there exists a Radon measure μ and a μ-measurable function $\eta : X \to \mathbb{R}^k$ with $|\eta(x)| = 1$ for μ-almost every $x \in X$, such that*

$$\phi(f) = \int_X \langle f, \eta \rangle \, d\mu \quad \textit{for all } f \in C_c^0(X, \mathbb{R}^k). \tag{6.3}$$

The pair μ, η with (6.3) is unique, and μ is the variation measure $|\phi|$.

We first address the uniqueness, which requires the following approximation statement.

Theorem 6.5 (Lusin)**.** *Let μ be a Radon measure on (X,d), and let $A \subset X$ with $\mu(A) < \infty$. Then for any μ-measurable $g : X \to \mathbb{R}$ there exists a function $\tilde{g} \in C^0(X)$ such that*

$$\mu(\{x \in A : \tilde{g}(x) \neq g(x)\}) < \varepsilon \quad \text{and} \quad \|\tilde{g}\|_{C^0(X)} \leq \operatorname{ess\,sup}_{x\in A}|g(x)|.$$

We postpone the proof of Lusin's theorem, first we address the representation theorem.

Proof. (uniqueness) Assume that λ, ζ have also the representation property, that is

$$\phi(f) = \int_X \langle f, \zeta\rangle \, d\lambda \quad \text{for all } f \in C^0_c(X, \mathbb{R}^k).$$

If $\operatorname{spt} f \subset U$ und $|f| \leq 1$, then $\phi(f) \leq \lambda(U)$ and hence $\mu(U) = |\phi|(U) \leq \lambda(U)$. Approximating from outside we see that $\mu \leq \lambda$. For the reverse inequality let $K \subset X$ be compact. For $U \supset K$ open with $\overline{U}$ compact and for $\varepsilon > 0$ there exists a function $\tilde{\zeta} \in C^0(X, \mathbb{R}^k)$ with the property

$$\lambda(E) < \varepsilon \text{ where } E = \{x \in U : \tilde{\zeta}(x) \neq \zeta(x)\} \quad \text{and} \quad \|\tilde{\zeta}\|_{C^0(X)} \leq \operatorname{ess\,sup}_{x\in U}|\zeta| \leq 1.$$

This follows directly by Theorem 6.5 in the case $k = 1$, i.e. ζ real-valued. For ζ vector-valued we apply Theorem 6.5 to each component ζ^i, and then obtain $\tilde{\zeta}$ by projecting onto the unit ball in $\mathbb{R}^k$. Let $\chi \in C^0_c(X)$, $0 \leq \chi \leq 1$, with $\operatorname{spt}\chi \subset U$ and $\chi \equiv 1$ auf K. We now estimate

$$\begin{aligned}
\mu(U) &\geq \phi(\chi\tilde{\zeta}) && (\mu \text{ is the variation measure})\\
&= \int_U \langle \chi\tilde{\zeta}, \zeta\rangle \, d\lambda && (\text{the representation property})\\
&= \int_U \chi \, d\lambda - \int_U \chi(\langle\tilde{\zeta}, \zeta\rangle - 1) \, d\lambda\\
&\geq \lambda(K) - 2\lambda(E)\\
&\geq \lambda(K) - 2\varepsilon.
\end{aligned}$$

Letting $\varepsilon \searrow 0$ and $U \searrow K$ we obtain $\mu(K) \geq \lambda(K)$, and hence $\mu \geq \lambda$ on all Borel sets by inner approximation. By Borel regularity, any $E \subset X$ has Borel hulls $B, B' \supset E$ for μ respectively λ. We may assume $B = B'$, otherwise we pass to $B \cap B'$. We now conclude that $\mu(E) = \lambda(E)$ for all sets E. Next for given $v \in \mathbb{R}^k$ and any $f \in C^0_c(X)$ we have $fv \in C^0_c(X, \mathbb{R}^k)$, therefore

$$\int_X \langle fv, \eta\rangle \, d\mu = \int_X \langle fv, \zeta\rangle \, d\mu \quad \text{for all } f \in C^0_c(X).$$

Now $C^0_c(X)$ is dense in $L^1(\mu)$[1], hence we have

$$\int_X f\varphi \, d\mu = 0 \quad \text{for all } f \in L^1(\mu), \text{ where } \varphi = \langle v, \eta - \zeta\rangle.$$

Taking $f = \chi_K \operatorname{sign}\varphi$ for K compact yields $\varphi = 0$. We finally choose for v the standard basis vectors and conclude $\eta = \zeta$, which finishes the proof of uniqueness. □

Proof. (existence) We take μ as the variation measure $|\phi|$. For any $v \in \mathbb{R}^k$, $|v| = 1$, let

$$\phi_v : C^0_c(X) \to \mathbb{R},\ \phi_v(f) = \phi(fv).$$

[1] Aufgabe 1, Serie 4

Our goal is to show that ϕ_v extends to a continuous linear functional on $L^1(\mu)$. By duality $L^1(\mu)' = L^\infty(\mu)$, we then obtain functions $\eta_i \in L^\infty(\mu)$ such that for all $f \in C^0_c(X, \mathbb{R}^k)$

$$\phi(f) = \sum_{i=1}^k \phi(f_i e_i) = \sum_{i=1}^k \phi_{e_i}(f_i) = \sum_{i=1}^k \int_X f_i \eta_i \, d\mu = \int_X \langle f, \eta \rangle \, d\mu.$$

Then we consider $\tilde{\mu} = \mu \llcorner |\eta|$ and

$$\tilde{\eta}(x) = \begin{cases} \eta(x)/|\eta(x)| & \text{if } \eta(x) \neq 0, \\ 0 & \text{if } \eta(x) = 0. \end{cases}$$

It follows that $|\tilde{\eta}(x)| = 1$ for $\tilde{\mu}$-almost all $x \in X$, and

$$\phi(f) = \int_X \langle f, \eta \rangle \, d\mu = \int_X \langle f, \tilde{\eta} \rangle \, |\eta| \, d\mu = \int_X \langle f, \tilde{\eta} \rangle \, d\tilde{\mu}.$$

This proves existence. Moreover, the proof of uniqueness now implies $\tilde{\mu} = |\phi| = \mu$, hence $|\eta| = 1$ μ-almost everywhere, and μ, η solve the representation problem.

We now address the extension problem. To estimate ϕ_v we introduce the functional

$$\varphi : C^0_c(X, \mathbb{R}^+_0) \to \mathbb{R}^+_0, \; \varphi(f) = \sup\{\phi(g) : g \in C^0_c(X, \mathbb{R}^k), \; |g| \leq f\}.$$

We claim that

$$\mu(U) = \sup\{\varphi(\chi) : \chi \in C^0_c(X, \mathbb{R}^+_0), \; \operatorname{spt}\chi \subset U, \; \chi \leq 1\}. \tag{6.4}$$

Namely for $g \in C^0_c(X, \mathbb{R}^k)$ with $\operatorname{spt} g \subset U$ and $|g| \leq 1$ we have

$$\phi(g) \leq \varphi(|g|) \leq \sup\{\varphi(\chi) : \chi \in C^0_c(X, \mathbb{R}^+_0), \; \operatorname{spt}\chi \subset U, \; \chi \leq 1\}.$$

On the other hand, for $\chi \in C^0_c(X, \mathbb{R}^+_0)$ with $\chi \leq 1$ we have

$$\varphi(\chi) = \sup\{\phi(g) : g \in C^0_c(X, \mathbb{R}^k), \; |g| \leq \chi\} \leq \mu(U).$$

Claim 1: φ is a half-linear functional on $C^0_c(X)$, which means that

$$\begin{aligned} \varphi(\alpha f) &= \alpha \varphi(f) && \text{for } f \in C^0_c(X, \mathbb{R}^+_0), \; \alpha \geq 0, \\ \varphi(f_1 + f_2) &= \varphi(f_1) + \varphi(f_2) && \text{for } f_{1,2} \in C^0_c(X, \mathbb{R}^+_0). \end{aligned}$$

Proof. The first line is by definition. For the second let $g_{1,2} \in C^0_c(X, \mathbb{R}^k)$ with $|g_i| \leq f_i$ and $\phi(g_i) \geq \varphi(f_i) - \varepsilon$. Then we have, choosing the sign appropriately,

$$\varphi(f_1) + \varphi(f_2) - 2\varepsilon \leq |\phi(g_1)| + |\phi(g_2)| = |\phi(g_1 \pm g_2)| \leq \varphi(f_1 + f_2).$$

For the reverse inequality let $g \in C^0_c(X, \mathbb{R}^k)$ be given with $|g| \leq f_1 + f_2$. Consider

$$g_i = \begin{cases} \frac{f_i}{f_1 + f_2} g & \text{falls } f_1 + f_2 > 0, \\ 0 & \text{sonst.} \end{cases}$$

Then $|g_i| \leq f_i$, in particular $g_i \in C^0_c(X, \mathbb{R}^k)$. As $g = g_1 + g_2$ we have

$$|\phi(g)| \leq |\phi(g_1)| + |\phi(g_2)| \leq \varphi(f_1) + \varphi(f_2),$$

hence $\varphi(f_1 + f_2) \leq \varphi(f_1) + \varphi(f_2)$.

Claim 2: We have $\varphi(f) = \int_X f \, d\mu$ for all $f \in C^0_c(X, \mathbb{R}^+_0)$.

Proof. For $\varepsilon > 0$ we take numbers $0 = t_0 < \ldots < t_N < \infty$ with

$$|t_i - t_{i-1}| < \varepsilon, \quad \max f \in (t_{N-1}, t_N) \quad \text{and} \quad \mu(f^{-1}\{t_i\}) = 0 \text{ für } i = 1, \ldots, N.$$

As $\mu(\mathrm{spt} f) < \infty$, the set of $t > 0$ with $\mu(f^{-1}\{t\}) > 0$ is at most countable, so that the choice of the t_i is possible. We put $U_i = f^{-1}(t_{i-1}, t_i)$ for $i = 1, \ldots, N$, these are open sets.

Estimate of $\varphi(f)$ from below:
Let $\chi_i \in C^0_c(X, \mathbb{R}^+_0)$ with $\mathrm{spt}\chi_i \subset U_i$, $\chi_i \leq 1$ for $i = 1, \ldots, N$. Then $\sum_{i=1}^N t_{i-1}\chi_i \leq f$. Now φ is monotone by definition, hence we obtain

$$\sum_{i=1}^N t_{i-1}\varphi(\chi_i) = \varphi\Big(\sum_{i=1}^N t_{i-1}\chi_i\Big) \leq \varphi(f).$$

We take the supremum with respect to the χ_i. From (6.4) we get $\sum_{i=1}^N t_{i-1}\mu(U_i) \leq \varphi(f)$, and further

$$\int_X f\,d\mu \leq \sum_{i=1}^N t_i\mu(U_i) \leq \sum_{i=1}^N (t_{i-1} + \varepsilon)\mu(U_i) \leq \varphi(f) + \varepsilon\mu(\mathrm{spt} f).$$

Estimate of $\varphi(f)$ from above:
For $i = 1, \ldots, N$ choose $V_i \supset \overline{U}_i$ open with $\mu(V_i) \leq \mu(\overline{U}_i) + \frac{\varepsilon}{N}$. There exist $\chi_i \in C^0_c(V_i)$ such that $\mathrm{spt}\chi_i \subset V_i$, $0 \leq \chi_i \leq 1$ und $\chi_i \equiv 1$ on $\overline{U}_i$. Then $f \leq \sum_{i=1}^N t_i\chi_i$ μ-almost everywhere, hence we can estimate

$$\begin{aligned}
\varphi(f) &\leq \sum_{i=1}^N t_i\varphi(\chi_i) \\
&\leq \sum_{i=1}^N (t_{i-1} + \varepsilon)\mu(V_i) \quad \text{(by (6.4))} \\
&\leq \sum_{i=2}^N (t_{i-1} + \varepsilon)\big(\mu(\overline{U}_i) + \frac{\varepsilon}{N}\big) + \varepsilon\big(\mu(\overline{U}_1) + \frac{\varepsilon}{N}\big) \\
&\leq \sum_{i=2}^N t_{i-1}\,\mu(\overline{U}_i) + \varepsilon\sum_{i=1}^N \mu(\overline{U}_i) + \frac{\varepsilon}{N}\sum_{i=2}^N t_{i-1} + C\varepsilon^2 \\
&\leq \int f\,d\mu + \varepsilon\,\mu(\mathrm{spt} f) + \varepsilon\,\|f\|_{C^0(X)} + C\varepsilon.
\end{aligned}$$

Claim 2 follows by letting $\varepsilon \searrow 0$ in both estimates. Finally, for $f \in C^0_c(X)$ we have $\phi_v(f) = \phi_v(f^+) - \phi_v(f^-)$. We conclude

$$|\phi_v(f)| \leq |\phi_v(f^+)| + |\phi_v(f^-)| \leq \varphi(f^+) + \varphi(f^-) = \int_X |f|\,d\mu.$$

Thus ϕ_v extends to a continuous functional on $L^1(\mu)$, which completes the proof. □

It remains for us to prove Lusin's theorem. We will apply th following classical extension result.

Lemma 6.6 (Tietze)**.** *Let $C \subset (X,d)$ be closed and $f : C \to \mathbb{R}$ be continuous. Then there exists an extension $\tilde{f} \in C^0(X)$ such that $\|\tilde{f}\|_{C^0(X)} = \sup_{x \in C} |f(x)|$.*

Proof. We may assume $1 \leq f \leq 2$, otherwise consider $2 + \frac{2}{\pi} \arctan f$. We now define

$$\tilde{f}(x) = \inf_{y \in C} f(y) \frac{d(x,y)}{d(x,C)} \quad \text{if } x \in X \backslash C, \text{ or equivalently } d(x,C) > 0.$$

We have $\inf f \leq \tilde{f} \leq \sup f$: the lower bound follows since $d(x,y)/d(x,C) \geq 1$, the upper bound is obtained by choosing $y \in C$ with $d(x,y) < (1+\varepsilon)d(x,C)$. We show the continuity of $\tilde{f}$ first for a point $x_0 \in \partial C$. Let $x \in B_\delta(x_0) \backslash C$ and $y \in C$ with $d(x,y) < (1+\alpha)d(x,C)$ where $\alpha \in (0,1]$. Then

$$d(x_0,y) \leq d(x_0,x) + d(x,y) < d(x_0,x) + (1+\alpha)d(x,C) < (2+\alpha)d(x_0,x) \leq 3\delta.$$

Putting $\varepsilon(\delta) = \sup_{|y-x_0| \leq \delta} |f(y) - f(x_0)|$ we obtain the estimate

$$f(x_0) - \varepsilon(3\delta) \leq f(y) \frac{d(x,y)}{d(x,C)} \leq (1+\alpha) f(x_0) + 2\varepsilon(3\delta).$$

For $d(x,y) \geq 2d(x,C)$ we have the trivial bound

$$f(y) \frac{d(x,y)}{d(x,C)} \geq 2 \inf f \geq f(x_0).$$

Taking $\alpha = 1$ we obtain the lower bound

$$\tilde{f}(x) \geq f(x_0) - \varepsilon(3\delta) \to f(x_0) \quad \text{as } \delta \searrow 0.$$

On the other hand, for any $\alpha > 0$ there exists $y \in C$ such that $d(x,y) < (1+\alpha)d(x,C)$, so that by letting $\alpha \searrow 0$ we obtain the upper bound

$$\tilde{f}(x) \leq f(x_0) + 2\varepsilon(3\delta) \to f(x_0) \quad \text{as } \delta \searrow 0.$$

The continuity on $X \backslash C$ follows easily from the triangle inequlity: using $1 \leq f \leq 2$ we have

$$\inf_{y \in C} f(y) d(x_1,y) \leq \inf_{y \in C} f(y) d(x_2,y) + 2d(x_1,x_2).$$

Thus $\inf_{y \in C} f(y) d(x,y)$ is Lipschitz on $X \backslash C$. The same holds for $d(x,C)$, and therefore $\tilde{f}$ is continuous on $X \backslash C$ as quotient of continuous functions. ☐

Proof. (*Theorem 6.5*) Consider the sets

$$A_{j,k} = \{x \in A : \frac{k}{j} \leq f(x) < \frac{k+1}{j}\} \quad \text{für } j \in \mathbb{N}, k \in \mathbb{Z}.$$

As $\mu \llcorner A$ is a Radon measure, we can choose compact sets $K_{j,k} \subset A_{j,k}$ such that $\mu(A_{j,k} \backslash K_{j,k}) < 2^{-j-|k|}\varepsilon/3$. Hence

$$\lim_{N \to \infty} \mu\Big(\bigcup_{k=-N}^{N} K_{j,k}\Big) = \mu\Big(A \backslash \bigcup_{k=-\infty}^{\infty} K_{j,k}\Big) < \sum_{k=-\infty}^{\infty} 2^{-j-|k|}\varepsilon/3 = 2^{-j}\varepsilon.$$

For $N_j \in \mathbb{N}$ sufficiently large we then have $\mu(K_j) < 2^{-j}\varepsilon$ where $K_j = \bigcup_{k=-N_j}^{N_j} K_{j,k}$. Thus for $K = \bigcap_{j=1}^{\infty} K_j$ we obtain

$$\mu(A \backslash K) \leq \mu\Big(\bigcup_{j=1}^{\infty} A \backslash K_j\Big) \leq \sum_{j=1}^{\infty} \mu(A \backslash K_j) < \varepsilon.$$

Now consider the functions $f_j : A \to \mathbb{R}$, $f_j(x) = \frac{j}{k}$ for $x \in A_{j,k}$. By compactness the sets $K_{j,k} \subset A_{j,k}$ are at positive distance, so that f_j is locally constant on $K_j \supset K$, in particular continuous. But $|f(x) - f_j(x)| \leq \frac{1}{j} \to 0$ as $j \to \infty$, hence the uniform limit $f|_K$ is continuous. Using the Tietze extension from Lemma 6.6 the theorem follows. □

We denote by $C_c^0(X)'$ the space of linear functionals Λ on $C_c^0(X)$, i.e. Λ is linear and for any compact set $K \subset X$ we have

$$|\Lambda(f)| \leq C(K)\, \|f\|_{C^0(X)} \quad \text{for all } f \in C_c^0(X) \text{ with } \operatorname{spt} f \subset K.$$

Definition 6.7. *Let $\Lambda_k, \Lambda \in C_c^0(X)'$. We say that $\Lambda_k \to \Lambda$ weak* in $C_c^0(X)'$ if*

$$\Lambda_k(f) \to \Lambda(f) \quad \textit{for all } f \in C_c^0(X).$$

Any Radon measure induces a (nonnegative) linear functional by integration. The corresponding weak convergence of Radon measures, notation $\mu_k \to \mu$, is given by

$$\int_X f\, d\mu_k \to \int_X f\, d\mu \quad \text{for all } f \in C_c^0(X).$$

A simple approximation argument shows that if μ_k converges weakly then the weak limit μ is unique. Moreover then $\mu_k(K)$ is bounded for any compact set $K \subset X$.

Lemma 6.8. *Let X be a locally compact, separable metric space. For Radon measures μ_k and μ on X the following statements are equivalent:*

(1) *$\mu_k \to \mu$ weakly as Radon measures,*

(2) *For all open U and all compact K we have*

$$\mu(U) \leq \liminf_{k\to\infty} \mu_k(U) \quad \textit{and} \quad \mu(K) \geq \liminf_{k\to\infty} \mu_k(K),$$

(3) *For any bounded Borel set B with $\mu(\partial B) = 0$ we have $\mu(B) = \lim_{k\to\infty} \mu_k(B)$.*

Proof. The claims are proved step by step.
(1) ⇒ (2): Let U be open. For compact $K \subset U$ choose $\chi \in C_c^0(X)$ with $0 \leq \chi \leq 1$, $\operatorname{spt}\chi \subset U$ and $\chi = 1$ on K. Then

$$\mu(K) \leq \int_X \chi\, d\mu = \lim_{k\to\infty} \int_X \chi\, d\mu_k \leq \liminf_{k\to\infty} \mu_k(U).$$

Taking the supremum over all such K the lower semicontinuity for open U follows. Now for given compact K we choose $U \supset K$ open and then χ as above, we obtain

$$\mu(U) \geq \int_X \chi\, d\mu = \lim_{k\to\infty} \int_X \chi\, d\mu_k \geq \limsup_{k\to\infty} \mu_k(K).$$

Taking the infimum over all such U proves the upper semicontinuity for compact K.

(2) $\Rightarrow$ (3): Note that $\overline{B}$ is compact. Therefore we can estimate

$$\mu(B) = \mu(\operatorname{int} B) \leq \liminf_{k\to\infty} \mu_k(\operatorname{int} B) \leq \limsup_{k\to\infty} \mu_k(\overline{B}) \leq \mu(\overline{B}) = \mu(B).$$

(3) $\Rightarrow$ (1): We may assume $f \in C^0_c(X, \mathbb{R}^+_0)$, otherwise we decompose $f = f^+ - f^-$. Let $B = B_R(x_0)$ such that $\operatorname{spt} f \subset B$ and $\mu(\partial B) = 0$. Then choose $0 = t_0 < \ldots < t_N$ with $|t_i - t_{i-1}| < \varepsilon$, $t_N > \max f$ and $\mu(\{f = t_i\}) = 0$ for $i = 1, \ldots, N$. Putting $A_i = \{t_{i-1} \leq f \leq t_i\}$ for $i = 2, \ldots, N$ we have

$$\sum_{i=2}^N t_{i-1}\chi_{A_i} \leq f \leq t_1\chi_B + \sum_{i=2}^N t_i\chi_{A_i}, \qquad \text{thus}$$

Integrating with respect to μ_k yields

$$\sum_{i=2}^N t_{i-1}\mu_k(A_i) \leq \int f\, d\mu_k \leq t_1\mu_k(B) + \sum_{i=2}^N t_i\mu_k(A_i).$$

Letting $k \to \infty$ we see that

$$\limsup_{k\to\infty} \int f\, d\mu_k \leq t_1\mu(B) + \sum_{i=2}^N t_i\mu(A_i) \leq \sum_{i=2}^N t_{i-1}\mu(A_i) + 2\varepsilon\, \mu(B) \leq \int f\, d\mu + 2\varepsilon\, \mu(B).$$

The lower bound follows similarly,

$$\liminf_{k\to\infty} \int f\, d\mu_k \geq \sum_{i=2}^N t_{i-1}\mu_k(A_i) \geq \sum_{i=2}^N t_i\mu_k(A_i) + t_1\mu(B) - 2\varepsilon\, \mu(B) \geq \int f\, d\mu - 2\varepsilon\, \mu(B).$$

Now (1) follows by letting $\varepsilon \searrow 0$ in both estimates. $\square$

The following is an important application of the Riesz representation theorem.

Theorem 6.9 (Compactness for Radon measures). *Let (X, d) be a locally compact, separable metric space, and let μ_k be a sequence of Radon measures on X such that*

$$\sup_{k\in\mathbb{N}} \mu_k(K) < \infty \quad \textit{for all compact } K \subset X.$$

Then there exists a Radon measure μ and a subsequence $\mu_{k'}$ such that $\mu_{k'} \to \mu$ weakly as Radon measures on X.

To prove the theorem we need the following fact.

Lemma 6.10. *Let (X, d) be locally compact and separable. Then $C^0_c(X)$ is separable.*

Proof. We first assume that X is compact. For $\varrho > 0$ we choose a covering $B_\varrho(x_j)$, $1 \leq j \leq N$, and define the partition of unity

$$\chi_j = \frac{\tilde{\chi}_j}{\sum_{j=1}^N \tilde{\chi}_j} \qquad \text{where } \tilde{\chi}_j(x) = \Big(1 - \frac{\operatorname{dist}\big(x, B_\varrho(x_j)\big)}{\varrho}\Big)^+$$

For given $f \in C^0(X)$ and any $x \in X$ we estimate, observing $\chi_j(x) = 0$ for $d(x, x_j) \geq 2\varrho$,

$$\Big| f(x) - \sum_{j=1}^{N} f(x_j)\chi_j(x) \Big| = \Big| \sum_{j=1}^{N} \big(f(x) - f(x_j)\big)\chi_j(x) \Big| \leq \operatorname{osc}(f, 2\varrho).$$

Taking $\varrho_k = \frac{1}{k}$ we obtain functions $\chi_{j,k}$, $1 \leq j \leq N_k$. Linear combinations of the $\chi_{j,k}$ with coefficients in $\mathbb{Q}$ are then dense. For general X the result follows by choosing an exhaustion by a sequence of compact subsets. $\square$

Proof. (of Theorem 6.9) We assume that X is compact, thus we have by assumption

$$C := \sup_{k \in \mathbb{N}} \mu_k(X) < \infty.$$

Choose a dense set of functions $\varphi_j \in C^0(X)$, $j \in \mathbb{N}$. We have

$$\sup_{k \in \mathbb{N}} \Big| \int_X \varphi_j \, d\mu_k \Big| \leq C \, \|\varphi_j\|_{C^0(X)}.$$

Taking successive subsequences and then passing to the diagonal sequence we obtain

$$\exists \lim_{k \to \infty} \int_X \varphi_j \, d\mu_k \quad \text{for all } j \in \mathbb{N}.$$

On $D = \operatorname{Span}\{\varphi_j : j \in \mathbb{N}\}$ we obtain the function

$$\Lambda : D \to \mathbb{R},\ \Lambda(\varphi) = \lim_{k \to \infty} \int_X \varphi \, d\mu_k.$$

Λ is well-defined, linear and satisfies the bound

$$|\Lambda(\varphi)| = \lim_{k \to \infty} \Big| \int_X \varphi \, d\mu_k \Big| \leq C \, \|\varphi\|_{C^0(X)}.$$

By density, Λ has a unique extension to a linear functional, also denoted by Λ, in $C^0(X)'$ with norm $\|\Lambda\| \leq C$. Now by Theorem 6.4 there exists a Radon measure μ and a μ-measurable function $\sigma : X \to \{\pm 1\}$ such that

$$\Lambda(\varphi) = \int_X \varphi \, \sigma d\mu \quad \text{for all } \varphi \in C^0(X).$$

We claim that $\sigma = 1$ μ-almost everywhere. If $K \subset \{\sigma = -1\}$ is compact, then

$$\varphi(x) = \Big(1 - \frac{\operatorname{dist}(x, K)}{\varepsilon}\Big)^+ \quad \Rightarrow \quad 0 \leq \Lambda(\varphi) \leq -\mu(K) + \mu\big(U_\varepsilon(K) \backslash K\big).$$

Letting $\varepsilon \searrow 0$ we have $\mu\big(U_\varepsilon(K) \backslash K\big) \to 0$ and hence $\mu(K) = 0$. As $\{\sigma = -1\}$ is μ-measurable it must be a null set, i.e.

$$\Lambda(\varphi) = \int_X \varphi \, d\mu \quad \text{for all } \varphi \in C^0(X).$$

Finally, for any $\varphi \in C^0(X)$ and $\varphi_0 \in D$ we have

$$\Big|\Lambda(\varphi) - \int \varphi \, d\mu_k\Big| \leq \Big|\Lambda(\varphi) - \Lambda(\varphi_0)\Big| + \Big|\Lambda(\varphi_0) - \int \varphi_0 \, d\mu_k\Big| + \Big|\int \varphi_0 \, d\mu_k - \int \varphi \, d\mu_k\Big|.$$

Letting $k \to \infty$ we conclude

$$\limsup_{k\to\infty} \Big|\Lambda(\varphi) - \int \varphi \, d\mu_k\Big| \leq 2C \, \|\varphi - \varphi_0\|_{C^0(X)}.$$

By density of the set D we conclude that $\mu_k \to \mu$ weakly as Radon measures. $\square$

It would habe been clearer to state two results: first the sequential compactness in the space $C^0_c(X)'$, and second the fact that if the sequence is induced by Radon measures, then the limit is again a Radon measure.

Chapter 7

Lipschitz functions

In this short section we discuss two basic results about Lipschitz functions.

Theorem 7.1 (Lipschitz extension)**.** *Let* (X,d) *be a metric space. Assume that* $f : A \to \mathbb{R}^k$ *is Lipschitz where* $A \subset X$*. Then there exists a Lipschitz map* $\tilde{f} : X \to \mathbb{R}^k$ *such that* $\tilde{f} = f$ *on* A *and* $\mathrm{Lip}(\tilde{f}) \leq \sqrt{k}\,\mathrm{Lip}(f)$*.*

Proof. We first consider the case $k = 1$, and define

$$\tilde{f} : X \to \mathbb{R},\ \tilde{f}(x) = \inf_{a \in A} f_a(x) \quad \text{ where } f_a(x) = f(a) + \mathrm{Lip}(f)\, d(x,a).$$

We have $\mathrm{Lip}(f_a) = \mathrm{Lip}(f)$, this implies for any $a \in A$

$$\tilde{f}(x) \leq f_a(x) \leq f_a(y) + \mathrm{Lip}(f)\, d(x,y).$$

Taking the infimum with respect to a shows $\mathrm{Lip}(\tilde{f}) \leq \mathrm{Lip}(f)$. Now for $b \in A$ we have

$$f(b) \leq f(a) + \mathrm{Lip}(f)\, d(b,a) = f_a(b) \quad \text{ for all } a \in A,$$

and equality is attained for $a = b$. This hows $\tilde{f} = f$ on A. In the vector-valued case, we apply this extension to each component f_i, and conclude

$$|\tilde{f}(x) - \tilde{f}(y)| = \Big(\sum_{i=1}^k \big(\tilde{f}_i(x) - \tilde{f}_i(y) \big)^2 \Big)^{\frac{1}{2}} \leq \sqrt{k}\,\mathrm{Lip}(f)\, d(x,y).$$

$\square$

In the case $X = \mathbb{R}^n$ a result of Kirszbraun asserts the existence of an extension with the same Lipschitz constant also in the vector-valued case. However this is not needed in the sequel.

Theorem 7.2 (Rademacher)**.** *Let* $f : \mathbb{R}^n \to \mathbb{R}$ *be locally Lipschitz. Then* f *has a classical derivative* $Df(x)$ *at* $\mathcal{L}^n$*-almost every point* $x \in \mathbb{R}^n$*.*

The following fact is of independant interest.

Lemma 7.3. *If* $f : \mathbb{R}^n \to \mathbb{R}$ *is locally Lipschitz, then* f *has weak derivatives* $\partial_i f \in L^\infty_{\mathrm{loc}}(\mathbb{R}^n)$*.*

Proof. We may assume that $\mathrm{Lip}(f) = L < \infty$. We consider the difference quotient operator

$$\partial_i^h f(x) = \frac{f(x+he_i) - f(x)}{h} \quad \text{for } i = 1, \ldots, n.$$

By substitution we infer the integration by parts formula, for $\zeta \in C_c^\infty(\mathbb{R}^n)$ say,

$$\int_{\mathbb{R}^n} \partial_i^h f(x)\zeta(x)\,dx = \int_{\mathbb{R}^n} f(x)\partial_i^{-h}\zeta(x)\,dx.$$

Now $\|\partial_i^h f\|_{L^\infty(\mathbb{R}^n)} \leq L$ by assumption, thus we find a sequence $h_k \to 0$ such that $\partial_i^h f \to g_i$ weak$*$ in $L^\infty(\mathbb{R}^n) = L^1(\mathbb{R}^n)'$, moreover $\|g_i\|_{L^\infty(\mathbb{R}^n)} \leq L$ by lower semicontiuity of the dual norm. Passing to the limit in the integration by parts formula yields

$$\int_{\mathbb{R}^n} g_i(x)\,\zeta(x)\,dx = -\int_{\mathbb{R}^n} f(x)\,\partial_i\zeta(x)\,dx.$$

The means that f has the weak derivative $\partial_i f = g_i$. □

In the one-dimensional case, for given $a < b$ and $h > 0$ small, we can take as test function

$$\zeta_h(x) = \begin{cases} \frac{1}{h}(x-a) & \text{on } [a, a+h], \\ \frac{1}{h}(b-x) & \text{on } [b-h, b], \\ 1 & \text{on } [a+h, b-h]. \end{cases}$$

Denoting by $g = f'$ the weak derivative, we obtain

$$\int_a^b g(x)\,\zeta_h(x)\,dx = \frac{1}{h}\Big(\int_{b-h}^b f(x)\,dx - \int_a^{a+h} f(x)\,dx\Big).$$

Letting $h \searrow 0$ we deduce the fundamental theorem of calculus, in the form

$$\int_a^b g(x)\,dx = f(b) - f(a).$$

By Lebesgue differentiation, see (5.1), we conclude for almost every $x \in \mathbb{R}$

$$\begin{aligned} \Big|\frac{f(x+h) - f(x)}{h} - g(x)\Big| &= \frac{1}{|h|}\Big|\int_x^{x+h} (g(y) - g(x))\,dy\Big| \\ &\leq \frac{1}{|h|}\int_{|y-x|\leq|h|} |g(y) - g(x)|\,dy \overset{h\to 0}{\longrightarrow} 0. \end{aligned}$$

This shows Rademacher's theorem in the case $n = 1$.

Proof. (*Theorem 7.2*) We assume $L := \mathrm{Lip}(f) < \infty$. For $v \in \mathbb{R}^n$ with $|v| = 1$ we introduce the upper and lower derivatives, taking values in $[-L, L]$,

$$\overline{D}_v f(x) = \limsup_{t\to 0} \frac{f(x+tv) - f(x)}{t} \quad \text{and} \quad \underline{D}_v f(x) = \liminf_{t\to 0} \frac{f(x+tv) - f(x)}{t}.$$

As supremum/infimum of continuous functions, $\overline{D}_v f$ and $\underline{D}_v f$ are lower/upper semicontinuous, in particular Borel measurable. We introduce the bad set

$$E_v = \{x \in \mathbb{R}^n : \underline{D}_v f(x) < \overline{D}_v f(x)\}.$$

By the one-dimensional case, the set $E_v \cap (y + \mathbb{R}v)$ has zero $\mathcal{L}^1$ measure for any $y \in v^\perp$. Using Fubini's theorem we obtain

$$\mathcal{L}^n(E_v) = \int_{y \in v^\perp} \mathcal{L}^1\big(E_v \cap (y + \mathbb{R}v)\big)\, d\mathcal{L}^{n-1}(y) = 0.$$

We next show that $D_v f = \sum_{i=1}^n v_i D_{e_i} f$ almost everywhere. For this we use that both functions are weak derivatives, in fact for any $\zeta \in C_c^\infty(\mathbb{R}^n)$ we have by dominated convergence

$$\begin{aligned}
\int_{\mathbb{R}^n} D_v f(x)\zeta(x)\, dx &= \lim_{t \searrow 0} \int_{\mathbb{R}^n} \frac{f(x+tv) - f(x)}{t}\, \zeta(x)\, dx \\
&= \lim_{t \searrow 0} \int_{\mathbb{R}^n} f(x) \frac{\zeta(x-tv) - \zeta(x)}{t}\, dx \\
&= -\int_{\mathbb{R}^n} f(x) D_v \zeta(x)\, dx \\
&= -\sum_{i=1}^n v_i \int_{\mathbb{R}^n} f(x) D_{e_i}\zeta(x)\, dx \\
&= \int_{\mathbb{R}^n} \Big(\sum_{i=1}^n v_i D_{e_i} f(x) \Big)\, \zeta(x)\, dx.
\end{aligned}$$

Our claim follows by the fundamental lemma of the calculus of variations. Note that the calculation also shows that the pointwise derivative $D_v f$ is also the weak derivative. Now let v_k, $k \in \mathbb{N}$, be dense in $\mathbb{S}^{n-1}$ and define G as the set of $x \in \mathbb{R}^n$ with the following conditions:

- $D_{e_i} f(x)$ exists for $i = 1, \ldots, n$,
- $D_{v_k} f(x)$ exists for all $k \in \mathbb{N}$,
- $D_{v_k} f(x) = \sum_{i=1}^n v_k^i D_{e_i} f(x)$ for all $k \in \mathbb{N}$.

We have proved that $\mathbb{R}^n \backslash G$ is a null set. We claim that $Df(x) = g(x)$ for $x \in G$ in the classical sense, where $g(x) = \sum_{i=1}^n D_{e_i} f(x) e_i$. Put

$$Q(v,t) = \frac{f(x+tv) - (f(x) + \langle g(x), tv \rangle}{t} \qquad \text{for any } v \in \mathbb{S}^{n-1}.$$

The equation $Df(x) = g(x)$ follows if we show that

$$\sup_{|v|=1} |Q(v,t)| \to \qquad \text{as } t \to 0. \tag{7.1}$$

In fact then we can estimate

$$\frac{|f(x+h) - (f(x) + \langle g(x), h \rangle)|}{|h|} = \Big| Q\Big(\frac{h}{|h|}, |h|\Big) \Big| \leq \sup_{|v|=1} |Q(v,|h|) \to 0 \qquad \text{as } h \to 0.$$

For $x \in G$ we have $Q(v_i, t) \to 0$ as $t \to 0$. Moreover

$$|Q(v,t) - Q(w,t)| \leq \frac{|f(x+tv) - f(x+tw)|}{|t|} + |\langle g(x), v-w \rangle| \leq (1 + \sqrt{n})\, L\, |v-w|.$$

Given $\varrho > 0$ there exists an $N \in \mathbb{N}$ such that $\mathbb{S}^{n-1} \subset \bigcup_{i=1}^N B_\varrho(v_i)$. For $v \in \mathbb{S}^{n-1}$ we choose $i \in \{1, \dots, N\}$ with $v \in B_\varrho(v_i)$ and estimate

$$|Q(v,t)| \leq |Q(v_i,t)| + (1+\sqrt{n})\, L\, |v - v_i| \leq \max_{1 \leq j \leq N} |Q(v_j,t)| + (1+\sqrt{n}) L \varrho.$$

Taking the supremum with respect to $v \in \mathbb{S}^{n-1}$ and then letting $t \to 0$ we obtain

$$\limsup_{t \to 0} \sup_{|v|=1} |Q(v,t)| \leq (1+\sqrt{n}) L \varrho.$$

Letting now $\varrho \searrow 0$ shows (7.1), and the theorem is proved. □

Corollary 7.4. *The following statements hold.*

(1) *If $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz, then $Df(x) = 0$ a. e. on $\{f = 0\}$.*

(2) *If $f, g : \mathbb{R}^n \to \mathbb{R}^n$ are Lipschitz, then $Dg(f(x))Df(x) = \mathrm{Id}$ a. e. on $\{g \circ f = \mathrm{id}\}$.*

Proof. Let $\partial_i f \in L^\infty(\mathbb{R}^n)$ be the weak derivatives, see Lemma 7.3. We claim that

$$\partial_i f^+ = \chi_{\{f>0\}}\, \partial_i f \quad \text{and} \quad \partial_i f^- = -\chi_{\{f<0\}}\, \partial_i f.$$

Fom this we get almost everywhere

$$\partial_i f = \partial_i f^+ - \partial_i f^- = 0 \quad \text{on the set } \{f = 0\}.$$

To compute $\partial_i f^+$ we approximate by $\chi_\varepsilon \circ f$ where

$$\chi_\varepsilon(s) = \begin{cases} \sqrt{s^2 + \varepsilon^2} - \varepsilon & \text{for } s \geq 0, \\ 0 & \text{for } s \leq 0. \end{cases}$$

Note that $\chi_\varepsilon \in C^1(\mathbb{R})$ with derivative

$$\chi_\varepsilon'(s) = \begin{cases} \frac{s}{\sqrt{s^2+\varepsilon^2}} & \text{for } s \geq 0, \\ 0 & \text{for } s \leq 0. \end{cases}$$

Using mollification of f we verify the weak chain rule, for any $\varphi \in C^\infty_c(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} \chi_\varepsilon \circ f\, \partial_i \varphi = -\int_{\mathbb{R}^n} \chi_\varepsilon' \circ f\, (\partial_i f) \varphi.$$

Letting $\varepsilon \searrow 0$ we conclude

$$\int_{\mathbb{R}^n} f^+\, \partial_i \varphi = -\int_{\mathbb{R}^n} \chi_{\{f>0\}}\, \partial_i f\, \varphi.$$

The formula for f^- follows by using $f^- = (-f)^+$, which completes the proof of claim (1).

In the second claim, we know by Rademacher that the sets E_f and E_g where the derivatives don't exist are null sets. By the classical chain rule, we have

$$D(g \circ f)(x) = Dg(f(x)Df(x) \quad \text{for all } x \notin E_f \cup \{x : f(x) \in E_g\}.$$

But $g(f(x)) = x$ and $f(x) \in E_g$ implies $x \in g(E_g)$, which is also a null set. (2) now follows from (1), applied to the function $(g \circ f)(x) - x$. □

Chapter 8

The area formula

In this section we consider Lipschitz maps $f : U \to \mathbb{R}^m$ where $U \subset \mathbb{R}^n$ and $n \leq m$. The goal is the area formula, which computes the $\mathcal{H}^n$-measure of the image, counted appropriately with multiplicities, in terms of the Jacobian integral.

Definition 8.1. *Let $f : U \to \mathbb{R}^m$ where $U \subset \mathbb{R}^n$ is open and $n \leq m$. If f is differentiable at $x \in U$ then the Jacobian is defined by*

$$Jf(x) = \sqrt{\det\, Df(x)^{\mathrm{T}} Df(x)}.$$

The matrix $Df(x)^{\mathrm{T}} Df(x) \in \mathbb{R}^{n \times n}$ is symmetric and positive semi-definit. In fact, with respect to the standard scalar products on $\mathbb{R}^n$ and $\mathbb{R}^m$ we have

$$\langle Df(x)^{\mathrm{T}} Df(x) v, v\rangle = |Df(x)v|^2 \quad \text{for any } v \in \mathbb{R}^n.$$

In particular $Jf(x) > 0$ if and only if $Df(x)$ has rank n.

Lemma 8.2. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $f(x) = y_0 + Lx$, be affine-linear where $n \leq m$. Then*

$$\mathcal{H}^n(f(A)) = JL\, \mathcal{L}^n(A) \quad \textit{for all } A \subset \mathbb{R}^n.$$

Proof. We assume $y_0 = 0$. The $\mathcal{H}^n$-measure is invariant under $S \in \mathbb{O}(m)$, and

$$J(SL) = \sqrt{\det (SL)^{\mathrm{T}} SL} = \sqrt{\det L^{\mathrm{T}} S^{\mathrm{T}} SL} = \sqrt{\det L^{\mathrm{T}} L} = JL.$$

Therefore we can assume $L = IM$ where $M \in \mathrm{End}(\mathbb{R}^n)$ and $I : \mathbb{R}^n \to \mathbb{R}^n \times \{0\} \subset \mathbb{R}^m$ is the inclusion map. In particular

$$JL = \sqrt{\det (IM)^{\mathrm{T}} IM} = \sqrt{\det M^{\mathrm{T}} I^{\mathrm{T}} IM} = \sqrt{\det M^{\mathrm{T}} M} = |\det M|.$$

By definition of Hausdorff measure, we see that $IM(A)$ and $M(A)$ have the same $\mathcal{H}^n$-measure. Using $\mathcal{H}^n = \mathcal{L}^n$ on $\mathbb{R}^n$, see Theorem 3.11, and the transformation formula we get

$$\mathcal{H}^n(IM(A)) = \mathcal{H}^n(M(A)) = \mathcal{L}^n(M(A)) = |\det M|\, \mathcal{L}^n(A).$$

$\square$

Lemma 8.3. *Let $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ where $n \leq m$, and let $x_0 \in \mathbb{R}^n$ with $Jf(x_0) > 0$. Then there exists a neighborhood U of x_0 with the following properties:*

(1) *$f|_U$ is injective,*

(2) *$f(A)$ is Borel for any Borel set $A \subset U$,*

(3) *$\mathcal{H}^n(f(A)) = \int_A Jf\, d\mathcal{L}^n$ for $A \subset U$ Borel.*

Proof. We first assume that f is a graph over $U \subset \mathbb{R}^n$, more precisely

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

Then f is trivially injective, moreover we have the diffeomorphism

$$F : U \times \mathbb{R}^{m-n} \to U \times \mathbb{R}^{m-n}, F(x,y) = (x, y + u(x)),$$

As $f(A) = F(A \times \{0\}$ we see that f maps Borel sets to Borel sets. Now let

$$\mu(E) = \mathcal{H}^n(f(E)) \quad \text{ for } E \subset U.$$

Clearly μ is an outer measure on U. For $A, B \subset U$ we have $\operatorname{dist}(f(A), f(B)) \geq \operatorname{dist}(A, B)$, thus $\operatorname{dist}(A, B) > 0$ implies (see also Lemma 3.2)

$$\mu(A \cup B) = \mathcal{H}^n(f(A) \cup f(B)) = \mathcal{H}^n(f(A)) + \mathcal{H}^n(f(B)) = \mu(A) + \mu(B).$$

To construct a Borel hull for $E \subset U$, we choose $B \supset f(E)$ Borel with $\mathcal{H}^n(B) = \mathcal{H}^n(f(E)) = \mu(E)$. We can assume $B \subset f(U)$, otherwise we pass to $B \cap f(U)$ which is again Borel. Now $\pi(B)$ is Borel and contains $\pi(f(E)) = E$, furthermore

$$\mu\big(\pi(B)\big) = \mathcal{H}^n(f(\pi(B)) = \mathcal{H}^n(B) = \mu(E).$$

Finally, for $K \subset U$ compact and $E \subset K$ we have, using Lemma 3.3 and Theorem 3.11,

$$\mu(E) = \mathcal{H}^n(f(E)) \leq L^n\, \mathcal{H}^n(E) = L^n\, \mathcal{L}^n(E) \quad \text{ where } L = \operatorname{Lip}(f|_K).$$

Thus μ is finite on compact subsets, and it is absolutely continuous with respect to $\mathcal{L}^n$. The lemma follows by Radon-Nikodym, Theorem 5.3, if we show that

$$D_{\mathcal{L}^n}\mu(x) = Jf(x) \quad \text{ for all } x \in U. \tag{8.1}$$

For fixed $x \in U$ we consider the remainder function

$$\varphi : U \to \mathbb{R}^{m-n},\ \varphi(y) = u(y) - \big(u(x) + Du(x)(y-x)\rangle\big).$$

Furthermore we define $\phi, \psi : U \times \mathbb{R}^{m-n} \to U \times \mathbb{R}^{m-n}$ by

$$\phi(y,z) = \big(y, z + \varphi(y)\big) \quad \text{ and } \quad \psi(y,z) = \big(y, z - \varphi(y)\big).$$

Clearly $\phi \circ \psi = \psi \circ \phi = \mathrm{id}_{U \times \mathbb{R}^{m-n}}$. Moreover

$$D\phi(y,z) = \begin{pmatrix} \mathrm{E}_n & 0 \\ D\varphi(y) & \mathrm{E}_{m-n} \end{pmatrix} \quad \text{ and } \quad D\psi(y,z) = \begin{pmatrix} \mathrm{E}_n & 0 \\ -D\varphi(y) & \mathrm{E}_{m-n} \end{pmatrix}.$$

As $D\varphi(x) = 0$ we have $D\phi(x,z) = D\psi(x,z) = \mathrm{E}_m$. Moreover given $\varepsilon > 0$ we can choose $\varrho > 0$ such that

$$|D\varphi(y)| = |Du(y) - Du(x)| < \varepsilon \quad \text{for all } y \in B_\varrho(x).$$

By the Lipschitz estimate, Lemma 3.3, we have

$$\frac{1}{(1+\varepsilon)^n}\,\mathcal{H}^n\big(\psi(f(B_\varrho(x)))\big) \leq \mathcal{H}^n\big(f(B_\varrho(x))\big) \leq (1+\varepsilon)^n\,\mathcal{H}^n\big(\psi(f(B_\varrho(x)))\big).$$

Now observe that $\psi \circ f$ is affine, in fact

$$\psi(f(y)) = (y, u(y) - \varphi(y)) = \big(y, u(x) + Du(x)(y-x)\big).$$

We compute $D(\psi \circ f)(x) = D\psi(x, u(x))Df(x) = Df(x)$. Using Lemma 8.2 we obtain

$$\mathcal{H}^n\big(\psi(f(B_\varrho(x)))\big) = Jf(x)\,\mathcal{L}^n(B_\varrho(x)).$$

Combining with the inequalities above shows (8.1).

It remains to give the reduction to the graphical case. By passing to $S(f - f(x_0))$ where $S \in \mathbb{O}(m)$, we can arrange that $f(x_0) = 0$ and $Df(x_0) = IM$ where $M \in \mathrm{GL}(\mathbb{R}^n)$ and $I : \mathbb{R}^n \to \mathbb{R}^m$ is the inclusion. Let $\pi : \mathbb{R}^m \to \mathbb{R}^n$ be the projection, then

$$D(\pi \circ f)(x_0) = \pi\, Df(x_0) = M.$$

By the inverse function theorem, there exists a neighborhood U of x_0 and a $\varrho > 0$ such that $\varphi = (\pi \circ f) : U \to B_\varrho(0)$ is a diffeomorphism. By definition

$$\pi \circ f \circ \varphi^{-1}(y) = y \quad \text{for } y \in B_\varrho(0).$$

Let $\pi^\perp : \mathbb{R}^m \to \mathbb{R}^{m-n}$ be the projection onto the last $m-n$ coordinates, and define

$$u : B_\varrho(0) \to \mathbb{R}^{m-n},\ u(y) = \pi^\perp \circ f \circ \varphi^{-1}(y).$$

It follows that $f \circ \varphi^{-1}(y) = (y, u(y)) =: g(y)$ for all $y \in B_\varrho(0)$. As $f = g \circ \varphi$ it is injective and maps Borel sets to Borel sets, moreover for $A \subset U$ we compute using $Jf(x) = Jg(\varphi(x))\,|\det D\varphi(x)|$ and the transformation formula

$$\mathcal{H}^n(f(A)) = \mathcal{H}^n(g(\varphi(A)) = \int_{\varphi(A)} Jg(y)\,d\mathcal{L}^n(y) = \int_A Jf(x)\,d\mathcal{L}^n(x).$$

□

Lemma 8.4. *Let $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ where $n \leq m$, and let $A \subset \{Jf > 0\}$ be Borel. Then the muliplicity $\mathcal{H}^0(A \cap f^{-1}\{y\})$ is Borel measurable on $\mathbb{R}^m$, and*

$$\int_A Jf(x)\,d\mathcal{L}^n(x) = \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\,d\mathcal{H}^n(y).$$

Proof. For $x \in A$ choose $U(x)$ as in Lemma 8.3. It is a fact that A is then covered by a countable collection $U(x_j)$, $j \in \mathbb{N}$ (for example, this follows by Vitali's theorem). Define

$$A_j = A \cap U(x_j) \backslash \bigcup_{i=1}^{j-1} U(x_i).$$

As $f|_{A_j}$ is injective we have $\mathcal{H}^0(A_j \cap f^{-1}\{y\}) = \chi_{f(A_j)}(y)$. The set $f(A_j)$ is Borel, hence the mulitiplicity of A_j is Borel measurable. We now compute by Lemma 8.3 and monotone convergence

$$\begin{aligned}
\int_A Jf\, d\mathcal{H}^n &= \sum_{j=1}^{\infty} \int_{A_j} Jf\, d\mathcal{H}^n \\
&= \sum_{j=1}^{\infty} \int_{\mathbb{R}^n} \mathcal{H}^0(A_j \cap f^{-1}\{y\})\, d\mathcal{H}^n(y) \\
&= \int_{\mathbb{R}^n} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n(y).
\end{aligned}$$

□

Lemma 8.5. *Let $f \in C^1(R^n, \mathbb{R}^m)$ where $n \leq m$. Then the image of the set $\{Jf = 0\}$ has $\mathcal{H}^n$-measure zero.*

Proof. We consider the C^1 immersion $f_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$, $f_\varepsilon(x) = (\varepsilon x, f(x))$. We compute

$$Df_\varepsilon(x)^{\mathrm{T}} Df_\varepsilon(x) = \varepsilon^2\, \mathrm{E_n} + Df(x)^{\mathrm{T}} Df(x).$$

The eigenvalues of $Df(x)^{\mathrm{T}} Df(x)$ are nonnegative, we denote them by $\lambda_1 \leq \ldots \leq \lambda_n$. Then $\lambda_i \leq |Df(x)|^2$ for $i = 1, \ldots, n$, moreover since $Jf(x) = 0$ we have $\lambda_1 = 0$. Thus

$$0 < \varepsilon^n \leq Jf_\varepsilon = \sqrt{\det\left(\varepsilon^2 \mathrm{E}_n + Df^{\mathrm{T}} Df\right)} \leq \left(\varepsilon^2 + |Df|^2\right)^{\frac{n-1}{2}} \varepsilon.$$

Now $f = \pi \circ f_\varepsilon$ where π is the projection onto $\mathbb{R}^m$. For $A = \{Jf = 0\} \cap B_R(0)$ we obtain by Lemma 3.3 and Lemma 8.4

$$\mathcal{H}^n(f(A)) \leq \mathcal{H}^n(f_\varepsilon(A)) \leq \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f_\varepsilon^{-1}\{y\})\, d\mathcal{H}^n(y) = \int_A Jf_\varepsilon\, d\mathcal{L}^n.$$

As $Jf_\varepsilon \to 0$ uniformly on A for $\varepsilon \searrow 0$ the claim of the lemma follows. □

Theorem 8.6 (area formula)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $n \leq m$, be locally Lipschitz, and let $A \subset \mathbb{R}^n$ be $\mathcal{L}^n$-measurable. Then $\mathcal{H}^0(A \cap f^{-1}\{y\})$ is $\mathcal{H}^n$-measurable on $\mathbb{R}^m$, and we have*

$$\int_A Jf(x)\, d\mathcal{L}^n(x) = \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n(y).$$

Proof. Assume first $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ and A Borel. Let $A_+ = A \cap \{Jf > 0\}$ and $A_0 = A \cap \{Jf = 0\}$. Then by Lemmas 8.4 and 8.5 we know that

- $\mathcal{H}^0(A_+ \cap f^{-1}\{y\})$ is Borel measurable,

- $\mathcal{H}^0(A_0 \cap f^{-1}\{y\}) = 0 \quad$ for $\mathcal{H}^n$-almost every $y \in \mathbb{R}^m$.

In fact, $A_0 \cap f^{-1}\{y\} \neq \emptyset$ implies $y \in f(A_0)$, which is a $\mathcal{H}^n$ null set. Now

$$\begin{aligned} \int_{A_+} Jf(x)\,d\mathcal{L}^n(x) &= \textstyle\int_{\mathbb{R}^m} \mathcal{H}^0(A_+ \cap f^{-1}\{y\})\,d\mathcal{H}^n(y), \\ \int_{A_0} Jf(x)\,d\mathcal{L}^n(x) &= \textstyle\int_{\mathbb{R}^m} \mathcal{H}^0(A_0 \cap f^{-1}\{y\})\,d\mathcal{H}^n(y). \end{aligned}$$

The area formula follows by adding the identities. To generalize the formula to Lipschitz maps we apply the following C^1 extension result.

Theorem 8.7 (Whitney). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz. Then for any $\varepsilon > 0$ there exists a function $\tilde{f} \in C^1(\mathbb{R}^n)$ such that*

$$\mathcal{L}^n\big(\{\tilde{f} \neq f\} \cup \{D\tilde{f} \neq Df\}\big) < \varepsilon.$$

The proof of the Whitney extension is involved, we refer to sections 6.5. and 6.6 in Evans-Gariepy. To continue with the area formula, assume now that A is $\mathcal{L}^n$-measurable, and let $f : \mathbb{R}^n \to \mathbb{R}^m$ be locally Lipschitz. Choose $f_k \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ such that

$$\mathcal{L}^n\big(\{f_k \neq f\} \cup \{Df_k \neq Df\}\big) \leq 2^{-k}.$$

Let $C_k \supset \{f_k \neq f\} \cup \{Df_k \neq Df\}$ be Borel sets also with $\mathcal{L}^n(C_k) \leq 2^{-k}$, and put

$$B_k = \bigcup_{j=k}^{\infty} C_j, \qquad \text{hence } \mathcal{L}^n(B_k) \leq 2^{1-k}.$$

Finally choose $B \subset A$ Borel with $\mathcal{L}^n(A \backslash B) = 0$. The sequence B_k is decreasing, furthermore $f_k = f$ and $Df_k = Df$ on $B \backslash B_k$. As $\bigcap_{k=1}^{\infty} B_k$ and $A \backslash B$ are $\mathcal{L}^n$ null sets and f is locally Lipschitz, the sets $f\big(\bigcap_{k=1}^{\infty} B_k\big)$ and $f(A \backslash B)$ are $\mathcal{H}^n$ null sets. It follows that for $\mathcal{H}^n$-almost all $y \in \mathbb{R}^m$ we have

$$\mathcal{H}^0\Big(\bigcap_{k=1}^{\infty} B_k \cap f^{-1}\{y\}\Big) = 0 \quad \text{and} \quad \mathcal{H}^0\big(A \backslash B \cap f^{-1}\{y\}\big) = 0.$$

Therefore we have, again for $\mathcal{H}^n$ almost all $y \in \mathbb{R}^m$,

$$\begin{aligned} \mathcal{H}^0(A \cap f^{-1}\{y\}) &= \mathcal{H}^0\Big(B \backslash \bigcap_{k=1}^{\infty} B_k \cap f^{-1}\{y\}\Big) \\ &= \lim_{k \to \infty} \mathcal{H}^0\big(B \backslash B_k \cap f^{-1}\{y\}\big) \quad \text{(continuity of } \mathcal{H}^0 \text{ measure)} \\ &= \lim_{k \to \infty} \mathcal{H}^0\big(B \backslash B_k \cap f_k^{-1}\{y\}\big) \quad (f = f_k \text{ on } B \backslash B_k). \end{aligned}$$

It follows that $\mathcal{H}^0(A \cap f^{-1}\{y\})$ is $\mathcal{H}^n$ measurable, and we conclude

$$\begin{aligned}
\int_A Jf\, d\mathcal{L}^n &= \lim_{k\to\infty} \int_{B\backslash B_k} Jf\, d\mathcal{L}^n \quad \text{(monotone convergence)}\\
&= \lim_{k\to\infty} \int_{B\backslash B_k} Jf_k\, d\mathcal{L}^n \quad (Jf = Jf_k \text{ on } B\backslash B_k)\\
&= \lim_{k\to\infty} \int \mathcal{H}^0\big(B\backslash B_k \cap f_k^{-1}\{y\}\big)\, d\mathcal{H}^n(y) \quad (f_k \in C^1,\ B\backslash B_k \text{ Borel})\\
&= \lim_{k\to\infty} \int \mathcal{H}^0\big(B\backslash B_k \cap f^{-1}\{y\}\big)\, d\mathcal{H}^n(y) \quad (f = f_k \text{ on } B\backslash B_k)\\
&= \int_{\mathbb{R}^m} \mathcal{H}^0(A \cap f^{-1}\{y\})\, d\mathcal{H}^n(y) \quad \text{(monotone convergence)}.
\end{aligned}$$

□

Corollary 8.8. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $n \le m$, be locally Lipschitz. Then for any $\mathcal{L}^n$-measurable function $g : \mathbb{R}^n \to [0,\infty)$ the function $\sum_{x\in f^{-1}\{y\}} g(x)$ is $\mathcal{H}^n$-measurable, and*

$$\int_{\mathbb{R}^n} g(x) Jf(x)\, d\mathcal{L}^n(x) = \int_{\mathbb{R}^m} \sum_{x\in f^{-1}\{y\}} g(x)\, d\mathcal{H}^n(y). \tag{8.2}$$

Proof. For $g = \chi_A$ where $A \subset \mathbb{R}^n$ is $\mathcal{L}^n$-measurable the statement follows from Theorem 8.6. The general case is then deduced by monotone approximation with step functions. □

Up to now we assumed $n \le m$, but of course the case when $n \ge m$ is also of interest. The Jacobian of a map $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ is then defined by

$$Jf(x) = \sqrt{\det Df(x) Df(x)^{\mathrm{T}}}.$$

A wellknown case is the so-called onion formula, where $f : \mathbb{R}^n \to \mathbb{R}$ is given by $f(x) = |x|$, with $Jf(x) = 1$ for all $x \ne 0$. We have

$$\mathcal{L}^n(A) = \int_0^\infty \mathcal{H}^{n-1}(A \cap f^{-1}\{r\})\, dr.$$

This is a special case of the following theorem.

Theorem 8.9 (coarea formula)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $n \ge m$, be locally Lipschitz. If $A \subset \mathbb{R}^n$ is $\mathcal{L}^n$-measurable, then the function $\mathcal{H}^{n-m}(A \cap f^{-1}\{y\})$ is $\mathcal{L}^m$-measurable on $\mathbb{R}^m$, and*

$$\int_A Jf(x)\, d\mathcal{L}^n(x) = \int_{\mathbb{R}^m} \mathcal{H}^{n-m}(A \cap f^{-1}\{y\})\, d\mathcal{L}^m(y). \tag{8.3}$$

For the proof we refer to section 3.4 in Evans-Gariepy. We note the following consequence, the proof is left to the reader.

Corollary 8.10 (C^1-Sard)**.** *Let $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ where $n \ge m$. Then for $\mathcal{L}^m$-almost all $y \in \mathbb{R}^m$ the set $f^{-1}\{y\}$ is a union of a $(n-m)$-dimensional C^1 submanifold and a closed set of $\mathcal{H}^{n-m}$-measure zero.*

Chapter 9

Rectifiable sets

In geometric measure theory, the class of rectifiable sets generalizes the class of C^1 submanifolds. In particular we will introduce a measure-theoretic notion of tangent space. Throughout the section we assume that $n \leq m$.

Definition 9.1 (rectifiable set)**.** *A set $M \subset \mathbb{R}^m$ is called countably n-rectifiable, if there exist functions $f_j \in \operatorname{Lip}(\mathbb{R}^n, \mathbb{R}^m)$, $j \in \mathbb{N}$, such that*

$$\mathcal{H}^n\Big(M \backslash \bigcup_{j=1}^{\infty} f_j(\mathbb{R}^n)\Big) = 0. \tag{9.1}$$

Lemma 9.2. *A set $M \subset \mathbb{R}^m$ is countably n-rectifiable if and only if there exist n-dimensional C^1-submanifolds N_j, $j \in \mathbb{N}$, such that*

$$\mathcal{H}^n\Big(M \backslash \bigcup_{j=1}^{\infty} N_j\Big) = 0. \tag{9.2}$$

Proof. Any C^1 submanifold $N \subset \mathbb{R}^m$ is a countable union of Lipschitz graphs. In fact, for any $x \in N$ there exists $\varrho(x) \in (0,1]$ such that $N \cap B_{5\varrho(x)}(x)$ is a Lipschitz graph. By Vitali N is covered by balls $B_{5\varrho(x_i)}(x_i)$, $i \in I$, where the balls $B_{\varrho(x_i)}(x_i)$ are disjoint. In particular, the set I is countable.

For the reverse direction, we may assume $M \subset f(\mathbb{R}^n)$ where $f \in \operatorname{Lip}(\mathbb{R}^n, \mathbb{R}^m)$. By Lusin's theorem, see Theorem 6.5, there exist $f_j \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ such that $\mathcal{L}^n(\{f_j \neq f\}) < \frac{1}{j}$. Put $C_j = \{Jf_j = 0\}$, then we can write

$$f_j(\mathbb{R}^n \backslash C_j) = \bigcup_{k=1}^{\infty} N_{j,k},$$

where $N_{j,k}$ are C^1 submanifolds. This follows by a covering argument similar to the above, using that f_j is locally an embedding on $\mathbb{R}^n \backslash C_j$. The remaining set E of all $x \in \mathbb{R}^n$ with $f(x) \notin f_j(\mathbb{R}^n \backslash C_j)$ for any j is contained in $E_0 \cup E_1$ where

$$E_0 = \bigcap_{j=1}^{\infty} \{f_j \neq f\} \quad \text{and} \quad E_1 = \bigcup_{j=1}^{\infty} \{f = f_j\} \cap C_j.$$

Clearly $\mathcal{L}^n(E_0) = 0$ and hence $\mathcal{H}^n(f(E_0)) = 0$. For E_1 we estimate by the area formula

$$\mathcal{H}^n\big(f(\{f = f_j\} \cap C_j)\big) \leq \mathcal{H}^n(f_j(C_j)) \leq \int_{C_j} Jf_j \, d\mathcal{L}^n = 0.$$

□

We now introduce a measure-theoretic concept of tangent space. We denote by $G(n,m)$ the set of all n-dimensional vector subspaces $E \subset \mathbb{R}^m$. We may identify E with the orthogonal projection P_E onto E, then $G(n,m)$ becomes the set

$$G(n,m) = \{P \in \mathbb{R}^{m \times m} : P^2 = P = P^{\mathrm{T}},\, \mathrm{tr}\, P = n\}.$$

$P^2 = P$ means that P is a projection, that is $\mathbb{R}^m = \ker P \oplus \mathrm{im}\, P$ and $P = \mathrm{Id}$ on $\mathrm{im}\, P$. The condition $P^{\mathrm{T}} = P$ yields that $\ker P$ and $\mathrm{im}\, P$ are orthogonal. Finally, the equation $\mathrm{tr}\, P = n$ implies that P has rank n. Note that

$$|P|^2 = \mathrm{tr}\,(P^{\mathrm{T}} P) = \mathrm{tr}\, P = n,$$

hence $G(n,m)$ is contained in the sphere of radius $\sqrt{n}$ in the space of symmetric matrices. In particular $G(n,m)$ is compact.

Definition 9.3 (approximate tangent space)**.** *Let μ be a Radon measure on $\mathbb{R}^m$ and $x \in \mathbb{R}^m$. Then μ has approximate tangent space $P \in G(n,m)$ at x with multiplicity $\theta > 0$, if*

$$\mu_{x,\lambda} \to \theta\, \mathcal{H}^n \llcorner P \text{ in } C^0_c(\mathbb{R}^m)' \text{ as } \lambda \searrow 0 \quad \text{where } \mu_{x,\lambda}(A) = \lambda^{-n} \mu(x + \lambda A). \tag{9.3}$$

It is sometimes useful to write $\mu_{x,\lambda} = \lambda^{-n} \eta_{x,\lambda}(\mu)$ where $\eta_{x,\lambda}(y) = \frac{y-x}{\lambda}$. In fact then

$$\eta_{x,\lambda}(\mu)(A) = \mu(\eta^{-1}_{x,\lambda}(A)) = \mu(x + \lambda A).$$

Of course we need to check that the concept is well-defined. If $\mu = \mathcal{H}^n \llcorner M$ where M is an n-dimensional C^1 submanifold, then the approximate tangent space at $x \in M$ should be the classical tangent space $T_x M$, with multiplicity $\theta = 1$. More generally, let M be the image of an immersion with k sheets passing through $x \in M$. If the tangent spaces of these sheets are all equal to $P \in G(n,m)$, then $\mathcal{H}^n \llcorner M$ should have approximate tangent space P at x with multiplicity $\theta = k$. Otherwise, the approximate tangent space should not exist. The verification of these facts is left to the reader.

Lemma 9.4. *Assume that μ has approximate tangent space $P \in G(n,m)$ at the point x with multiplicity $\theta > 0$. Then the following holds:*

(1) $\theta^n(\mu, x) = \theta$,

(2) *P and θ are uniquely determined.*

Proof. Let $f,\, g \in C^0_c(\mathbb{R}^m)$ such that $f \leq \chi_{B_1(0)} \leq g$. Then we have for $n = \dim P$

$$\begin{aligned}
\theta \int_P f \, d\mathcal{H}^n &= \lim_{\lambda \searrow 0} \int f \, d\mu_{x,\lambda} \leq \liminf_{\lambda \searrow 0} \mu_{x,\lambda}(B_1(0)) = \alpha_n \liminf_{\lambda \searrow 0} \frac{\mu(B_\lambda(x))}{\alpha_n \lambda^n}, \\
\theta \int_P g \, d\mathcal{H}^n &= \lim_{\lambda \searrow 0} \int g \, d\mu_{x,\lambda} \geq \limsup_{\lambda \searrow 0} \mu_{x,\lambda}(B_1(0)) = \alpha_n \limsup_{\lambda \searrow 0} \frac{\mu(B_\lambda(x))}{\alpha_n \lambda^n}.
\end{aligned}$$

Letting $f \nearrow \chi_{B_1(0)}$ and $g \searrow \chi_{B_1(0)}$ the left hand side goes to $\theta\alpha_n$, which proves claim (1). For (2) we first observe that the dimension n of the tangent space and the multiplicity is determined by (1). The existence of an n-dimensional tangent space means in particular that $\mu_{x,\lambda}$ converges to a measure μ. Moreover if P is the tangent space then $P = \operatorname{spt}\mu$. $\square$

We now come to the main result of this section. We show that the existence of tangent spaces, which is an infinitesimal information, implies the local property of rectifiabilty.

Theorem 9.5 (rectifiability of measures). *Let μ be a Radon measure on $\mathbb{R}^m$. Denote by M the set of $x \in \mathbb{R}^m$ at which μ has an approximate tangent space $T_x\mu \in G(n,m)$, for some multiplicity $\theta(x) > 0$, and put $\theta = 0$ on $\mathbb{R}^m \backslash M$. If $\mu(\mathbb{R}^m \backslash M) = 0$ then the following holds:*

(1) *M is $\mathcal{H}^n$-measurable and countably n-rectifiable.*

(2) *θ is $\mathcal{H}^n$-measurable and $\mu = \mathcal{H}^n \llcorner \theta$, thus in fact $\theta \in L^1_{\mathrm{loc}}(\mathcal{H}^n)$.*

Proof. We assume $\operatorname{spt}\mu$ is compact. By Corollary 5.9 we know that if $\mu(E) = 0$ then

$$\theta^n(\mu, x) = 0 \quad \text{for } \mathcal{H}^n\text{-almost all } x \in E. \tag{9.4}$$

Using this for $E = \mathbb{R}^m \backslash M$ yields a $\mathcal{H}^n$ null set $Z \subset \mathbb{R}^m \backslash M$ such that $\theta^n(\mu, x) = \theta(x)$ for all $x \in \mathbb{R}^m \backslash Z$. As the upper/lower desities are Borel measurable, we see that θ and also M are both μ and $\mathcal{H}^n$ measurable.

Our goal is to find pieces of M which are Lipschitz graphs over the approximate tangent spaces. For this we introduce some notation. Let $k = m - n$ be the codimension. For $\pi \in G_k(\mathbb{R}^n)$ and $0 < \alpha \leq 1$ we consider the vertical cone

$$X_\alpha(\pi, x) = \{y \in \mathbb{R}^m : |\pi(y - x)| \geq \alpha |y - x|\}.$$

The opening angle of this cone is $\arccos\alpha \in [0, \frac{\pi}{2})$. For $x \in M$ and $\pi = T_x\mu^\perp$ we compute

$$\begin{aligned}
\overline{\theta}^n(\mu, X_{\frac{1}{2}}(\pi, x), x) &= \limsup_{r \searrow 0} \frac{\mu(x + rA)}{r^n} \qquad \text{where } A = X_{\frac{1}{2}}(\pi, 0) \cap B_1(0) \\
&= \limsup_{r \searrow 0} \frac{1}{\alpha_n} \mu_{x,r}(A) \\
&\leq \frac{\theta(x)}{\alpha_n} \mathcal{H}^n \llcorner T_x\mu(A) = 0.
\end{aligned}$$

Thus we have

$$\overline{\theta}^n(\mu, X_{\frac{1}{2}}(\pi, x), x) = 0 \quad \text{for all } x \in M. \tag{9.5}$$

For $\lambda > 0$ let $M_\lambda = \{x \in M : \theta(x) \geq \lambda\}$. Choosing $\lambda > 0$ sufficiently small we have

$$\mu(\mathbb{R}^m \backslash M_\lambda) \leq \frac{1}{4}\, \mu(\mathbb{R}^m). \tag{9.6}$$

Next consider the functions $f_k, q_k : M \to [0, \infty)$ given by

$$f_k(x) = \inf_{0 < r < \frac{1}{k}} \frac{\mu(B_r(x))}{\alpha_n r^n} \qquad \text{and} \qquad q_k(x) = \sup_{0 < r < \frac{1}{k}} \frac{\mu\big(X_{\frac{1}{2}}(T_x\mu^\perp, x) \cap B_r(x)\big)}{\alpha_n r^n}.$$

For $k \to \infty$ we know that $f_k(x) \to \theta(x)$ and $q_k(x) \to 0$. By Egorov there is a μ-measurable set $E \subset M_\lambda$, with $\mu(M_\lambda \backslash E) \leq \frac{1}{4}\,\mu(\mathbb{R}^m)$, such that the convergence is uniform on E. Thus for $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $x \in E$ and all $r \in (0, \delta]$ we have the following:

$$\begin{aligned} \frac{\mu(B_r(x))}{\alpha_n r^n} &> \lambda - \varepsilon, \\ \frac{\mu\big(X_{\frac{1}{2}}(T_x\mu^\perp, x) \cap B_r(x)\big)}{\alpha_n r^n} &< \varepsilon, \\ \mu(\mathbb{R}^n \backslash E) &\leq \frac{1}{2}\mu(\mathbb{R}^m). \end{aligned} \tag{9.7}$$

Now choose $\pi_1, \ldots, \pi_N \in G_k(\mathbb{R}^m)$ such that

$$G_k(\mathbb{R}^m) \subset \bigcup_{j=1}^N B_{\frac{1}{16}}(\pi_j).$$

Here the balls are defined using the Hilbert-Schmidt norm. It follows that we have a covering

$$E \subset \bigcup_{j=1}^N E_j \quad \text{ where } E_j = \{x \in E : T_x\mu^\perp \in B_{\frac{1}{16}}(\pi_j)\}.$$

Claim 1. *For $\varepsilon > 0$ small and $\delta > 0$ with (9.7) we have*

$$X_{\frac{3}{4}}(\pi_j, x) \cap E_j \cap B_{\frac{\delta}{2}}(x) = \{x\} \quad \text{ for all } x \in E_j. \tag{9.8}$$

Let $y \in X_{\frac{3}{4}}(\pi_j, x) \cap E_j \cap B_{\frac{\delta}{2}}(x)$, and assume by contradiction that $\varrho := |y - x| > 0$. As $2\varrho \leq \delta$ we may apply (9.7) to get

$$\frac{\mu\big(X_{\frac{1}{2}}(T_x\mu^\perp, x) \cap B_{2\varrho}(x)\big)}{\alpha_n (2\varrho)^n} < \varepsilon.$$

Now we show that

$$B_{\frac{\varrho}{8}}(y) \subset \big(X_{\frac{1}{2}}(T_x\mu^\perp, x) \cap B_{2\varrho}(x)\big).$$

In fact for $z \in B_{\frac{\varrho}{8}}(y)$ we have $|z - x| \leq |z - y| + |y - x| \leq \frac{9}{8}\varrho$, and we calculate

$$\begin{aligned} |\pi_{T_x\mu^\perp}(z - x)| &\geq |\pi_{T_x\mu^\perp}(y - x)| - |\pi_{T_x\mu^\perp}(z - y)| \\ &\geq |\pi_j(y - x)| - \frac{1}{16}|y - x| - |z - y| \\ &\geq \frac{3}{4}\varrho - \frac{1}{16}\varrho - \frac{1}{8}\varrho \\ &= \frac{9}{16}\varrho \geq \frac{1}{2}|z - x|. \end{aligned}$$

Applying again (9.7), but now for $y \in E_j \subset E$, we obtain

$$\begin{aligned} (\lambda - \varepsilon)\,\alpha_n \Big(\frac{\varrho}{8}\Big)^n &< \mu(B_{\frac{\varrho}{8}}(y)) \\ &\leq \mu\big(X_{\frac{1}{2}}(T_x\mu^\perp, x) \cap B_{2\varrho}(x)\big) \\ &< \varepsilon\,\alpha_n (2\varrho)^n. \end{aligned}$$

We conclude $\lambda < \left(2^{4n}+1\right)\varepsilon$. Thus for $\varepsilon \leq (2^{4n}+1)^{-1}\lambda$ we arrive at a contradiction, and then (9.8) must hold for these $\varepsilon > 0$. This finishes the proof of claim 1.

Claim 2. *For any $x_0 \in E_j$ there exists a Euclidean motion Q of $\mathbb{R}^m$ with $Q(\mathbb{R}^k) = x_0 + \pi_j$ and a function $u \in \operatorname{Lip}(\mathbb{R}^n, \mathbb{R}^k)$ such that*

$$E_j \cap B_{\frac{\delta}{4}}(x_0) \ \subset\ Q(\operatorname{graph} u).$$

To show this we assume for simplicity that $x_0 = 0$. Consider two different points $x_{1,2}$ in $E_j \cap B_{\frac{\delta}{4}}(0)$. Decompose $x_{1,2} = y_{1,2} + z_{1,2}$ where $y_{1,2} \in \pi_j^\perp$ and $z_{1,2} \in \pi_j$. As $x_2 \in B_{\frac{\delta}{2}}(x_1)$ we obtain from claim 1, see equation (9.8),

$$|z_1 - z_2| = |\pi_j(x_1 - x_2)| \leq \frac{3}{4}\,|x_1 - x_2| \leq \frac{3}{4}\left(|y_1 - y_2| + |z_1 - z_2|\right).$$

We conclude that $|z_1 - z_2| \leq 3\,|y_1 - y_2|$, in particular $y_1 \neq y_2$. Let A be the projection of $E_j \cap B_{\frac{\delta}{4}}(0)$ onto $\pi_j^\perp$. For $y = \pi_j^\perp(x) \in A$ we define $u(y) = z$ where $x = y + z$. This is well-defined by the above, moreover u is Lipschitz with constant $\operatorname{Lip}(u) \leq 3$. Claim 2 follows using the extension from Theorem 7.1.

Now recall that $\delta > 0$ in (9.7) is independent of the point $x \in E$. By Vitali, the set E_j is covered by a family of balls of radius $\frac{\delta}{4}$ as in claim 2, such that the concentric balls of radius $\frac{\delta}{20}$ are disjoint. As all these balls intersect the compact set $\operatorname{spt}\mu$, the family is actually finite. As E is covered by $E_1, \ldots, E_N$, we obtain

$$E \subset \bigcup_{i=1}^{L} Q_i(\operatorname{graph} u_i) =: G_1,$$

where $u_i \in \operatorname{Lip}(\mathbb{R}^n, \mathbb{R}^k)$ and the Q_i are Euclidean motions. Furthermore (9.7) yields

$$\mu\big(\mathbb{R}^m \backslash G_1\big) \leq \frac{1}{2}\,\mu(\mathbb{R}^m).$$

We now repeat the argument by considering the Radon measure $\mu_1 = \mu \llcorner \big(\mathbb{R}^m \backslash G_1\big)$. As G_0 is closed, the measure μ_1 has approximate tangent space $T_x\mu$ with multiplicity $\theta(x)$ for all $x \in \mathbb{R}^m \backslash G_0$, and trivially $\mu_1(G_0) = 0$. Therefore we can iterate the whole argument. For $i = 1, 2, \ldots$ we obtain sets G_i, each a finite union of sets of the form $Q(\operatorname{graph} u)$ where $u \in \operatorname{Lip}(\mathbb{R}^n, \mathbb{R}^k)$ and Q is a Euclidean motion, such that

$$\mu\Big(\mathbb{R}^m \backslash \bigcup_{i=1}^{j} G_i\Big) \leq 2^{-j}\,\mu(\mathbb{R}^m).$$

The set $M_0 = M \backslash \bigcup_{i=1}^\infty G_i$ has μ-measure zero. As $\theta^n(\mu, x) = \theta(x) > 0$ on $M \supset M_0$, we obtain from (9.4) that M_0 is also a $\mathcal{H}^n$ null set, and conclude that M is countably n-rectifiable.

We now turn to the proof of (2). We first claim that μ is absolutely continuous with respect to $\mathcal{H}^n$. For this let $M^t = \{x \in M : \theta(x) \leq t\}$ where $t < \infty$. Then $\overline{\theta}^n(\mu, x) \leq t$ for $\mathcal{H}^n$-almost all $x \in M^t$, and Theorem 5.8(2) yields

$$\mu(A) \leq 2^n t\,\mathcal{H}^n(A) \quad \text{ for all } A \subset M^t.$$

Thus if $\mathcal{H}^n(A) = 0$ then $\mu(A \cap M^t) = 0$ for all $t < \infty$, and hence $\mu(A) = 0$. Now according to Lemma 9.2 there exist C^1-submanifolds N_j, $j \in \mathbb{N}$, such that

$$\mathcal{H}^n\Big(M \backslash \bigcup_{j=1}^{\infty} N_j\Big) = 0.$$

As $\mu(N_j \backslash M) = 0$ by assumption, we get from Corollary 5.9 and Lemma 9.4

$$\lim_{\varrho \searrow 0} \frac{\mu(B_\varrho(x))}{\alpha_n \varrho^n} = \begin{cases} 0 \, = \, \theta(x) & \text{for } \mathcal{H}^n\text{-a. e. } x \in N_j \backslash M, \\ \theta(x) & \text{for all } x \in N_j \cap M. \end{cases}$$

Using (9.10) we conclude for $\mathcal{H}^n$-almost every $x \in N_j$ that

$$\lim_{\varrho \searrow 0} \frac{\mu(N_j \cap B_\varrho(x))}{\mathcal{H}^n(N_j \cap B_\varrho(x))} = \lim_{\varrho \searrow 0} \Big(\frac{\mu(N_j \cap B_\varrho(x))}{\alpha_n \varrho^n} \cdot \frac{\alpha_n \varrho^n}{\mathcal{H}^n(N_j \cap B_\varrho(x))}\Big) = \theta(x).$$

For any $x \in N_j$ there is an open neighborhood $U_x \subset \mathbb{R}^m$ such that $N_j \cap U_x$ is properly embedded, for instance a graph. Then $\mu \llcorner N_j$ and $\mathcal{H}^n \llcorner N_j$ are Radon measures in U_x. By Radon-Nikodym, see Theorem 5.3, we get that $\mu = \mathcal{H}^n \llcorner \theta$ on $N_j \cap U_x$, and hence on all of N_j. Now for any Borel set $B \subset \mathbb{R}^m$ put

$$B_j = B \cap N_j \backslash \bigcup_{i=1}^{j-1} N_i \quad \text{for } j \in \mathbb{N}.$$

By construction the B_j are pairwise disjoint. Let

$$B_0 := B \backslash \bigcup_{j=1}^{\infty} B_j \, \subset \, \mathbb{R}^m \backslash \bigcup_{j=1}^{\infty} N_j.$$

We have $\mathcal{H}^n(B_0 \cap M) = 0$ and hence $\mu(B_0) = 0$, as well as $\mathcal{H}^n \llcorner \theta(B_0) = 0$. Thus

$$\mu(B) = \sum_{j=1}^{\infty} \mu(B_j) = \sum_{j=1}^{\infty} \mathcal{H}^n \llcorner \theta(B_j) = \mathcal{H}^n \llcorner \theta(B).$$

By Borel regularity we conclude $\mu = \mathcal{H}^n \llcorner \theta$. The theorem is proved. $\square$

The next result is kind of converse, asserting the existence of approximate tangent spaces.

Theorem 9.6 (Existence of approximate tangent space). *Let $M \subset \mathbb{R}^m$ be $\mathcal{H}^n$-measurable and countably n-rectifiable, and let $\theta \in L^1_{\mathrm{loc}}(\mathcal{H}^n)$ be nonnegative with $M = \{\theta > 0\}$. Then for $\mathcal{H}^n$-almost all $x \in M$ the Radon measure $\mu = \mathcal{H}^n \llcorner \theta$ has an approximate tangent space $T_x\mu$ with multiplicity $\theta(x)$.*

Before entering the proof we recall some basic facts about measures with density. If μ is an outer measure on a set X and $\theta : X \to [0, \infty]$ is μ-measurable, then one defines

$$\mu \llcorner \theta(A) = \int_A \theta \, d\mu \quad \text{if } A \text{ is } \mu\text{-measurable.}$$

The measure is extended to all sets E by approximating

$$\mu \llcorner \theta(E) = \inf\{\mu \llcorner \theta(A) : A \ \mu\text{-measurable}, \ E \subset A\}.$$

It is easy to see that $\mu \llcorner \theta$ is an outer measure, moreover it has the following properties:

- any μ-measurable set E is also $\mu\llcorner\theta$-measurable,
- $\mu(E) = 0$ implies $\mu\llcorner\theta(E) = 0$,
- $\int f\, d(\mu\llcorner\theta) = \int f\, \theta d\mu$, whenever $f : X \to [0,\infty]$ is μ-measurable,
- if μ is Borel regular then $\mu\llcorner\theta$ is also Borel regular.

All these assertions follow easily using the monotone convergence theorem.

Proof. (*of Theorem 9.6*) By Lemma 9.2 there exist C^1-submanifolds N_j, $j \in \mathbb{N}$, such that $M_0 := M\setminus\bigcup_{j=1}^{\infty} N_j$ is a $\mathcal{H}^n$ null set. We claim that $\mu = \mathcal{H}^n\llcorner\theta$ has approximate tangent space

$$T_x\mu = T_xN_j, \text{ with multiplicity } \theta(x), \quad \text{for } \mathcal{H}^n\text{-almost all } x \in M \cap N_j. \tag{9.9}$$

Here T_xN_j is the classical tangent space of the submanifold N_j. Let $f \in C_c^0(\mathbb{R}^m)$ be fixed, with $\operatorname{spt} f \subset B_R(0)$ and $\sup|f| \le C$. For given $x \in M_j$ we write

$$\int f(z)\, d\mu_{x,\lambda}(z) = \int_{\frac{1}{\lambda}(N_j - x)} f(z)\, d\mu_{x,\lambda}(z) + \int_{\mathbb{R}^m\setminus\frac{1}{\lambda}(N_j-x)} f(z)\, d\mu_{x,\lambda}(z).$$

In the first integral we substitute $z = \frac{y-x}{\lambda}$ to obtain further

$$\begin{aligned}
&\int_{\frac{1}{\lambda}(N_j - x)} f(z)\, d\mu_{x,\lambda}(z)\\
= \;& \frac{\theta(x)}{\lambda^n}\int_{N_j} f\Big(\frac{y-x}{\lambda}\Big)\, d\mathcal{H}^n(y) + \frac{1}{\lambda^n}\int_{N_j} f\Big(\frac{y-x}{\lambda}\Big)\big(\theta(y) - \theta(x)\big)\, d\mathcal{H}^n(y)\\
= \;& \theta(x)\int_{\frac{1}{\lambda}(N_j - x)} f(z)\, d\mathcal{H}^n(z) + \frac{1}{\lambda^n}\int_{N_j} f\Big(\frac{y-x}{\lambda}\Big)\big(\theta(y) - \theta(x)\big)\, d\mathcal{H}^n(y).
\end{aligned}$$

Now clearly

$$\lim_{\lambda\searrow 0}\int_{\frac{1}{\lambda}(N_j - x)} f(z)\, d\mathcal{H}^n(z) = \int f(z)\, d(\mathcal{H}^n\llcorner T_xN_j)(z).$$

It is sufficient to prove (9.9) for $\mathcal{H}^n$-almost all $x \in M_j \cap U$, where U is an open set such that $\mu(U) < \infty$ and $\mathcal{H}^n\llcorner N_j(U) < \infty$. Moreover we can assume that

$$\mathcal{H}^n\llcorner N_j(B_\varrho(x)) \le C\varrho^n \quad \text{for all } x \in N_j \cap U,\ \varrho \in (0,\varrho_0].$$

For the first integral we estimate

$$\begin{aligned}
\Big|\int_{\mathbb{R}^m\setminus\frac{1}{\lambda}(N_j-x)} f(z)\, d\mu_{x,\lambda}(z)\Big| \;&\le\; C\,\mu_{x,\lambda}\Big(B_R(0)\setminus\frac{1}{\lambda}(N_j - x)\Big)\\
&\le\; \frac{C}{\lambda^n}\,\mu\big(B_{\lambda R}(x)\setminus N_j\big).
\end{aligned}$$

The right hand side goes to zero as $\lambda \searrow 0$ for $\mathcal{H}^n$-almost all $x \in N_j \cap U$. Namely since $\mu\llcorner U(\mathbb{R}^m\setminus N_j) < \infty$, the density formula (9.10) implies

$$\lim_{\varrho\searrow 0}\frac{\mu(B_\varrho(x)\setminus N_j)}{\alpha_n\varrho^n} = \theta^n(\mu\llcorner U, \mathbb{R}^m\setminus N_j, x) = 0 \quad \text{for } \mathcal{H}^n\text{-almost all } x \in N_j \cap U. \tag{9.10}$$

The remaining second integral is estimated by

$$\frac{1}{\lambda^n}\Big|\int_{N_j} f\Big(\frac{y-x}{\lambda}\Big)\big(\theta(y)-\theta(x)\big)\,d\mathcal{H}^n(y)\Big| \;\leq\; \frac{C}{\lambda^n}\int_{B_{\lambda R}(x)} |\theta(y)-\theta(x)|\,d(\mathcal{H}^n\llcorner N_j)(y).$$

We use the Lebesgue point property (5.1) to conclude that the right hand side goes to zero as $\lambda \searrow 0$, for $\mathcal{H}^n$-almost all points $x \in N_j \cap U$. Note that $\mathcal{H}^n\llcorner N_j$ is a Radon measure on U, moreover by the above $\mathcal{H}^n\llcorner N_j(B_{\lambda R}(x)) \leq C(\lambda R)^n$. Therefore (5.1) applies to the function $\theta \in L^1_{\mathrm{loc}}(\mathcal{H}^n\llcorner N_j)$ for $x \in N_j \cap U$, and the proof is finished. □

Corollary 9.7 (measurability of Gauß map)**.** *Let μ be a Radon measure on $\mathbb{R}^m$, and let M be the set of $x \in \mathbb{R}^m$ with an approximate tangent space $T_x\mu \in G(n,m)$. Define the Gauß map*

$$G : M \to G(n,m),\ G(x) = T_x\mu.$$

If $\mu(\mathbb{R}^m\backslash M) = 0$ then $G^{-1}(B)$ is $\mathcal{H}^n$-measurable for any Borel set $B \subset G(n,m)$.

Proof. It suffices to show that $G^{-1}(U)$ is $\mathcal{H}^n$-measurable for any open $U \subset G(n,m)$. By Theorem 9.5 the set M and the multiplicity function θ (with $\theta = 0$ on $\mathbb{R}^m\backslash M$) satisfy the assumptions of Theorem 9.6. Let N_j, $j \in \mathbb{N}$, be C^1-submanifolds which cover M, up to a $\mathcal{H}^n$ null set, and denote by $G_j : N_j \to G(n,m)$ the Gauß maps of the N_j. Then by (9.9) we know that $G(x) = G_j(x)$ for $\mathcal{H}^n$-almost all $x \in M \cap N_j$, hence

$$G^{-1}(U) \cap N_j = G_j^{-1}(U) \cap M \quad \text{ up to a } \mathcal{H}^n \text{ null set.}$$

As $G_j^{-1}(U)$ is open in N_j, the set $G_j^{-1}(U) \cap M$ is $\mathcal{H}^n$ measurable. Now $G^{-1}(U)$ is the union of the sets $G^{-1}(U) \cap N_j$, $j \in \mathbb{N}$, and another $\mathcal{H}^n$ null set. □

Roughly speaking the last results say hat one has a reasonable notion of tangent space for any $\mathcal{H}^n$-measurable, countably n-rectifiable set. This summarizes as follows.

Corollary 9.8. *Let $M \subset \mathbb{R}^m$ be $\mathcal{H}^n$-measurable and countably n-rectifiable. There exists a map $G_M : M \to G(n,m)$, $G_M(x) =: T_xM$, with the following properties:*

(a) *$G_M^{-1}(B)$ is $\mathcal{H}^n$-measurable for any Borel set $B \subset G(n,m)$.*

(b) *For any measure $\mu = \mathcal{H}^n\llcorner\theta$, with $\theta \in L^1_{loc}(\mathcal{H}^n)$ and $M = \{\theta > 0\}$, one has $T_x\mu = T_xM$ with multiplicity $\theta(x)$ for $\mathcal{H}^n$-almost all $x \in M$.*

(c) *For any n-dimensional C^1-submanifold $N \subset \mathbb{R}^m$ one has $T_xN = T_xM$ for $\mathcal{H}^n$-almost all $x \in M \cap N$.*

Proof. We first construct a function $\theta(x)$ as in (b). For this we assume that M is covered, up to a $\mathcal{H}^n$ null set M_0, by C^1 submanifolds N_j for $j \in \mathbb{N}$. We can assume $\mathcal{H}^n(N_j) \leq 1$, otherwise we pass to a covering of N_j by subsets with this condition. Now let

$$\theta(x) = 2^{-j} \quad \text{ for } x \in M \cap N_j \backslash \bigcup_{i=1}^{j-1} N_i.$$

By Theorem 9.6 the space $T_x\mu$ exists $\mathcal{H}^n$-almost everywhere, and from (9.9) we know that

$$T_x\mu = T_xN_j \quad \text{ for } \mathcal{H}^n\text{-almost every } x \in M \cap N_j.$$

We define $G_M(x) = T_x\mu$, then (a) is proved in Corollary 9.7. Now if $\tilde{\mu} = \mathcal{H}^n \llcorner \tilde{\theta}$ is another measure as in (b), then Theorem 9.6 can be applied again, and (9.9) yields that $T_x\tilde{\mu} = T_x N_j = G_M(x)$ with multiplicity $\tilde{\theta}(x)$ for $\mathcal{H}^n$-almost all $x \in M \cap N_j$, which proves claim (b). Finally if $N \subset \mathbb{R}^m$ is a C^1 submanifold of dimension n, then we can simply add N to the collection N_j and obtain (c) again from (9.9). $\square$

A set $M \subset \mathbb{R}^m$ with $0 < \mathcal{H}^n(M) < \infty$ is purely n-unrectifiable if it has no subset of positive $\mathcal{H}^n$-measure which is countably n-rectifiable. A deep theorem of Besicovitch ($n = 1$, $m = 2$) and Federer ($n < m$ arbitrary) asserts that then $P(M)$ has zero $\mathcal{L}^n$ measure for almost every $P \in G(n, m)$. An example is the triangular Cantor set considered in series 8 of the homework assignments.

Chapter 10

Varifolds

Definition 10.1. *Let $U \subset \mathbb{R}^m$ be open. By definition, an n-varifold V on U is a Radon measure on $G_n(U) = U \times G(n,m)$. Notation: $V \in \mathbb{V}_n(U)$.*

Recall that $G(n,m)$ is identified with the subset of $L(\mathbb{R}^m, \mathbb{R}^m) \cong \mathbb{R}^{m \times m}$ consisting of all orthogonal projections of rank n. Then the Hilbert-Schmidt norm induces a metric on $G(n,m)$. For $P \in G(n,m)$ we have For $P \in G(n,m)$ we have $|P|^2 = \operatorname{tr}(P^{\mathrm{T}}P) = n$, therefore $G(n,m)$ is compact.

Definition 10.2. *Let $\pi : G_n(U) = U \times G(n,m) \to U,\ \pi(x,P) = x$. The weight measure of $V \in \mathbb{V}_n(U)$ is the Radon measure $\mu_V = \pi(V)$ on U. Alternative notation: $\mu_V = |V|$.*

The projection π is proper, in fact for compact $K \subset U$ the set $\pi^{-1}(K) = K \times G(m,n)$ is also compact. Moreover $G(n,m)$ is of course separable. The Radon measure property then follows from Theorem 2.17. By the transformation formula we have for any Borel function $\varphi : U \to [0,\infty)$

$$\int_U \varphi(x)\, d\mu_V(x) = \int_{G_n(U)} \varphi(x)\, dV(x,P). \tag{10.1}$$

Theorem 10.3 (Disintegration)**.** *For any $V \in \mathbb{V}_n(U)$ there exists a family V^x, $x \in U$, of Radon measures on $G(n,m)$ such that for every Borel function $f : G_n(U) \to [0,\infty)$ the following holds:*

(1) *the function $x \mapsto \int_{G(n,m)} f(x,P)\, dV^x(P)$ is μ_V measurable, and*

(2) *$\int_{G_n(U)} f(x,P)\, dV(x,P) = \int_U \int_{G(n,m)} f(x,P)\, dV^x(P)\, d\mu_V(x)$.*

Two families of Radon measures with (1) and (2) coincide up to a μ_V null set.

Remark 10.4. *Taking $f(x,P) = \varphi(x)$ in (2) we infer using (10.1)*

$$\int_U \varphi(x) V^x(G(n,m))\, d\mu_V(x) = \int_{G_n(U)} \varphi(x)\, dV(x,P) = \int_U \varphi(x)\, d\mu_V(x).$$

We conclude that $V^x(G(n,m)) = 1$ for μ_V-almost all $x \in U$, i.e. V^x is a probability measure.

Proof. (Theorem 10.3) We first check the uniqueness. Take in (2) the product function $\chi_A(x) f(P)$, where A is a Borel set and $f \in C^0(G(n,m))$ is nonnegative. This yields

$$\int_A \int_{G(n,m)} f(P)\, dV^x(P)\, d\mu_V(x) = \int_{G_n(A)} f(P)\, dV(x,P) = \pi(V \llcorner f)(A).$$

The Radon measure $\pi(V \llcorner f)$ is absolutely continuous with respect to $\mu_V = \pi(V)$. The equation says that the corresponding Radon Nikodym density is

$$x \mapsto \Lambda^x(f) := \int_{G(n,m)} f(P)\, dV^x(P).$$

The density is determined μ_V-almost everywhere. Thus for two such families $\Lambda^x_{1,2}$ there exists a μ_V null set Z_f such that $\Lambda^x_1(f) = \Lambda^x_2(f)$ for all $x \in U \backslash Z_f$. Now choose a dense set, see for instance Lemma 6.10,

$$\{f_j\}_{j \in \mathbb{N}} \subset \{f \in C^0(G(n,m)) : f \geq 0\}.$$

Then $Z = \bigcup_{j=1}^{\infty} Z_{f_j}$ is a μ_V null set, and $\Lambda^x_1(f_j) = \Lambda^x_2(f_j)$ for all $j \in \mathbb{N}$, $x \in U \backslash Z$. But the linear functionals $f \mapsto \Lambda^x_{1,2}(f)$ are continuous, in fact we have $\|\Lambda^x_{1,2}\| \leq 1$. By density we conclude that $\Lambda^x_1(f) = \Lambda^x_2(f)$ for all f, and all $x \in U \backslash Z$.

To prove existence, we consider for $\varrho > 0$ the functionals $\Lambda_{x,\varrho} \in C^0(G(n,m))'$ given by

$$\Lambda_{x,\varrho}(f) = \fint_{G_n(B_\varrho(x))} f(P)\, dV(x,P) = \frac{\pi(V \llcorner f)(B_\varrho(x))}{\mu_V(B_\varrho(x))}.$$

By Radon Nikodym, there exists a null set Z_f such that

$$\exists \lim_{\varrho \searrow 0} \Lambda_{x,\varrho}(f) =: \Lambda_x(f) \quad \text{for all } x \in U \backslash Z_f. \tag{10.2}$$

Let f_j be the dense set as above, and $Z = \bigcup_{j=1}^{\infty} Z_{f_j}$. Then for $f \in C^0(G(n,m))$, $f \geq 0$, and $\sigma, \varrho > 0$ we estimate for $x \in U \backslash Z$

$$|\Lambda_{x,\sigma}(f) - \Lambda_{x,\varrho}(f)| \leq |\Lambda_{x,\sigma}(f_j) - \Lambda_{x,\varrho}(f_j)| + 2\,\|f - f_j\|_{C^0}.$$

Letting first $\sigma, \varrho \searrow 0$ and then taking the infimum among all f_j we conclude

$$\Lambda_{x,\varrho} \to \Lambda_x \quad \text{in } C^0(G(n,m))' \quad \text{for all } x \in U \backslash Z.$$

For $x \in U \backslash Z$ we obtain by Riesz a Radon measure V^x on $G(n,m)$ such that

$$\Lambda_x(f) = \int_{G(n,m)} f(P)\, dV^x(P) \quad \text{for all } f \in C^0(G(n,m)),\ x \in U \backslash Z.$$

The functions $x \mapsto \Lambda_{x,\varrho}(f)$ are Borel, hence $\Lambda_x(f)$ is μ_V measurable for $f \in C^0(G(n,m))$. We claim that for any Borel set $B \subset G(n,m)$, the function

$$\Lambda_x(\chi_B) = \int_{G(n,m)} \chi_B(P)\, dV^x(P) = V^x(B)$$

is μ_V measurable. Namely the sets with this property form a σ-algebra, and for $K \subset U$ compact the property follows using the monotone approximation $\chi_k(P) = \big(1 - k\,\mathrm{dist}(P,K)\big)^+$ of the characteristic function χ_K. Next by Radon Nikodym we know that for any function $\chi_A(x) f(P)$, where $A \subset U$ is Borel and $f \in C^0(G(n,m))$ is nonnegative, we have property (2), that is

$$\int_{G_n(A)} f(P)\, dV(x,P) = \int_A \int_{G(n,m)} f(P)\, dV^x(P)\, d\mu_V(x),$$

Repeating the argument just before, this holds also with f replaced by χ_B where $B \subset G(n,m)$ is Borel. The product sets $A \times B$ generate the Borel σ-algebra of $U \times G(n,m)$, therefore we now have (1) and (2) for all Borel sets. The theorem follows by using a final approximation by step functions. $\square$

Example 10.5. *Let $M \subset U$ be a C^1 submanifold of dimension n. Assume that M is properly embedded, this means $M \cap K$ is compact for any compact $K \subset U$. Using the Riesz representation theorem, we define the n-varifold V by the functional*

$$V(f) = \int_M f(x, T_x M)\, d\mathcal{H}^n(x) \quad \text{for all } f \in C^0_c\big(U \times G(n,m)\big).$$

Taking $f(x,P) = \varphi(x)$ where $\varphi \in C^0_c(U)$ we obtain from (10.1)

$$\int_U \varphi(x)\, d\mu_V(x) = \int_{G_n(U)} \varphi(x)\, dV(x,P) = V(\varphi) = \int_M \varphi(x)\, d\mathcal{H}^n(x).$$

Thus $\mu_V = \mathcal{H}^n \llcorner M$. Next take a product function $\varphi(x) f(P)$ and compute

$$\begin{aligned}
\int_M \varphi(x) f(T_x M)\, d\mathcal{H}^n(x) &= V(\varphi f) \\
&= \int_U \varphi(x) \int_{G(n,m)} f(P)\, dV^x(P)\, d\mu_V(x) \\
&= \int_M \varphi(x) \int_{G(n,m)} f(P)\, dV^x(P)\, d\mathcal{H}^n(x).
\end{aligned}$$

For μ_V-almost all $x \in U$ we conclude that

$$\int_{G(n,m)} f(P)\, dV^x(P) = f(T_x M) \quad \text{for all } f \in C^0(G(n,m)).$$

This means $V^x = \delta_{T_x M}$ for μ_V-almost all $x \in U$.

Our next issue is to define the pushforward of a varifold $V \in \mathbb{V}_n(U)$ under a map $\phi \in C^1(U,U')$. We have an induced map between the Graßmannians given by

$$G_n\phi : G_n^+(U) \to G_n(U'),\ G_n\phi(x,P) = (\phi(x), D\phi(x)P). \tag{10.3}$$

Here $G_n^+(U)$ is the set of (x,P) where $D\phi(x)|_P$ is injective, this restriction is obviously needed. While the definition views $G(n,m)$ as the set of subspaces, a description in terms of projections is needed for computations. Assume that $v_1, \ldots, v_n$ is a basis of P. Define the Gram matrix

$$G(x) \in \mathbb{R}^{n \times n},\ G_{ij}(x) = \langle D\phi(x) v_i, D\phi(x) v_j \rangle.$$

We have $G(x) > 0$ if and only if $(x,P) \in G_n^+(U)$. The projection onto $Q = D\phi(x)P$ is

$$P_Q w = \sum_{i,j=1}^n G^{ij}(x) \langle w, D\phi(x) v_i \rangle D\phi(x) v_j.$$

Note that $v_1, \ldots, v_n$ need not be orthonormal, thus for P' close to P we may use the basis $v_i' = P' v_i$. In particular, we see that $G_n^+(U)$ is open and $G_n\phi$ is continuous for ϕ in C^1. For any $P \in G(n,m)$ we define the Jacobian

$$J_P\phi(x) = \det\big((D\phi(x)|_P)^* D\phi(x)|_P\big)^{\frac12} = (\det G(x))^{\frac12}, \quad \text{if } v_1, \ldots, v_n \text{ is orthonormal.}$$

$J\phi : G_n(U) \to [0,\infty)$ is continuous, and $J_P\phi(x) > 0$ if and only if $(x,P) \in G_n^+(U)$.

Definition 10.6. *Let $V \in \mathbb{V}_n(U)$ be a varifold such that $\phi|_{\operatorname{spt}\mu_V}$ is proper. The pushforward $\phi_* V \in \mathbb{V}_n(U')$ is given by*

$$\phi_* V(B) = \int_{(G_n\phi)^{-1}(B)} J_P\phi(x)\, dV(x,P) \quad \text{for any } B \subset G_n(U').$$

In this lecture we only need the case when ϕ is a diffeomorphism. Then $G_n^+(u) = G_n(U)$, and the condition that ϕ is proper is automatic. We emphasize that the varifold pushforward is different from the measure pushforward under $G_n\phi$, because the Jacobian appears.

Example 10.7. *For $x \in U$ and $\lambda > 0$, let $V_{x,\lambda}$ be the pushforward of $V \in \mathbb{V}_n(U)$ by*

$$\eta_{x,\lambda} : U \to \mathbb{R}^m,\ \eta_{x,\lambda}(y) = \frac{y-x}{\lambda}.$$

We have $D\eta_{x,\lambda}(x) = \frac{1}{\lambda}\mathrm{Id}$, in particular $J_P\eta_{x,\lambda}(y) = \lambda^{-n}$ for all $(y,P) \in G_n(U)$. Moreover $G_n\eta_{x,\lambda}(y,P) = \big(\frac{y-x}{\lambda}, P\big)$. If $A \subset \mathbb{R}^m$ is bounded then $\eta_{x,\lambda}^{-1}(A) = x + \lambda A$ is contained in U for $\lambda > 0$ sufficiently small. Thus

$$\begin{aligned}
\mu_{V_{x,\lambda}}(A) &= V_{x,\lambda}(A \times G(n,m)) \\
&= \frac{1}{\lambda^n} V((x+\lambda A) \times G(n,m)) \\
&= \frac{1}{\lambda^n}\mu_V(x+\lambda A) \\
&= (\mu_V)_{x,\lambda}(A).
\end{aligned}$$

Furthermore, testing with a product set yields

$$\begin{aligned}
\int_{\mathbb{R}^m} \chi_A(z)(V_{x,\lambda})^z(B)\, d\mu_{V_{x,\lambda}}(z) &= V_{x,\lambda}(A \times B) \\
&= \lambda^{-n} V\big((x+\lambda A) \times B\big) \\
&= \lambda^{-n} \int_U \chi_{x+\lambda A}(y)\, V^y(B)\, d\mu_V(y) \\
&= \int_{\mathbb{R}^m} \chi_A(z) V^{x+\lambda z}(B)\, d\mu_{x,\lambda}(z).
\end{aligned}$$

In summary, the blowup varifold $V_{x,\lambda} = (\eta_{x,\lambda})_* V$ has $\mu_{V_{x,\lambda}} = \mu_{x,\lambda}$ and $(V_{x,\lambda})^z = V^{x+\lambda z}$.

Definition 10.8. *A varifold $V \in \mathbb{V}_n(U)$ is called rectifiable if the following holds:*

(1) *$T_x\mu_V$ exists for μ_V-almost all $x \in U$, with some multiplicity $\theta(x) \in (0,\infty)$.*

(2) *$V^x = \delta_{T_x\mu_V}$ for μ_V-almost all $x \in U$.*

Denote by M the set where $T_x\mu_V$ exists. Recalling Corollary 9.8 we obtain the representation

$$V(f) = \int_U f(x, T_x\mu_V)\, d\mu_V(x) = \int_M f(x, T_xM)\, d(\mathcal{H}^n \llcorner \theta)(x). \tag{10.4}$$

Reversely, let M be $\mathcal{H}^n$-measurable and countably n-rectifiable, and let $\theta \in L^1_{\mathrm{loc}}(\mathcal{H}^n)$ such that $M = \{\theta > 0\}$. We may define a varifold by the right hand side of (10.4), using the tangent space $T_x M = G_M(x)$ as defined in Corollary 9.8. The function $f(x, T_x M)$ is then $\mathcal{H}^n$-measurable, in fact for any product $A \times B \subset U \times G(n,m)$ of Borel sets we have

$$(\mathrm{id} \times G)^{-1}(A \times B) = A \cap G^{-1}(B).$$

This is $\mathcal{H}^n$-measurable by Corollary 9.8. Hence the varifold is well-defined by the formula

$$V(f) = \int_M f(x, T_x M)\, d(\mathcal{H}^n \llcorner \theta)(x).$$

Arguing as in Example 10.5 we see that $\mu_V = \mu$ and $V^x = \delta_{T_x \mu}$ for μ-almost all $x \in U$. We write $V = \mathbf{v}(M, \theta)$ if a varifold arises in this way.

We want to compute the pushforward in the case of a rectifiable varifold. For this we need an extension of the area formula to rectifiable sets resp. varifolds. Consider first an n-dimensional submanifold $M \subset U$ of class C^1. The notion of a C^1 map $f : M \to \mathbb{R}^p$ can be defined using charts. In particular the differential $Df(x) : T_x M \to \mathbb{R}^p$ is well-defined and linear. We introduce the following operators, where $\tau_1, \ldots, \tau_n$ is any orthonormal basis of $T_x M$:

- the gradient of $f \in C^1(M)$ is $\nabla^M f(x) = \sum_{i=1}^n (Df(x)\tau_i)\, \tau_i$. It is characterized by

  $$\langle \nabla^M f(x), v \rangle = Df(x)v \quad \text{for all } v \in T_x M.$$

  If f is C^1 on all of U then $\nabla^M f(x) = P_{T_x M} \nabla f(x)$, where $\nabla f(x)$ is the gradient in $\mathbb{R}^m$. In this context $\nabla^M f(x)$ is sometimes called the tangential gradient.

- for $f \in C^1(M, \mathbb{R}^p)$ the divergence on M is defined by $\mathrm{div}^M f(x) = \mathrm{tr}\left(P_{T_x M} Df(x)\right)$, or

  $$\mathrm{div}^M f(x) = \sum_{i=1}^n \langle Df(x)\tau_i, \tau_i \rangle.$$

- the Jacobian of $f \in C^1(M, \mathbb{R}^p)$ is given by $J_M f(x) = \sqrt{\det(Df(x)^* Df(x))}$, hence

  $$J^M f(x) = \sqrt{\det G(x)} \quad \text{where } G_{ij}(x) = \langle Df(x)\tau_i, Df(x)\tau_j \rangle.$$

Now let M be countably n-rectifiable and $\mathcal{H}^n$-measurable. For $f \in C^1(U, \mathbb{R}^p)$ we define

$$D^M f(x) = Df(x)|_{T_x M} \quad \text{for } \mathcal{H}^n\text{-a.e. } x \in M.$$

Here $T_x M$ is as in Corollary 9.8. In particular if N is an n-dimensional C^1 submanifold then $T_x M = T_x N$ for $\mathcal{H}^n$-almost all $x \in M \cap N$.

Lemma 10.9. *Let $f \in C^1(U, U')$ where $U \subset \mathbb{R}^m$, $U' \subset \mathbb{R}^p$ are open. Assume that $M \subset U$ is $\mathcal{H}^n$-measurable and countably n-rectifiable, and let $g : M \to [0, \infty)$ be $\mathcal{H}^n$-measurable. Then the function $y \mapsto \sum_{\{x \in M : f(x) = y\}} g(x)$ is $\mathcal{H}^n$ measurable on U' and*

$$\int_M g(x) J^M f(x)\, d\mathcal{H}^n(x) = \int_{U'} \sum_{x \in M : f(x) = y} g(x)\ d\mathcal{H}^n(y).$$

Proof. Assume first that M is a C^1 submanifold parametrized by $\varphi \in C^1(\Omega, \mathbb{R}^m)$. Then

$$D(f \circ \varphi)(x) = D^M f(\varphi(x)) D\varphi(x) \quad \text{where } D\varphi(x) : \mathbb{R}^n \to T_{\varphi(x)}M.$$

This implies

$$D(f \circ \varphi)(x)^* D(f \circ \varphi)(x) = D\varphi(x)^* D^M f(\varphi(x))^* D^M f(\varphi(x)) D\varphi(x).$$

To compute the Jacobian we choose an orthonormal basis $\mathcal{A} = \{\tau_1, \ldots, \tau_n\}$ of $T_{\varphi(x)}M$, and denote by $\mathcal{E} = \{e_1, \ldots, e_n\}$ the standard basis of $\mathbb{R}^n$. Then

$$\big(D(f \circ \varphi)(x)^* D(f \circ \varphi)(x)\big)_{\mathcal{E}\mathcal{E}} = D\varphi(x)^*_{\mathcal{E}\mathcal{A}} \big(D^M f)(\varphi(x))^* D^M f(\varphi(x)\big)_{\mathcal{A}\mathcal{A}} D\varphi(x)_{\mathcal{A}\mathcal{E}}.$$

Taking the determinant we obtain

$$J(f \circ \varphi)(x) = J^M f(\varphi(x))\, J\varphi(x).$$

Thus we get from the standard area formula

$$\begin{aligned}
\int_{U'} \sum_{p \in M : f(p) = y} g(p)\, d\mathcal{H}^n(y) &= \int_{U'} \sum_{x \in \Omega : f(\varphi(x)) = y} g(\varphi(x))\, d\mathcal{H}^n(y) \\
&= \int_{\varphi^{-1}(M)} g(\varphi(x))\, J(f \circ \varphi)(x)\, d\mathcal{L}^n(x) \\
&= \int_{\varphi^{-1}(M)} g(\varphi(x)) J^M f(\varphi(x)) J\varphi(x)\, d\mathcal{L}^n(x) \\
&= \int_M g(p) J^M f(p)\, d\mathcal{H}^n(p).
\end{aligned}$$

The formula extends to any C^1 submanifold by a partition of unity. Now let M be as in the theorem, and assume again that M is covered by C^1 submanifolds N_j, $j \in \mathbb{N}$, up to a $\mathcal{H}^n$ null set. Let $N_j' = N_j \backslash \bigcup_{i=1}^{j-1} N_i$. Then

$$\begin{aligned}
\int_{U'} \sum_{p \in M : f(p) = y} \chi_{N_j'}(p)\, g(p)\, d\mathcal{H}^n(y) &= \int_M \chi_{N_j'}(p)\, g(p)\, J^{N_j} f(p)\, d\mathcal{H}^n(p) \\
&= \int_M \chi_{N_j'}(p)\, g(p)\, J^M f(p)\, d\mathcal{H}^n(p).
\end{aligned}$$

Here we used $D^M f(p) = D^{N_j}(p)$ $\mathcal{H}^n$-a.e. on $M \cap N_j$. The lemma follows by adding up. □

Theorem 10.10 (rectifiable pushforward)**.** *Let $f \in C^1(U, U')$ where $U \subset \mathbb{R}^m$, $U' \subset \mathbb{R}^p$ are open. Let V be a rectifiable varifold in U such that $f|_{\operatorname{spt} \mu_V}$ is proper. Then $f_* V$ is rectifiable, moreover if V is represented by $V = \mathbf{v}(M, \theta)$, then $f_* V = \mathbf{v}(f(M), \theta_f)$ where*

$$\theta_f(y) = \sum_{x \in M : f(x) = y} \theta(x).$$

Proof. By Lemma 9.2 M is covered by C^1 submanifolds N_j, $j \in \mathbb{N}$, up to a $\mathcal{H}^n$ null set M_0. Thus $f(M)$ is covered by the $f(N_j)$ up to the null set $f(M_0)$, and is countably n-rectifiable by definition. Further $f(M) = \{\theta_f > 0\}$, so that $f(M)$ is also $\mathcal{H}^n$-measurable by Lemma 10.9 above. We compute for any Borel set $B \subset U'$

$$\begin{aligned}
\mu_{f_*V}(B) &= f_*V(B \times G(n,m)) \\
&= \int_{f^{-1}(B)\times G(n,m)} J_P f(x)\, dV(x,P) \\
&= \int_{f^{-1}(B)} J^M f(x)\, d\mu_V(x) \\
&= \int_M \chi_{f^{-1}(B)}(x)\, J^M f(x)\, \theta(x)\, d\mathcal{H}^n(x) \\
&= \int_{U'} \sum_{x\in M: f(x)=y} \theta(x)\, \chi_B(y)\, d\mathcal{H}^n(y) \\
&= (\mathcal{H}^n \llcorner \theta_f)(B).
\end{aligned}$$

Thus $\mu_{f_*V} = \mathcal{H}^n \llcorner \theta_f$. To compute $(f_*V)^y$ we need an extra consideration. We claim that for $\mathcal{H}^n$-almost all $y \in f(M)$ we have

$$Df(x)\, T_x M = T_y f(M) \quad \text{for all } x \in f^{-1}\{y\}. \tag{10.5}$$

Let $M^+ = \{x \in M : J^M f(x) > 0\}$. By the area formula the set $f(M\backslash M^+)$ is a $\mathcal{H}^n$ null set:

$$\mathcal{H}^n\big(f(M/M^+)\big) \leq \int_{U'} \mathcal{H}^0\big(M\backslash M^+ \cap f^{-1}\{y\}\big)\, d\mathcal{H}^n(y) = \int_{M\backslash M^+} J^M f(x)\, d\mathcal{H}^n(x) = 0.$$

By Corollary 9.8 $T_x M = T_x N_j$ for $\mathcal{H}^n$-almost all $x \in M^+ \cap N_j$. For these x the image $f(N_j)$ is locally a manifold, we have

$$Df(x) T_x M \quad = \quad Df(x) T_x N_j = T_{f(x)} f(N_j).$$

On the other hand, again by Corollary 9.8, locally

$$T_y f(M) \quad = \quad T_y f(N_j) \quad \text{for } \mathcal{H}^n\text{-a.e. } y \in f(M) \cap f(N_j).$$

Combining shows (10.5). Now we can calculate for a product set $B \times S$ in $U' \times G(n,m)$

$$\begin{aligned}
\int_B (f_*V)^y(S)\, d\mu_{f_*V}(y) &= (f_*V)(B \times S) \\
&= \int_{\{x\in f^{-1}(B): Df(x)P\in S\}} J_P f(x)\, dV(x,P) \\
&= \int_{f^{-1}(B)} \int_{Df(x)^{-1}(S)} J_P f(x)\, dV^x(P)\, d\mu_V(x) \\
&= \int_{f^{-1}(B)} \delta_{Df(x)T_xM}(S) J^M f(x)\, \theta(x) d\mathcal{H}^n(x) \\
&= \int_B \sum_{x\in M: f(x)=y} \delta_{Df(x)T_xM}(S) \theta(x)\, d\mathcal{H}^n(y) \\
&= \int_B \delta_{T_y f(M)}(S)\, \theta_f(y)\, d\mathcal{H}^n(y) \\
&= \int_B \delta_{T_y f(M)}(S)\, d\mu_{f_*V}(y).
\end{aligned}$$

Recalling again Corollary 9.8 we have $(f_*V)^y = \delta_{T_y f(M)} = \delta_{T_y(\mu_{f_*V})}$. $\square$

Chapter 11

The first variation

For his paper from 1973 Allard selected the title *The first variation of a varifold.* Previous applications of geometric measure theory were dealing with minimizers, either in the context of BV functions and Caccioppoli sets (Di Giorgi) or in the setting of area-minimizing currents (Federer-Fleming). The focus of Allard is on results for critical points instead of minimizers, and this is pointed out in his title.

Let $U \subset \mathbb{R}^m$ be open, and let $\phi \in C^2(U \times (-\delta,\delta), U)$, $\phi_t = \phi(\,\cdot\,,t)$, be a family of maps with the following properties:

(1) $\phi(\,\cdot\,,0) = \mathrm{id}_U$.

(2) there is a compact set $K \subset U$ such that $\phi(\,\cdot\,,t) = \mathrm{id}$ on $U \backslash K$, for all $t \in (-\varepsilon,\varepsilon)$.

Such a family is called a variation with compact support in U. The associated velocity field is

$$X(x) = \frac{\partial \phi}{\partial t}(x,0), \qquad \text{hence spt}\, X \subset K.$$

Theorem 11.1 (first variation)**.** *Let $V \in \mathbb{V}_n(U)$ be an n-varifold with finite mass in U. Then for any variation ϕ as above*

$$\frac{d}{dt}\,\mu_{(\phi_t)_* V}(U)|_{t=0} = \int_{G_n(U)} \langle DX(x), P\rangle\, dV(x,P). \tag{11.1}$$

If $v_1,\dots,v_n$ is an orthonormal basis of P then

$$\langle DX(x), P\rangle = \sum_{i=1}^n \langle D_{v_i} X(x), v_i\rangle = \mathrm{div}_P X(x).$$

Proof. By definition of the pushforward we have

$$\mu_{(\phi_t)_* V}(U) = (\phi_t)_* V(U \times G(n,m)) = \int_{U \times G(n,m)} J_P \phi_t(x)\, dV(x,P).$$

We compute using an orthonormal basis $v_1,\dots,v_n$ of P

$$\frac{\partial}{\partial t} J_P \phi_t(x)|_{t=0} = \frac{\partial}{\partial t}\big(\det\langle D\phi_t(x)v_i, D\phi_t(x)v_j\rangle\big)^{\frac{1}{2}}|_{t=0} = \sum_{i=1}^n \langle D_{v_i} X(x), v_i\rangle.$$

The assumptions allow to differentiate under the integral (check), we get

$$\frac{d}{dt}\mu_{(\phi_t)_*V}(U)|_{t=0} = \int_{G_n(U)} \langle DX(x), P\rangle \, dV(x,P).$$

The function $J_P\phi_t(x)$ is of class C^1 on $U \times (-\delta, \delta)$, and $J_P\phi_t(x) = 1$ on $U\backslash K$. Therefore we can differentiate under the integral to obtain the result. $\square$

Example 11.2. *Let $v_1, \ldots, v_k$ be unit vectors in $\mathbb{R}^m$. Denote by P_i the projection onto $\mathbb{R}v_i$, and consider the rectifiable 1-varifold*

$$V(f) = \sum_{i=1}^{k} \int_0^\infty f(sv_i, P_i)\, ds.$$

It follows that the first variation is

$$\begin{aligned}
\int_{G_1(\mathbb{R}^m)} \langle DX(x), P\rangle \, dV(x,P) &= \sum_{i=1}^{k} \int_0^\infty \langle DX(sv_i)v_i, v_i\rangle \, ds \\
&= \sum_{i=1}^{k} \int_0^\infty \frac{d}{ds}\langle X(sv_i), v_i\rangle \, ds \\
&= -\Big\langle X(0), \sum_{i=1}^{k} v_i \Big\rangle.
\end{aligned}$$

The first variation vanishes for all X if and only if $\sum_{i=1}^k v_i = 0$.

Example 11.3. *Let $M \subset U$ be a compact embedded submanifold of class C^2, possibly with boundary ∂M. As discussed the induced varifold V has weight measure $\mu_V = \mathcal{H}^n \llcorner \theta$ and vertical measures $V^x = \delta_{T_xM}$, see Example 10.5. The second fundamental form of M is*

$$A(X,Y) = (D_X Y)^\perp \quad \text{for tangential vector fields } X, Y : M \to \mathbb{R}^m.$$

Here $\perp$ means the projection onto $(TM)^\perp$. For $X \in C^1_c(U, \mathbb{R}^m)$ we compute, using a local orthonormal tangential frame $\tau_1, \ldots, \tau_n$,

$$\begin{aligned}
\mathrm{div}^M(X^\perp) &= \sum_{i=1}^{n} \langle D_{\tau_i}(X^\perp), \tau_i\rangle \\
&= \sum_{i=1}^{n} \big(D_{\tau_i}\langle X^\perp, \tau_i\rangle - \langle X^\perp, D_{\tau_i}\tau_i\rangle \big) \\
&= -\langle X, \vec{H}\rangle.
\end{aligned}$$

Here $\vec{H} = \sum_{i=1}^n A(\tau_i, \tau_i)$ is the mean curvature vector. Thus the first variation integral becomes

$$\begin{aligned}
\int_{G_n(U)} \mathrm{div}_P X(x)\, dV(x,P) &= \int_M \mathrm{div}_{T_xM} X(x)\, d\mathcal{H}^n(x) \\
&= \int_M \mathrm{div}^M X^\top \, d\mathcal{H}^n + \int_M \mathrm{div}^M X^\perp \, d\mathcal{H}^n \\
&= -\int_{\partial M} \langle X, \eta\rangle \, d\mathcal{H}^{n-1} - \int_M \langle X, \vec{H}\rangle \, d\mathcal{H}^n.
\end{aligned}$$

Here η denotes the interior unit conormal along ∂M, we used the theorem of Gauß on M for the tangential filed $X^\top$. The mean curvature term is absolutely continuous with respect $\mu_V = \mathcal{H}^n \llcorner M$, whereas the boundary term is singular.

We consider the right hand side of the first variation formula as a linear functional

$$\delta V : C^1_c(U, \mathbb{R}^m) \to \mathbb{R},\, \delta V(X) = \int_{G_n(U)} \langle DX(x), P \rangle \, dV(x,P). \tag{11.2}$$

Definition 11.4. *$V \in \mathbb{V}_n(U)$ has locally bounded first variation if for all compact $K \subset U$*

$$|\delta V|(K) = \sup \big\{ \delta V(X) : X \in C^1_c(U, \mathbb{R}^m),\, \operatorname{spt} X \subset K,\, |X| \leq 1 \big\} < \infty.$$

If this holds, then δV extends to a continuous functional on $C^0_c(U, \mathbb{R}^m)$ by density, and the Riesz representation theorem applies. $|\delta V|$ is a Radon measure, and there is a $|\delta V|$ measurable function $\eta_V : U \to \mathbb{R}^m$ with $|\eta_V| = 1$ such that

$$\delta V(X) = -\int_U \big\langle X(x), \eta_V(x) \big\rangle \, d|\delta V|(x).$$

The choice of sign is for convenience. Now by Radon-Nikodym for any Borel set A

$$|\delta V|(A) = \int_A D_{\mu_V} |\delta V| \, d\mu_V + (|\delta V| \llcorner Z)(A).$$

Here $Z = \{x \in U : D_{\mu_V} |\delta V|(x) = \infty\}$, and $\mu_V(Z) = 0$. In order to arrive at a notation analogous to manifolds, we put

$$\vec{H}_V = (D_{\mu_V} |\delta V|)\, \eta_V \quad \text{ and } \quad \sigma_V = |\delta V| \llcorner Z.$$

Then $\vec{H}_V \in L^1_{\mathrm{loc}}(\mu_V, \mathbb{R}^m)$, and the formula becomes

$$\delta V(X) = -\int_U \langle X, \vec{H}_V \rangle \, d\mu_V - \int_U \langle X, \eta_V \rangle \, d\sigma_V.$$

For a compact submanifold with boundary, $\vec{H}_V = \vec{H}$ is the mean curvature vector, $\sigma_V = \mathcal{H}^{n-1} \llcorner \partial M$ is the boundary measure and η_V is the interior conormal along ∂M.

Definition 11.5. *Let $V = \mathbf{v}(M, \theta)$ be a rectifiable n-varifold in $U \subset \mathbb{R}^m$. We say that V has weak mean curvature $H \in L^p_{\mathrm{loc}}(\mu_V)$ where $p \in [1, \infty]$, if*

$$\int_U \operatorname{div}^M X \, d\mu_V = -\int_U \langle \vec{H}, X \rangle \, d\mu_V \quad \textit{for all } X \in C^1_c(U, \mathbb{R}^m). \tag{11.3}$$

A rectifiable n-varifold V with weak mean curvature $\vec{H} = 0$ is called stationary.

If V has locally bounded first variation with singular part $\sigma_V = 0$, then V has weak mean curvature $\vec{H}_V = (D_{\mu_V} |\delta V|) \eta_V \in L^1_{\mathrm{loc}}(\mu_V)$. It is easy to see that the reverse implication is also valid. The following application relates to a maximum principle for classical surfaces.

Theorem 11.6 (inclusion principle)**.** *Let V be a rectifiable n-varifold in $\mathbb{R}^m$ with* $\operatorname{spt} \mu$ *compact and $\vec{H} \in L^1(\mu)$. Assume that*

$$\langle \vec{H}(x), x \rangle > -n \quad \textit{for all } x \in \mathbb{R}^m \backslash B_R(0).$$

Then $\operatorname{spt} \mu \subset \overline{B_R(0)}$.

Proof. Let $\gamma \in C^1([0,\infty)$ be monotonically increasing with $\gamma(r) = 0$ for $r \in [0,R]$. For the vector field $X(x) = \gamma(r)x$ where $r = |x|$ we compute

$$\operatorname{div}^M X = n\gamma(r) + r\gamma'(r)|\nabla^M r|^2.$$

X is admissible in the first variation formula since $\operatorname{spt}\mu$ is compact. We conclude

$$0 \leq \int r\gamma'(r)|\nabla^M r|^2\,dr = \int \big(\operatorname{div}^M X - n\gamma(r)\big)\,d\mu = -\int \gamma(r)\big(\langle \vec{H}(x), x\rangle + n\big)\,d\mu.$$

Choosing $\gamma(r) > 0$ for $r > R$ we conclude $\mu(\mathbb{R}^m \backslash B_r(0)) = 0$ and hence $\operatorname{spt}\mu \subset \overline{B_R(0)}$. $\square$

Corollary 11.7 (convex hull property)**.** *Let V be a rectifiable n-varifold with compact support, and assume that V is stationary in $\mathbb{R}^m\backslash K$ where K is compact. Then the support of μ is contained in the convex hull of K.*

Proof. Assume that $K \subset B_R(x_0)$. By assumption we have

$$\langle \vec{H}(x), x - x_0\rangle = 0 > -n \quad \text{ for all } x \in \mathbb{R}^m\backslash B_R(x_0).$$

The inclusion principle implies that $\operatorname{spt}\mu \subset B_R(x_0)$. It is an elementary fact that for K compact, the intersection of all balls containing K yields the convex hull. $\square$

We now come to the fundamental monotonicity formula. The original proof of Allard, see [**?**], is for general varifolds with locally bounded variation, it employs the method of slicing a varifold. Our version is taken from Simon's book [**?**] and is restricted to rectifiable varifolds.

Theorem 11.8 (monotonicity formula)**.** *Let V be a rectifiable n-varifold in $B_R(x_0) \subset \mathbb{R}^m$ with weak mean curvature $\vec{H} \in L^1_{\mathrm{loc}}(\mu)$ where $\mu = \mu_V$. Then for all $0 < \sigma \leq \varrho < R$ we have*

$$\begin{aligned} \Big[\frac{\mu(B_r(x_0))}{r^n}\Big]_{r=\sigma}^{r=\varrho} &= \int_{B_\varrho(x_0)\backslash B_\sigma(x_0)} \frac{|\nabla r(x)^\perp|^2}{|x-x_0|^n}\,d\mu(x) \\ &\quad + \int_\sigma^\varrho \frac{1}{r^{n+1}} \int_{B_r(x_0)} \langle \vec{H}(x), x - x_0\rangle\,d\mu(x)\,dr. \end{aligned} \tag{11.4}$$

Proof. In the first variation we use as vector field $X(x) = \gamma(r)(x - x_0)$ where $\gamma \in C^1_c\big([0,\varrho)\big)$ and $r = |x - x_0|$. We compute using $|\nabla r| = 1$

$$\operatorname{div}^M X(x) = n\gamma(r) + r\gamma'(r)\,|\nabla^M r|^2 = n\gamma(r) + r\gamma'(r)\big(1 - |(\nabla r)^\perp|^2\big).$$

Thus we have the identity

$$\int \big(n\gamma(r) + r\gamma'(r)\big)\,d\mu = \int r\gamma'(r)|(\nabla r)^\perp|^2\,d\mu - \int \gamma(r)\langle \vec{H}, x - x_0\rangle\,d\mu. \tag{11.5}$$

We let $\gamma(r) = \phi\big(\frac{r}{\varrho}\big)$ where $\phi \in C^1_c([0,1)$, and consider the functions

$$\begin{aligned} I(\varrho) &= \int \phi\Big(\frac{r}{\varrho}\Big)\,d\mu, \\ J(\varrho) &= \int \phi\Big(\frac{r}{\varrho}\Big)\,|(\nabla r)^\perp|^2\,d\mu, \\ L(\varrho) &= \int \phi\Big(\frac{r}{\varrho}\Big)\,\langle \vec{H}, x - x_0\rangle\,d\mu. \end{aligned}$$

Differentiating under the integrl we infer, using $r\gamma'(r) = \frac{r}{\varrho}\,\phi'\big(\frac{r}{\varrho}\big)$,

$$\begin{aligned}
I'(\varrho) &= \int \phi'\Big(\frac{r}{\varrho}\Big)\,\Big(-\frac{r}{\varrho^2}\Big)\,d\mu = -\frac{1}{\varrho}\int r\gamma'(r)\,d\mu,\\
J'(\varrho) &= \int \phi'\Big(\frac{r}{\varrho}\Big)\,\Big(-\frac{r}{\varrho^2}\Big)|(\nabla r)^\perp|^2\,d\mu = -\frac{1}{\varrho}\int r\gamma'(r)|(\nabla r)^\perp|^2\,d\mu.
\end{aligned}$$

Using (11.5) we calculate

$$\begin{aligned}
\frac{d}{d\varrho}\big(\varrho^{-n}I(\varrho)\big) &= -\varrho^{-n-1}\int n\gamma(r) + r\gamma'(r)\big)\,d\mu\\
&= -\varrho^{-n-1}\Big(\int r\gamma'(r)|(\nabla r)^\perp|^2\,d\mu - \int \gamma(r)\langle \vec{H}, x-x_0\rangle\,d\mu\Big)\\
&= \varrho^{-n}J'(\varrho) + \varrho^{-n-1}L(\varrho).
\end{aligned}$$

A pointwise differential inequality cannnot be integrated easily, therefore we now pass to a weak formulation. For any text function $\eta \in C^1_c((0,R))$ we have

$$-\int \varrho^{-n}I(\varrho)\eta'(\varrho)\,d\varrho = \int \varrho^{-n}J'(\varrho)\eta(\varrho)\,d\varrho + \int \varrho^{-n-1}L(\varrho)\eta(\varrho)\,d\varrho. \tag{11.6}$$

We choose $\phi_\varepsilon \in C^1_c([0,1)$ with $0 \leq \phi_\varepsilon \leq 1$, such that

$$\phi_\varepsilon(s) = 1 \;\text{ for } 0 \leq s \leq 1-\varepsilon \quad \text{ and } \quad |\phi'_\varepsilon(s)| \leq \frac{C}{\varepsilon} \text{ for all } s.$$

The functions $\gamma_\varepsilon(r) = \phi_\varepsilon\big(\frac{r}{\varrho}\big)$ converge pointwise everywhere to the characteristic function of the (open) ball $B_\varrho(x_0)$. As $\varepsilon \searrow 0$ we have by dominated convergence

$$\begin{aligned}
I_\varepsilon(\varrho) &= \int \phi_\varepsilon\Big(\frac{r}{\varrho}\Big)\,d\mu \to \mu\big(B_\varrho(x_0)\big),\\
L_\varepsilon(\varrho) &= \int \phi_\varepsilon\Big(\frac{r}{\varrho}\Big)\langle \vec{H}, x-x_0\rangle\,d\mu \to \int_{B_\varrho(x_0)} \langle \vec{H}, x-x_0\rangle\,d\mu.
\end{aligned}$$

Here we used that $\vec{H} \in L^1_{\text{loc}}(\mu)$. We have further again by dominated convergence

$$\begin{aligned}
\int \varrho^{-n}I_\varepsilon(\varrho)\,\eta'(\varrho)\,d\varrho &\to \int \varrho^{-n}\mu\big(B_\varrho(x_0)\big)\,\eta'(\varrho)\,d\varrho,\\
\int \varrho^{-n-1}L_\varepsilon(\varrho)\,\eta(\varrho)\,d\varrho &\to \int \varrho^{-n-1}\int_{B_\varrho(x_0)} \langle \vec{H}, x-x_0\rangle\,\eta(\varrho)\,d\varrho.
\end{aligned}$$

For the integral $J_\varepsilon(\varrho)$ we argue differently. For any $\sigma \in (0,\varrho)$ we have for $\varepsilon > 0$ small

$$\begin{aligned}
&\Big|\varrho^{-n}J'_\varepsilon(\varrho) - \frac{d}{d\varrho}\int_{\{|x-x_0|\geq\sigma\}} \phi_\varepsilon\big(\frac{r}{\varrho}\big)\,\frac{|(\nabla r)^\perp|^2}{|x-x_0|^n}\,d\mu(x)\Big|\\
= &\;\Big|\int_{B_\varrho(x_0)\backslash B_{(1-\varepsilon)\varrho}(x_0)} \phi'_\varepsilon\Big(\frac{r}{\varrho}\Big)\,\Big(-\frac{r}{\varrho^2}\Big)\,|(\nabla r)^\perp|^2\,\big(\varrho^{-n} - r^{-n}\big)\,d\mu\Big|\\
\leq &\;\frac{C}{\varepsilon\varrho^{n+1}}\big(\underbrace{(1-\varepsilon)^{-n} - 1}_{\leq C\varepsilon}\big)\,\mu\big(B_\varrho(x_0)\backslash B_{(1-\varepsilon)\varrho}(x_0)\big) \xrightarrow{\varepsilon\searrow 0} 0.
\end{aligned}$$

By dominated convergence, this implies further

$$\int \Big(\varrho^{-n} J_\varepsilon'(\varrho) - \frac{d}{d\varrho} \int_{\{|x-x_0|\geq\sigma\}} \phi_\varepsilon\Big(\frac{r}{\varrho}\Big) \frac{|(\nabla r)^\perp|^2}{|x-x_0|^n}\, d\mu(x)\Big)\, \eta(\varrho)\, d\varrho \to 0.$$

On the other hand partial integration yields

$$-\int \frac{d}{d\varrho} \int_{\{|x-x_0|\geq\sigma\}} \phi_\varepsilon\Big(\frac{r}{\varrho}\Big) \frac{|(\nabla r)^\perp|^2}{|x-x_0|^n}\, d\mu(x)\, \eta(\varrho)\, d\varrho \overset{\varepsilon\searrow 0}{\longrightarrow} \int \int_{B_\varrho(x_0)} \frac{|(\nabla r)^\perp|^2}{|x-x_0|^n}\, d\mu(x)\, \eta'(\varrho)\, d\varrho.$$

Collecting terms we obtain for $\varrho \in (\sigma, R)$ the weak differential equality

$$\frac{d}{d\varrho}\Big(\frac{\mu\big(B_\varrho(x_0)\big)}{\varrho^n} - \int_{B_\varrho(x_0)\setminus B_\sigma(x_0)} \frac{|(\nabla r)^\perp|^2}{|x-x_0|^n}\, d\mu(x)\Big) = \frac{1}{\varrho^{n+1}} \int_{B_\varrho(x_0)} \langle \vec{H}, x - x_0\rangle\, d\mu(x). \tag{11.7}$$

To integrate the equation we first observe the continuity of the function

$$r \longmapsto \frac{\mu\big(B_r(x_0)\big)}{r^n} - \int_{B_r(x_0)\setminus B_\sigma(x_0)} \frac{|\nabla r(x)^\perp|^2}{|x-x_0|^n}\, d\mu(x).$$

As $B_\varrho(x_0)$ is the open ball the continuity for $r \nearrow \varrho$ is clear. For $r \searrow \varrho$ we observe that

$$\nabla^M r(x) = 0 \quad \text{ for } \mathcal{H}^{n-1}\text{-almost all } x \in M \cap \partial B_r(x).$$

To see this let M be covered by C^1 submanifolds N_j up to a $\mathcal{H}^n$ null set. By (9.9) we have $\nabla^M r(x) = \nabla^{N_j} r(x)$ for $\mathcal{H}^n$-almost all $x \in M \cap N_j$. But $\{x \in N_j \cap \partial B_r(x_0) : \nabla^{N_j} r(x) \neq 0\}$ is an $(n-1)$-dimensional submanifold and hence a $\mathcal{H}^n$ null set, the claim follows. Now the right hand side in (11.7) is locally bounded on $(0, R)$, hence its integral is locally Lipschitz. Thus we obtain (11.4) for all $\varrho \in [\sigma, R)$, up to an integration constant. By right continuity at $r = \sigma$, the constant is zero and the theorem is proved. □

Remark 11.9. *Put $h(x) = \langle \vec{H}(x), x - x_0\rangle$. Using Fubini the mean curvature term in (11.4) can be transformed as follows:*

$$\begin{aligned}
\int_\sigma^\varrho \frac{1}{r^{n+1}} \int_{B_r(x_0)} h(x)\, d\mu(x)\, dr &= \int_\sigma^\varrho \frac{1}{r^{n+1}} \int_{B_\varrho(x_0)} h(x) \chi_{\{|x-x_0|<r\}}\, d\mu(x)\, dr \\
&= \int_{B_\varrho(x_0)} h(x) \int_\sigma^\varrho \frac{1}{r^{n+1}} \chi_{\{|x-x_0|<r\}}\, dr\, d\mu(x) \\
&= \int_{B_\varrho(x_0)} h(x) \Big(\frac{1}{\max(|x-x_0|, \sigma)^n} - \frac{1}{\varrho^n}\Big)\, d\mu(x).
\end{aligned}$$

Lemma 11.10. *Let V be a rectifiable n-varifold in $B_R(x_0) \subset \mathbb{R}^m$ sstisfying*

$$\|\vec{H}\|_{L^p(\mu)} \leq \Gamma \quad \text{for some } p > n. \tag{11.8}$$

Then for $0 < \sigma \leq \varrho < R$ we have

$$\Big[\Big(\frac{\mu(B_r(x_0))}{r^n} \Big)^{\frac{1}{p}} + \frac{\Gamma}{p-n}\, r^{1-\frac{n}{p}} \Big]_{r=\sigma}^{r=\varrho} \geq 0.$$

Proof. From (11.7) we know that

$$\frac{d}{dr}\underbrace{\frac{\mu(B_r(x_0))}{r^n}}_{=:I(r)} \geq -\frac{1}{r^n}\int_{B_r(x_0)} |\vec{H}(x)|\, d\mu(x) \geq -\Gamma\Big(\frac{\mu(B_r(x_0)}{r^n}\Big)^{1-\frac{1}{p}} r^{-\frac{n}{p}}.$$

This implies further

$$\frac{d}{dr}I(r)^{\frac{1}{p}} = \frac{1}{p}I(r)^{\frac{1}{p}-1}I'(r) \geq -\frac{\Gamma}{p}\, r^{-\frac{n}{p}} = -\frac{d}{dr}\,\frac{\Gamma}{p-n}\, r^{1-\frac{n}{p}}.$$

The inequality holds weakly, and the lemma follows by integration. □

Theorem 11.11 (existence of density)**.** *Let V be a rectifiable n-varifold in $U \subset \mathbb{R}^m$ with weak mean curvature $\vec{H} \in L^p_{\mathrm{loc}}(\mu_V)$ for some $p > n$. Then the density*

$$\theta^n(\mu_V, x) = \lim_{r \searrow 0} \frac{\mu_V(B_r(x))}{\alpha_n r^n}$$

exists for all points $x \in U$. Moreover the function $\theta^n(\mu_V, \cdot\,)$ is upper semicontinuous.

Proof. We put $\mu = \mu_V$ and assume that $\Gamma = \|\vec{H}\|_{L^p(\mu)} < \infty$. The existence of the density is then immediate from Lemma 11.10. Let $B_\sigma(y) \subset B_\varrho(x)$. Then $r = \mathrm{dist}(y, \partial B_\varrho(x)) \in [\sigma, \varrho]$, and we estimate by Lemma 11.10

$$\begin{aligned}\Big(\frac{\mu(B_\sigma(y))}{\alpha_n \sigma^n}\Big)^{\frac{1}{p}} &\leq \Big(\frac{\mu(B_r(y))}{\alpha_n r^n}\Big)^{\frac{1}{p}} + \frac{\Gamma}{p-n} r^{1-\frac{n}{p}}\\ &\leq \Big(\frac{\varrho}{r}\Big)^{\frac{1}{p}}\Big(\frac{\mu(B_\varrho(x))}{\alpha_n \varrho^n}\Big)^{\frac{1}{p}} + \frac{\Gamma}{p-n}\varrho^{1-\frac{n}{p}}.\end{aligned}$$

For $\sigma \searrow 0$ we obtain

$$\theta^n(\mu, y) \leq \Big(\frac{\varrho}{r}\Big)^{\frac{1}{p}}\Big(\frac{\mu(B_\varrho(x))}{\alpha_n \varrho^n}\Big)^{\frac{1}{p}} + \frac{\Gamma}{p-n}\varrho^{1-\frac{n}{p}}.$$

Letting now $y \to x$, hence $r \to \varrho$, we get

$$\limsup_{y \to x}\, \theta^n(\mu, y)^{\frac{1}{p}} \leq \Big(\frac{\mu(B_\varrho(x))}{\alpha_n \varrho^n}\Big)^{\frac{1}{p}} + \frac{\Gamma}{p-n}\varrho^{1-\frac{n}{p}}.$$

Finally we let $\varrho \searrow 0$ and obtain

$$\limsup_{y \to x}\, \theta^n(\mu, y)^{\frac{1}{p}} \leq \theta^n(\mu, x)^{\frac{1}{p}}.$$

This is the upper semicontinuity. □

Lemma 11.12. *Let V be a rectifiable n-varifold in $B_R(x_0) \subset \mathbb{R}^m$. Assume that*

(1) *$\theta^n(\mu, x_0) \in (0, \infty)$ exists.*

(2) *$V_{x_0, \lambda_j} \to Y$ for some sequence $\lambda_j \searrow 0$, where Y is a stationary rectifiable n-varifold.*

Then Y is a cone around the origin, that is $Y_{0,\lambda} = Y$ for all $\lambda > 0$.

Proof. For any Radon measure γ and $\lambda > 0$ we let $\gamma_\lambda = \lambda^{-n}\eta_\lambda(\gamma)$ where $\eta_\lambda(x) = \frac{x}{\lambda}$. Consider

$$I(\lambda) = \int h(y)\,d\gamma_\lambda(y) = \lambda^{-n}\int h\Big(\frac{x}{\lambda}\Big)\,d\gamma(x).$$

Differentiating at $\lambda = 1$ yields

$$I'(1) = -n\int h(x)\,d\gamma(x) - \int \langle dh(x), x\rangle\,d\gamma(x).$$

To compute the derivative for all $\lambda > 0$, we apply the formula when γ is replaced by γ_λ. Note

$$(\gamma_\lambda)_\sigma = \sigma^{-n}\eta_\sigma\big(\lambda^{-n}\eta_\lambda(\gamma)\big) = (\lambda\sigma)^{-n}\eta_{\lambda\sigma}(\gamma) = \gamma_{\lambda\sigma}.$$

Using this we calculate

$$\lambda I'(\lambda) = \frac{d}{d\sigma}I(\lambda\sigma)|_{\sigma=1} = \frac{d}{d\sigma}\int h\,d\gamma_{\lambda\sigma}|_{\sigma=1} = \frac{d}{d\sigma}\int h\,d\big((\gamma_\lambda)_\sigma\big)|_{\sigma=1}.$$

Thus we obtain

$$\lambda I'(\lambda) = -\int \big(nh(x) + \langle\nabla h(x), x\rangle\big)\,d\gamma_\lambda(x). \tag{11.9}$$

Now let $\mu = \mu_V$ and $\gamma = \mu_Y$ be the weight measures of V and Y. Assumption (2) implies $\mu_{\lambda_j} \to \gamma$ as $j \to \infty$. From assumption (1) we obtain for all $\varrho > 0$ (except a countable set)

$$\frac{\gamma\big(B_\varrho(x_0)\big)}{\alpha_n\varrho^n} = \lim_{j\to\infty}\frac{\mu_{\lambda_j}\big(B_\varrho(x_0)\big)}{\alpha_n\varrho^n} = \lim_{j\to\infty}\frac{\mu\big(B_{\lambda_j\varrho}(x_0)\big)}{\alpha_n(\lambda_j\varrho)^n} = \theta^n(\mu, x_0).$$

As Y is stationary by assumption, the monotonicity formula (11.4) implies

$$\int_{B_\varrho(x_0)\setminus B_\sigma(x_0)} \frac{|(\nabla r)^\perp|^2}{|x-x_0|^n}\,d\gamma(x) = 0 \quad \text{for all } 0 < \sigma \le \varrho < R. \tag{11.10}$$

Using once more that Y is stationary we get

$$\begin{aligned} 0 &= \int \operatorname{div}^\top(h(x)x)\,d\gamma_\lambda(x) \\ &= \int \big(nh(x) + \langle(\nabla h)(x)^\top, x\rangle\big)\,d\gamma_\lambda(x) \\ &= \int \big(nh(x) + \langle(\nabla h)(x), x\rangle\big)\,d\gamma_\lambda(x). \end{aligned}$$

In the last step we used that by (11.10) we have $x \in T_x\gamma$ for γ-almost all $x \in \mathbb{R}^m$. Thus $I'(\lambda) = 0$ in (11.9) for all $\lambda > 0$, and we conclude $\gamma_\lambda = \gamma$ for all $\lambda > 0$. As Y is assumed to be rectifiable, this implies $Y_{0,\lambda} = Y$ for all $\lambda > 0$, which proves the lemma. □

Theorem 11.13 (tangent cones)**.** *Let V be a rectifiable n-varifold in $U \subset \mathbb{R}^m$ with weak mean curvature $\vec{H} \in L^p_{\mathrm{loc}}(\mu)$ for some $p > n$. Then for any $x_0 \in U$ the sequence V_{x_0,λ_j} where $\lambda_j \searrow 0$ subconverges to a stationary, rectifiable cone Y.*

Proof. We know already from Theorem 11.11 that the density $\theta^n(\mu, x_0)$ exists and is positive on $\operatorname{spt}\mu$. Putting $\Gamma = \|\vec{H}\|_{L^p(\mu)}$ we have further by Lemma 11.10 for $\lambda < \frac{R_0}{R}$

$$\begin{aligned}\mu_{x_0,\lambda}\big(B_R(0)\big) &= \lambda^{-n}\mu\big(B_{\lambda R}(x_0)\big)\\ &= R^n\,\frac{\mu\big(B_{\lambda R}(x_0)\big)}{(\lambda R)^n}\\ &\leq R^n\Big(\Big(\frac{\mu(B_{R_0}(x_0))}{R_0^n}\Big)^{\frac{1}{p}} + \frac{\Gamma}{p-n}R_0^{1-\frac{n}{p}}\Big)^p.\end{aligned}$$

By passing to a subsequence we have $V_{x_0,\lambda_j} \to Y$ and $\mu_{x_0,\lambda} \to \gamma$ where $\gamma = \mu_Y$. Now for $\phi \in C^1_c(\mathbb{R}^m, \mathbb{R}^m)$ we compute (see Example 10.7)

$$\begin{aligned}\delta V_{x_0,\lambda}(\phi) &= \int_{G_n(\mathbb{R}^m)} \langle D\phi(x), P\rangle\, dV_{x_0,\lambda}(x,P)\\ &= \lambda^{-n}\int_{G_n(B_R(x_0))} \Big\langle D\phi\Big(\frac{y-x_0}{\lambda}\Big), P\Big\rangle\, dV(y,P)\\ &= \lambda^{1-n}\int_{B_R(x_0)} \operatorname{div}^{\top}\big(\phi\circ\eta_{x_0,\lambda}\big)(y)\, d\mu(y)\\ &= -\lambda^{1-n}\int_{B_R(x_0)} \big\langle \vec{H}(y), \phi\circ\eta_{x_0,\lambda}(y)\big\rangle\, d\mu(y)\\ &= -\lambda\int_{\mathbb{R}^m} \big\langle \vec{H}(x_0+\lambda x), \phi(x)\big\rangle\, d\mu_{x_0,\lambda}(x).\end{aligned}$$

Thus $V_{x_0,\lambda}$ has mean curvature $\vec{H}_{x_0,\lambda}(x) = \lambda\,\vec{H}(x_0+\lambda x)$, and

$$\Big(\int_{B_R(0)} |\vec{H}_{x_0,\lambda}|^p\, d\mu_{x_0,\lambda}\Big)^{\frac{1}{p}} = \lambda^{1-\frac{n}{p}}\Big(\int_{B_{\lambda R}(x_0)} |\vec{H}|^p\, d\mu\Big)^{\frac{1}{p}}.$$

Assuming $\operatorname{spt}\phi \subset B_R(0)$ and $|\phi| \leq 1$ we conclude

$$\begin{aligned}\big|\delta V_{x_0,\lambda}(\phi)\big| &\leq \|\vec{H}_{x_0,\lambda}\|^{\frac{1}{p}}_{L^p\left(\mu_{x_0,\lambda};B_R(0)\right)}\,\mu_{x_0,\lambda}\big(B_R(0)\big)^{1-\frac{1}{p}}\\ &\leq C\,\|\vec{H}\|_{L^p(\mu;B_R(x_0))}\,\lambda^{1-\frac{n}{p}}R^n \to 0 \quad \text{as } \lambda \searrow 0.\end{aligned}$$

It follows that $\delta Y(\phi) = \lim_{j\to\infty}\delta V_{x_0,\lambda_j}(\phi) = 0$, hence Y is stationary. Applying Lemma 11.12 we conclude that Y is a stationary cone. Actually for this we need to know in addition that Y is rectifiable, this will be proved in the next section. □