Part III — Introduction to Geometric Measure Theory

Based on lectures by S Becker-Kahn
Notes taken by Daniel Ng

Lent 2020

Contents

1 Measure Theory

Carathéodory's Construction

Definition 1.1. Let X be a set. An *outer measure* on X is a function $\mu\colon 2^X \to [0,+\infty]$ such that (i) $\mu(\emptyset) = 0$, (ii) (*subadditivity*) $\mu(B) \le \sum_{j=1}^\infty A_j$ whenever $B \subset \cup_{j=1}^\infty A_j$.

We 'll describe a general method for building outer measures on a metric space (X, d). Fix $\mathcal{F} \subset 2^X$ and a function $\zeta\colon \mathcal{F} \to [0,+\infty]$. For $A \subset X$ and $\delta > 0$, define

$$\mathcal{C}_\delta(A) = \{\{S_j\}_{j\in J} : A \subset \cup_{j\in J} S_j, J \text{ countable}, \operatorname{diam} S_j < \delta, \forall j\}$$

The diameter is $\operatorname{diam} S_j = \sup_{x,y\in S_j} d(x,y)$. Then for $\delta > 0$, define

$$\mu_\delta(A) = \inf \left\{ \sum_{j\in J} \zeta(S_j) : (S_j)_{j\in J} \in \mathcal{C}_\delta \right\}.$$

If $0 < \delta_1 < \delta_2 \le \infty$, then $\mu_{\delta_2}(A) \le \mu_{\delta_1}(A)$. This means $\mu(A) = \lim_{\delta\downarrow 0} \mu_\delta(A) = \sup_{\delta>0} \mu_\delta(A)$ exists. [Exercise: both μ_δ and μ are outer measures.]

Example 1.2. Lebesgue Outer Measure. Let $\mathcal{F}$ be all boxes in $\mathbb{R}^n$. Let $B = \prod_{i=1}^n I_i$ where $I_i = [a_i, b_i]$ are intervals and $\zeta(B) = \prod_{i=1}^n (b_i - a_i)$.

s-dimensional Hausdorff measure. In $\mathbb{R}^n$, $s > 0$, take $\mathcal{F} = 2^{\mathbb{R}^n}$ and $\zeta(A) = \omega_s(\operatorname{diam} A/2)^s$, where ω_s is a normalising constant. We pick ω_s so that if, for example $s = n$, then $\mathcal{H}^n = \mathscr{L}^n$, i.e. we need $\mathcal{H}^n$-measure of a ball of radius 1 to be $\pi^{n/2}\Gamma(n/2+1)^{-1}$, where $\Gamma(q) = \int_0^\infty t^{q-1} \exp(-t)\, dt$. You get the same measure if you take $\mathcal{F}$ to be all closed sets or all open sets [exercise].

Definition 1.3. Let μ be an outer measure on a set X. We say $A \subset X$ is *μ-measurable* if

$$\mu(C) = \mu(C \cap A) + \mu(C \cap A^c) \qquad \forall C \subset X.$$

To remember whehther A or C is the test set, remember that we want to define measurable sets A to have nice boundaries.

Remark 1.4. Note that A measurable iff A^c measurable. From subadditivity, we always have $\le$. So we just need to check

$$\mu(C) \ge \mu(C \cap A) + \mu(C \cap A^c) \qquad \forall C \subset X.$$

Theorem 1.5. Let μ be an outer measure on X and $(A_j)_{j=1}^\infty$ be μ-measurable sets. Then

(i) $\bigcup_{j=1}^\infty A_j$ and $\bigcap_{j=1}^\infty A_j$ are measurable;

(ii) If $(A_j)_{j=1}^\infty$ are disjoint, then $\mu(\bigcup_{j=1}^\infty A_j) = \sum_{j=1}^\infty A_j$.

(iii) If $A_1 \subset A_2 \subset ...$ is an increasing sequence of sets, then $\lim_{j\to\infty} \mu(A_j) = \mu(\bigcup_{j=1}^\infty A_j)$.

(iv) If $A_1 \supset A_2 \supset \ldots$ is a decreasing sequence of sets with $\mu(A_1) < \infty$, then $\lim_{j\to\infty} \mu(A_j) = \mu(\bigcup_{j=1}^\infty A_j)$.

Proof. (i) We first show that measurability is closed under finite unions and intersections. For any $C \subset X$,

$$\begin{aligned}\mu(C) &= \mu(C \cap A_1) + \mu(C \cap A_1^c)\\ &= \mu(C \cap A_1) + \mu(C \cap A_1^c \cap A_2) + \mu(C \cap A_1^c \cap A_2^c)\\ &\geq \mu(C \cap (A_1 \cup A_2)) + \mu(C \cap (A_1 \cup A_2)^c),\end{aligned}$$

because $A_1 \cup A_2 \subset A_1 \cup (A_1^c \cap A_2)$ and $A_1^c \cap A_2^c = (A_1 \cup A_2)^c$. By induction, measurability is closed under finite unions and intersections.

(ii) If $(A_j)_{j=1}^\infty$ is disjoint, then set $B_N = \bigcup_{j=1}^N A_j$. Then

$$\mu(B_{N+1}) = \mu(B_{N+1} \cap A_{N+1}) + \mu(B_{N+1} \cap A_{N+1}^c) = \mu(A_{N+1}) + \mu(B_N).$$

By induction, $\mu(\bigcup_{j=1}^N A_j) = \sum_{j=1}^N \mu(A_j)$. By subadditivity, $\sum_{j=1}^N \mu(A_j) = \mu(\bigcup_{j=1}^N A_j) \leq \mu(\bigcup_{j=1}^\infty A_j)$. Take $N \to \infty$ to one inequality. Use subadditivity to obtain the other inequality.

(iii) If A_j are increasing, then write their union as disjoint annuli:

$$\begin{aligned}\mu\left(\bigcup_{j=1}^\infty A_j\right) &= \mu\left(A_1 \cup \bigcup_{j=1}^\infty (A_{j+1} \setminus A_j)\right)\\ &= \mu(A_1) + \sum_{j=1}^\infty \mu(A_{j+1} \setminus A_j)\\ &= \lim_{N\to\infty} \mu(A_{N+1}).\end{aligned}$$

(iv) If A_j are decreasing, then $\mu(A_1) \leq \mu(\bigcap_{j=1}^\infty A_j) + \mu(A_1 \setminus \bigcap_{j=1}^\infty A_j)$. So

$$\begin{aligned}\mu(A_1) - \mu\left(\bigcap_{j=1}^\infty A_j\right) &\leq \mu\left(A_1 \setminus \bigcap_{j=1}^\infty A_j\right)\\ &= \mu\left(\bigcup_{j=1}^\infty (A_1 \setminus A_j)\right)\\ &= \lim_{N\to\infty} \mu(A_1 \setminus A_N)\\ &= \mu(A_1) - \lim_{N\to\infty} \mu(A_N).\end{aligned}$$

Since $\mu(A_1) < \infty$, we can cancel it from both sides.

(v) We go back to showing (i) for countable unions and intersections. Take $C \subset X$. If $\mu(C) = \infty$ then we automatically have $\mu(C) \geq \mu(C \cap \bigcup_{j=1}^\infty A_j) + \mu(C \cap (\bigcup_{j=1}^\infty A_j)^c)$. So WLOG $\mu(C) < \infty$. Consider the outer measure $\mu\lfloor C$ given by $(\mu\lfloor C)(B) = \mu(C \cap B)$. Easy to see all A_j are $(\mu\lfloor C)$-measurable. Now $\mu(C \cap \bigcup_{j=1}^\infty A_j) = \lim_{N\to\infty} (\mu\lfloor C)(\bigcup_{j=1}^N A_j)$ and $\mu(C \cap (\bigcup_{j=1}^\infty A_j)^c) =$

$\lim_{N\to\infty}(\mu\lfloor C)((\bigcup_{j=1}^{N} A_j)^c)$. The proof is completed by adding these lines together and using the fact that the finite union is $(\mu\lfloor C)$-measurable. □

Definition 1.6. An outer measure μ on X is said to be *regular* if for every $C \subset X$ there exists a μ-measurable set $A \supset C$ with $\mu(A) = \mu(C)$.

Definition 1.7. An outer measure μ on a topological space X is said to be *Borel* if every Borel set is μ-measurable; and is said to be *Borel regular* if it is Borel and for every $C \subset X$ there exists a Borel set $A \supset C$ with $\mu(A) = \mu(C)$. Note being Borel regular is not the same as being Borel and regular.

Proposition 1.8. Let μ be a regular outer measure on X. If $A_1 \subset A_2 \subset ... \subset X$, then $\lim_{j\to\infty} \mu(A_j) = \mu(\cup_{j=1}^{\infty} A_j)$.

Remark 1.9. The A_j need not be μ-measurable.

Proof. For each j, pick measurable $A'_j \supset A_j$ with $\mu(A'_j) = \mu(A_j)$. Then set $B_N = \cup_{j=N}^{\infty} A'_j$. Notice $A_N \subset B_N$. And then notice $\mu(A_N) \leq \mu(B_N) \leq \mu(A'_N) = \mu(A_N)$. Therefore

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) \leq \mu\left(\bigcup_{j=1}^{\infty} B_j\right) = \lim_{j\to\infty} \mu(B_j) = \lim_{j\to\infty} \mu(A_j).$$

□

Definition 1.10. An outer measure μ on $\mathbb{R}^n$ is said to be a *Radon measure* if it is Borel regular and $\mu(K) < \infty$ for all compact $K \subset \mathbb{R}^n$.

Example 1.11. Let $B_1(0)$ be the unit ball in $\mathbb{R}^{n+k}$ where $k \geq 1$.Then [exercise] $\mathcal{H}^n(B_1(0)) = +\infty$. So $\mathcal{H}^n$ is not a Radon measure on $\mathbb{R}^{n+k}$. But a typical kind of example will be the following type of thing: let P be some n-dimensional subspace of $\mathbb{R}^{n+k}$. Then $\mathcal{H}^n\lfloor P$ is a Radon measure on $\mathbb{R}^{n+k}$.

Lemma 1.12. Let μ be a Borel regular measure on $\mathbb{R}^n$ and let $A \subset \mathbb{R}^n$ be a μ-measurable set with $\mu(A) < \infty$. Then $\nu := \mu\lfloor A$ is a Radon measure.

Proof. For any $C \subset \mathbb{R}^n$, $\nu(C) = \mu(C \cap A) \leq \mu(A) < \infty$. Now we need to check ν is Borel regular. Since μ is Borel regular, there exists Borel $B \supset A$ with $\mu(B) = \mu(A)$. We claim $\mu\lfloor A = \mu\lfloor B$. Given $C \subset \mathbb{R}^n$, $(\mu\lfloor B)(C) = \mu(B \cap C) = \mu(B\cap C\cap A)+\mu(B\cap C\cap A^c) \leq \mu(C\cap A)+\mu(B\cap A^c) = (\mu\lfloor A)(C)+\mu(B)-\mu(A) = (\mu\lfloor A)(C)$. And since $A \subset B$, we have $(\mu\lfloor A)(C) \leq (\mu\lfloor B)(C)$.

Now we will prove $\mu\lfloor B$ is Borel regular. Take $C \subset \mathbb{R}^n$. We know there exists Borel $E \supset B \cap C$ with $\mu(E) = \mu(B \cap C)$. So $E \cup B^c$ is Borel and contains C. Now $\nu(C) = \mu(C \cap B) \leq \mu(E \cap B) = \nu(E \cup B^c)$ and $\nu(E \cup B^c) = \mu(E \cap B) \leq \mu(E) = \mu(B \cap C) = \nu(C)$. □

Lemma 1.13. Let μ be a Borel regular outer measure on $\mathbb{R}^n$. For any Borel set B with $\mu(B) < \infty$ and $\epsilon > 0$, there exists closed set $C \subset B$ with $\mu(B \setminus C) < \epsilon$.

Proof. Let $\mathcal{F} = \mathcal{F}_B$ be the collection of all μ-measurable sets A with the property that for any $\epsilon > 0$, there exists closed $C \subset A$ such that $(\mu\lfloor B)(A \setminus C) < \epsilon$. Let's check that $\mathcal{F}$ is closed under countable unions and intersections. Suppose

$(A_j)_{j=1}^\infty \subset \mathcal{F}$. For each j, pick closed $C_j \subset A_j$ with $(\mu\lfloor B)(A_j \setminus C_j) < \epsilon 2^{-j}$. Certainly $C := \cap_{j=1}^\infty C_j$ is closed and $(\mu\lfloor B)(\cap_{j=1}^\infty A_j \setminus \cap_{j=1}^\infty C_j) \leq (\mu\lfloor B)(\cup_{j=1}^\infty (A_j \setminus C_j)) \leq \sum_{j=1}^\infty (\mu\lfloor B)(A_j \setminus C_j) \leq \sum_{j=1}^\infty \epsilon 2^{-j} = \epsilon$. And

$$(\mu\lfloor B)(\bigcup_{j=1}^\infty A_j \setminus \cup_{j=1}^\infty C_j) \leq (\mu\lfloor B)(\bigcup_{j=1}^\infty (A_j \setminus C_j)) \leq \sum_{j=1}^\infty (\mu\lfloor B)(A_j \setminus C_j) \leq \epsilon.$$

Since $\mu(B) < \infty$, we know from Theorem 1.5 that

$$(\mu\lfloor B)\left(\bigcup_{j=1}^\infty A_j \setminus \bigcup_{j=1}^\infty C_j\right) = \lim_{N\to\infty} (\mu\lfloor B)\left(\bigcup_{j=1}^\infty A_j \setminus \bigcup_{j=1}^N C_j\right).$$

So pick $m \geq 1$ such that $(\mu\lfloor B)\left(\bigcup_{j=1}^\infty A_j \setminus \bigcup_{j=1}^m C_j\right) \leq 2\epsilon$.

Then let $\mathcal{G} = \{A \in \mathcal{F} : A^c \in \mathcal{F}\}$. You can check from the previous work that $\mathcal{G}$ is a σ-algebra. It's clear that $\mathcal{F}$ contains closed sets. Every open set is the countable union of closed sets, so $\mathcal{F}$ also contains open sets. This means $\mathcal{G}$ contains open sets, and therefore Borel sets, including B itself. □

Lemma 1.14 (1.14)**.** Let μ be a Radon measure $\mathbb{R}^n$. For any Borel set B and $\epsilon > 0$, there exists open set $U \supset B$ with $\mu(U \setminus B) < \epsilon$.

Proof. Let $B_j(0)$ be a ball of radius j centred at the origin. Since μ is Radon, we know that $\mu(B_j(0) \setminus B)$ is finite. By applying Lemma 1.13, there exists closed $C_j \subset B_j(0) \setminus B$ with $\mu(B_j(0) \setminus B \setminus C_j) < \epsilon/2^j$. Let $U = \bigcup_{j=1}^\infty (B_j(0) \setminus C_j)$. Now $B = \bigcup_{j=1}^\infty (B \cap B_j(0)) \subset \cup_{j=1}^\infty (B_j(0) \setminus C_j)$ [Remember that $C_j \subset B_j(0) \cap B^c$, so $B \cup B_j(0)^c \subset C_j^c$, which implies $B \cap B_j(0) \subset B_j(0) \cap C_j^c = B_j(0) \setminus C_j$.] And now

$$\mu(U \setminus B) = \mu\left(\bigcup_{j=1}^\infty (B_j(0) \setminus C_j) \setminus \bigcup_{j=1}^\infty (B_j(0) \cap B)\right) \leq \sum_{j=1}^\infty \frac{\epsilon}{2^j} \leq \epsilon.$$

□

Theorem 1.15 (Inner and Outer Regularity)**.** Let μ be a Radon measure on $\mathbb{R}^n$.

(i) For any $A \subset \mathbb{R}^n$,

$$\mu(A) = \inf\{\mu(U) : U \text{ open}, U \supset A\};$$

(ii) For μ-measurable $A \subset \mathbb{R}^n$,

$$\mu(A) = \sup\{\mu(K) : K \text{ compact}, K \subset A\}.$$

Proof. For (i), the $\leq$ direction is immediate and if $\mu(A) = +\infty$, equality must hold. Since μ is Borel regular, there exists Borel $B \supset A$ with $\mu(B) = \mu(A)$. Then by Lemma 1.14, there exists open $U \supset B$ with $\mu(U \setminus B) < \epsilon$. So $\mu(U) < \mu(B) + \epsilon = \mu(A) + \epsilon$. Now take infimum over all such U.

For (ii), to begin with, we assume $\mu(A) < \infty$. In particular, $\mu := \mu\lfloor A$ is Radon (by Lemma 1.12). By (i) there exists an open set $U \supset A^c$ with $\nu(U \setminus A^c) < \epsilon$.

Of course, $U \setminus A^c = U \cap A = A \setminus U^c$, and U^c is a closed set with $U^c \subset A$. And we have that $\nu(A) < \nu(U^c) + \epsilon$. So $\mu(A) = \nu(A) < \nu(U^c) + \epsilon \leq \mu(U^c) + \epsilon$.

Now if $\mu(A) = +\infty$, set $D_j := B_j(0) \setminus B_{j-1}(0)$, so that A is the disjoint union $\bigcup_{j=1}^{\infty}(A \cap D_j)$. Since $\mu(A \cap D_j) < \infty$, apply previous work to get closed $C_j \subset A \cap D_j$ with $\mu((A \cap D_j) \setminus C_j) < \epsilon/2^j$. So now $\lim_{N\to\infty} \mu(\bigcup_{j=1}^{N} C_j) = \mu(\bigcup_{j=1}^{\infty} C_j) = \sum_{j=1}^{\infty} \mu(C_j) \geq \sum_{j=1}^{\infty}(\mu(A \cap D_j) - \epsilon 2^{-j}) \geq \mu(A) - \epsilon$. □

Definition 1.16. Let μ be an outer measure on $\mathbb{R}^n$. We say that $f\colon \mathbb{R}^n \to \mathbb{R}^k$ is *μ-measurable* if for every Borel set $B \subset \mathbb{R}^k$, the set $f^{-1}(B)$ is μ-measurable.

Theorem 1.17 (Lusin)**.** Let μ be a Radon measure on $\mathbb{R}^n$, and let $f\colon \mathbb{R}^n \to \mathbb{R}^k$ be a μ-measurable function. Then for any μ-measurable $A \subset \mathbb{R}^n$ with finite measure and any $\epsilon > 0$ there exists a compact $K \subset A$ with $\mu(A \setminus K) < \epsilon$, and $f|_K$ continuous. [Remark: as $\epsilon \downarrow 0$, the set K becomes wilder and wilder.]

Proof. For each $i = 1, 2, 3, ...$, write $\mathbb{R}^k$ as a disjoint union $\bigcup_{j=1}^{\infty} B_{ij}$ with B_{ij} Borel sets with $\operatorname{diam} B_{ij} < 1/i$. (e.g. chop the codomain up into half-open cubes of width $1/i$.) Then $f^{-1}(B_{ij})$ are disjoint and $A_{ij} = f^{-1}(B_{ij}) \cap A$ forms a partition of A into μ-measurable sets. Next, by inner regularity, let K_{ij} be compact sets with $K_{ij} \subset A_{ij}$ with $\mu(A_{ij}/K_{ij}) < \epsilon 2^{-i-j-1}$, so that $\mu(A \setminus \bigcup_{j=1}^{\infty} K_{ij}) < \epsilon 2^{-i-1}$. Pick $N = N(i)$ with $\mu(A \setminus \bigcup_{j=1}^{N(i)} K_{ij}) < \epsilon 2^{-i}$. The set $K_i = \bigcup_{j=1}^{N(i)}$ is compact. Pick $b_{ij} \in B_{ij}$, and define $f_i\colon K_i \to \mathbb{R}^k$ by $f_i(x) = b_{ij} \mathbb{1}_{x \in K_{ij}, j \leq N(i)}$. Set $K = \bigcap_{i=1}^{\infty} K_i$. Then by construction

$$\sup_{x \in K} |f_i(x) - f(x)| < \frac{1}{i},$$

and $\mu(A \setminus K) \leq \epsilon$, and f is the uniform limit of the continuous functions f_i on K. □

There is also *Egoroff's Theorem*:

Theorem 1.18 (Egoroff)**.** Let μ be an outer measure on $\mathbb{R}^n$ and let $A \subset \mathbb{R}^n$ be a μ-measurable subset with $\mu(A) < \infty$. Suppose $f_j\colon A \to \mathbb{R}^k$ is a sequence of μ-measurable functions converging pointwise μ-a.e. on A to the function $f\colon A \to R^k$. Given any $\epsilon > 0$, there exists (μ-measurable?) $B \subset A$ with $\mu(A \setminus B) < \epsilon$ such that $f_j|_B$ converges uniformly to $f|_B$. [Apart from some ϵ, pointwise is just uniform??]

Proof. Write

$$E_{j,m} := \{x \in A : \exists i > j, |f_i(x) - f(x)| > 1/m\}.$$

Notice the $E_{j,m}$ are decreasing sets in j, and on $\bigcap_{j=1}^{\infty} E_{j,m}$, f_j doesn't converge to f, so this set has zero measure by the assumption on pointwise convergence. Pick $J = J(m)$ so that $E_{J,m}$ has measure $\epsilon 2^{-m}$. If x is not in $\bigcup_{m=1}^{\infty} E_{J,m}$, then $\forall m \geq 1$, such that $\forall i > J(m)$, $|f_i(x) - f(x)| \leq 1/m$. □

2 Covering Theorems

This is a very GMT section. We will talk about Besicovitch Covering Theorem, which is just a statement involving balls of $\mathbb{R}^n$ with no mention on measures.

Theorem 2.1 (Besicovitch Covering Theorem)**.** Let $n \geq 1$. Then there exists $N = N(n) \geq 1$ such that the following is true. Let $A \subset \mathbb{R}^n$ and let $\mathcal{F}$ be a collection of balls in $\mathbb{R}^n$ with centres in A for which every point of A is the centre of some ball in $\mathcal{F}$, and $D := \sup_{B \in \mathcal{F}} \operatorname{diam} B < \infty$. Then there exists countable subcollections $\mathcal{G}_1, ..., \mathcal{G}_N \subset \mathcal{F}$ each of which consists of pairwise disjoint balls and such that $A \subset \bigcup_{i=1}^{N} \bigcup \mathcal{G}_i$.

Remark. This says that $\mathcal{G} = \bigcup_{i=1}^{N} \mathcal{G}_i$ is a countable collection of balls from $\mathcal{F}$ such that each point of A is contained in at most $N(n)$ balls. More analytically, this implies

$$\sum_{B \in \mathcal{G}} \mathbb{1}_B(x) \leq N \qquad \forall x \in \bigcup \mathcal{G}.$$

Assuming the theorem is true, we have

$$\mu(\bigcup_{j=1}^{\infty} B_j) \leq \sum_{j=1}^{\infty} \mu(B_j) = \sum_{j=1}^{\infty} \int_{\bigcup_{j=1}^{\infty} B_j} \mathbb{1}_{B_j} \,\mathrm{d}\mu \leq \int_{\bigcup_{j=1}^{\infty} B_j} \sum_{j=1}^{\infty} \mathbb{1}_{B_j} \,\mathrm{d}\mu \leq N\mu(\bigcup_{j=1}^{\infty} B_j).$$

Proof. Let's assume A is bounded. We will construct a list $\{B_{r_j}(a_j)\}_{j=1}^{J}$, where possibly $J = +\infty$ and such that:

(i) $j > i \implies r_i \geq \frac{3}{4} r_j$ [the radii do not grow too quickly];

(ii) $\{B_{r_j/3}(a_j)\}_{j=1}^{J}$ is disjoint;

(iii) $A \subset \bigcup_{j=1}^{J} B_{r_j}(a_j)$.

Take $B_1 := B_{r_1}(a_1)$ with $a_1 \in A$ and $r_1 \geq \frac{3}{4}\frac{D}{2}$. If we have already found inductively $B_{r_1}(a_1), ..., B_{r_{j-1}}(a_{j-1})$, set $A_j := A \setminus \bigcup_{i=1}^{j-1} B_{r_i}(a_i)$ and pick $B_j = B_{r_j}(a_j)$ with $a_j \in A_j$ and $r_j \geq \frac{3}{4} \sup\{r : B_r(a) \in \mathcal{F}, a \in A_j\}$. If $A_j = \emptyset$, then $J = j - 1$ and we stop.

Check (i): notice A_j is decreasing and so if $j > i$ then $a_j \in A_i$ and $r_i \geq \frac{3}{4} \sup\{r : B_r(a) \in \mathcal{F}, a \in A_i\} \geq \frac{3}{4} r_j$.

Check (ii): For $j > i$, $|a_i - a_j| \geq r_i = \frac{1}{3} r_i + \frac{2}{3} r_i \geq \frac{1}{3} r_i + \frac{2}{3}\frac{3}{4} r_j > \frac{1}{3} r_i + \frac{1}{3} r_j$.

Check (iii): If $J < \infty$, it's immediate. Otherwise, notice from (ii) we get that $r_i \to 0$ as $i \to \infty$. Now pick some $\bar{a} \in A$. Then there is $B_{\bar{r}}(\bar{a}) \in \mathcal{F}$. And eventually $r_i < \frac{3}{4}\bar{r}$ which implies that $\bar{r} > \sup\{r : B_s(r) \in \mathcal{F}, a \in A_i\}$ and so $\bar{a} \notin A_i$, i.e. $a \in \bigcup_{i'=1}^{i-1} B_{r_{i'}}(a_{i'})$.

Claim. Fix $k \in \{1, ..., J\}$. Write $I = I_k := \{j < k : B_j \cap B_k \neq \emptyset\}$. We claim there exists $N = N(n)$ with $|I| \leq N$.

Let's see why the claim suffices. For $i = 1, ..., N$, let $\mathcal{G}_i^N = \{B_{r_i}(a_i)\}$. Having constructed $\mathcal{G}_i^\ell$ for $i = 1, ..., N$ and $\ell \leq L$ such that each $\mathcal{G}_i^\ell$ is a disjoint collection of balls from $\{B_{r_j}(a_j)\}_{j=1}^{J}$, to construct $\mathcal{G}_i^{L+1}$, proceed as follows: since

$$|\{j < L + 1 : B_{r_j}(a_j) \cap B_{r_{L+1}}(a_{L+1}) \neq \emptyset\}| < N,$$

there exists $i' \in \{1, ..., N\}$ such that $B \cap B_{r_{L+1}}(a_{L+1}) = \emptyset$ for all $B \in \mathcal{G}_{i'}^L$. So set

$$\mathcal{G}_{i'}^{L+1} = \mathcal{G}_i^L \cup \{B_{r_{L+1}}(a_{L+1})\} \quad \text{and} \quad \mathcal{G}_i^{L+1} = \mathcal{G}_i^L \quad \text{for } i \neq i'.$$

Then the collections $\mathcal{G}_i = \bigcup_{\ell=1}^{\infty} \mathcal{G}_i^L$ works as in the statement.

Proof of claim. Write $K := \{j \in I : r_j \leq 3r_k\}$. If $i \in K$ then $|a_i - a_k| < r_i + r_k$. So take $x \in B_{r_i/3}(a_i)$. We have: $|x - a_k| \leq |x - a_i| + |a_i - a_k| \leq \frac{1}{3}r_i + r_i + r_k \leq 5r_k$. So $B_{r_i/3}(a_i) \subset B_{5r_k}(a_k)$. Since $k > i$, $r_k \leq 4/3r_i$, so $3/4r_k \leq r_i$.

$$\sum_{i \in K} \left(\frac{r_k}{4}\right)^n \leq \sum_{i \in K} \left(\frac{r_i}{3}\right)^n \leq (5r_k)^n.$$

Cancel the the r_k^n to get $|K| \leq 20^n$.

So we are left to bound $I \setminus K$. Pick $i, j \in I \setminus K$. Let $\theta \in [0, \pi]$ be the angle between $a_i - a_k$ and $a_j - a_k$.

Final claim. We claim there exists $\theta_0 = \theta_0(n) > 0$ such that $\theta \geq \theta_0$.

This suffices because there exists $r_0 \in (0, 1)$ such that if $x \in \partial B_1(0)$ and $y, z \in B_{r_0}(x)$, then angle between y and z is $< \theta_0$. And there is some constant $L = L(n)$ such that $\partial B_1(0)$ can be covered by L balls of radius r_0 centred on $\partial B_1(0)$ but not $L - 1$ such balls. So by rescaling/translating, $\partial B_{r_k}(a_k)$ can be covered by L balls of radius $r_0 r_k$ centred on ∂B_k but not $L - 1$ such balls. So since the rays through a_i and a_j from a_k have angle bounded below by θ_0, we will conclude $|I \setminus K| \leq L$.

Proof of final claim. Translate a_k to origin. We know that $3r_k < r_i < |a_i - 0| < r_i + r_k$, $3r_k < r_j < |a_j - 0| < r_j + r_k$ and we can WLOG $|a_i| \leq |a_j|$. If $|a_i - a_j| \geq |a_j|$, then:

$$\cos\theta = \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|} \leq \frac{1}{2}\frac{|a_i|}{|a_j|} \leq \frac{1}{2}.$$

Now suppose $|a_i - a_j| \leq |a_j|$. We are also free to assume $\cos\theta > 5/6$. Then

$$\begin{aligned}
\frac{5}{6} &< \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|} \\
&\leq \frac{1}{2} + \frac{(|a_j| - |a_i - a_j|)(|a_j| + |a_i - a_j|)}{2|a_i||a_j|} \\
&< \frac{1}{2} + \frac{|a_j| - |a_i - a_j|}{|a_i|} \\
&< \frac{1}{2} + \frac{r_j + r_k - |a_i - a_j|}{3r_k} \\
&= \frac{5}{6} + \frac{r_j - |a_i - a_j|}{3r_k},
\end{aligned}$$

so $|a_i - a_j| < r_j$. This tells us $a_i \in B_{r_j}(a_j)$. So we know that $j > i$, which means $r_i \leq |a_i - a_j|$ and we can deduce $r_j \leq \frac{4}{3}r_i$.

Now

$$\begin{aligned}\cos\theta &= \frac{|a_i|^2 + |a_j|^2 - |a_i - a_j|^2}{2|a_i||a_j|} \\ &= \frac{(|a_i| - |a_j|)^2 - |a_i - a_j|^2}{2|a_i||a_j|} + 1 \\ &= \frac{(|a_i| - |a_j| - |a_i - a_j|)(|a_i| - |a_j| + |a_i - a_j|)}{2|a_i||a_j|} + 1 \\ &= 1 - \frac{(-|a_i| + |a_j| + |a_i - a_j|)(|a_i| - |a_j| + |a_i - a_j|)}{2|a_i||a_j|}\end{aligned}$$

We can bound the second term: using $|a_i| > r_i$, $|a_j| < r_j + r_k$, $|a_i - a_j| > r_i$,

$$\begin{aligned}\frac{(-|a_i| + |a_j| + |a_i - a_j|)(|a_i| - |a_j| + |a_i - a_j|)}{2|a_i||a_j|} &\geq \frac{r_i[r_i - r_j - r_k + r_i]}{2(r_i + r_k)|a_j|} \\ &\geq \frac{r_i[\frac{1}{2}r_j - r_k]}{2(r_i + r_k)|a_j|} \\ &\geq \frac{r_i \frac{1}{6} r_j}{2(r_i + r_k)|a_j|} \\ &= \frac{1}{12} \frac{r_i}{r_i + r_k} \frac{r_j}{|a_j|}.\end{aligned}$$

Notice $r_i + r_k < \frac{4}{3} r_i$, and $|a_j| < r_j + r_k < \frac{4}{3} r_j$, so above $\geq \frac{1}{12}\frac{3}{4}\frac{3}{4} = \frac{3}{64}$. So $\theta > \arccos(\frac{61}{64}) =: \theta_0 > 0$. □

Corollary 2.2. Let $A \subset \mathbb{R}^n$ and let μ be a Radon measure on $\mathbb{R}^n$ with $\mu(A) < \infty$. Given an open set $U \subset \mathbb{R}^n$ and a collection $\mathcal{F}$ of closed balls with $\inf_{B_r(a) \in \mathcal{F}} r = 0$, $\forall a \in A$, there exists countable disjoint subcollection $\mathcal{G} \subset \mathcal{F}$ with

- $\bigcup_{B \in \mathcal{G}} B \subset U$;
- $\mu((A \cap U) \setminus \bigcup_{B \in \mathcal{G}} B) = 0$.

Proof. Take $\mathcal{F}' := \{B_r(a) \in \mathcal{F} : a \in A \cap U, r \leq 1, B_r(a) \subset U\}$. Apply Besicovitch to get countable disjoint subcollections $\mathcal{G}_1, ..., \mathcal{G}_N$ with $A \cap U \subset \bigcup_{i=1}^N \bigcup_{B \in \mathcal{G}_i} B$. This implies that $\mu(A \cap U) \leq \sum_{i=1}^N \mu(A \cap U \cap \bigcup_{B \in \mathcal{G}_i} B)$. There exists $i_0 \in \{1, ..., N\}$ with $\mu(A \cap U \cap \bigcup_{B \in \mathcal{G}_{i_0}} B) \geq \frac{1}{N}\mu(A \cap U)$. Choose $\theta \in (0, \frac{1}{N})$. So there are disjoint balls $B_1, .., B_{M_1} \in \mathcal{G}_{i_0}$ with $\mu(A \cap U \cap \bigcup_{j=1}^{M_1} B_j) > \theta\mu(A \cap U)$. This implies $\mu(A \cap U \setminus \bigcup_{j=1}^{M_1} B_j) < (1 - \theta)\mu(A \cap U)$. We can inductively repeat this process to get for each $\ell \geq 1$, a disjoint union of balls $\bigcap_{j=1}^{M_\ell} B_j$ with

$$\mu\left(A \cap U \setminus \bigcup_{j=1}^{M_\ell} B_j\right) < (1 - \theta)^\ell \mu(A \cap U).$$

So in the end $\bigcup_{j=1}^{\infty} B_j$ works. □

3 Densities and Differentiation of Radon Measures

Definition 3.1. Let μ, ν be Radon measures on $\mathbb{R}^n$. The *upper density* of ν with respect to ν is

$$\bar{D}_\mu \nu(x) := \begin{cases} \limsup_{r \downarrow 0} \frac{\nu(\overline{B_r(x)})}{\mu(\overline{B_r(x)})} & \text{if } \forall r > 0, \mu(\overline{B_r(x)}) > 0 \\ +\infty & \text{if } \exists r > 0, \mu(\overline{B_r(x)}) = 0, \end{cases}$$

The *lower density* of ν with respect to μ is

$$\underline{D}_\mu \nu(x) := \begin{cases} \liminf_{r \downarrow 0} \frac{\nu(\overline{B_r(x)})}{\mu(\overline{B_r(x)})} & \text{if } \forall r > 0, \mu(\overline{B_r(x)}) > 0 \\ +\infty & \text{if } \exists r > 0, \mu(\overline{B_r(x)}) = 0, \end{cases}$$

We say ν is *differentiable* with respect to μ at x if $\bar{D}_\mu \nu(x) = \underline{D}_\mu \nu(x)$.

Lemma 3.2. Let μ, ν be Radon measures on $\mathbb{R}^n$ and let $\alpha \in (0, \infty)$. If

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

then $\nu(A) \leq \alpha \mu(A)$. And if

$$A \subset \{x \in \mathbb{R}^n : \bar{D}_\mu \nu(x) \geq \alpha\},$$

then $\nu(A) \geq \alpha \mu(A)$.

Proof. We will prove the first statement in detail. By restricting to compact sets, we will assume μ and ν are finite. Fix $\epsilon > 0$ and an open set $U \supset A$. For each $a \in A$, there exists arbitrarily small radii $r > 0$ for which $\nu(\overline{B_r(a)}) < (\alpha + \epsilon)\mu(\overline{B_r(a)})$. So, we can consider

$$\mathcal{F} := \{\overline{B_r(a)} \subset U : a \in A, \nu(\overline{B_r(a)}) < (\alpha + \epsilon)\mu(\overline{B_r(a)})\}.$$

By Corollary 2.2, there exists countable, disjoint subcollection $\mathcal{G} \subset \mathcal{F}$ with $\bigcup_{B \in \mathcal{G}} B \subset U$ and $\nu(A \setminus \bigcup_{B \in \mathcal{G}} B) = 0$. Now $\nu(A) = \nu(\bigcup_{B \in \mathcal{G}} B) = \sum_{B \in \mathcal{G}} \nu(B) \leq (a + \epsilon) \sum_{B \in \mathcal{G}} \mu(B) = (\alpha + \epsilon)\mu(\bigcup_{B \in \mathcal{G}} B) \leq (\alpha + \epsilon)\mu(U)$. By outer regularity and arbitrariness of ϵ, we are done. □

Theorem 3.3. Let μ, ν be Radon measures on $\mathbb{R}^n$. Then $D_\mu \nu$:

(i) exists μ-a.e.;

(ii) is finite μ-a.e.;

(iii) is μ-measurable.

Proof. Once again, assume μ and ν are finite. Write $I := \{x \in \mathbb{R}^n : \bar{D}_\mu \nu = +\infty\}$. Then for any $\alpha > 0$, we have $I \subset \{\bar{D}_\mu \nu \geq \alpha\}$, and so by Lemma 3.2, $\mu(I) \leq \frac{1}{\alpha}\nu(I)$. Since $\nu(I) < \infty$, we deduce $\mu(I) = 0$.

For $a, b \in \mathbb{Q}$ with $a < b$, let $R_{a,b} := \{\underline{D}_\mu \nu < a < b < \bar{D}_\mu \nu\}$. Now $\{D_\mu \nu \text{ does not exist}\} \subset \bigcup_{a < b, a, b \in \mathbb{Q}} R_{a,b}$, and from Lemma 3.2, we have

$$b\mu(R_{a,b}) \leq \nu(R_{a,b}) \leq a\mu(R_{a,b}).$$

But $a < b$, so we must have $\mu(R_{a,b}) = 0$. Observe now that $x \mapsto \mu(\overline{B_r(x)})$ is Borel measurable for any fixed $r > 0$ and any Radon measure μ (this follows from the fact that it is upper semicontinuous). Then for fixed $k \in \mathbb{N}$,

$$x \mapsto \begin{cases} \frac{\nu(\overline{B_{1/k}(x)})}{\mu(\overline{B_{1/k}(x)})} & \text{if } \mu(\overline{B_{1/k}(x)}) > 0, \\ 0 & \text{otherwise} \end{cases}$$

is Borel measurable. Now it follows $D_\mu \nu$ is Borel measurable, just by taking lim inf, lim sup. □

Theorem 3.4 (Radon-Nikodym Derivatives)**.** Let μ, ν be Radon measures on $\mathbb{R}^n$ and suppose $\nu \ll \mu$. Then, for every μ-measurable set $A \subset \mathbb{R}^n$, we have:

$$\nu(A) = \int_A D_\mu \nu(x) \, \mathrm{d}\mu(x).$$

Recall that $\nu \ll \mu$ means ν is absolutely continuous with respect to μ, whose definition is $\mu(A) = 0 \implies \nu(A) = 0$ for all μ-measurable $A \subset \mathbb{R}^n$.

Proof. We claim the sets $I = \{D_\mu \nu = +\infty\}$, $Z = \{D_\mu \nu = 0\}$ and $U = \{\underline{D}_\mu \nu < \bar{D}_\mu \nu\}$ are all ν-null. $\mu(I) = \mu(U) = 0$ from Theorem 3.3, so by hypothesis, $\nu(I) = \nu(U) = 0$. And $\forall \epsilon > 0$, $Z \subset \{\underline{D}_\mu \nu \leq \epsilon\}$. So by the lemma, $\nu(Z) \leq \epsilon \mu(Z)$. Assuming for now that μ, ν are finite, we deduce $\nu(Z) = 0$. Fix μ-measurable $A \subset \mathbb{R}^n$. For $m \in \mathbb{Z}$ and $t > 1$, write $A_m := \{x \in A : t^m < D_\mu \nu(x) \leq t^{m+1}\}$. These are all μ-measurable. This implies ν-measurable: take Borel $B \supset A_m$ with $\mu(B) = \mu(A_m)$. Then $\mu(B \setminus A_m) = 0$. So $\nu(B \setminus A_m) = 0$. So $B \setminus A_m$ is ν-measurable. And $A_m = B \setminus (B \setminus A_m)$ so is ν-measurable. Now:

$$\begin{aligned} \int_A D_\mu \nu \, \mathrm{d}\mu &= \sum_{m \in \mathbb{Z}} \int_{A_m} D_\mu \nu \, \mathrm{d}\mu + \int_{Z \cap A} D_\mu \nu \, \mathrm{d}\mu \leq t^{m+1} \sum_{m \in \mathbb{Z}} \mu(A_m) \\ &< t \sum_{m \in \mathbb{Z}} \nu(A_m) = t\nu(\bigcup_{m \in \mathbb{Z}} A_m) = t\nu(A), \end{aligned}$$

because $\nu(I \cup Z \cup U) = 0$. Also

$$t\nu(A) = t \sum_{m \in \mathbb{Z}} \nu(A_m) \leq t^{m+2} \sum_{m \in \mathbb{Z}} \mu(A_m) \leq t^2 \sum_{m \in \mathbb{Z}} \int_{A_m} D_\mu \nu \, \mathrm{d}\mu = t^2 \int_A D_\mu \nu \, \mathrm{d}\mu.$$

Now let $t \downarrow 1$ to complete. □

Definition 3.5. We say two Radon measures μ, ν on $\mathbb{R}^n$ are *mutually singular*, and write $\mu \perp \nu$, if there exists a Borel set $B \subset \mathbb{R}^n$ such that $\mu(B) = \nu(B^c) = 0$.

Theorem 3.6 (Lebesgue Decomposition)**.** Let μ, ν be Radon measures on $\mathbb{R}^n$. There exists Radon measures ν_{ac} and ν_s with: (i) $\nu = \nu_{ac} + \nu_s$; (ii) $\nu_{ac} \ll \mu$; (iii) $\mu_s \perp \mu$; (iv) $D_\mu \nu = D_\mu \nu_{ac}$ μ-a.e.; (v) $D_\mu \nu_s = 0$ μ-a.e. [So, by Theorem 3.4: $\nu(A) = \int_A D_\mu \nu \, \mathrm{d}\mu + \nu_s(A)$ for all μ-measurable $A \subset \mathbb{R}^n$.]

Proof. Again assume μ, ν are finite. Let $\mathcal{E} = \{\text{Borel } B : \mu(B^c) = 0\}$. Let $\{B_k\}_{k=1}^\infty \in \mathcal{E}$ be such that $\nu(B_k) \leq \inf_{A \in \mathcal{E}} \nu(A) + \frac{1}{k}$. Then $B = \bigcap_{k=1}^\infty B_k$ is Borel, and $\mu(B^c) \leq \sum_{k=1}^\infty \mu(B_k^c) = 0$. And $\nu(B) = \inf_{A \in \mathcal{E}} \nu(A)$. So write $\nu \lfloor B^c =: \nu_s$ and $\nu_{ac} := \nu \lfloor B$. Clearly $\nu_s(B) = 0$.

We check (i)-(iii). Take μ-measurable C with $\mu(C) = 0$. Suppose (for contradiction) that $\nu_{ac}(C) = \nu(C \cap B) > 0$. Take Borel $S \supset B \setminus (C \cap B)$ with $\nu_{ac}(S) = \nu_{ac}(B \setminus (C \cap B)) < \nu(B)$. And $\mu((S \cap B)^c) \leq \mu(S^c) + \mu(B^c) \leq \mu(C) + 2\mu(B^c) = 0$. So we violate the fact that $\nu(B) = \inf_{A \in \mathcal{E}} \nu(A)$.

It is enough to show (v) as (iv) is equivalent. With $T := \{x \in B : D_\mu \nu_s \geq \alpha > 0\}$, we have $\nu_s(T) \leq \nu_s(B) = 0$ and $\mu(T) \leq \frac{1}{\alpha}\nu_s(T) = 0$. And since $\mu(B^c) = 0$, we deduce $D_\mu \nu_s = 0$ μ-a.e. Then by additivity of density with respect to μ (up to sets of μ-measure zero), we deduce (iv). □

Theorem 3.7 (Lebesgue-Besicovitch Differentiation)**.** Let μ be a Radon measure on $\mathbb{R}^n$. If $f \in L^1_{loc}(\mathbb{R}^n; \mu)$, then for μ-a.e. $x \in \mathbb{R}^n$ we have

$$\lim_{r \downarrow 0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f(y) \, \mathrm{d}\mu(y) = f(x).$$

And for $p \in [1, \infty)$, if $f \in L^p_{loc}(\mathbb{R}^n; \mu)$, then for μ-a.e. $x \in \mathbb{R}^n$,

$$\lim_{r \downarrow 0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |f(x) - f(y)|^p \, \mathrm{d}\mu(y) = 0.$$

Proof. Let $f^\pm$ be positive and negative parts of f. For Borel $B \subset \mathbb{R}^n$, define $\nu^\pm(B) := \int_B f^\pm \, \mathrm{d}\mu$. And for general $A \subset \mathbb{R}^n$, define

$$\nu^\pm(A) := \inf\{\nu^\pm(B) : B \text{ Borel, } B \supset A\}.$$

One checks now that $\nu^\pm$ are Radon measures, absolutely continuous with respect to μ. So by Radon-Nikodym, for every Borel $B \subset \mathbb{R}^n$, we have $\nu^\pm(B) = \int_B f^\pm \, \mathrm{d}\mu = \int_B D_\mu \nu^\pm \, \mathrm{d}\mu$. We deduce that $D_\mu \nu^\pm = f^\pm$ μ-a.e.. Now, for μ-a.e. x,

$$\begin{aligned} \int_{B_r(x)} f(y) \, \mathrm{d}\mu(y) &= \int_{B_r(x)} f^+(y) \, \mathrm{d}\mu(y) - \int_{B_r(x)} f^-(y) \, \mathrm{d}\mu(y) \\ &= \nu^+(B_r(x)) - \nu^-(B_r(x)). \end{aligned}$$

Divide both sides by $\mu(B_r(x))$ and send $r \downarrow 0$. Then

$$\lim_{r \downarrow 0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} f(y) \, \mathrm{d}\mu(y) = D_\mu \nu^+(x) - D_\mu \nu^-(x) = f^+(x) - f^-(x) = f(x).$$

Note: the lecturer mixed up closed balls and open balls, so replace all open balls with closed balls.

For the second part, let $\{r_j\}_{j=1}^\infty$ be dense subsets of $\mathbb{R}$ and apply the first statement to $x \mapsto |f(x) - r_j|^p \in L^1_{loc}(\mathbb{R}^n, \mu)$ for each j. So now there exists a μ-null $A \subset \mathbb{R}^n$ such that if $x \notin A$ then for all j,

$$\lim_{r \downarrow 0} \frac{1}{\mu(B_r(x))} \int_{B_r(x)} |f(y) - r_j|^p \, \mathrm{d}\mu(y) = |f(x) - r_j|^p.$$

So pick r_j with $|r_j - f(x)| < \epsilon$. Using $|f(y) - f(x)|^p \leq 2^p(|f(y) - r_j|^p + |f(x) - r_j|^p)$, and send $\epsilon \downarrow 0$. □

Remark. With Lebesgue on $\mathbb{R}^n$, apply to $\mathbb{1}_E$ for Lebesgue measurable E. Then for almost every $x \in E$,

$$\lim_{r \downarrow 0} \frac{\mathrm{Leb}(B_r(x) \cap E^c)}{\mathrm{Leb}(B_r(x))} = 0.$$

4 Area and Co-area

Developing a rigorous notion of area in $\mathbb{R}^n$ was one of the main driving forces for developing Geometric Measure Theory. For a few minutes I will treat you like Calculus 101 students.

Let D be the open disk in $\mathbb{R}^2$. Then to evaluate $Area(D) = \int_{B_1(0)} \mathrm{d}x\,\mathrm{d}y$ we use change of variables. Let's use polar coordinates. There's this thing called r and θ, and we can integrate over $r \in (0,1]$ and $\theta \in [0, 2\pi)$. But it's dangerous to show them a picture of a rectangle in the (r,θ) plane. It's confusing for some people. It's not the same as the area of the rectangle $\int_0^{2\pi}\int_0^1 \mathrm{d}r\,\mathrm{d}\theta$. At the heart of this example of course is some function $F(r,\theta) = (r\cos\theta, r\sin\theta)$. What's the area of F(rectangle)? We work out the Jacobian. This gives us $\int_0^{2\pi}\int_0^1 r\,\mathrm{d}r\,\mathrm{d}\theta$. We generalise this to F almost everywhere differentiable. We will focus on F being Lipschitz.

Proposition 4.1. A continuous increasing function $f\colon [a,b] \to \mathbb{R}$ is differentiable almost everywhere.

Proof. With $\mathcal{F} = \{(c,d) : (c,d) \subset [a,b]\}$ and $\zeta((c,d)) = f(d) - f(c)$, let $\mathscr{L}_f$ be the outer measure produced by Carathéodory's general construction. This is a Radon measure. So by decomposing $\mathscr{L}_f$ with respect to $\mathscr{L}^1$, we get an $\mathscr{L}^1$-measurable function $g\colon [a,b] \to \mathbb{R}$, a Radon measure ν and a set $S \subset [a,b]$ such that $\nu([a,b]\setminus S) = 0$, $\mathscr{L}^1(S) = 0$ and such that for every $\mathscr{L}^1$-measurable $A \subset [a,b]$, we have $\mathscr{L}_f(A) = \int_A g\,\mathrm{d}\mathscr{L}^1 + \nu(A)$. Since f is continuous, we can check that $\mathscr{L}_f((c,d)) = f(d) - f(c)$. Now, look at

$$\begin{aligned}
\left|\frac{f(x+h)-f(x)}{h} - g(x)\right| &= \left|\frac{\mathscr{L}_f((x,x+h))}{h} - g(x)\right| \\
&= \left|\frac{1}{h}\int_x^{x+h} g(y)\,\mathrm{d}\mathscr{L}^1(y) - g(x) + \frac{1}{h}\nu((x,x+h))\right| \\
&\le \frac{1}{h}\int_x^{x+h} |g(y)-g(x)|\,\mathrm{d}\mathscr{L}^1(y) + \frac{\nu((x,x+h))}{h} \\
&\le 2\frac{1}{2h}\int_{x-h}^{x+h} |g(y)-g(x)|\,\mathrm{d}\mathscr{L}^1(y) + 2\frac{\nu([x-h,x+h])}{2h}.
\end{aligned}$$

As $h \downarrow 0$, this goes to zero for almost every x (first term by Lebesgue differentiation, second term because $D_{\mathscr{L}^1}\nu = 0$ $\mathscr{L}^1$-a.e. □

Remark. ν is the failure of the Fundamental Theorem of Calculus – it is called the *Cantor part of the derivative*. See the *Devil Staircase*.

Proposition 4.2. Let $f\colon [a,b] \to \mathbb{R}$ be Lipschitz. Then f is differentiable $\mathscr{L}^1$-a.e., $f' \in L^\infty$, and $f(x) = f(a) + \int_a^x f'(t)\,\mathrm{d}t$ for all $x \in [a,b]$.

Proof. Write

$$V_f(x) := \sup\left\{\sum_{j=1}^n |f(x_j) - f(x_{j-1})| : a \le x_0 < x_1 < \ldots < x_n \le x\right\}.$$

Notice $\sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| \leq \mathrm{Lip}(f)(b-a)$. This function is continuous and increasing [exercise]. And so is $V_f(x) - f(x)$, so that we can use this function to express f as the difference of two continuous, increasing functions. So by Proposition 4.1, f is differentiable $\mathscr{L}^1$-a.e. Since $|f(x+h) - f(x)| \leq \mathrm{Lip}(f)|h|$, we get $\|f'\|_{L^\infty} \leq \mathrm{Lip}(f)$.

Write $F(x) = \int_a^x f'(t)\,\mathrm{d}\mathscr{L}^1(t)$. So now:

$$\frac{1}{2h}\int_{x-h}^{x+h} f'(t)\,\mathrm{d}t = \frac{1}{2}\left(\frac{F(x+h)-F(x)}{h} + \frac{F(x)-F(x-h)}{h}\right).$$

Since F is Lipschitz, it is differentiable almost everywhere, and using Lebesgue differentiation on the left, we deduce, $f'(x) = F'(x)$ $\mathscr{L}^1$-a.e.

Now $g = F - f$ is Lipschitz with $g' = 0$ $\mathscr{L}^1$-a.e. [We want to conclude that g is constant, and we need to use the Lipschitz condition. This step is surprisingly hard. As far as the lecturer knows, we need to use a covering lemma.]

Let $E \subset (a,b)$ be the set with $\mathscr{L}^1((a,b) \setminus E) = 0$ and such that g is differentiable at each point of E with derivative zero. Fix $x \in (a,b)$ and $\epsilon > 0$. Using Corollary to Besicovitch, there exists a countable disjoint collection $[x_j - h_j, x_j + h_j] \subset (a,x)$ for $j = 1, 2, \ldots$ with $|g(x_j + h_j) - g(x_j - h_j)| \leq 2h_j\epsilon$ and $\mathscr{L}^1\left(E \cap (a,x) \setminus \bigcup_{j=1}^{\infty}[x_j - h_j, x_j + h_j]\right) = 0$. Fix $N \geq 1$ such that $\mathscr{L}^1\left(E \cap (a,x) \setminus \bigcup_{j=1}^{N}[x_j - h_j, x_j + h_j]\right) < \epsilon$, and write the intervals in order. So now, since

$$|x_1 - h_1 - a| + \sum_{j=1}^{N-1} |(x_{j+1} - h_{j+1}) - (x_j + h_j)| + |x - (x_N + h_N)| \leq \epsilon,$$

we know

$$|f(x_1 - h_1) - f(a)| + \sum_{j=1}^{N-1} |g(x_{j+1} - h_{j+1}) - g(x_j + h_j)| + |g(x) - g(x_N + h_N)| \leq 2\,\mathrm{Lip}(f)\epsilon.$$

So now,

$$|g(a) - g(x)| \leq 2\,\mathrm{Lip}(f)\epsilon + \sum_{j=1}^{N} |g(x_j + h_j) - g(x_j - h_j)| \leq 2\,\mathrm{Lip}(f)\epsilon + 2(b-a)\epsilon.$$

So since $\epsilon > 0$ was arbitrary, we can deduce $g(x) = g(a)$ for all $x \in (a,b]$. □

Theorem 4.3 (Rademacher's Theorem)**.** Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be Lipschitz. Then f is differentiable $\mathscr{L}^n$-a.e.

Proof. For any $x \in \mathbb{R}^n$ and 'direction' $\omega \in S^{n-1} = \partial B_1(0)$, the function $t \mapsto f(x + t\omega)$ is a Lipschitz function on the line $\ell_{x,w} := \{x + t\omega : t \in \mathbb{R}\}$ and hence for $\mathscr{L}^1$-a.e. $t \in \ell_{x,\omega}$, the derivative $\frac{d}{dt}f(x + t\omega)$ exists. So let A_ω denote the set of points $x \in \mathbb{R}^n$ at which the derivative $\frac{d}{dt}\big|_{t=0} f(x + t\omega)$ exists. Then $A_\omega^c \cap \ell_{y,\omega}$ has $\mathscr{L}^1$-measure zero for any $y \in \mathbb{R}^n$. So, by Fubini's Theorem, $\mathscr{L}^n(\mathbb{R}^n \setminus A_\omega) = 0$ for every $\omega \in \mathbb{R}^n$.

Next, we check that $D_\omega f(x) := \frac{d}{dt}\big|_{t=0} f(x+t\omega)$ is equal to $\sum_{j=1}^n \omega^j D_j f(x)$ a.e. Fix $\xi \in C_c^\infty(\mathbb{R}^n)$ and consider

$$\int_{\mathbb{R}^n} \frac{f(x+t\omega)-f(x)}{t}\xi(x)\,\mathrm{d}x = -\int_{\mathbb{R}^n} \frac{\xi(x)-\xi(x-t\omega)}{t} f(x)\,\mathrm{d}x.$$

By Dominated Convergence Theorem,

$$\begin{aligned}\int_{\mathbb{R}^n} D_\omega f(x)\xi(x)\,\mathrm{d}x &= -\int_{\mathbb{R}^n} \sum_{j=1}^n \omega^j D_j \xi(x) f(x)\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} \sum_{j=1}^n \omega^j D_j f(x)\xi(x)\,\mathrm{d}x.\end{aligned}$$

Since this holds for arbitrary $\xi \in C_c^\infty(\mathbb{R}^n)$, we have $D_\omega f(x) = \sum_{j=1}^n \omega^j D_j f(x)$ $\mathscr{L}^n$-a.e.

Finally, let $\omega_1, \omega_2, \ldots$ be dense in S^{n-1} and write

$$Q(\omega,h)(x) := \frac{f(x+h\omega)-f(x)}{h} - D_\omega f(x).$$

For each $k \in \mathbb{N}$, let A_k be the set of x for which $D_{\omega_k} f(x)$ exists, $D_1 f(x), \ldots, D_n f(x)$ exists and $D_{\omega_k} f(x) = \sum_{j=1}^n \omega_k^j D_j f(x)$. Then let $A = \bigcap_{k=1}^\infty A_k$. We know that $\mathscr{L}^n(\mathbb{R}^n \setminus A) = 0$ and for each $x \in A$ and $k \in \mathbb{N}$, we have $Q(\omega_k, h)(x) \to 0$ as $h \to 0$.

Fix $x_0 \in A$ and $\epsilon > 0$. There exists K such that $S^{n-1} \subset \bigcup_{j=1}^K B_\epsilon(\omega_j)$, then let $\bar{h}$ be such that $|h| < \bar{h} \implies \max_{j=1,\ldots,K} |Q(\omega_j,h)(x_0)| < \epsilon$. And now $|Q(\omega,h)(x_0)| \le |Q(\omega,h)(x_0) - Q(\omega_i,h)(x_0)| + |Q(\omega_i,h)(x_0)|$, (where $|\omega - \omega_i| < \epsilon$), so that for $|h| < \bar{h}$,

$$\begin{aligned}|Q(\omega,h)(x_0)| &\le \left|\frac{f(x+h\omega)-f(x+h\omega_i)}{h}\right| + \left|\sum_{j=1}^n \omega^j D_j \xi(x) f(x)\,\mathrm{d}x\right| + \epsilon \\ &\le (2\,\mathrm{Lip}(f)+1)\epsilon.\end{aligned}$$

Think carefully as to why bound of this form suffices. □

Area Formula

Given $f\colon \mathbb{R}^n \to \mathbb{R}^{n+k}$ Lipschitz, write Df_x for the linear map from $\mathbb{R}^n \to \mathbb{R}^{n+k}$ with matrix $(D_j f^i(x))_{i=1,\ldots,n+k, j=1,\ldots,n}$. And write

$$\mathcal{J} f(x) := \sqrt{\det((Df_x)^* \circ Df_x)},$$

called the *Jacobian matrix.*

Theorem 4.4 (Area Formula)**.** Let $A \subset \mathbb{R}^n$ be $\mathscr{L}^n$-measurable and let $f\colon A \to \mathbb{R}^{n+k}$ be Lipschitz. Then

- $\mathcal{H}^n(f(A)) = \int_A \mathcal{J} f(x)\,\mathrm{d}\mathscr{L}^n(x)$ if f is injective.

- In general

$$\int_{\mathbb{R}^{n+k}} \mathcal{H}^0(A \cap f^{-1}(y)) \, \mathrm{d}\mathcal{H}^n(y) = \int_A \mathcal{J}f(x) \, \mathrm{d}\mathscr{L}^n(x).$$

- And if $u \colon A \to \mathbb{R}$ is $\mathscr{L}^n$-integrable, we have

$$\int_{\mathbb{R}^{n+k}} \sum_{x \in f^{-1}(y)} u(x) \, \mathrm{d}\mathcal{H}^n(y) = \int_A u(x)\mathcal{J}f(x) \, \mathrm{d}\mathscr{L}^n(x).$$

Remark. (i) If $A \subset \mathbb{R}^n$ and $f \colon A \to \mathbb{R}$ Lipschitz, then

$$\bar{f}(x) := \inf_{z \in A} [f(z) + \mathrm{Lip}(f)|x - z|]$$

is a Lipschitz function on $\mathbb{R}^n$ with $\bar{f}|_A = f$ and $\mathrm{Lip}(\bar{f}|_A) = \mathrm{Lip}(f)$.

(ii) Notice that (for the purposes of proving the theorem), we can assume that f is differentiable on all of A.

To prove the first statement in the theorem, you approximate by linear functions.

Example. For $f \colon [0,1] \to \mathbb{R}^k$ Lipschitz, $\mathcal{J}f = |(\frac{d}{dt}f^1, ..., \frac{d}{dt}f^k)| = |\dot{f}(t)|$, so $\mathcal{H}^1(f([0,1])) = \int_0^1 |\dot{f}(t)| \, \mathrm{d}t$ holds for injective parametrisation of a curve, cf *Analyst's Travelling Salesman Problem.*

Coarea Formula

Given $f \colon \mathbb{R}^{n+k} \to \mathbb{R}^n$ Lipschitz, write Df_x for the linear map from $\mathbb{R}^{n+k} \to \mathbb{R}^n$ with matrix $(D_j f^i(x))_{i=1,\dots,n, j=1,\dots,n+k}$. And write

$$\mathcal{J}f(x) := \sqrt{\det(Df_x \circ (Df_x)^*)}.$$

Another Calculus 101 digression. Let $f \colon D \to [0,1]$ be the function from the disk in $\mathbb{R}^2$ to $\mathbb{R}$ by $f(x,y) = \sqrt{x^2 + y^2}$. If we have $t \in [0,1]$, then $f^{-1}(\{t\})$ is a circle. For each level set, find the length of the circle then integrate over t:

$$\int_0^1 \mathcal{H}^1(f^{-1}(\{t\})) \, \mathrm{d}t = \int_0^1 2\pi t \, dt = \pi = \mathrm{Area}(D) = \int_D 1 \, \mathrm{d}\mathcal{H}^2.$$

Theorem 4.5 (Coarea Formula)**.** Let $A \subset \mathbb{R}^{n+k}$ be $\mathscr{L}^{n+k}$-measurable and $f \colon A \to \mathbb{R}^n$ be Lipschitz. Then:

(i) $\int_{\mathbb{R}^n} \mathcal{H}^k(A \cap f^{-1}(y)) \, \mathrm{d}\mathcal{H}^n(y) = \int_A \mathcal{J}f(x) \, \mathrm{d}\mathscr{L}^{n+k}(x)$.

(ii) And if $u \colon A \to \mathbb{R}$ is $\mathscr{L}^{n+k}$-integrable,

$$\int_{\mathbb{R}^n} \left(\int_{f^{-1}(y)} u(x) \, \mathrm{d}\mathcal{H}^k(x) \right) \mathrm{d}\mathcal{H}^n(y) = \int_A u(x)\mathcal{J}f(x) \, \mathrm{d}\mathscr{L}^{n+k}(x).$$

Remark. For appropriate f, you will sometimes want to take $u = 1/\mathcal{J}f$. If $n = 1$, we have $\int_{\mathbb{R}} \int_{f^{-1}(y)} \frac{1}{|\nabla f|} \, \mathrm{d}\mathcal{H}^k \, \mathrm{d}\mathcal{H}^1 = \mathscr{L}^{n+1}(A)$.

5 Rectifiability and C^1 submanifolds

Theorem 5.1. If $f\colon \mathbb{R}^n \to \mathbb{R}$ is Lipschitz, then given $\epsilon > 0$ there exists C^1 function $g\colon \mathbb{R}^n \to \mathbb{R}$ with $\mathscr{L}^n(\{f(x) \neq g(x)\} \cup \{Df(x) \neq Dg(x)\}) < \epsilon$.

Theorem 5.2 (Whitney Extension — C^1 case)**.** Let $A \subset \mathbb{R}^n$ be closed and let $f\colon A \to \mathbb{R}$ be a continuous function. Suppose $\nu\colon A \to \mathbb{R}^n$ is a continuous function for which

$$\lim_{\delta \downarrow 0} \sup_{\substack{x,y \in K \\ 0 < |x-y| < \delta}} \frac{f(y) - f(x) - \nu(x)(x-y)}{|x-y|} = 0$$

for all compact $K \subset A$. Then there exists a C^1 function $g\colon \mathbb{R}^n \to \mathbb{R}$ with $g|_A = f$ and $Dg|_A = \nu$.

Remark. If A has an interior, then the condition on ν implies that the derivative of f is already ν.

Proof. We will describe the *Whitney decomposition* of A^c. Define the kth dyadic mesh as

$$\mathscr{M}_k := \left\{ \left[\frac{q_1}{2^k}, \frac{q_1+1}{2^k}\right] \times \dots \times \left[\frac{q_n}{2^k}, \frac{q_n+1}{2^k}\right] \subset \mathbb{R}^n : q_1, \dots q_n \in \mathbb{Z} \right\}.$$

The key property is that if $Q, Q' \in \bigcup_{k \in \mathbb{Z}} \mathscr{M}_k$ and $\operatorname{int} Q \cap \operatorname{int} Q' \neq \emptyset$, then either $Q \subset Q'$ or $Q' \subset Q$. Then define kth layer from A as

$$\Omega_k := \left\{ x \in \mathbb{R}^n : (2\sqrt{n})2^{-k} < \operatorname{dist}(x, A) < (2\sqrt{n})2^{-k+1} \right\}.$$

Now set

$$\mathscr{F}_0 := \bigcup_{k \in \mathbb{Z}} \{Q \in \mathscr{M}_k : Q \cap \Omega_k \neq \emptyset\}.$$

Notice that each cube $Q \in \mathscr{F}_0$ intersects at most 2 different dyadic layers. So if $Q, Q' \in \mathscr{F}_0$ are such that $\operatorname{int} Q \cap \operatorname{int} Q' \neq \emptyset$, then if $Q' \supset Q$, we have $\operatorname{diam}(Q') \leq 4 \operatorname{diam} Q$. So for each $Q \in \mathscr{F}_0$, $\{Q' \in \mathscr{F}_0 : Q' \supset Q\}$ is actually finite. So let $\mathscr{F}$ be a subcollection of $\mathscr{F}_0$ which is maximal with respect to inclusions (i.e. for each cube $Q \in \mathscr{F}_0$, take biggest cube containing it). We see now that

(i) $\operatorname{int} Q \cap \operatorname{int} Q' = \emptyset$ for all $Q, Q' \in \mathscr{F}$;

(ii) $A^c = \bigcup_{Q \in \mathscr{F}} Q$.

But much more is true. If $Q \in \mathscr{F} \cap \mathscr{M}_k$ then there is a point $x_0 \in Q \cap \Omega_k$ and this means $\operatorname{dist}(x_0, A) \leq (2\sqrt{n})2^{-k+1} = 4 \operatorname{diam} Q$. So $\operatorname{dist}(Q, A) \leq 4 \operatorname{diam} Q$. And, for any $a \in A, x \in Q$, we have

$$(2\sqrt{n})2^{-k} \leq |x_0 - a| \leq |x_0 - x| + |x - a| \leq \operatorname{diam}(Q) + |x - a|,$$

so $|x - a| \geq \operatorname{diam}(Q)$, i.e. $\operatorname{diam} Q \leq \operatorname{dist}(Q, A)$. So

$$\operatorname{diam}(Q) \leq \operatorname{dist}(Q, A) \leq 4 \operatorname{diam}(Q).$$

Next, if $Q, Q' \in \mathscr{F}$ are adjacent cubes ($Q \cap Q' \neq \emptyset$), then $\operatorname{diam}(Q) \leq \operatorname{dist}(Q, A) \leq \operatorname{dist}(Q', A) + \operatorname{diam}(Q') \leq 5 \operatorname{diam}(Q')$. In fact this means $\operatorname{diam}(Q) \leq 4 \operatorname{diam}(Q')$.

A given cube $Q \in \mathscr{M}_k$ intersects 3^n cubes in $\mathscr{M}_k$, and each cube in $\mathscr{M}_k$ determines 4^n cubes in $\mathscr{M}_{k+2}$. So at most 12^n cubes in $\mathscr{F}$ intersect Q.

Now, if $Q, Q' \in \mathcal{F}$ are not adjacent, i.e. $Q \cap Q' = \emptyset$, then there is some other cube Q'' such that $\operatorname{dist}(\frac{9}{8}Q', Q) \geq e(Q'') - \frac{1}{8}e(Q') > 0$ where e is the edge length. So $\frac{9}{8}Q' \cap Q \neq \emptyset \implies Q' \cap Q \neq \emptyset$. So now for any $x_0 \in A^c$, $x_0 \in Q$ for some $Q \in \mathcal{F}$. And $|\{\frac{9}{8}Q' : \frac{9}{8}Q' \ni x_0\}| \leq |\{\frac{9}{8}Q' : \frac{9}{8}Q' \cap Q \neq \emptyset\}| \leq |\{\frac{9}{8}Q' : Q' \cap Q \neq \emptyset\}| \leq C(n)$. So $\{\frac{9}{8}Q' : Q' \in \mathcal{F}\}$ has bounded overlap, in the sense that

$$\sum_{Q \in \mathcal{F}} 1_{\frac{9}{8}Q} \leq C(n).$$

Now the next part we are going to be sketchy. Let Q_0 denote the unit cube at the origin. Let $\varphi \in C_c^\infty(\frac{9}{8}Q_0)$ with $\varphi \equiv 1$ on Q_0, $0 \leq \varphi \leq 1$, and the derivatives are bounded by a constant that only depends on the dimension of the space and the power of the derivative, i.e. $|D^\alpha \varphi| \leq C(n, |\alpha|)$. Then set

$$\varphi_Q(x) \coloneqq \varphi\left(\frac{x - x_Q}{e(Q)}\right),$$

where x_Q is the centre of the cube Q. Then

$$\varphi_Q^* \coloneqq \frac{\varphi_Q}{\sum_{Q' \in \mathcal{F}} \varphi_{Q'}}$$

is a locally finite partition of unity (the sum is only ever a finite sum). For each cube $Q \in \mathcal{F}$ we also pick $p_Q \in A$ with $\operatorname{dist}(Q, A) = \operatorname{dist}(Q, p_Q)$. Finally, set, for $x \in A^c$,

$$g(x) = \sum_{Q \in \mathcal{F}} \varphi_Q^*(x)\big(f(p_Q) + \nu(p_Q)(x - p_Q)\big).$$

It is now quite a lot of checking to verify that this works. □

Proof of Theorem 5.1. There exists $A_1 \subset \mathbb{R}^n$ such that $\mathscr{L}^n(\mathbb{R}^n \setminus A_1) = 0$ and f is differentiable on A_1 (Rademacher). By Lusin, there exists $A_2 \subset A_1$ such that $\mathscr{L}^n(A_1 \setminus A_2) < \epsilon/4$ and $Df|_{A_2}$ is continuous. For $k \in \mathbb{N}$, let

$$\eta_k(x) \coloneqq \sup_{y \in B_{1/k}(x) \setminus \{x\}} \frac{f(y) - f(x) - Df(x)(y - x)}{|x - y|}.$$

We know $\eta_k \to 0$ pointwise on A_2. By Egoroff's Theorem $\exists A_3 \subset A_2$ with $\mathscr{L}^n(A_2 \setminus A_3) < \epsilon/4$ and such that $\eta_k \to 0$ locally uniformly on A_3. And then by inner regularity, there exists a closed $A \subset A_3$ with $\mathscr{L}^n(A_3 \setminus A) < \epsilon/4$. Now apply Whitney extension on A with $\nu = Df$. □

Definition. A set $M \subset \mathbb{R}^{n+k}$ is said to be *countably n-rectifiable* (or often just n-rectifiable or rectifiable) if $\mathcal{H}^n(M \cap K) < \infty$ for all compact $K \subset \mathbb{R}^{n+k}$ and $M \subset M_0 \cup \bigcup_{j=1}^\infty F_j(\mathbb{R}^n)$, where $\mathcal{H}^n(M_0) = 0$ and $F_j \colon \mathbb{R}^n \to \mathbb{R}^{n+k}$ are Lipschitz functions.

Exercise. If M is n-rectifiable then $M \subset M_0 \cup \bigcup_{j=1}^\infty N_j$, where $\mathcal{H}^(M_0) = 0$ and N_j are embedded n-dimensional C^1-submanifolds of $\mathbb{R}^{n+k}$.

Definition. Let $M \subset \mathbb{R}^{n+k}$ be a set with the following property: $\forall y \in M$, $\exists$ open sets $\mathcal{U} \subset \mathbb{R}^{n+k}$, $V \subset \mathbb{R}^n$ and a C^1 map $\Psi \colon V \to U$ such that

(i) $y \in \mathcal{U}$;

(ii) $\Psi(V) = M \cap \mathcal{U}$;

(iii) Ψ is injective;

(iv) $D\Psi(x)$ is injective for every $x \in V$;

(v) $\Psi^{-1}(K)$ is compact in V whenever K is compact in $\mathcal{U}$ [properness].

Then we say that M is a *(properly) embedded n-dimensional C^1-submanifold* of $\mathbb{R}^{n+k}$.

Remark. Notice if $A \subset \mathcal{U}$ is Borel, then $\mathcal{H}^n(A \cap M) = \int_{\Psi^{-1}(A)} \mathcal{J}\,\Psi \,\mathrm{d}\mathcal{H}^n$, by the area formula. And

$$\mathcal{J}\,\Psi = \det(D_i\Psi \cdot D_j\Psi)^{1/2},$$

i.e. in the language of differential geometry, Riemannian volume on M is given by $\sqrt{\det g}$, where $g_{ij} = D_i\Psi \cdot D_j\Psi$ is the metric induced by coordinates given by $\Psi^{-1}|_{M\cap\mathcal{U}}$.

Definition. Let $(\mu_j)_{j=1}^{\infty}$ and μ be Radon measures on the metric space X. We say $\mu_j \to \mu$ in the sense of Radon measures if $\mu_j(f) \to \mu(f)$ for all $f \in C_c(X)$. [In probability, this is weak convergence of measures.] Space of Radon measures on X is $C_c(X)^*$ and so this is a w^*-convergence in the functional analysis sense.

Definition. We say that $M \subset \mathbb{R}^{n+k}$ has an *approximate tangent plane P* at $x_0 \in M$ if $\mathcal{H}^n\lfloor\eta_{x_0,\rho}(M) \to \mathcal{H}^n\lfloor P$ as $\rho \to 0$ in the sense of Radon measures, where $\eta_{x_0,\rho}(x) = \frac{x-x_0}{\rho}$.

Exercise. If M is n-rectifiable then it has an approximate tangent plane at $\mathcal{H}^n$-almost every point of M.

Theorem. Suppose $M \subset \mathbb{R}^{n+k}$ has $\mathcal{H}^n(M \cap K) < \infty$ for all compact $K \subset \mathbb{R}^{n+k}$. If M has an approximate tangent plane at $\mathcal{H}^n$-a.e. point of M, then it is n-rectifiable. By taking $f \in C_c(\mathbb{R}^{n+k})$ with $0 \le f \le 1$, $f \equiv 0$ on $B_{1+\epsilon}(0)^c$ and $f \equiv 1$ on $B_1(0)$, we have

$$\frac{\mathcal{H}^n(M \cap B_\rho(x_0))}{\omega_n \rho^n} = \omega_n^{-1}\mathcal{H}^n(\eta_{x_0,\rho}(M) \cap B_1(0)) = \omega_n^{-1}\int_{\eta_{x_0,\rho}(M)} f \,\mathrm{d}\mathcal{H}^n + O(\epsilon),$$

and so if M has approximate tangent plane at x_0, then $\lim_{\rho\downarrow 0} \frac{\mathcal{H}^n(M\cap B_\rho(x_0))}{\omega_n\rho^n} = 1$. This is the start of a very long story.