GRADUATE REAL ANALYSIS

DAVID BATE

Part 1. General measure theory

1. Measures

We wish to assign a value to the size of subsets of some given space, such as the length, area or volume of subsets of $\mathbb{R}^m$.

Definition 1.1. A *measure* μ on a set X is a function

$$\mu\colon \{A : A \subset X\} \to [0, \infty]$$

such that

(1) $\mu(\emptyset) = 0$;
(2) $\mu(A) \subset \mu(B)$ whenever $A \subset B \subset X$;
(3) $\mu(\bigcup_{i\in\mathbb{N}} A_i) \leq \sum_{i\in\mathbb{N}} \mu(A_i)$ whenever $A_1, A_2, \ldots \subset X$.

A function satisfying (2) is said to be *monotonic* and a function satisfying (3) is said to be *countably sub-additive*.

Definition 1.2. Let μ be a measure on a set X. A set $A \subset X$ is *μ-measurable* if, for every $E \subset X$,

$$\mu(E) = \mu(E \cap A) + \mu(E \setminus A). \tag{1.1}$$

Remark 1.3.
(1) Since a measure is countably sub-additive, it is sufficient to check the $\geq$ inequality in (1.1).
(2) In particular, it suffices to check (1.1) for $E \subset X$ with $\mu(E) < \infty$.
(3) If A is μ-measurable then so is $X \setminus A$.
(4) If $\mu(A) = 0$ then A is μ-measurable.

Definition 1.4. If μ is a measure on a set X and $S \subset X$, the *restriction* of μ to A is defined as

$$\mu|_S(A) := \mu(S \cap A).$$

Lemma 1.5. *Let μ be a measure on a set X and $S \subset X$. Then $\mu|_S$ is a measure on X and any μ-measurable set is also $\mu|_S$-measurable.*

Proof. The fact that $\mu|_S$ is a measure follows immediately from the fact that μ is a measure. If $A \subset X$ is μ-measurable, then for any $E \subset X$,

$$\begin{aligned}\mu|_S(E) = \mu(E \cap S) &= \mu(E \cap S \cap A) + \mu(E \cap S \setminus A)\\ &= \mu|_S(E \cap A) + \mu|_S(E \setminus A),\end{aligned}$$

as required. □

Theorem 1.6. *Let μ be a measure on a set X and let $\mathcal{M}$ be the set of μ-measurable subsets of X.*

(1) *If $A_1, A_2, \ldots \in \mathcal{M}$ then $\bigcup_{i\in\mathbb{N}} A_i \in \mathcal{M}$ and $\bigcap_{i\in\mathbb{N}} A_i \in \mathcal{M}$.*

Date: Autumn 2021.

(2) μ *is* countably additive on $\mathcal{M}$. *That is, if* $A_1, A_2 \ldots \in \mathcal{M}$ *are disjoint then*

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \sum_{i\in\mathbb{N}} \mu(A_i).$$

(3) *If* $A_1 \subset A_2 \subset \ldots \in \mathcal{M}$ *then*

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) = \lim_{i\to\infty} \mu(A_i).$$

(4) *If* $A_1 \supset A_2 \supset \ldots \in \mathcal{M}$ *and* $\mu(A_1) < \infty$ *then*

$$\mu\left(\bigcap_{i\in\mathbb{N}} A_i\right) = \lim_{i\to\infty} \mu(A_i).$$

Proof. We first prove (1) for finite unions and intersections. If $A, B \in \mathcal{M}$ then for every $E \subset X$,

$$\begin{aligned}\mu(E) &= \mu(E\cap A) + \mu(E\setminus A)\\ &= \mu(E\cap A) + \mu((E\setminus A)\cap B) + \mu(E\setminus(A\cup B))\\ &\geq \mu(E\cap(A\cup B)) + \mu(E\setminus(A\cup B))\end{aligned}$$

by sub-additivity. Thus $A\cup B$ is μ-measurable and induction gives finite unions. Taking complements gives finite intersections.

To prove (2) note that the inequality $\leq$ is given by sub-additivity. For the other inequality, for each $i\in\mathbb{N}$ let $A_i \in \mathcal{M}$ be disjoint and for each $j \in \mathcal{M}$ let

$$B_j = \bigcup_{i=1}^{j} A_i,$$

which is measurable by (1). Note that

$$B_j = B_{j-1} \cup A_j$$

and that this union is disjoint. Therefore, since A_j is μ-measurable,

$$\begin{aligned}\mu(B_j) &= \mu(B_j\cap A_j) + \mu(B_j\setminus A_j)\\ &= \mu(A_j) + \mu(B_{j-1}),\end{aligned}$$

since the A_i are all disjoint. Therefore, by induction, $\mu(B_j) = \sum_{i=1}^{j} \mu(A_i)$ for each $j\in\mathbb{N}$. Finally, for each $j\in\mathbb{N}$, since μ is monotonic,

$$\mu\left(\bigcup_{i\in\mathbb{N}} A_i\right) \geq \mu(B_j) = \sum_{i=1}^{j} \mu(A_i)$$

and so letting $j\to\infty$ gives (2).

(3) follows by applying (2) to the disjoint measurable sets $B_j = A_j \setminus A_{j-1}$.

(4) follows from (3) by setting $B_j = A_1\setminus A_j$, so that

$$A_1 = \bigcap_{i\in\mathbb{N}} A_i \cup \bigcup_{i\in\mathbb{N}} B_i$$

and the B_j increase. By sub-additivity,

$$\begin{aligned}\mu(A_1) &\leq \mu\left(\bigcap_{i\in\mathbb{N}} A_i\right) + \lim_{j\to\infty} \mu(B_j) \\ &= \mu\left(\bigcap_{i\in\mathbb{N}} A_i\right) + \lim_{j\to\infty} \mu(A_1) - \mu(A_j),\end{aligned}$$

by applying (1) for finite unions. Since $\mu(A_1) < \infty$, (4) follows.

Finally, to prove (1) for countable unions, for each $j \in \mathbb{N}$ let

$$B_j = \bigcup_{i=1}^{j} A_i,$$

an increasing sequence, and let $E \subset X$ with $\mu(E) < \infty$. Since the B_j are μ-measurable,

$$\begin{aligned}\mu(E) &= \lim_{j\to\infty} \mu(E \cap B_j) + \lim_{j\to\infty} \mu(E \setminus B_j) \\ &= \mu\left(E \cap \bigcup_{i\in\mathbb{N}} B_i\right) + \mu\left(E \setminus \bigcup_{i\in\mathbb{N}} B_i\right) \\ &= \mu\left(E \cap \bigcup_{i\in\mathbb{N}} A_i\right) + \mu\left(E \setminus \bigcup_{i\in\mathbb{N}} A_i\right),\end{aligned}$$

using the fact that the B_j are $\mu|_E$-measurable in the second equality. Taking complements shows that countable intersections of measurable sets are measurable. $\square$

Definition 1.7. A collection Σ of subsets of a set X is a *σ-algebra* if

(1) $\emptyset \in \Sigma$;
(2) $A \in \Sigma \Rightarrow X \setminus A \in \Sigma$;
(3) $A_1, A_2, \ldots \in \Sigma \Rightarrow \bigcup_{i\in\mathbb{N}} A_i \in \Sigma$.

Theorem 1.6 shows that the set of μ-measurable sets is a σ-algebra.

For Ω a set of subset of a set X, the *σ-algebra generated by* Ω is

$$\Sigma(\Omega) := \bigcap \{\Sigma' : \Sigma' \supset \Omega,\ \Sigma' \text{ a } \sigma\text{-algebra}\}.$$

By Exercise 1.2, it is a σ-algebra.

The *Borel σ-algebra* of a topological space X is the σ-algebra generated by the open (respectively closed) subsets of X. It will be denoted by $\mathcal{B}(X)$ and its elements called the *Borel* subsets of X.

A measure for which all Borel sets are measurable is a *Borel measure.* It is *Borel regular* if for every $A \subset X$ there exists a Borel $B \supset A$ with $\mu(B) = \mu(A)$.

Theorem 1.8 (Carathéodory criterion)**.** *Let (X, d) be a metric space and μ a measure on X which is additive on separated sets. That is, whenever $A, B \subset X$ with*

$$\inf\{d(x, y) : x \in A,\ y \in B\} > 0,$$

we have

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

Then μ is a Borel measure.

Proof. Let $C \subset X$ be closed and $E \subset X$ with $\mu(E) < \infty$. We need to show

$$\mu(E) \geq \mu(E \cap C) + \mu(E \setminus C).$$

For each $j \in \mathbb{N}$ let

$$E_j = \{x \in E : \frac{1}{j+1} < \operatorname{dist}(x, C) \leq \frac{1}{j}\}$$

and

$$E_0 = \{x \in E : \operatorname{dist}(x, C) > 1\}.$$

Since C is closed,

$$E \setminus C = E_0 \cup \bigcup_{j \in \mathbb{N}} E_j.$$

Moreover, the E_j with j odd are pairwise separated so

$$\mu(E) \geq \mu(\bigcup_{j \text{ odd}} E_j) = \sum_{j \text{ odd}} \mu(E_j)$$

and so the sum is convergent. Similarly the sum over even indices is convergent and so

$$\sum_{j \geq n} \mu(E_j) \to 0 \quad \text{as } n \to \infty.$$

Therefore

$$\begin{aligned}
\mu(E) &\geq \mu\left(E \cap C \cup \bigcup_{j=0}^{n} E_j\right) \\
&= \mu(E \cap C) + \mu\left(\bigcup_{j=0}^{n} E_j\right) \\
&\geq \mu(E \cap C) + \mu(E \setminus C) - \sum_{j > n} \mu(E_j) \\
&\to \mu(E \cap C) + \mu(E \setminus C),
\end{aligned}$$

using the additivity on separated sets for the equality and countable sub-additivity for the second inequality. □

Definition 1.9 (Carathéodory construction)**.** Let (X, d) be a metric space, F a set of subsets of X and $\zeta \colon F \to [0, \infty]$. For each $\delta > 0$ and $A \subset X$ define

$$\psi_\delta(A) = \inf \sum_{S \in G} \zeta(S),$$

where the infimum is taken over all countable

$$G \subset \{S \in F : \operatorname{diam}(S)\} < \delta$$

such that

$$A \subset \bigcup_{S \in G} S.$$

Finally, define $\psi(A) = \sup_{\delta > 0} \psi_\delta(A)$.

For any $\delta > 0$, ψ_δ is a measure, as is ψ. Theorem 1.8 shows that ψ is a Borel measure on X. Indeed, if $\operatorname{dist}(A, B) > \delta$ then

$$\psi_\delta(A \cup B) \geq \psi_\delta(A) + \psi_\delta(B).$$

If F consists only of Borel sets then ψ is Borel regular.

Remark 1.10. The fact that $\psi_{\delta'} \leq \psi_\delta$ whenever $\delta' \geq \delta$ implies that

$$\psi(A) = \lim_{\delta \to 0} \psi_\delta(A).$$

Definition 1.11. We define some properties of a measure μ on a topological space X.

(1) μ is *locally finite* if every point in X has a neighbourhood of finite measure.
(2) μ is *σ-finite* if there exist measurable $X_i \subset X$ with $\mu(X_i) < \infty$ and $X = \bigcup_{i\in\mathbb{N}} X_i$.
(3) μ is *finite* if $\mu(X) < \infty$.
(4) A Borel regular measure μ is a *Radon* measure if
  (a) $\mu(K) < \infty$ for all compact $K \subset X$,
  (b) $\mu(A) = \sup\{\mu(K) : K \subset A \text{ compact}\}$ for all Borel $A \subset X$.
  (c) $\mu(A) = \inf\{\mu(U) : U \supset A \text{ open}\}$ for all Borel $A \subset X$.

Definition 1.12. Let μ be a measure on a set X. A property of points in X holds *μ almost everywhere* (or μ-a.e.) if the set of points for which the property doesn't hold has μ measure zero.

Definition 1.13. Let X, Y be sets, μ be a measure on X and let $f\colon X \to Y$. The *push forward* of μ under f, written $f_\#\mu$ is defined by

$$f_\#\mu(S) = \mu(f^{-1}(S)).$$

Definition 1.14. The *Lebesgue measure* on $\mathbb{R}^n$, denoted $\mathcal{L}^n$, is defined using the Carathéodory construction with F the set of cubes and $\zeta(Q) = \mathrm{vol}(Q)$. Its measurable sets are called the *Lebesgue measurable* subsets of $\mathbb{R}^n$.

The following lemma is very useful.

Lemma 1.15. *Let μ be a finite measure on a set X and let $\mathcal{S}$ be a set of μ-measurable subsets of X. There exists disjoint $S_i \in \mathcal{S}$ such that any $S \in \mathcal{S}$ with*

$$S \subset X \setminus \bigcup_{i\in\mathbb{N}} S_i$$

satisfies $\mu(S) = 0$.

In particular, if each μ-measurable subset of X of positive measure contains an element of $\mathcal{S}$ of positive measure then we can decompose almost all of X into countably many disjoint elements of $\mathcal{S}$.

Proof. We find the S_i by induction. First let $\mathcal{S}^1 \subset \mathcal{S}$ be countable and disjoint such that

$$\mu(\cup\mathcal{S}^1) \geq \sup\{\mu(\cup\mathcal{S}') : \mathcal{S}' \subset \mathcal{S} \text{ countable and disjoint}\} - 1/1.$$

Now let $\mathcal{M}_2$ be the set of all $\mathcal{S}' \subset \mathcal{S}$ that are countable, disjoint and disjoint from $\cup\mathcal{S}^1$. Let $\mathcal{S}^2 \in \mathcal{M}_2$ be such that

$$\mu(\cup\mathcal{S}^2) \geq \sup\{\mu(\cup\mathcal{S}') : \mathcal{S}' \in \mathcal{M}_2\} - 1/2.$$

Inductively, given countable, disjoint $\mathcal{S}^1, \dots, \mathcal{S}^{i-1}$ such that each $\mathcal{S}^j$ and $\mathcal{S}^k$ are disjoint for $k < j$, let $\mathcal{M}_i$ be the set of all $\mathcal{S}' \subset \mathcal{S}$ that are countable, disjoint and disjoint from $\mathcal{S}^1 \cup \ldots \cup \mathcal{S}^{i-1}$. Let $\mathcal{S}^i \in \mathcal{M}_i$ be such that

$$\mu(\cup\mathcal{S}^i) \geq \sup\{\mu(\cup\mathcal{S}') : \mathcal{S}' \in \mathcal{M}_i\} - 1/i.$$

We claim that any $S \in \mathcal{S}$ with

$$S \subset X \setminus \bigcup_{i\in\mathbb{N}} \bigcup \mathcal{S}^i$$

satisfies $\mu(S) = 0$. If not, let $i \in \mathbb{N}$ be such that $1/i < \mu(S)$. Then $\mathcal{T} := \mathcal{S}^i \cup \{S\} \in \mathcal{M}_i$ and

$$\mu(\cup\mathcal{T}) > \sup\{\mu(\cup\mathcal{S}') : \mathcal{S}' \in \mathcal{M}_i\} - 1/i + 1/i,$$

a contradiction. $\square$

1.1. Exercises.

Exercise 1.1. Usually in measure theory, a measure is defined as a countably additive function defined on a σ-algebra. However, using our definition is simply a convenience rather than a restriction.

Indeed, suppose μ is a countably additive function defined on a σ-algebra Σ of X with $\mu(\emptyset) = 0$. Show that it can be extended to the power set of X by

$$\overline{\mu}(A) = \inf\{\mu(B) : A \subset B \in \Sigma\}$$

and that any $B \in \Sigma$ is μ-measurable. What about

$$\underline{\mu}(A) = \sup\{\mu(B) : A \supset B \in \Sigma\}?$$

Conversely, any measure is countably additive when restricted to any σ-algebra of measurable sets.

Exercise 1.2. Let Ω be a set of subsets of a set X. Show that $\Sigma(\Omega)$ is a σ-algebra. Note that it is the smallest σ-algebra of X containing Ω.

Exercise 1.3. Show that the following sets are Borel subsets of $\mathbb{R}$: $\mathbb{Q}$, $[0, 1)$, the set of points in $[0, 1]$ whose first decimal is even.

Let $f : [0, 1] \to [0, 1]$. Show that the set of points where f is continuous is a Borel set. What about the set of points where f is differentiable?

Exercise 1.4. Let X be a set and $x \in X$. The *Dirac measure at x* is defined as $\delta_x(A) = 1$ if $x \in A$, $\delta_x(A) = 0$ otherwise. Show that δ_x is a measure on X. What are its measurable sets?

Define the *counting measure* on X to be the cardinality (finite or ∞) of any subset of X. Show that this is a measure. What are its measurable sets?

Exercise 1.5. For (X, d) a metric space and $s \geq 0$ the *s-dimensional Hausdorff measure* on X, denoted $\mathcal{H}^s$, is defined using the Carathéodory construction with F the set of all sets and $\zeta(S) = \operatorname{diam}(S)^s$.

(1) Show that $\mathcal{L}^n$ and $\mathcal{H}^n$ are non-zero, translation invariant and n-homogenous measures. That is, for any $A \subset \mathbb{R}^n$, $x \in \mathbb{R}^n$ and $t > 0$, $\mathcal{L}^n(A + x) = \mathcal{L}^n(A)$ and $\mathcal{L}^n(tA) = t^n\mathcal{L}^n(A)$ (and similarly for $\mathcal{H}^n$).
(2) On $\mathbb{R}^n$ show that there exists a $C > 0$ such that $\mathcal{H}^n/C \leq \mathcal{L}^n \leq C\mathcal{H}^n$.
(3) Let $f : X \to Y$ be an L-Lipschitz function between two metric spaces. Show that for any $s \geq 0$ and $A \subset X$,

$$\mathcal{H}^s(f(A)) \leq L^s\mathcal{H}^s(A).$$

(4) For any metric space X, show that $\mathcal{H}^0$ is the counting measure on X.
(5) For $0 \leq s < t < \infty$, suppose that $\mathcal{H}^s(A) < \infty$. Show that $\mathcal{H}^t(A) = 0$. Hence there exists a single $0 \leq s \leq \infty$ for which $\mathcal{H}^t(A) = 0$ for all $t > s$ and $\mathcal{H}^t(A) = \infty$ for all $t < s$. This t is called the *Hausdorff dimension* of A, denoted $\dim_{\mathrm{H}} A$.

Exercise 1.6. The *Cantor set* $K \subset [0,1]$ is defined as follows. Let $K_0 = [0,1]$ and for each $i \in \mathbb{N}$ let K_i be obtained from deleting the "middle third" open interval from each of the intervals in K_{i-1}. That is, $K_1 = [0,1/3] \cup [2/3,1]$,

$$K_2 = [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1],$$

etc. Define $K = \bigcap_{i\in\mathbb{N}} K_i$. Note that K is compact and hence Borel.

(1) Show that K is uncountable.
(2) Let $s = \log 2/\log 3$. Show that $0 < \mathcal{H}^s(K) < \infty$.

In particular, K is an uncountable subset of $\mathbb{R}$ with $\mathcal{L}^1(K) = 0$.

Exercise 1.7. Give an examples of $S \subset \mathbb{R}^2$ with $\dim_{\mathrm{H}} S = 1$ for which $\mathcal{H}^1|_S$ is:

(1) finite,
(2) σ-finite but not finite,
(3) not σ-finite.

Exercise 1.8. The fundamental properties of measures are those given in Theorem 1.6, in particular countable additivity. It is necessary for us to only require this to be true for measurable sets, as can be seen from the existence of non-measurable sets.

Define a *Vitali set* as follows. Consider the equivalence relation $\sim$ on $\mathbb{R}$ defined by $x \sim y$ iff $x - y \in \mathbb{Q}$. By the density of $\mathbb{Q}$ in $\mathbb{R}$, each equivalence class V_x intersects $[0,1]$. Therefore, by the axiom of choice(!), we may construct a set $\mathcal{V} \subset [0,1]$ consisting of exactly one member of each equivalence class.

Show:

(1) If $p \neq q$ are rational then $p + \mathcal{V}$ and $q + \mathcal{V}$ are disjoint.
(2) $[0,1] \subset \bigcup\{q + \mathcal{V} : q \in \mathbb{Q} \cap [-1,1]\} \subset [-1,2]$.
(3) Show that $\mathcal{L}^1(\mathcal{V}) \neq 0$.
(4) Deduce that $\mathcal{V}$ is not Lebesgue measurable.

Exercise 1.9. Show that the two extensions given in Exercise 1.1 may not agree. For example, after extending Lebesgue measure (restricted to the Borel sets), what are the values of a Vitali set?

Exercise 1.10. Let μ be a finite Borel measure on a metric space X. Prove that for every Borel $B \subset X$,

$$\mu(B) = \sup\{\mu(C) : C \subset B \text{ closed}\} \tag{1.2}$$

and

$$\mu(B) = \inf\{\mu(U) : U \supset B \text{ open}\}. \tag{1.3}$$

Property (1.2) is called *inner regularity by closed sets* and (1.3) is called *outer regularity by open sets.*

Hint: observe that it suffices to show that all Borel sets satisfy (1.2). Show that the set

$$\{B \subset X : B \text{ and } X \setminus B \text{ satisfy } (1.2)\}$$

is a σ-algebra that contains all closed subsets of X.

Show that a σ-finite μ is inner regular by closed sets. Show that a σ-finite μ is outer regular by open sets if there exist open sets $U_i \subset X$ with $\mu(U_i) < \infty$ for all $i \in \mathbb{N}$ and $X = \bigcup_{i\in\mathbb{N}} U_i$. Give an example of a σ-finite μ that is *not* outer regular by open sets.

Exercise 1.11. Let X be a complete and separable metric space. Show that any finite Borel measure on X is a Radon measure.

Hint: a metric space is compact if and only if it is complete and totally bounded.

2. Integration

Definition 2.1. Let μ be a measure on a set X. A *simple function* is any function of the form

$$\sum_{i=1}^{m} a_i \chi_{A_i},$$

where each $a_i \in \mathbb{R}$ and the $A_i \subset X$ are disjoint μ-measurable sets. We treat $0\cdot\infty = 0$.

Definition 2.2. Let μ be a measure on a set X and let $f\colon X \to \mathbb{R}^+$. The (lower) *integral of f with respect to μ* is

$$\int f \,\mathrm{d}\mu := \sup \left\{ \sum_{i=1}^{m} a_i \mu(A_i) : s = \sum_{i=1}^{m} a_i \chi_{A_i} \leq f,\ s \text{ simple} \right\}.$$

Definition 2.3. Let μ be a measure on a set X. A function $f\colon X \to \mathbb{R}$ is *μ-measurable* if $f^{-1}((a,\infty))$ is μ-measurable for every $a \in \mathbb{R}$.

For $f\colon X \to \mathbb{R}$ measurable, let $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$ (both μ-measurable), so that $f = f^+ - f^-$ and $|f| = f^+ + f^-$. If one of $\int_X f^+ \,\mathrm{d}\mu$ and $\int_X f^- \,\mathrm{d}\mu$ are finite, we say that f is *μ-integrable* and we define the *integral of f with respect to μ* to be

$$\int_X f \,\mathrm{d}\mu = \int_X f^+ \,\mathrm{d}\mu - \int_X f^- \,\mathrm{d}\mu.$$

If only $\int_X f^- \,\mathrm{d}\mu < \infty$ (respectively $\int_X f^+ \,\mathrm{d}\mu < \infty$) we write $\int_X f \,\mathrm{d}\mu = \infty$ (respectively $\int_X f \,\mathrm{d}\mu = -\infty$).

Let X be a topological space. A function $f\colon X \to \mathbb{R}$ is a *Borel function* if $f^{-1}((a,\infty))$ is a Borel set for every $a \in \mathbb{R}$.

There are some simple properties of the integral to check, such as linearity and monotonicity. See Exercise 2.3.

Linear combinations of measurable functions are measurable, as are limits of measurable functions. See Exercise 2.2.

Theorem 2.4 (Fatou's lemma)**.** *Let μ be a measure on a set X and $f_k\colon X \to [0,\infty]$ μ-measurable. Then*

$$\int_X \liminf_{k\to\infty} f_k \,\mathrm{d}\mu \leq \liminf_{k\to\infty} \int_X f_k \,\mathrm{d}\mu.$$

Proof. Let

$$s = \sum_{i=1}^{m} a_i \chi_{A_i}$$

be a simple function with

$$s \leq \liminf f_k.$$

for each $x \in A_i$ and each $1 \leq i \leq m$ and let $0 < t < 1$. For each $1 \leq i \leq m$, the sets

$$G_{k,i} := \{x \in A_i : f_k(x) \geq t a_i \text{ for all } j \geq k\}$$

monotonically increase to A_i as k increases. Therefore

$$\int f_k \,\mathrm{d}\mu \geq \sum_{i=1}^{m} t a_i \mu(G_{k,i}) \to \sum_{i=1}^{n} t a_i \mu(A_i).$$

and hence
$$\sum_{i=1}^{n} ta_i\mu(A_i) \leq \liminf_{k\to\infty} \int f_k \,\mathrm{d}\mu.$$
Since $0 < t < 1$ is arbitrary, the conclusion follows. $\square$

Remark 2.5 (Reverse Fatou). Suppose that there exists $g \geq 0$ with $\int g \,\mathrm{d}\mu < \infty$ and $f_k \leq g$ for all k. Then
$$\limsup_{k\to\infty} f_k \,\mathrm{d}\mu \geq \limsup_{k\to\infty} \int_X f_k \,\mathrm{d}\mu.$$
Indeed, this follows by applying Fatou's lemma to $g - f_k$.

Theorem 2.6 (Monotone convergence theorem)**.** *Let μ be a measure on a set X and $f_k\colon X \to [0,\infty]$ μ-measurable. Suppose that for every $x \in X$ and all $k \in \mathbb{N}$, $f_{k+1}(x) \geq f_k(x)$. Then*
$$\lim_{k\to\infty} \int f_k \,\mathrm{d}\mu = \int \lim_{k\to\infty} f_k \,\mathrm{d}\mu.$$

Proof. The monotonicity of the integral gives $\leq$ whilst Fatou's lemma gives $\geq$. $\square$

Theorem 2.7. *Let μ be a measure on X and $f_n\colon X \to \mathbb{R}$ μ-measurable such that $f_n \to f$ pointwise. Suppose that there exists measurable $g\colon X \to [0,\infty]$ with $\int g \,\mathrm{d}\mu < \infty$ such that $|f_n(x)| \leq g(x)$ for all $x \in X$. Then*
$$\int f_n \,\mathrm{d}\mu \to \int f \,\mathrm{d}\mu.$$

Proof. Observe that for all $n \in \mathbb{N}$, $|f - f_n| \leq 2g$ and that $\limsup |f - f_n| = 0$. Then by the reverse Fatou lemma,
$$\left| \int f \,\mathrm{d}\mu - \int f_n \,\mathrm{d}\mu \right| \leq \int |f - f_n| \,\mathrm{d}\mu \to 0.$$
$\square$

2.1. **Exercises.**

Exercise 2.1. For μ a measure on a set X, let $f\colon X \to \mathbb{R}$ be measurable, respectively Borel. Show that the pre-image of any Borel $B \subset \mathbb{R}$ is μ-measurable, respectively Borel. Compare this to the definition of a continuous function.

Exercise 2.2. Let μ be a measure on X and for each $i \in \mathbb{N}$ let $f_i\colon X \to \mathbb{R}$ be μ-measurable. Show that the functions
$$\limsup_{i\to\infty} f_i \quad \text{and} \quad \liminf_{i\to\infty} f_i$$
are μ-measurable.

Show that a linear combination of μ-measurable functions is μ-measurable. Show that a countable (pointwise) sum of μ-measurable functions is μ-measurable.

Exercise 2.3. There are some simple properties of the integral to check:

(1) If $f \leq g$ μ-a.e. then
$$\int f d\mu \leq \int g \,\mathrm{d}\mu;$$
(2) The integral with respect to μ is a linear operator;

(3) If $S \subset X$ is μ-measurable then
$$\int_X f\,\mathrm{d}\mu = \int_S f\,\mathrm{d}\mu + \int_{X\setminus S} f\,\mathrm{d}\mu;$$
(4) $|\int f\,\mathrm{d}\mu| \leq \int |f|\,\mathrm{d}\mu$;
(5) etc...

Exercise 2.4. Show that $f\colon X \to \mathbb{R}$ is μ-measurable if and only if
$$\mu(E) \geq \mu(E \cap f^{-1}((-\infty, a))) + \mu(E \cap f^{-1}((b, \infty)))$$
for every $E \subset X$ and $a < b \in \mathbb{Q}$.

Exercise 2.5. State and prove a reverse monotone convergence theorem for monotonically decreasing sequences of functions.

Exercise 2.6. Show that the Fatou lemma is false if the functions are not uniformly bounded below.

Show that the reverse Fatou lemma is false if the sequence is not bounded above by an integrable g.

Show that the monotone convergence theorem is false if the sequence does not monotonically increase.

Exercise 2.7. Let X, Y be sets, μ a measure on X and $f\colon X \to Y$. Show that $f_{\#}\mu$ is a measure on Y. If X, Y are topological spaces and μ, f are Borel, show that $f_{\#}\mu$ is a Borel measure on Y.

3. Some standard theorems

Theorem 3.1 (Egorov's theorem)**.** *Let μ be a finite measure on a set X and $f_n\colon X \to \mathbb{R}$ a sequence of μ-measurable functions such that $f_n \to f$ pointwise μ-a.e. Then for every $\epsilon > 0$ there exists a measurable $G \subset X$ with $\mu(X \setminus G) < \epsilon$ and $f_n \to f$ uniformly on G.*

Proof. Fix $\epsilon > 0$, $k \in \mathbb{N}$ and for each $n \in \mathbb{N}$ let
$$B_{n,k} = \{x \in X : |f_n(x) - f(x)| > 1/k \text{ for some } m \geq n\}.$$
By assumption, the $B_{n,k}$ are measurable sets that monotonically decrease to a μ-null set as $n \to \infty$. Therefore, there exists $n \in \mathbb{N}$ such that $\mu(B_{n,k}) < \epsilon 2^{-k}$. Let
$$G_k = X \setminus B_{n,k} \quad \text{and} \quad G = \bigcap_{k\in\mathbb{N}} G_k.$$
Then $\mu(X \setminus G) < \epsilon$ and, for each $x \in G$ and each $k \in \mathbb{N}$, $G \subset G_k$, so there exists $n \in \mathbb{N}$ such that
$$|f(x) - f_m(x)| < 1/k$$
for all $m \geq n$. That is, $f_m \to f$ uniformly on G, as required. □

Theorem 3.2 (Lusin's theorem)**.** *Let μ be a finite Borel measure on a metric space X and let $f\colon X \to \mathbb{R}$ be μ-measurable. Then for every $\epsilon > 0$ there exists a closed $C \subset X$ with $\mu(X \setminus C) < \epsilon$ such that $f|_C$ is continuous.*

Proof. Fix $\epsilon > 0$ and for each $i \in \mathbb{Z}$ let
$$X_i = f^{-1}([i\epsilon, (i+1)\epsilon)),$$

a collection of disjoint Borel sets which cover X. Since $\mu(X) < \infty$, there exists $n \in \mathbb{N}$ such that
$$\mu\left(X \setminus \bigcup_{i=1}^{n} X_i\right) < \epsilon.$$
Since μ is Borel, for each $1 \le i \le n$ there exists a closed $C_i \subset X_i$ with $\mu(X_i \setminus C_i) < \epsilon/n$.

For a moment fix $1 \le i \le n$ and let
$$D = \bigcup_{1 \le j \ne i \le n} C_j.$$
Since D is closed, the sets $B(D,\delta)$ monotonically decrease to D as $\delta \to 0$, which is disjoint from C_i. Therefore there exists $\delta_i > 0$ such that
$$C_i' := C_i \setminus B(D,\delta_i)$$
satisfies $\mu(C_i \setminus C_i') < \epsilon/n$. Note that C_i' is closed.

Let $\delta := \min_{1\le i \le n} \delta_i > 0$ and $C_\epsilon = \bigcup_{i=1}^n C_i'$, a closed set. For any $1 \le i \ne j \le n$ and $x \in C_i'$ and $y \in C_j'$, $d(x,y) \ge \delta$. Therefore, if $x, y \in C_\epsilon$ with $d(x,y) < \delta$, $|f(x) - f(y)| < \epsilon$. Repeat this for each $k \in \mathbb{N}$ with $\epsilon_k = 2^{-k}\epsilon/3$ and let $C = \bigcap_{k \in \mathbb{N}} C_{\epsilon_k}$, so that $\mu(X \setminus C) < \epsilon$. Then f is continuous on C. □

Definition 3.3. Let μ, ν be measures on a set X. We say ν is *absolutely continuous* with respect to μ, written $\nu \ll \mu$, if for every $S \subset X$, $\mu(S) = 0 \implies \nu(S) = 0$. We say that ν is *singular* with respect to μ, written $\nu \perp \mu$ if there exists $A \subset X$ with $\mu(X \setminus A) = 0 = \nu(A)$.

Let μ be a measure on a set X and let $f \colon X \to \mathbb{R}^+$. The set valued function
$$\nu(S) = \int_S f \, \mathrm{d}\mu$$
defines a measure on X. Note also that $\nu \ll \mu$. The Radon-Nikodym theorem provides the converse to this statement.

Theorem 3.4 (Radon–Nikodym)**.** *Let μ, ν be finite measures on a set X such that $\nu \ll \mu$. There exists a μ-measurable $f \colon X \to \mathbb{R}^+$ such that*
$$\nu(S) = \int_S f \, \mathrm{d}\mu$$
for all μ-measurable $S \subset X$. The function f is called the Radon-Nikodym derivative *of ν with respect to μ.*

Proof. Let $\mathcal{F}$ be the set of all μ-measurable $f \colon X \to \mathbb{R}^+$ such that
$$\int_S f \, \mathrm{d}\mu \le \nu(S)$$
for all μ-measurable $S \subset X$. Note that $0 \in \mathcal{F}$ and
$$f, g \in \mathcal{F} \Rightarrow \max\{f, g\} \in \mathcal{F}. \tag{3.1}$$
Let
$$M = \sup\left\{\int f \, \mathrm{d}\mu : f \in \mathcal{F}\right\},$$
so that $0 \le M \le \nu(X) < \infty$, and let $f_i \in \mathcal{F}$ be such that
$$\int f_i \, \mathrm{d}\mu \to M.$$

Equation (3.1) implies that we may suppose the f_i monotonically increase. Let $f\colon X \to \mathbb{R}^+$ be the pointwise limit of the f_i. Then f is μ-measurable and, by the monotone convergence theorem, $f \in \mathcal{F}$ and $\int f \,\mathrm{d}\mu \geq M$. Thus

$$\int f \,\mathrm{d}\mu = M. \tag{3.2}$$

We claim that f satisfies the conclusion of the proposition. Indeed, suppose that $B \subset X$ is μ-measurable but

$$\nu(B) > \int_B f \,\mathrm{d}\mu$$

and let $\epsilon > 0$ be such that

$$\nu(B) > \int_B f + \epsilon \,\mathrm{d}\mu. \tag{3.3}$$

Let $\mathcal{S}$ be the collection of all μ-measurable $S \subset B$ such that

$$\nu(S) \leq \int_S f + \epsilon \,\mathrm{d}\mu.$$

We claim that there exists a μ-measurable $G \subset B$ of positive μ-measure such that each μ-measurable $G' \subset G$ of positive μ-measure is not contained in $\mathcal{S}$. Indeed, if not, then each $G \subset B$ of positive μ-measure contains an element of $\mathcal{S}$ of positive μ-measure. Thus Lemma 1.15 gives a countable disjoint decomposition of μ-almost all of B into elements of $\mathcal{S}$. Since $\nu \ll \mu$ this implies

$$\nu(B) \leq \int_B f + \epsilon \,\mathrm{d}\mu,$$

contradicting (3.3).

Note that $f + \epsilon G \in \mathcal{F}$. Indeed, if $S \subset X$ is μ-measurable,

$$\begin{aligned}\int_S f + \epsilon G \,\mathrm{d}\mu &= \int_{S\setminus G} f + \int_{S\cap G} f + \epsilon \,\mathrm{d}\mu \\ &\leq \nu(S \setminus G) + \nu(S \cap G) \\ &= \nu(S).\end{aligned}$$

On the other hand, since $\mu(G) > 0$,

$$\int f + \epsilon G \,\mathrm{d}\mu = M + \epsilon\mu(G) > M,$$

contradicting the definition of M. □

Theorem 3.5 (Lebesgue decomposition theorem)**.** *Let μ, ν be finite measures on a set X. There exists a ν-measurable $A \subset X$ with $\mu(X \setminus A) = 0$ such that, for all $S \subset A$, $\mu(S) = 0 \Rightarrow \nu(S) = 0$. That is, $\nu = \nu_{ac} + \nu_\perp$ with $\nu_{ac} \ll \mu$ and $\nu_\perp \perp \mu$.*

Proof. Let $\mathcal{S}$ be the set of all ν-measurable $S \subset X$ with $\mu(S) = 0$. By Lemma 1.15, there exists $S_i \in \mathcal{S}$ such that each $S \in \mathcal{S}$ with

$$S \subset A := X \setminus \bigcup_{i\in\mathbb{N}} S_i$$

satisfies $\nu(S) = 0$. Since $\mu(X \setminus A) = 0$, this is the required decomposition. □

3.1. **Exercises.**

Exercise 3.1. Let μ be a Borel measure on a metric space X, $f\colon X \to \mathbb{R}$ μ-integrable and $\epsilon > 0$.

(1) Show that there exists a simple function s such that
$$\int_X |f - s| \, \mathrm{d}\mu < \epsilon.$$
(2) If f is positive show that we may require $0 \leq s \leq f$ in the previous point.
(3) Show that if μ is finite, there exists $g \in C(X)$ with
$$\int_X |f - g| \, \mathrm{d}\mu < \epsilon.$$
(4) Show that the previous point may fail if μ is only σ-finite.

Exercise 3.2. Prove the following variant of Lusin's theorem for the case that μ is not finite but f is μ-integrable: for every $\epsilon > 0$ there exists a closed $C \subset X$ with
$$\int_{X \setminus C} |f| \, \mathrm{d}\mu < \epsilon$$
such that $f|_C$ is continuous.

Exercise 3.3. Let μ, ν be measures on a set X. Show that if $A \subset X$ is $\mu + \nu$ measurable then it is also μ-measurable.

Exercise 3.4. Prove the Radon-Nikodym and Lebesgue decomposition theorems for σ-finite measures.

Exercise 3.5. Let μ, ν be finite measures on a set X and suppose $\nu \ll \mu$. For any μ, ν-measurable $g\colon X \to \mathbb{R}^+$ show that
$$\int g \, \mathrm{d}\nu = \int g f \, \mathrm{d}\nu,$$
where f is the Radon-Nikodym derivative of ν with respect to μ.

Exercise 3.6. For a measure μ on a set X, we say a function $f\colon X \to \mathbb{R}^n$ is μ-measurable if each component of f is μ-measurable and define the integral of f component by component.

An n-dimensional *vector valued measure* is a function
$$\nu\colon \{A : A \subset X\} \to \mathbb{R}^n$$
for which there exists a measure μ on X and a μ-measurable $f\colon X \to \mathbb{S}^{n-1}$ such that
$$\nu(A) = \int_A f \, \mathrm{d}\mu$$
for each $A \subset X$.

(1) Show that if
$$\nu = \int f \, \mathrm{d}\mu = \int f' \, \mathrm{d}\mu'$$
are two representations of a vector valued measure then $\mu = \mu'$ (when restricted to the set of μ-measurable sets), and hence $f = f'$ μ-a.e. We denote this unique measure by $|\nu|$. It is called the *total variation* of ν.
(2) Show that the set of all vector valued measures form a normed vector space when equipped with
$$\|\nu\| = |\nu|(X).$$

(3) Show that this space is complete.

A *signed measure* is a 1-dimensional vector valued measure.

Exercise 3.7. Let Σ be a σ-algebra on X. The standard definition of a signed measure on Σ is a countably additive function

$$\mu\colon \Sigma \to \mathbb{R}.$$

The *Hahn decomposition theorem* states that there exist disjoint $P, N \in \Sigma$ with $X = P \cup N$ such that:

- For every $S \in \Sigma$ with $S \subset P$, $\mu(P) \geq 0$ and
- For every $S \in \Sigma$ with $S \subset N$, $\mu(P) \leq 0$.

That is, $\mu|_P$ and $-\mu|_N$ are (positive) measures.

Use the Lebesgue decomposition and Radon–Nikodym theorems to show that the two definitions of a signed measure agree.

4. The Riesz representation theorem for $C_0(X)$

Let's consider the set $C(X)$ of all real valued bounded continuous functions on a metric space X, equipped with the supremum norm. For any Borel measure μ on X,

$$T(f) := \int_X f \, \mathrm{d}\mu$$

defines an element of $C(X)'$. Moreover, it is *positive*: for all $f \in C(X)$ with $f \geq 0$, $T(f) \geq 0$.

In this section we prove a suitable converse to this statement and hence give a precise characterisation of measures in terms of the dual of a Banach space of continuous functions. However, $C(X)$, and hence its dual, is far too large for this purpose and we must carefully choose the correct space of functions. As will be shown in Exercise 4.1, the only possible choice is $C_0(X)$, the set of continuous $f\colon X \to \mathbb{R}$ such that, for each $\epsilon > 0$, there exists a compact $K \subset X$ with $|f(x)| < \epsilon$ for all $x \notin K$.

Theorem 4.1 (Riesz representation theorem)**.** *Let X be a locally compact metric space and $T \in C_0(X)'$. Then there exists a unique Radon measure μ such that*

$$T(f) = \int_X f \, \mathrm{d}\mu$$

for all $f \in C_0(X)$.

Intuitively, for $S \subset X$ we wish to define $\mu(S) = T(\chi_S)$. Of course, χ_S may not be continuous and so we approximate it by continuous functions. We begin with the following observation. Recall that $C_c(X)$ is the set of continuous functions on X with compact support.

Observation 4.2. *Let X be a metric space and $f \in C_0(X)$ with $f \geq 0$. For any $\epsilon > 0$ there exists $f^\epsilon \in C_c(X)$ with such that $0 \leq f^\epsilon \leq f$ and $\|f - f^\epsilon\|_\infty < \epsilon$.*

Proof. Fix $\epsilon > 0$. There exists a compact $K \subset X$ with $|f(x)| < \epsilon/2$ for all $x \notin K$. In particular,

$$C := \{x \in X : f(x) \geq \epsilon/2\}$$

is a closed subset of K, and hence is compact. Let

$$W = \{x \in X : f(x) \geq \epsilon\}.$$

Then

$$\theta(x) = \frac{\operatorname{dist}(x, X \setminus C)}{\operatorname{dist}(x, X \setminus C) + \operatorname{dist}(x, W)}$$

is continuous, $0 \leq \theta \leq 1$ and $\theta(x) = 1$ if $x \in W$ and $\theta(x) = 0$ if $x \notin C$. In particular, $\|f - \theta f\|_\infty < \epsilon$ and so $f^\epsilon := \theta f$ has the required properties. □

Definition 4.3. Let X be a metric space and $T\colon C_0(X) \to \mathbb{R}$. For an open set $U \subset X$ define

$$\mu_T(U) = \sup\{T(f) : 0 \leq f \leq \chi_U,\ f \in C_0(X)\}.$$

Since μ_T is monotonic, it may be extend to arbitrary $A \subset X$ by defining

$$\mu_T(A) = \inf\{\mu_T(U) : U \supset A \text{ open}\}.$$

We now show that μ_T satisfies the conclusions of the Riesz representation theorem.

Lemma 4.4. *For any metric space X and any $T \in C_0(X)'$, μ_T is a finite measure on X.*

Proof. Certainly μ_T satisfies $\mu(\emptyset) = 0$ and is monotonic. If $f \in C_0(X)$ with $0 \leq f \leq \chi_X$,

$$T(f) \leq \|T\|\|f\|_\infty \leq \|T\|,$$

and so μ_T is finite.

To see that μ is countably sub-additive, let $\epsilon > 0$, let $A_1, A_2, \ldots \subset X$ and for each $i \in \mathbb{N}$ let $U_i \supset A_i$ be open with

$$\mu_T(U_i) \leq \mu_T(A_i) + \epsilon 2^{-i}. \tag{4.1}$$

Also let $U := \bigcup_{i\in\mathbb{N}} U_i$ and $f \in C_0(X)$ with $0 \leq f \leq \chi_U$. By Observation 4.2, there exists a $f^\epsilon \in C_c(X)$ with $\|f - f^\epsilon\|_\infty < \epsilon$. In particular, there exists $N \in \mathbb{N}$ such that $\operatorname{spt} f^\epsilon \subset \bigcup_{i=1}^N U_i =: U'$.

For each $1 \leq i \leq N$ set $\eta_i = \operatorname{dist}(\cdot, X \setminus U_i)$, so that $\sum_{i=1}^N \eta_i$ is positive and finite everywhere in U' and equals zero outside U'. For each $i \in \mathbb{N}$ set

$$\tilde{\eta}_i = \eta_i / \sum_{i=1}^N \eta_i,$$

so that $\sum_{i=1}^N \tilde{\eta}_i = \chi_{U'}$ and hence $\sum_{i=1}^N \tilde{\eta}_i f^\epsilon = f^\epsilon$. Also, for every $1 \leq i \leq N$, $\tilde{\eta}_i f^\epsilon \in C_c(X)$ with $0 \leq \tilde{\eta}_i f^\epsilon \leq \chi_{U_i}$. Indeed, $\tilde{\eta}_i f^\epsilon$ is continuous at $x \in U'$ because locally it is the product of continuous functions. It is continuous at $x \notin U'$ because $f^\epsilon = 0$ on a neighbourhood of x.

Then

$$T(f) - \|T\|\epsilon \leq T(f^\epsilon) = T\left(f^\epsilon \sum_{i=1}^N \tilde{\eta}_i\right) = \sum_{i=1}^N T(f^\epsilon \tilde{\eta}_i) \leq \sum_{i=1}^N \mu_T(U_i).$$

Taking the supremum over all such f gives

$$\mu(U) \leq \sum_{i\in\mathbb{N}} \mu_T(U_i) + \|T\|\epsilon$$

and hence, since $A \subset U$,

$$\mu(A) \leq \mu(U) \leq \sum_{i\in\mathbb{N}} \mu_T(U_i) + \|T\|\epsilon \leq \sum_{i\in\mathbb{N}} \mu_T(A_i) + \epsilon + \|T\|\epsilon$$

by (4.1). Since $\epsilon > 0$ is arbitrary, μ is countably sub-additive. □

Lemma 4.5. *For any metric space X and any $T \in C_0(X)'$, μ_T is a finite Radon measure on X.*

Proof. We know that μ_T is a finite measure by Lemma 4.4. To see that μ_T is Borel, let $A, B \subset X$ with

$$\inf\{d(x,y) : x \in A,\ y \in B\} = \delta > 0.$$

Let $N_A = B(A, \delta/2)$ and $N_B = B(A, \delta/2)$. For any open $U \supset A \cup B$, let $V := U \cap N_A \supset A$ and $W := U \cap N_B \supset B$, open sets with

$$U \supset V \cup W.$$

Thus, since μ_T is monotonic,

$$\mu_T(U) \geq \mu_T(V \cup W). \tag{4.2}$$

Let $f, g \in C_0(X)$ with $0 \leq f \leq \chi_V$ and $0 \leq g \leq \chi_W$. Since V and W are disjoint, $0 \leq f + g \leq \chi_{V \cup W}$ and so

$$\mu_T(V) + \mu_T(W) = \sup_{0 \leq f \leq \chi_V} T(f) + \sup_{0 \leq g \leq \chi_W} T(g) = \sup_{f,g} T(f+g) \leq \mu_T(V \cup W).$$

Combining with (4.2) shows that $\mu_T(U) \geq \mu_T(V) + \mu_T(W)$. Thus Theorem 1.8 shows that μ_T is a Borel measure.

Since X is a metric space, Exercise 1.10 implies that μ_T is outer regular by open sets and inner regular by closed sets. Thus, to see that μ_T is Radon, we need to show that it is Borel regular and inner regular by compact sets.

By definition, $\mu_T(A) = \inf \mu_T(U)$ such that $U \supset A$ is open. Let $U_i \supset A$ be open such that $\mu_T(U_i) \to \mu_T(A)$ and for each $n \in \mathbb{N}$ let $V_i = \bigcap_{i<n} U_i \supset A$. Then the V_i are open, monotonically decrease and, by the monotonicity of μ_T, $\mu_T(V_n) \to \mu_T(A)$. In particular,

$$\mu_T\left(\bigcap_{n \in \mathbb{N}} V_n\right) = \mu_T(A),$$

so that μ_T is Borel regular.

Let $S \subset X$ be Borel and $\epsilon > 0$. By inner and outer regularity, there exist $U \supset S$ open and $C \subset S$ closed with $\mu(U \setminus C) < \epsilon$. By the definition of μ_T, there exists $f \in C_0(X)$ with $0 \leq f \leq \chi_U$ and

$$\mu_T(U) \leq T(f) + \epsilon.$$

By Observation 4.2, there exists $f^\epsilon \in C_c(X)$ such that $f^\epsilon \geq 0$ and $\|f - f^\epsilon\|_\infty < \epsilon$. In particular,

$$\mu_T(U) \leq T(f^\epsilon) + (1 + \|T\|)\epsilon.$$

Note that if $V \supset \operatorname{spt} f^\epsilon$ is open then

$$\mu_T(V) \geq T(f^\epsilon)$$

and so

$$\mu_T(\operatorname{spt} f^\epsilon) \geq \mu_T(U) - (1 + \|T\|)\epsilon.$$

In particular,

$$\mu_T(\operatorname{spt} f^\epsilon \cap C) \geq \mu_T(U) - (2 + \|T\|)\epsilon.$$

Since $\operatorname{spt} f^\epsilon \cap C$ is compact, this completes the proof. ☐

Lemma 4.6. *Let X be a metric space and let $T \in C_0(X)'$ be positive. Then for any $f \in C_0(X)$,*

$$T(f) = \int_X f \, \mathrm{d}\mu_T.$$

Proof. Fix $f \in C_0(X)$. It suffices to show that

$$T(f) \geq \int_X f \, \mathrm{d}\mu_T.$$

Fix $\epsilon > 0$. Since $f(X)$ is a compact interval, there exists a finite $I \subset \mathbb{Z}$ such that

$$X = \bigcup_{i \in \mathbb{N}} \{x \in X : i\epsilon \leq f(x) < (i+1)\epsilon\} =: X_i.$$

Since each X_i is a Borel set, there exists compact $K_i \subset X_i$ with $\mu_T(X_i \setminus K_i) < \epsilon/|I|$. The K_i are a finite family of disjoint compact sets and so are separated. Therefore there exist disjoint open $U_i \supset K_i$ with $\mu_T(U_i \setminus K_i) < \epsilon/|I|$. Moreover, by reducing the U_i if necessary, by the continuity of f,

$$(i-1)\epsilon < f(x) < (i+1)\epsilon \tag{4.3}$$

for each $x \in U_i$ and each $i \in I$.

For each $i \in I$ let $g_i \in C_0(X)$ satisfy $0 \leq h_i \leq \chi_{U_i}$ and

$$T(g_i) \geq \mu_T(U_i) - \epsilon/|I|.$$

Note that, by (4.3),

$$f \geq \sum_{i \in I} (i-1)\epsilon g_i.$$

Therefore, since T is positive,

$$\begin{aligned}
T(f) &\geq T\left(\sum_{i \in I} (i-1)\epsilon g_i\right) \\
&= \sum_{i \in I} (i-1)\epsilon T(g_i) \\
&\geq \sum_{i \in I} (i-1)\epsilon(\mu_T(U_i) - \epsilon/|I|) \\
&\geq \sum_{i \in I} (i-1)\epsilon \mu_T(U_i) - \epsilon\|f\|_\infty(\mu_T(X) + \epsilon) \\
&\geq \int_{\cup_{i \in I} U_i} (f - \epsilon) \, \mathrm{d}\mu_T - \epsilon\|f\|_\infty(\mu_T(X) + \epsilon) \\
&= \int_X f \, \mathrm{d}\mu_T - \int_{X \setminus \cup_{i \in I} U_i} f - \epsilon\|f\|_\infty(\mu_T(X) + \epsilon) \\
&\geq \int_X f \, \mathrm{d}\mu_T - \|f\|_\infty \mu_T(X \setminus \cup_{i \in I} K_i) - \epsilon\|f\|_\infty(\mu_T(X) + \epsilon) \\
&\geq \int_X f \, \mathrm{d}\mu_T - \epsilon\|f\|(\mu_T(X) + 1 + \epsilon).
\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, this completes the proof. ☐

Lemma 4.7. *Let X be a locally compact metric space. If μ, ν are Radon measures on X such that*

$$\int_X f \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\nu \tag{4.4}$$

for all $f \in C_0(X)$, then $\mu = \nu$.

Proof. Since μ and ν are both Radon measures on X, it suffices to prove that, for all compact $K \subset X$, $\mu(K) = \nu(K)$.

To this end, let $K \subset X$ be compact. Since X is locally compact, there exist a sequence of open sets $U_n \supset K$ each with compact closure such that U_n decreases to K. In particular,

$$\mu(U_n) \to \mu(K) \quad \text{and} \quad \nu(U_n) \to \nu(K). \tag{4.5}$$

For each $n \in \mathbb{N}$ let

$$f_n(x) = \frac{\operatorname{dist}(x, X \setminus U_n)}{\operatorname{dist}(x, X \setminus U_n) + \operatorname{dist}(x, K)},$$

so that $f_n \in C_c(X)$ and $\chi_K \leq f_n \leq \chi_{U_n}$. In particular,

$$\mu(K) \leq \int_X f_n \,\mathrm{d}\mu \leq \mu(U_n)$$

and

$$\nu(K) \leq \int_X f_n \,\mathrm{d}\nu \leq \nu(U_n)$$

for each $n \in N$. Combining this with (4.5) gives

$$\mu(K) = \lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu \quad \text{and} \quad \nu(K) = \lim_{n\to\infty} \int_X f_n \,\mathrm{d}\nu$$

and so (4.4) completes the proof. □

Proof of Theorem 4.1. This is simply given by combining Lemmas 4.4 to 4.7. □

The Riesz representation theorem allows us to identify the set of finite Radon measures on a locally compact space X with the set of positive elements of $C_0(X)'$. For B a Banach space, a sequence $T_n \in B'$ *weak* converges* to $T \in B'$ if $T_n(x) \to T(x)$ for every $x \in B$. In $C_0(X)'$, this translates to the following.

Definition 4.8. Let X be a locally compact metric space. We say that a sequence μ_n of finite Radon measures on X *weak* converges* to a finite Radon measure μ if, for every $f \in C_0(X)$,

$$\int_X f \,\mathrm{d}\mu_n \to \int_X f \,\mathrm{d}\mu.$$

By the Banach-Alaoglu theorem (see Exercise 4.4), the unit ball of B' is weak* compact. Since the weak* limit of a sequence of positive operators on $C_0(X)$ is positive, we have the following.

Theorem 4.9. *Let X be a locally compact metric space and μ_n a sequence of Radon measures with uniformly bounded total measures. There exists a finite Radon measure μ on X and subsequence μ_{n_k} that weak* converges to μ.*

4.1. **Exercises.**

Exercise 4.1. Let (X, d) be a metric space and suppose that $\mathcal{C}$ is a collection of continuous functions on X such that, for every $T \in \mathcal{C}'$ there exists a Borel measure μ on X such that $T = T_\mu$. We consider $\mathcal{C}$ equipped with the supremum norm.

For $\delta > 0$, suppose that

$$S = \{x_1, x_2, \ldots\} \subset X$$

is such that $d(x_i, x_j) > \delta$ for each $i \neq j$. Let c be the set of all $f\colon S \to \mathbb{R}$ such that $\lim_j f(x_j)$ exists and define $T\colon c \to \mathbb{R}$ by $T(f) = \lim_j f(x_j)$.

(1) Observe that T is linear, continuous on c and hence can be extended by the Hahn-Banach theorem to an element of $\mathcal{C}$. (Such an extension is called a *Banach limit.*)
(2) Show that any finite Borel measure on S is a convergent sum of Dirac masses.
(3) Hence show that there is no Borel measure μ on X such that $T_\mu = T$.
(4) Deduce that, for any $\epsilon > 0$ and $f \in \mathcal{C}$, the set
$$\{x \in X : |f(x)| > \epsilon\}$$
is totally bounded.

Any metric space isometrically embeds into its completion, so insisting that we work with complete metric spaces is not a loss of generality (and is of course very natural in analysis). In this case, (4) shows that we *must* work with $\mathcal{C} = C_0(X)$ if we wish to represent every element of $\mathcal{C}'$ as a measure.

Exercise 4.2. Let X be a metric space and suppose that $x \in X$ has no neighbourhood that is compact. Show that, for any $f \in C_0(X)$, $f(x) = 0$. Deduce that Lemma 4.7 is not true in X.

Exercise 4.3. Let X be a metric space and $T \in C_0(X)'$. For $f \geq 0$ define
$$S(f) = \sup_{0 \leq g \leq f} T(f).$$
and define $T^+ \in C_0(X)$ by $T^+(f) = S(f^+) - S(f^-)$ and $T^- = -(T - T^+)$ (recall **??**).

(1) Show that T^+, T^- belong to $C_0(X)'$ and are positive.
(2) Show that $\|T\| = \|T^+\| + \|T^-\|$.
(3) By (1), any $T \in C_0(X)$ can be identified with two measures μ^+ and μ^- and hence with a signed measure (recall Exercise 3.6). Show that this identification is an isometric isomorphism.

Exercise 4.4. Let B be a separable Banach space and $\mathcal{D}$ a countable dense subset of B. Suppose that $T_n \in B'$ satisfy $\|T_n\| \leq M$ for some $M > 0$.

(1) Show that there exists a subsequence T_{n_j} and a $T \in B'$ such that $T_{n_j}(d) \to T(d)$ for each $d \in \mathcal{D}$.
(2) Deduce that, for any $x \in B$, $T_{n_j}(x) \to T(x)$.

That is, closed and bounded subsets of B' are weak* compact.

5. Fubini's theorem

Definition 5.1. Let μ, ν be measures on sets X, Y respectively and let
$$\mathcal{R} = \{A \subset X : A\ \mu\text{-measurable}\} \otimes \{B \subset Y : B\ \nu\text{-measurable}\}.$$
The *product measure* $\mu \times \nu$ on $X \times Y$ is defined by
$$\mu \times \nu(S) = \inf \sum_{i \in \mathbb{N}} \mu(A_i)\nu(B_i),$$
where the infimum is taken over all countable collections of rectangles $A_i \times B_i \in \mathcal{R}$ with
$$S \subset \bigcup_{i \in \mathbb{N}} A_i \times B_i.$$
This is a measure, see Exercise 5.1.

Lemma 5.2. *Let X, Y be sets and μ, ν measures on X, Y respectively. Then $\mu \times \nu$ is equivalently defined by the formula*

$$\mu \times \nu(S) = \inf \sum_{i \in \mathbb{N}} \mu(A_i)\nu(B_i),$$

where the infimum is taken over all disjoint *countable collections of rectangles $A_i \times B_i \in \mathcal{R}$ with*

$$S \subset \bigcup_{i \in \mathbb{N}} A_i \times B_i.$$

Proof. If A, C are μ-measurable and B, D are ν-measurable then

$$\begin{aligned}(A \times B) \setminus (C \times D) &= [(A \setminus C) \times B] \cup [(A \cap C) \times (B \setminus D)] \\ &:= A_1 \times B_1 \cup A_2 \times B_2\end{aligned}$$

is a decomposition into disjoint rectangles. Since A, C and B, D are μ and ν measurable respectively,

$$\begin{aligned}\mu(A_1)\nu(B_1) + \mu(A_2)\nu(B_2) &= [\mu(A) - \mu(A \cap C)]\nu(B) + \mu(A \cap C)[\nu(B) - \nu(B \cap D)] \\ &= \mu(A)\nu(B) - \mu(A \cap C)\nu(B \cap D) \\ &\leq \mu(A)\nu(B).\end{aligned}$$

Thus the two formulae agree. $\square$

Theorem 5.3 (Fubini's theorem)**.** *Let X, Y be sets and μ, ν σ-finite measures on X, Y respectively.*

(1) *If A is μ-measurable and B is ν-measurable then $A \times B$ is $\mu \times \nu$-measurable and*

$$\mu \times \nu(A \times B) = \mu(A)\nu(B).$$

(2) *If S is $\mu \times \nu$-measurable then*

$$S^y := \{x \in X : (x, y) \in S\}$$

is μ-measurable for ν-a.e. $y \in Y$, $y \mapsto \mu(S^y)$ is ν-measurable;

$$S_x := \{y \in Y : (x, y) \in S\}$$

is ν-measurable for μ-a.e. $x \in X$, $x \mapsto \nu(S_x)$ is μ-measurable; and

$$\mu \times \nu(S) = \int_Y \mu(S^y)\, \mathrm{d}\nu(y) = \int_X \nu(S_x)\, \mathrm{d}\mu(x).$$

(3) *If $f \colon X \times Y \to \mathbb{R}^+$ is $\mu \times \nu$-measurable or $f \colon X \times Y \to \mathbb{R}$ is $\mu \times \nu$ integrable then*

$$\int_{X \times Y} f \, \mathrm{d}\mu \times \nu = \int_X \int_Y f \, \mathrm{d}\nu(y)\, \mathrm{d}\mu(x) = \int_Y \int_X f \, \mathrm{d}\mu(x)\, \mathrm{d}\nu(y).$$

Proof. Note that (3) follows from (2) by the monotone convergence theorem. We prove the theorem for finite μ, ν.

To begin, let

$$\mathcal{U} := \left\{ \bigcup_{i \in \mathbb{N}} R_i : R_i \in \mathcal{R} \text{ pairwise disjoint} \right\}$$

and, for $S \subset X \times Y$, define $\sigma_S \colon Y \to \mathbb{R}$ by

$$\sigma_S(y) = \mu(S^y).$$

Further, let

$$\mathcal{P} := \{S \subset X \times Y : \sigma_S \text{ is } \nu\text{-measurable}\}$$

and for any $S \in \mathcal{P}$ define

$$\rho(S) := \int_Y \sigma_S \,\mathrm{d}\nu.$$

Observe that σ_S and hence $\rho(S)$ are monotonic in S.

If $S = A \times B \in \mathcal{R}$ then $\sigma_S = \mu(A)\chi_B$ is ν-measurable and $\rho(S) = \mu(A)\nu(B)$. If $U \in \mathcal{U}$ with

$$U = \bigcup_{i \in \mathbb{N}} A_i \times B_i$$

a disjoint union, then

$$\sigma_U = \sum_{i \in \mathbb{N}} \mu(A_i)\chi_{B_i}$$

is a countable sum of ν-measurable functions and so $U \in \mathcal{P}$. Moreover,

$$\rho(U) = \int_Y \sigma_U \,\mathrm{d}\nu = \sum_{i \in \mathbb{N}} \mu(A_i)\nu(B_i). \tag{5.1}$$

Thus, for any $S \subset X \times Y$,

$$\mu \times \nu(S) = \inf\{\rho(U) : S \subset U \in \mathcal{U}\}. \tag{5.2}$$

To prove (1), let A be μ-measurable and B be ν-measurable. By definition, $\rho(A \times B) = \mu(A)\nu(B)$ and, since ρ is monotonic, $\rho(A \times B) \leq \rho(U)$ whenever $A \times B \subset U \in \mathcal{U}$. Thus, by (5.2),

$$\mu \times \nu(A \times B) = \mu(A)\nu(B).$$

Let $E \subset X \times Y$ and $E \subset U \in \mathcal{U}$. Observe

$$U \cap (A \times B) \quad \text{and} \quad U \setminus (A \times B)$$

are disjoint members of $\mathcal{U}$. Therefore, by (5.1) and (5.2),

$$\begin{aligned}
\rho(U) &= \rho(U \cap (A \times B)) + \rho(V \setminus (A \times B)) \\
&\geq \mu \times \nu(E \cap (A \times B)) + \mu \times \nu(E \setminus (A \times B)).
\end{aligned}$$

When taking the infimum over all such U, the left hand side converges to $\mu \times \nu(E)$, and so $A \times B$ is $\mu \times \nu$-measurable. This also implies that all elements of $\mathcal{U}$ are measurable.

Let $S \subset X \times Y$ and suppose that $U_1, U_2, \ldots \in \mathcal{U}$ are such that $\rho(U_i) \to \mu \times \nu(S)$. Since the intersection of any two rectangles is a rectangle,

$$V_i := U_i \cap U_{i-1} \cap \ldots U_1 \in \mathcal{U}$$

for each $i \in \mathbb{N}$. Moreover, $\rho(V_i)$ monotonically decreases to $\mu \times \nu(S)$. Let $W = \bigcap_{i \in \mathbb{N}} V_i \supset S$. Note that σ_{V_i} monotonically decreases to σ_W, so that $W \in \mathcal{P}$. Since μ, ν are finite, the monotone convergence theorem implies that $\rho(V_i) \to \rho(W)$ and hence $\mu \times \nu(S) = \rho(W)$. Since $\mu \times \nu(W) \geq \mu \times \nu(S)$, the monotonicity of ρ implies $\mu \times \nu(W) = \rho(W)$.

To prove (2), if S is $\mu \times \nu$-measurable, then $\mu \times \nu(W \setminus S) = 0$. By (5.2) there exists $Z \supset W \setminus U$ with $\rho(Z) = 0$. That is, $\mu(W^y) = \mu(S^y)$ for ν-a.e. $y \in Y$, and hence the first conclusion of (2). Moreover, $\rho(S) = \rho(W) = \mu \times \nu(S)$, which concludes the proof. $\square$

5.1. **Exercises.**

Exercise 5.1. Let X, Y be sets and μ, ν measures on X, Y respectively. Prove that $\mu \times \nu$ is a measure on $X \times Y$.

Exercise 5.2. Let X, Y be separable metric spaces. Show that

$$\mathcal{B}(X \times Y) = \Sigma(\mathcal{B}(X) \otimes \mathcal{B}(Y)).$$

Exercise 5.3. Let X, Y be separable metric spaces and μ, ν finite Borel measures on X, Y respectively. Show that $\mu \times \nu|_{\mathcal{B}(X \times Y)}$ is the unique countably additive set function on $\mathcal{B}(X \times Y)$ satisfying $\mu(A \times B) = \mu(A) \times \nu(B)$ for all $A \times B \in \mathcal{B}(X) \otimes \mathcal{B}(Y)$.

Exercise 5.4. Prove Fubini's theorem and Exercises 5.2 and 5.3 for σ-finite measures μ, ν.

Exercise 5.5. Show that Theorem 5.3 (3) may fail if

(1) $f\colon X \times Y \to \mathbb{R}$ is measurable but not integrable; Hint: consider μ, ν the counting measure on $\mathbb{N}$. Exploit the "identity"

$$(1-1) + (1-1) + \ldots = 1 + (-1+1) + (-1+1) + \ldots$$

(2) $f\colon X \times Y \to \mathbb{R}^+$ is measurable but μ is not σ-finite. Hint: consider $\mu = \mathcal{L}^1$, ν the counting measure on $\mathbb{R}$.

Prove Theorem 5.3 (3) for $f\colon X \times Y \to \mathbb{R}$ integrable, even if μ, ν are not σ-finite.

Exercise 5.6. Note in Theorem 5.3 (2) we must exclude a set of measure zero. Indeed, if $\mathcal{V}$ is a Vitali set, note that $\mathcal{V} \times \{0\} \subset \mathbb{R}^2$ is $\mathcal{L}^2$-measurable.

Exercise 5.7. Let X be a separable metric space and $f\colon X \to [0, \infty)$ a Borel function. Prove that

$$\int_X f \, \mathrm{d}\mu = \int_0^\infty \mu(\{x \in X : f(x) \geq t\}) \, \mathrm{d}t.$$

Hint: consider

$$A = \{(x, t) : f(x) \geq t\}.$$

Exercise 5.8. A measure μ on a metric space X is *uniformly distributed* if there exists a $g\colon (0, \infty) \to (0, \infty)$ such that $\nu(B(x, r)) = g(r)$ for all $x \in X$ and $r > 0$. Let μ, ν be uniformly distributed Borel regular measures on a separable metric space X (with functions g and h respectively). Let $U \subset X$ be open.

(1) Observe that, for any $x \in U$,

$$\lim_{r \to 0} \frac{\nu(U \cap B(x, r))}{h(r)} = 1$$

for every $x \in U$.

(2) Deduce that

$$\mu(U) \leq \liminf_{r \to 0} h(r)^{-1} \int_U \nu(U \cap B(x, r)) \, \mathrm{d}\mu(x).$$

(3) Deduce that

$$\mu(U) \leq \liminf_{r \to 0} h(r)^{-1} \int_U \mu(U \cap B(y, r)) \, \mathrm{d}\nu(y) = \liminf_{r \to 0} \frac{g(r)}{h(r)} \nu(U).$$

Deduce that $\mu = c\nu$ for some $c > 0$.

6. Covering theorems

We will use $B(x,r)$ to denote the closed ball in a metric space X centred at $x \in X$ with radius $r \geq 0$. Since the centre and radius of a ball are not uniquely defined by its elements, formally by a "ball" we mean a pair $(x,r) \in X \times (0,\infty)$, but in practice we mean the set of its elements.

Lemma 6.1 (Vitali covering lemma)**.** *Let X be a metric space and $\mathcal{B}$ an arbitrary collection of closed balls of uniformly bounded radii. There exists a disjoint sub-collection $\mathcal{B}' \subset \mathcal{B}$ such that any $B \in \mathcal{B}$ intersects a ball $B' \in \mathcal{B}'$ with*

$$\operatorname{rad} B' \geq \operatorname{rad} B/2.$$

In particular,

$$\bigcup_{B \in \mathcal{B}'} 5B \supset \bigcup_{B \in \mathcal{B}} B.$$

Here, $5B$ denotes the ball with the same centre as B and 5 times the radius.

Proof. For each $n \in \mathbb{Z}$ let

$$\mathcal{B}_n = \{B \in \mathcal{B} : 2^n \leq \operatorname{rad} B < 2^{n+1}\}.$$

Since the balls in $\mathcal{B}$ have uniformly bounded radii, the $\mathcal{B}_n$ are empty for all $n > N$, for some $N \in \mathbb{N}$. Let $\mathcal{B}'_N$ be a maximal disjoint sub-collection of $\mathcal{B}_N$. That is, the elements of $\mathcal{B}'_N$ are disjoint elements of $\mathcal{B}_N$ and if $B \in \mathcal{B}_N$, there exists a $B' \in \mathcal{B}'_N$ with $B \cap B' \neq \emptyset$. (In general such a maximal collection exists by Zorn's lemma. See also Exercise 6.1.) Let $\mathcal{B}'_{N-1}$ be a maximal collection such that $\mathcal{B}'_N \cup \mathcal{B}'_{N-1}$ is a disjoint collection. Repeat this for each $i \in \mathbb{N}$, obtaining a maximal collection $\mathcal{B}'_{N-i}$ such that $\mathcal{B}'_N \cup \ldots \cup \mathcal{B}'_{N-i}$ is a disjoint collection, and set $\mathcal{B}' = \bigcup_{n \leq N} \mathcal{B}'_n$.

Now suppose that $B \in \mathcal{B}$, say $B \in \mathcal{B}_n$. Then by construction there exists $B' \in \mathcal{B}'_m$ for some $m \geq n$ with $B \cap B' \neq \emptyset$. In particular, $\operatorname{rad} B' \geq \operatorname{rad} B/2$.

The final statement of the lemma follows from the triangle inequality. □

Definition 6.2. Let X be a metric space and $S \subset X$. A *Vitali cover* of S is a collection $\mathcal{B}$ of closed balls such that, for each $x \in S$ and each $\epsilon > 0$, there exists a ball $B \in \mathcal{B}$ with $\operatorname{rad} B < \epsilon$ and $x \in B$.

Proposition 6.3. *Let X be a metric space, $S \subset X$ and suppose that $\mathcal{B}$ is a Vitali cover of S. Then there exists a disjoint $\mathcal{B}' \subset \mathcal{B}$ such that, for every finite $I \subset \mathcal{B}'$,*

$$S \setminus \bigcup_{B \in I} B \subset \bigcup_{B \in \mathcal{B}' \setminus I} 5B.$$

In particular, if $\mathcal{B}' = \{B_1, B_2, \ldots\}$ is countable (for example, if X is separable), then

$$S \setminus \bigcup_{i=1}^{n} B_i \subset \bigcup_{i>n} 5B_i$$

for each $n \in \mathbb{N}$.

Proof. Note that we may suppose $\mathcal{B}$ consists of balls with uniformly bounded radii. Let $\mathcal{B}'$ be a disjoint sub-collection of $\mathcal{B}$ obtained from Lemma 6.1. If $I \subset \mathcal{B}'$ is finite then

$$C := \bigcup_{B \in I} B$$

is closed. Therefore, if $x \in S \setminus C$, since $\mathcal{B}$ is a Vitali cover of S, there exists $B \in \mathcal{B}$ with $x \in B$ such that $B \cap C = \emptyset$. However, B must intersect some $B' \in \mathcal{B}'$ with $\operatorname{rad} B' \geq \operatorname{rad} B/2$, and so $x \in 5B'$. That is, x belongs to

$$\bigcup_{B \in \mathcal{B}' \setminus I} 5B,$$

as required. □

Definition 6.4. A Borel measure μ on a metric space X is a *doubling measure* if there exists a $C_\mu \geq 1$ such that

$$0 < \mu(2B) \leq C_\mu \mu(B) < \infty$$

for all balls $B \subset X$.

Remark 6.5. Note that, for any $m \geq 2$,

$$\mu(mB) \leq C_\mu^{\log_2 m} \mu(B).$$

Lebesgue measure is a doubling measure.

Theorem 6.6 (Vitali covering theorem)**.** *Let μ be a doubling measure on a metric space X and let $\mathcal{B}$ be a Vitali cover of a set $S \subset X$. There exists a countable disjoint $\mathcal{B}' \subset \mathcal{B}$ such that*

$$\mu\left(S \setminus \bigcup_{B \in \mathcal{B}'} B\right) = 0.$$

Proof. First note that it suffices to prove the result for S bounded, say S is contained in some ball $\tilde{B}$. We may also suppose that each $B \in \mathcal{B}$ is a subset of $2\tilde{B}$.

Let $\mathcal{B}'$ be a disjoint sub-collection of $\mathcal{B}$ obtained from Proposition 6.3. Note that $\mathcal{B}'$ is countable. Indeed, for each $m \in \mathbb{N}$, at most $m\mu(2\tilde{B})$ balls $B \in \mathcal{B}'$ can satisfy $\mu(B) > 1/m$.

Enumerate $\mathcal{B}' = \{B_1, B_2, \ldots\}$. Since the B_i are disjoint subsets of $2\tilde{B}$,

$$\sum_{i>n} \mu(B_i) \to 0.$$

By the conclusion of Proposition 6.3,

$$S \setminus \bigcup_{i=1}^{n} B_i \subset \bigcup_{i>n} 5B_i$$

for each $n \in \mathbb{N}$. Since μ is doubling, $\mu(5B_i) \leq C\mu(B_i)$ for each $i \in \mathbb{N}$ and so

$$\mu\left(S \setminus \bigcup_{i=1}^{n} B_i\right) \leq C \sum_{i>n} \mu(B_i) \to 0,$$

as required. □

Definition 6.7. Let μ be a Borel measure on a metric space X and $f \in L^1(\mu)$. Suppose that $0 < \mu(B(x,r)) < \infty$ for all $x \in X$ and all $r > 0$.

Define the *Hardy–Littlewood maximal function* of f by

$$Mf(x) = \sup_{r>0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f| \, \mathrm{d}\mu.$$

By Exercise 6.3, the maximal function is a Borel function.

Theorem 6.8 (Hardly–Littlewood maximal inequality)**.** *Let μ be a doubling measure on a metric space X. There exists a $C > 0$ such that, for any $f \in \mathcal{L}^1(\mu)$ and $\lambda > 0$,*

$$\mu(\{x : Mf(x) > \lambda\}) \leq \frac{C}{\lambda}\|f\|_1.$$

Proof. For $\lambda > 0$, let

$$S = \{x \in X : Mf(x) > \lambda\}$$

and let $\mathcal{B}$ be the collection of balls $B(x, r)$ with $x \in S$ and

$$\int_{B(x,r)} |f| \,\mathrm{d}\mu > \lambda\mu(B(x, r)).$$

Let $\mathcal{B}'$ satisfy the conclusion of Lemma 6.1. Then

$$\begin{aligned}
\int_X |f| \,\mathrm{d}\mu &\geq \sum_{B\in\mathcal{B}'} \int_B |f| \,\mathrm{d}\mu \\
&> \sum_{B\in\mathcal{B}'} \lambda\mu(B) \\
&\geq \frac{1}{C_\mu^{\log_2 5}} \sum_{B\in\mathcal{B}'} \lambda\mu(5B) \\
&\geq \frac{\lambda}{C}\mu(S).
\end{aligned}$$

□

Theorem 6.9 (Lebesgue differentiation theorem)**.** *Let μ be a doubling measure on a metric space X and $f \in L^1(\mu)$. For μ-a.e. $x \in X$,*

$$\frac{1}{\mu(B(x, r))} \int_{B(x,r)} |f - f(x)| \,\mathrm{d}\mu \to 0$$

as $r \to 0$. Such an x is called a Lebesgue point *of f.*

Proof. First note that the theorem is true if f is continuous.

Fix $\epsilon > 0$ and let $g \colon X \to \mathbb{R}$ be continuous with $\|f - g\|_1 < \epsilon$ (such a g exists by **??**). Let

$$B = \{x \in X : |f(x) - g(x)| \geq \sqrt{\epsilon}\},$$

so that $\mu(B) < \sqrt{\epsilon}$.

If

$$S = \{x : M(f - g) > \sqrt{\epsilon}\},$$

then by Theorem 6.8,

$$\mu(S) \leq \frac{C}{\sqrt{\epsilon}}\|f - g\|_1 < C\sqrt{\epsilon}.$$

Moreover, if $x \notin S$,

$$\frac{1}{\mu(B(x, r))} \int_{B(x,r)} |f - g| \,\mathrm{d}\mu \leq \sqrt{\epsilon}$$

for all $r > 0$. In particular, since g is continuous at x,

$$\limsup_{r\to 0} \frac{1}{\mu(B(x, r))} \int_{B(x,r)} |f - g(x)| \,\mathrm{d}\mu \leq \sqrt{\epsilon}.$$

Therefore, if $x \notin B$,

$$\limsup_{r\to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f - f(x)| \, \mathrm{d}\mu \leq 2\sqrt{\epsilon}.$$

We are now done; repeat the above for a countable collection of $\epsilon \to 0$. The corresponding $B \cup S$ monotonically decrease to a set of measure zero. The set of $x \in X$ that does not belong to infinitely many of the $B \cup S$ has full measure, and for such an x,

$$\lim_{r\to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f - f(x)| \, \mathrm{d}\mu = 0.$$

□

Corollary 6.10. *Let μ be a doubling measure on a metric space X and let $S \subset X$ be μ-measurable with $\mu(S) < \infty$. Then*

$$\lim_{r\to 0} \frac{\mu(S \cap B(x,r))}{\mu(B(x,r))}$$

equals 1 for μ-a.e. $x \in S$ and 0 for μ-a.e. $x \notin S$. Such an x for which the limit equals 1 is called a density point *of S.*

6.1. **Exercises.**

Exercise 6.1. Let X be a separable metric space. Show that for any collection of balls, there exists a maximal disjoint sub-collection.

Exercise 6.2. Show that the $5r$ covering Lemma may not be true if the radii are not uniformly bounded.

Exercise 6.3. Let μ be a finite Borel measure on a metric space (X, d) and let $x_n \to x \in X$ such that $d(x, x_n)$ is a decreasing sequence. Let $U(y, r)$ denote the open ball centred on y with radius r.

(1) Show that, for any $r > 0$,
$$U(x, r) \setminus U(x_n, r)$$
decreases to the empty set.
(2) Deduce that $y \mapsto \mu(U(y, r))$ is lower semi-continuous.
(3) Give an example to show that $y \mapsto \mu(U(y, r))$ may not be continuous.
(4) Show that, for any $y \in X$,
$$\mu(B(y, r)) = \lim_{\mathbb{Q}\ni q \downarrow r} \mu(U(y, q)).$$
(5) Deduce that, for any $r > 0$, $y \mapsto \mu(B(y, r))$ is a Borel function.
(6) Show that the Hardy–Littlewood maximal function is equivalently defined by taking the supremum over all *rational* $r > 0$.
(7) Deduce that the Hardy–Littlewood maximal function is a Borel function.

Exercise 6.4. Let μ be a doubling measure on a metric space X and let $S \subset X$ with $\mu(S) < \infty$. Suppose that there exists a μ-measurable $S' \supset S$ with $\mu(S') = \mu(S)$. Show that

$$\lim_{r\to 0} \frac{\mu(S \cap B(x,r))}{\mu(B(x,r))} = 1$$

for μ-a.e. $x \in S$.

Exercise 6.5. Prove that on $\mathbb{R}^n$, $\mathcal{L}^n = c\mathcal{H}^n$ for some $c > 0$. There are two ways to prove this (recall Exercise 1.5 items 1 and 2).

Let μ, ν be two finite Borel measures on a set X with $\mu \ll \nu$ and let $c > 0$. Suppose that ν is doubling and that

$$\frac{\mu(B(x,r))}{\nu(B(x,r))} \to c$$

as $r \to 0$ for ν-a.e. $x \in X$.

By either combining the Radon-Nikodym theorem and the Lebesgue differentiation theorem, or by using the Vitali covering theorem, to show that $\mu = c\nu$.

7. Differentiability of Lipschitz functions

The regularity of a Lipschitz function $f\colon \mathbb{R}^n \to \mathbb{R}$ is very interesting. Of course, Lipschitz functions are continuous, but they may not be differentiable everywhere. However, it is quite easy to convince yourself that they cannot be non-differentiable on quite a large set. The question to quantify how large the non-differentiability set of a Lipschitz function can be was one of the motivating questions of Lebesgue's development of measure theory.

Definition 7.1. Let $f\colon [a,b] \to \mathbb{R}$. The *total variation* of f, $Vf\colon [a,b] \to [0,\infty]$ is defined by

$$Vf(x) = \sup \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|$$

where the supremum ranges over all $a = t_0 < t_1 < \ldots < t_n = b$.

If $Vf(b) < \infty$, f is said to have *bounded variation (BV)*.

Definition 7.2. A function $f\colon [a,b] \to \mathbb{R}$ is *absolutely continuous (AC)* if for any $\epsilon > 0$ there exists a $\delta > 0$ such that, for any intervals $I_1, I_2 \ldots \subset [a,b]$ with $\sum_i \mathcal{L}^1(I_i) < \delta$, we have $\sum_i \mathcal{L}^1(f(I_i)) < \epsilon$.

Note that AC functions are BV, and Lipschitz functions are AC (see Exercise 7.1). Also, if f is BV then Vf and $Vf - f$ are non-decreasing. If f is AC then so are Vf and $Vf - f$, see Exercise 7.2.

Theorem 7.3 (Lebesgue)**.** *Let $f\colon [a,b] \to \mathbb{R}$ be absolutely continuous. Then f is differentiable L^1 almost everywhere. Moreover, for any $x > y \in [a,b]$,*

$$f(x) - f(y) = \int_y^x f' \,\mathrm{d}x.$$

Proof. By Exercise 7.2 it suffices to assume that f is non-decreasing. In this case define a measure μ on $[a,b]$ using the Carathéodory construction with F the set of compact intervals and $\zeta([c,d]) = f(d) - f(c)$. This defines a finite Borel measure such that $\mu([c,d]) = f(d) - f(c)$ for all intervals $[c,d] \subset [a,b]$.

Note that $\mu \ll \mathcal{L}^1$. Indeed, given $\epsilon > 0$, let $\delta > 0$ be given by the definition of f being absolutely continuous. If $\mathcal{L}^1(N) = 0$, we may cover N by countably many closed intervals I_i such that $\mathcal{L}^1(\cup_i I_i) < \delta$. In particular $\sum_i f(I_i) < \epsilon$ and hence $\mu(N) < \epsilon$. Therefore,

$$\mu = \int \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^1} \,\mathrm{d}\mathcal{L}^1,$$

with $\mathrm{d}\mu/\,\mathrm{d}\mathcal{L}^1 \in L^1(\mathcal{L}^1)$.

By the Lebesgue differentiation theorem, for any Lebesgue point x of $\mathrm{d}\mu/\mathrm{d}\mathcal{L}^1$,

$$\begin{aligned}\lim_{t\to 0}\frac{f(x+t)-f(x)}{t} &= \lim_{t\to 0}\frac{\mu([x+t,x])}{t}\\ &= \lim_{t\to 0}\frac{1}{t}\int_x^{x+t}\frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^1}\,\mathrm{d}\mathcal{L}^1\\ &= \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^1}(x).\end{aligned}$$

□

Theorem 7.4 (Rademacher)**.** *Any Lipschitz $f\colon \mathbb{R}^n\to\mathbb{R}$ is differentiable $\mathcal{L}^n$ almost everywhere.*

Proof. For notational simplicity, we prove the case $n=2$.

For each $y\in\mathbb{R}$, $x\mapsto f(x,y)$ is a Lipschitz function $\mathbb{R}\to\mathbb{R}$ and so is differentiable $\mathcal{L}^1$-a.e. That is, for every y, $\partial_1 f(x,y)$ exists for $\mathcal{L}^1$-a.e. x. By Fubini's theorem, $\partial_1 f$ exists $\mathcal{L}^2$-a.e. Similarly, $\partial_2 f$ exists almost everywhere too.

Fix $\epsilon>0$. For $D\in\mathbb{Q}^2$ and $j\in\mathbb{N}$ let

$$X_{D,j}=\{x : |f(x+he_i)-f(x)-D_ih|<\epsilon|h|,\ \forall 0<|h|<1/j,\ i=1,2\}.$$

These are Borel sets. Further, for $D\in\mathbb{Q}^2$, if

$$|\partial_1 f(x)-D_1|<\epsilon/2 \quad\text{and}\quad |\partial_2 f(x)-D_2|<\epsilon/2,$$

then $x\in X_{D,j}$ for sufficiently large j. That is,

$$X^\epsilon=\bigcup_{D\in\mathbb{Q}}\bigcup_{j\in\mathbb{N}}X_{D,j}$$

is a set of full measure.

Fix $D\in\mathbb{Q}^2$ and $j\in\mathbb{N}$. Let x be a density point of $X_{D,j}$. Let $R>0$ such that

$$\mathcal{L}^n(B(x,r)\cap X_{D,j})\geq(1-\epsilon^n)\mathcal{L}^n(B(x,r))$$

for all $0<r<R$. In particular, for every $y\in B(x,r)$ there exists $y'\in X_{D,j}$ with

$$\|y-y'\|<\epsilon\|y-x\|.$$

Now let $r<\min\{R,1/j\}$ and $\|x-y\|<r$. Set $h=y-x$, $\tilde{y}=x+\pi_1 y$ and $\tilde{\tilde{y}}\in X_{D,j}$ with

$$\|\tilde{y}-\tilde{\tilde{y}}\|<\epsilon\|x-\tilde{y}\|\leq\epsilon\|x-y\|.$$

Also let y',y'' lie on the same vertical line as $\tilde{\tilde{y}}$ such that $y',\tilde{y}$ have the same vertical component as do $y'',\tilde{y}$. Then, since $x\in X_{D,j}$,

$$|f(\tilde{y})-f(x)-D_1h_1|<\epsilon|h_1|=\epsilon\|x-\tilde{y}\|\leq\epsilon\|x-y\|; \tag{7.1}$$

Since f is Lipschitz,

$$|f(\tilde{y})-f(y')|\leq L\|\tilde{y}-y'\|\leq L\epsilon\|x-y\|; \tag{7.2}$$

Since $\tilde{\tilde{y}}\in X_{D,j}$,

$$|f(y'')-f(y')-D_2h_2|\leq\epsilon\|y'-y''\|\leq\epsilon\|x-y\|; \tag{7.3}$$

Since f is Lipschitz,

$$|f(y'')-f(y)|\leq L\|y''-y\|=L\|y'-\tilde{y}\|\leq L\|\tilde{\tilde{y}}-\tilde{y}\|\leq\epsilon L\|x-y\|. \tag{7.4}$$

By combining eqs. (7.1) to (7.4),

$$|f(y)-f(x)-D\cdot h|\leq 2(1+L)\epsilon\|x-y\|.$$

This is true for all y with $\|x - y\| < r$ and for any density point x of the full measure set X^ϵ. That is, for $\mathcal{L}^n$-a.e x. Taking a countable intersection over $\epsilon \to 0$ concludes the proof. $\square$

7.1. **Exercises.**

Exercise 7.1. Prove that Lipschitz functions are AC and that AC functions are BV.

Exercise 7.2. Let $f\colon [a,b] \to \mathbb{R}$ be BV. Show that Vf and $Vf - f$ are non-decreasing. If f is AC then show that Vf and $Vf - f$ are AC.

Exercise 7.3. Show that any monotonic $f\colon \mathbb{R} \to \mathbb{R}$ is continuous except at countably many points.

Exercise 7.4. In this exercise we will show that monotonic functions are differentiable almost everywhere.

Let $f\colon [a,b] \to \mathbb{R}$ be non-decreasing. For each $x \in (a,b)$ let

$$\underline{D}f(x) = \liminf_{h\to 0} \frac{f(x+h) - f(x)}{h} \quad \overline{D}f(x) = \limsup_{h\to 0} \frac{f(x+h) - f(x)}{h}.$$

Observe that the set of $x \in (a,b)$ where f is not differentiable at x is the countable union, over $p < q \in \mathbb{Q}$, of the sets

$$B_{p,q} := \{x \in (a,b) : \underline{D}f(x) < p < q < \overline{D}f(x)\}.$$

We now fix $p < q \in \mathbb{Q}$.

(1) Let
$$\mathcal{B} = \{[x, x+h] : f(x+h) - f(x) < ph\}.$$
Note that $\mathcal{B}$ satisfies the hypotheses of the Vitali covering theorem (recall Theorem 6.6). Let $\mathcal{B}'$ be a disjoint sub-cover obtained from the Vitali covering theorem with respect to Lebesgue measure and let $S = \cup\mathcal{B}'$. Prove that
$$\mathcal{L}^1(f(B_{p,q} \cap S)) \le p\mathcal{L}^1(B_{p,q} \cap S).$$
Note: this is the step where we require f to be monotonic.
(2) Similarly, prove that $\mathcal{L}^1(f(B_{p,q} \cap S)) \ge q\mathcal{L}^1(B_{p,q} \cap S)$.
(3) Deduce that f is differentiable almost everywhere.
(4) Deduce that a BV function is differentiable almost everywhere.

However, BV functions do not satisfy the fundamental theorem of calculus:

Exercise 7.5. Recall the definition of the Cantor set from Exercise 1.6. Define the *Cantor function* $f\colon [0,1] \to [0,1]$ as follows. For each $n \in \mathbb{N}$, define $f_n\colon [0,1] \to [0,1]$ by

$$f(x) = \left(\frac{3}{2}\right)^n \mathcal{L}^1([0,x] \cap C_n).$$

Show that the f_n converge uniformly on $[0,1]$ to a monotonic, continuous function f. For each $x \in [0,1] \setminus C$, show that $f'(x) = 0$.

Thus, f is monotonic and hence BV, has derivative 0 almost everywhere, but does not satisfy the fundamental theorem of calculus.

Exercise 7.6. In lectures we proved that the derivative of any AC function is an absolutely continuous measure. Prove the converse: for any finite, absolutely continuous measure μ on $[0,\infty)$, show that

$$f(x) := \int_0^x \frac{\mathrm{d}\mu}{\mathrm{d}\mathcal{L}^1}\,\mathrm{d}\mathcal{L}^1 = \mu([0,x])$$

defines an absolutely continuous function.

Up to now, we have considered points where functions *are* differentiable. We now consider points of non-differentiability (which are much more interesting).

Exercise 7.7. Show that the Cantor function is not differentiable at any point of the Cantor set.

Exercise 7.8. Let $N \subset [0,1]$ satisfy $\mathcal{L}^1(N) = 0$.

(1) For each $n \in \mathbb{N}$, iteratively construct a countable collection of open intervals $\mathcal{O}_n$ such that, for each $n \in \mathbb{N}$,
   - N is contained in the union of $\mathcal{O}_n$;
   - for every $I \in \mathcal{O}_n$ there exists $J \in \mathcal{O}_{n-1}$ with $I \subset J$;
   - for each $I \in \mathcal{O}_{n-1}$,
$$\mathcal{L}^1(I \cap \cup\{J : J \in \mathcal{O}_n\}) < 2^{-n}|I|.$$

(2) Let
$$S = \bigcap_{n\in\mathbb{N}} \bigcup_{m>n} \cup\{J : J \in \mathcal{O}_m\},$$
the "limsup" of the $\mathcal{O}_n$ (S is the set of points that are contained in infinitely many intervals from the $\mathcal{O}_n$). In particular, $S \supset N$.

For each $x \in [0,1] \setminus S$, let $N(x)$ be the largest n for which there exists $I \in \mathcal{O}_n$ with $x \in I$. Define $P(x) = 1$ if $N(x)$ is even, $P(x) = 0$ otherwise. Finally, for each $x \in [0,1]$ define
$$f(x) = \mathcal{L}^1(\{t \in [0,x] : P(t) = 1\}).$$

Show that f is Lipschitz, monotonic, and not differentiable at any point of N. Hint: show that $\underline{D}f(x) = 0$ and $\overline{D}f(x) = 1$ for each $x \in N$.

Part 2. Some topics in Geometric Measure Theory

8. Hausdorff measure and densities

Recall the definition of Hausdorff measure from Exercise 1.5.

We are interested in the measure $\mathcal{H}^s|_S$, for $S \subset X$ some $\mathcal{H}^s$-measurable set with $\mathcal{H}^s(S) < \infty$. In particular, we require some counterpart to the Lebesgue density theorem, but, of course, $\mathcal{H}^s|_X$ may not be locally finite.

Definition 8.1. Let X be a metric space, $A \subset X$ and $s \geq 0$. The upper and lower *Hausdorff densities* of A are

$$\Theta^{*,s}(A,x) = \limsup_{r\to 0} \frac{\mathcal{H}^s(A \cap B(x,r))}{(2r)^s}$$

and

$$\Theta^s_*(A,x) = \liminf_{r\to 0} \frac{\mathcal{H}^s(A \cap B(x,r))}{(2r)^s}.$$

Lemma 8.2. *Let X be a metric space, $s \geq 0$ and $A \subset X$ with $\mathcal{H}^s(A) < \infty$. Then*
$$2^{-s} \leq \Theta^{*,s}(A,x) \leq 1$$
for $\mathcal{H}^s$-a.e. $x \in A$.

Proof. The set of $x \in A$ with $\Theta^{*,s}(A,x) < 2^{-s}$ is a countable union countable of the sets
$$S_\delta := \{x \in A : \mathcal{H}^s(A \cap B(x,r)) < (1-\delta)r^s \ \forall 0 < r < \delta\}.$$
Thus, for the first inequality, it suffices to show that $\mathcal{H}^s(S_\delta) = 0$ for all $\delta > 0$.

Fix $\delta, \epsilon > 0$. We may cover S_δ by sets $E_1, E_2, \ldots$ such that, for each $i \in \mathbb{N}$, $\operatorname{diam} E_i < \epsilon$, $S_\delta \cap E_i \neq \emptyset$ and
$$\sum_{i \in \mathbb{N}} \operatorname{diam} E_i^s \leq \mathcal{H}^s(S_\delta) + \epsilon.$$
For each $i \in \mathbb{N}$ let $x_i \in S_\delta \cap E_i$ and set $r_i = \operatorname{diam} E_i$. Then
$$\begin{aligned}\mathcal{H}^s(S_\delta) \leq \sum_{i\in\mathbb{N}} \mathcal{H}^s(S_\delta \cap E_i) \leq \sum_{i\in\mathbb{N}} \mathcal{H}^s(A \cap B(x_i, r_i))& \\ \leq (1-\delta)\sum_{i\in\mathbb{N}} \operatorname{diam} E_i^s \leq (1-\delta)(\mathcal{H}^s(S_\delta)+\epsilon)&.\end{aligned}$$
Since $\epsilon > 0$ is arbitrary and $\delta > 0$, this implies $\mathcal{H}^s(S_\delta) = 0$, as required.

For the second inequality, since $\mathcal{H}^s$ is Borel regular (see Exercise 8.2), it suffices to assume that A is Borel. As before, given $\delta > 0$, it suffices to prove that
$$S := \{x \in A : \Theta^{*,s}(A,x) > 1+\delta\}$$
satisfies $\mathcal{H}^s(S) = 0$. Fix $\epsilon > 0$ and let $U \supset S$ be open with
$$\mathcal{H}^s(A \cap U) \leq \mathcal{H}^s(S) + \epsilon$$
(which exists by the outer regularity of the measure $\mathcal{H}^s|_A$). Let $\mathcal{B}_\epsilon$ be the collection of balls B centred at a point of S with $\operatorname{rad} B < \epsilon$ such that $B \subset U$ and
$$\mathcal{H}^s(A \cap B) > (1+\delta)(2\operatorname{rad} B)^s. \tag{8.1}$$
This is a Vitali cover of S. Let $\mathcal{B}'_\epsilon$ be obtained from Proposition 6.3.

Since $\mathcal{H}^s(S) < \infty$, S is separable (see Exercise 8.4) and so $\mathcal{B}'_\epsilon = \{B_1, B_2, \ldots\}$ is countable and the conclusion of Proposition 6.3 states that
$$S \setminus \bigcup_{i\in\mathbb{N}} B_i \subset \bigcup_{i>n} 5B_i$$
for each $n \in \mathbb{N}$. Since $\operatorname{diam} B_i < \epsilon$ for each $i \in \mathbb{N}$, the B_i and $5B_i$ may be used to estimate $\mathcal{H}^s_{10\epsilon}(S)$. For each $n \in \mathbb{N}$ we obtain
$$\begin{aligned}\mathcal{H}^s_{10\epsilon}(S) &\leq \sum_{i\in\mathbb{N}} (2\operatorname{rad} B_i)^s + \sum_{i>n} (10 \operatorname{rad} B_i)^s \\ &\leq \sum_{i\in\mathbb{N}} \frac{\mathcal{H}^s(A\cap B_i)}{1+\delta} + 5^s \sum_{i>n} \frac{\mathcal{H}^s(A \cap B_i)}{1+\delta}\end{aligned}$$
where the second inequality follows by (8.1). Since the B_i are disjoint and $\mathcal{H}^s(A) < \infty$, the second term converges to 0 as $n \to \infty$. Since the B_i are subsets of U we obtain
$$\mathcal{H}^s_{10\epsilon}(S) \leq \frac{\mathcal{H}^s(A \cap U)}{1+\delta} \leq \frac{\mathcal{H}^s(S)+\epsilon}{1+\delta}.$$

Since $\epsilon > 0$ is arbitrary, this implies $\mathcal{H}^s(S) \leq \mathcal{H}^s(S)/(1+\delta)$ and hence $\mathcal{H}^s(S) = 0$, as required. $\square$

Lemma 8.3. *Let X be a metric space, $s \geq 0$ and let $A \subset X$ be $\mathcal{H}^s$-measurable with $\mathcal{H}^s(A) < \infty$. Then*

$$\Theta^{*,s}(A, x) = 0$$

for $\mathcal{H}^s$-a.e. $x \notin A$.

Proof. It suffices to show that, for $t > 0$, the set

$$S = \{x \in X \setminus A : \Theta^{*,n}(A, x) > t\}$$

satisfies $\mathcal{H}^s(S) = 0$. Fix $\epsilon > 0$. Since A is $\mathcal{H}^s$-measurable, $\mathcal{H}^s|_A$ is Borel regular. Therefore, since $\mathcal{H}^s|_A(S) = 0$, there exists an open $U \supset S$ with

$$\mathcal{H}^s(A \cap U) = \mathcal{H}^s|_A(U) < \epsilon.$$

For each $x \in S$ and $\delta > 0$ there exists a ball B centred on x with $\operatorname{rad} B < \delta$ such that

$$\frac{\mathcal{H}^s(A \cap B)}{(2 \operatorname{rad} B)^s} > t.$$

By Lemma 6.1 there exists a disjoint collection $\mathcal{B}$ of such balls such that

$$S \subset \bigcup_{B \in \mathcal{B}} 5B.$$

Since $\mathcal{H}^s(A) < \infty$, A is separable and each of these balls contains a point of A, $\mathcal{B}$ is countable. Therefore

$$t\mathcal{H}^s_{5\delta}(S) \leq t \sum_{B \in \mathcal{B}} (2 \operatorname{rad} 5B)^s < 5^s \sum_{B \in \mathcal{B}} \mathcal{H}^s(A \cap B) \leq 5^s \mathcal{H}^s(A \cap U) < 5^s \epsilon.$$

Since $\delta, \epsilon > 0$ are arbitrary, this completes the proof. $\square$

8.1. **Exercises.**

Exercise 8.1. Let $\mathcal{V} \subset [0, 1]$ be a Vitali set as constructed in Exercise 1.8.

(1) Show that, for any Borel $B \subset \mathcal{V}$, $\mathcal{L}^1(B) = 0$.

(2) Deduce that $\mathcal{L}^1([0, 1] \setminus \mathcal{V}) = 1$ and hence, if C is a Borel set with

$$[0, 1] \setminus \mathcal{V} \subset C \subset [0, 1],$$

then $\mathcal{L}^1(C) = 1$.

(3) Hence show that $\mathcal{L}^1(\mathcal{V} \cap C) = \mathcal{L}^1(\mathcal{V}) > 0$.

Note however that we cannot deduce the value of $\mathcal{L}^1(\mathcal{V})$ from our construction in Exercise 1.8. Indeed, for any $\epsilon > 0$, that construction may produce a $\mathcal{V} \subset [0, \epsilon]$.

Exercise 8.2. Let X be a metric space and $s \geq 0$.

(1) Show that $\mathcal{H}^s$ is Borel regular. Hint: first show that in the definition of $\mathcal{H}^s$, we may take F to be the collection of *closed* sets.

(2) We are usually interested in $\mathcal{H}^s|_A$ for some $A \subset X$. Show that for *any* $A \subset X$, $\mathcal{H}^s|_A$ is a Borel measure.

(3) Now assume that $A \subset X$ is $\mathcal{H}^s$-measurable with $\mathcal{H}^s(A) < \infty$. Show that $\mathcal{H}^s|_A$ is Borel regular. Hint: show that there exist Borel sets $B \supset A \supset B'$ with $\mathcal{H}^s(B \setminus B') = 0$.

(4) Show that $\mathcal{H}^s|_A$ may not be Borel regular if A is not $\mathcal{H}^s$ measurable. Hint: consider Exercise 8.1.

Exercise 8.3. In this exercise we construct the four corner Cantor set. Let $K_0 = [0,1]^2$. Let K_1 be the "four corners" of K_0 of side length $1/4$. That is

$$K_1 = [0,1/4]^2 \cup [3/4,1]^2 \cup [0,1/4] \times [3/4,1] \cup [3/4,1] \cup [0,1/4].$$

Inductively, K_n is constructed by taking the four corners of side length $1/4^n$ of all the squares of K_{n-1}. Finally let $K = \bigcap_{n \in \mathbb{N}} K_n$, a compact set. Show that $0 < \mathcal{H}^1(K) < \infty$.

Exercise 8.4. For $s \geq 0$ let X be a metric space with $\mathcal{H}^s(X) < \infty$. Show that X is separable.

Exercise 8.5. Show that Lemmas 8.2 and 8.3 may be false if A has only σ-finite $\mathcal{H}^s$ measure.

Exercise 8.6. Recall from Exercise 5.8 and Exercise 6.5 that $\mathcal{L}^n = c_n \mathcal{H}^n$ on $\mathbb{R}^n$. We will show that $\mathcal{H}^n(B(0,1)) = 2^n = \operatorname{diam} B(0,1)^n$.

(1) Use the translation and scale invariance of $\mathcal{H}^n$, and the fact that
$$\Theta^{*,n}(B(0,1),x) \leq 1 \quad \text{for } \mathcal{H}^n\text{-a.e. } x \in B(0,1)$$
to deduce that $\mathcal{H}^n(B(0,1)) \leq 2^n$.

(2) The isodiametric inequality states that $\mathcal{L}^n(A) \leq c_n(\operatorname{diam} A/2)^n$ for any $A \subset \mathbb{R}^n$. Show that, for any sets $A_1, A_2, \ldots \subset \mathbb{R}^n$ with
$$B(0,1) \subset \bigcup_{i \in \mathbb{N}} A_i,$$
$1 \leq \sum_i (\operatorname{diam} A_i/2)^n$.

(3) Deduce that $\mathcal{H}^n(B(0,1)) \geq 2^n$.

9. Rectifiable sets and approximate tangent planes

Rectifiable sets are the measure theoretic counterpart to manifolds.

Definition 9.1. A $\mathcal{H}^n$-measurable set $E \subset \mathbb{R}^m$ is *n-rectifiable* if there exist Lipschitz $f_i \colon \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\mathcal{H}^n\left(E \setminus \bigcup_{i \in \mathbb{N}} f_i(\mathbb{R}^n)\right) = 0.$$

We will show that n-rectifiable sets possess a unique *approximate n-dimensional tangent plane* at almost every point.

Given $V \in G(m,n)$, $a \in \mathbb{R}^n$ and $0 < s < 1$ define the *cone* around V centred at a with aperture s as

$$C(a,V,s) = \{x \in \mathbb{R}^n : \operatorname{dist}(x-a, V) < s\|x-a\|\}.$$

Definition 9.2. Let $A \subset \mathbb{R}^m$ and $a \in A$. A $V \in G(m,n)$ is an *approximate tangent plane* to A at a if

$$\limsup_{r \to 0} \frac{\mathcal{H}^n(A \cap B(a,r))}{r^n} > 0 \tag{9.1}$$

and, for every $0 < s < 1$,

$$\lim_{r \to 0} \frac{\mathcal{H}^n(A \cap B(a,r) \setminus C(a,V,s))}{r^n} = 0. \tag{9.2}$$

Rademacher's theorem gives a candidate for the approximate tangent plane to a rectifiable set. There are three steps required to prove that the derivative is indeed an approximate tangent plane: show that the derivative has full rank at almost every point; prove the density condition (9.1); and show that the sets from other parametrisations of the rectifiable set do not destroy the approximation by a tangent plane at almost every point.

The second and third steps follow from the results of the previous section. For the first step we use the following.

Lemma 9.3 (Easy Sard's theorem)**.** *If $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz then*

$$\mathcal{H}^n(\{f(x) : \operatorname{rank} Df(x) < n\}) = 0.$$

Proof. Let $L = \operatorname{Lip} f$. Fix $0 < R < \infty$, $\delta, \epsilon > 0$ and let

$$A = \{x \in B(x, R) : \operatorname{rank} Df(x) < n\}.$$

For $x \in A$ let

$$W_x = f(x) + Df(x)(\mathbb{R}^n).$$

Then for sufficiently small $0 < r_x < \delta$,

$$f(B(x, r_x)) \subset B(f(x), Lr_x) \cap \{y : \operatorname{dist}(y, W_x) < \epsilon r_x\}.$$

Since $\operatorname{rank} Df(x) < n$, the set on the right hand side can be covered by $(L/\epsilon)^{n-1}$ cubes of side length ϵr_x.

Since A is covered by balls of the form $B(x, r_x/5)$, there exists a disjoint collection of balls $B(x_i, r_i/5)$ such that A is covered by the union of the $B(x_i, r_i)$. Then

$$f(A) \subset f\left(\bigcup_{i\in\mathbb{N}} B(x_i, r_i)\right) \subset \bigcup_{i\in\mathbb{N}} f(B(x_i, r_i)).$$

By the previous argument, each factor of the right hand side is covered by $(L/\epsilon)^{n-1}$ cubes of side length ϵr_i. Thus

$$\mathcal{H}^n_{2\delta}(f(A)) \leq \sum_{i\in\mathbb{N}} \left(\frac{L}{\epsilon}\right)^{n-1} (\epsilon r_i)^{n-1} = L^{n-1}\epsilon \sum_{i\in\mathbb{N}} r_i^n.$$

However, the $B(x_i, r_i/5)$ are disjoint subsets of $B(0, R+\delta) \subset \mathbb{R}^n$ and so

$$\sum_{i\in\mathbb{N}} \left(\frac{r_i}{5}\right)^n \leq (R+\delta)^n.$$

Since $\epsilon > 0$ is arbitrary, this implies that $\mathcal{H}^n_{2\delta}(f(A)) = 0$ and hence $\mathcal{H}^n(f(A)) = 0$. Taking a countable union over $R \to \infty$ completes the proof. □

Lemma 9.4. *Let $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz and $S \subset \mathbb{R}^n$. Suppose that there exists a $\delta > 0$ such that, for each $x, y \in S$,*

$$\|f(x) - f(y)\| \geq \delta \|x - y\|.$$

Then $f(S)$ has a unique approximate tangent plane at $\mathcal{H}^n$ almost every point.

Proof. By Lemma 9.3, we may suppose $\operatorname{rank} Df(x) = n$ for every $x \in S$. By Lemma 8.2, we may suppose $\Theta^{*,n}(E, f(x)) > 0$ for every $x \in S$. Fix $x \in S$ and $0 < s < 1$. There exists $\epsilon > 0$ such that

$$\|f(y) - f(x) - Df(x)(y - x)\| < s\|y - x\|$$

for all $y \in B(x, \epsilon) \cap S$. Moreover, if $y \in S \setminus B(x, \epsilon)$ then

$$\|f(y) - f(x)\| \geq \delta\epsilon.$$

That is, if $a = f(x)$ and $b = f(y)$ with $\|a - b\| \leq \delta\epsilon$ and $V = a + Df(x)(\mathbb{R}^n)$,

$$\operatorname{dist}(b - a, V) < s\|y - x\| \leq s\|b - a\|/\delta.$$

Therefore, V is an approximate tangent plane to $f(S)$ at a.

This approximate tangent plane is unique at any density point x of S. Indeed, if $V' \neq V$, let $v \in V \backslash V'$ and let $0 < s < 1$ be such that $C(f(x), \mathbb{R}v, s) \cap C(f(x), V', s) = \{0\}$. Since $\operatorname{rank} Df(x) = n$ and x is a density point of S, for sufficiently small $r > 0$ there exists $y \in S \cap B(x, r)$ such that $B(f(y), sr) \cap C(f(x), V', s) = \emptyset$ and

$$\frac{\mathcal{H}^n(B(f(y), sr) \cap S)}{r^n} \geq \delta s.$$

In particular, V' is not an approximate tangent to $f(S')$ at x. □

Theorem 9.5. *Let $E \subset \mathbb{R}^m$ be n-rectifiable with $\mathcal{H}^n(E) < \infty$. Then for $\mathcal{H}^n$-a.e. $x \in E$, E has a unique approximate tangent plane at x.*

Proof. Let $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be one of the Lipschitz functions as in the definition of a rectifiable set and let $S = f^{-1}(E)$. It suffices to prove that E has a unique approximate tangent plane at $f(x)$ for $\mathcal{L}^n$-a.e. $x \in S$.

By Lemma 9.3, we may suppose that $\dim Df(\mathbb{R}^n) = n$ for all $x \in S$. Fix such an x and let $0 < \epsilon < \|Df(x)^{-1}\|/2$. There exists $\delta > 0$ such that

$$\|f(y) - f(x) - Df(x)(y - x)\| < \epsilon\|y - x\|$$

for all $y \in B(x, \delta) \cap S$. In particular, by the triangle inequality,

$$\|f(y) - f(x)\| > \epsilon\|y - x\|/2.$$

Therefore, the sets

$$S_\epsilon := \{x \in G : \|f(y) - f(x)\| > \epsilon\|y - x\| \ \forall y \in B(x, \epsilon)\}$$

are Borel and monotonically increase to S as $\epsilon \to 0$. Therefore it suffices to prove the result for $\mathcal{L}^n$-a.e. x in some fixed S_ϵ. Cover S_ϵ by finitely many balls $B_1, B_2, \ldots, B_N$ of radius ϵ. It suffices to prove the result for $\mathcal{L}^n$-a.e. x in some fixed $S' := S_\eta \cap B_i$.

However, S' satisfies the hypotheses of Lemma 9.4 and so $f(S')$ has a unique approximate tangent at $\mathcal{H}^n$ almost every point. To see that this tangent is a unique approximate tangent to E at $\mathcal{H}^n$ almost every point, we simply use Lemma 8.3: for $\mathcal{H}^n$-a.e. $x \in f(S')$, $\Theta^{*,n}(E \setminus f(S'), x) = 0$. □

In Theorem 10.10 we will see that the converse to Theorem 9.5 holds.

For $V \in G(n, m)$, write π_V for the orthogonal projection onto V and equip $G(n, m)$ with the metric $d(V, W) = \|\pi_V - \pi_W\|$. We will consider $\mathcal{L}^n$ on an element of $G(n, m)$.

Lemma 9.6. *Let $f\colon \mathbb{R}^m \subset \mathbb{R}^n \to \mathbb{R}^m$ be Lipschitz and let $S \subset \mathbb{R}^n$ satisfy $\mathcal{L}^n(S) > 0$. For $\epsilon > 0$ suppose that there exists an invertible linear $L\colon \mathbb{R}^n \to \mathbb{R}^m$ such that, for all $x, y \in S$,*

$$\|f(x) - f(y) - L(x - y)\| < \frac{\epsilon}{\|L^{-1}\|}\|x - y\|.$$

Then for any $V \in G(n, m)$ with $\|(\pi_V|_{L(\mathbb{R}^n)})^{-1}\|^{-1} \geq 2\epsilon$, $\mathcal{L}^n(\pi_V(f(S))) > 0$.

Proof. For any $V \in G(n, m)$,

$$\|\pi_V(f(x) - f(y)) - \pi_V(L(x - y))\| < \epsilon\|L^{-1}\|^{-1}\|x - y\|$$

and so, if $\|(\pi_V|_{L(\mathbb{R}^n)})^{-1}\|^{-1} \geq 2\epsilon$,

$$\begin{aligned}\|\pi_V(f(x)-f(y))\| &\geq \|\pi_V(L(x-y))\| - \epsilon\|L^{-1}\|^{-1}\|x-y\| \\ &\geq \|(\pi_V|_{L(\mathbb{R}^n)})^{-1}\|^{-1}\|L(x-y)\| - \epsilon\|L(x-y)\| \\ &\geq \epsilon\|L(x-y)\| \\ &\geq \epsilon\|L^{-1}\|\|x-y\|.\end{aligned}$$

Thus $\pi_V \circ f$ has Lipschitz inverse on S and hence $\mathcal{L}^n(\pi_V(f(S))) > 0$ by Exercise 1.5. □

Corollary 9.7. *Let $E \subset \mathbb{R}^m$ be n-rectifiable with $\mathcal{H}^n(E) > 0$. Then there exists $W \in G(m-n,m)$ such that $\pi_V(E) > 0$ for all $V \in G(n,m)$ with $V \cap W = \{0\}$.*

Remark 9.8. The set of V that satisfy the conclusion of Corollary 9.7 is very large; try some examples in reasonable dimensions.

Proof. Since $E \subset \mathbb{R}^m$ is rectifiable with $\mathcal{H}^n(E) > 0$, there exists a Lipschitz $f\colon \mathbb{R}^n \to \mathbb{R}^m$ with $\mathcal{H}^n(E \cap f(\mathbb{R}^n)) > 0$. In particular, $S := f^{-1}(E)$ satisfies $\mathcal{L}^n(S) > 0$. By Lemma 9.3, for $\mathcal{L}^n$-a.e. $x \in S$, $Df(x)$ is injective.

Fix $\epsilon > 0$ For $M > 0$, the set of invertible $L\colon \mathbb{R}^n \to \mathbb{R}^m$ with $\|L^{-1}\| < M$ may be covered by countably many sets of diameter ϵ/M. Varying $M \in \mathbb{N}$, we see that $\mathcal{L}^n$ almost all of S is covered by countably many sets of the form

$$\{x \in S : \|Df(x) - L\| < \epsilon/2\|L^{-1}\|\}.$$

Moreover, each of these sets may be covered by countably many sets of the form

$$\{x \in S : \|f(x) - f(y) - L(x-y)\| < \epsilon\|x-y\|/\|L^{-1}\| \ \forall y \in B(x,\epsilon)\}.$$

Finally, these sets may be covered by countably many sets of diameter ϵ. Therefore, for each $j \in \mathbb{N}$ there exists $S_j \subset S$ and invertible $L_j\colon \mathbb{R}^n \to \mathbb{R}^m$ such that, for all $x, y \in S_j^\epsilon$,

$$\|f(x) - f(y) - L_j(x-y)\| < \epsilon\|x-y\|/\|L_j\|^{-1},$$

and $\mathcal{L}^n(S \setminus \bigcup_{j\in\mathbb{N}} S_j) = 0$.

Since $\mathcal{L}^n(S) > 0$, there exists $j \in \mathbb{N}$ with $\mathcal{L}^n(S_j) > 0$. Then S_j satisfies the hypotheses of Lemma 9.6 and so $\mathcal{L}^n(\pi_V(S)) \geq \mathcal{L}^n(\pi_V(S_j)) > 0$ for all $V \in G(n,m)$ with $\|(\pi_V|_{L_j(\mathbb{R}^n)})^{-1}\|^{-1} \geq 2\epsilon$. Let $L_\epsilon = L_j$. Repeat this for each $i \in \mathbb{N}$ with $\epsilon = 1/i$. The set $G(n,m)$ is compact and so we may suppose that $L_{1/i}(\mathbb{R}^n) \to W \in G(n,m)$. The only $V \in G(n,m)$ for which $\mathcal{L}^n(\pi_V(S)) = 0$ satisfy $\|(\pi_V|_{L_{1/i}(\mathbb{R}^n)})^{-1}\|^{-1} < 2/i$ and hence $\|(\pi_V|_W)^{-1}\|^{-1} < 2/i$ for each $i \in \mathbb{N}$. That is, $V \cap W^\perp \neq \{0\}$ as required. □

9.1. **Exercises.**

Exercise 9.1. Let X be a metric space, $Y \subset X$ and $f\colon Y \to \mathbb{R}$ L-Lipschitz. Define $\tilde{f}\colon X \to \mathbb{R}$ by

$$\tilde{f}(x) = \sup\{f(y) - Ld(x,y) : y \in Y\}.$$

(1) Show that $\tilde{f}$ is an L-Lipschitz extension of f to X. This is called the *McShane–Whitney extension theorem*
(2) If $f\colon Y \to \mathbb{R}^n$ is L-Lipschitz, show that there is a $\sqrt{n}L$-Lipschitz extension of f to X.

(3) The following example shows that the vector valued extension cannot have the same Lipschitz constant in general: Let

$$Y = \{(-1,1),(1,-1),(1,1)\} \subset \ell^2_\infty$$

and define

$$f(-1,1) = (-1,0), \quad f(1,-1) = (1,0), \quad f(1,1) = (0,\sqrt{3}).$$

Show that f is 1-Lipschitz but has no 1-Lipschitz extension to $Y \cup \{(0,0)\}$.

(4) However, the *Kirszbraun extension theorem* states that any Lipschitz map between any two Hilbert spaces may be extended whilst preserving the Lipschitz constant.

Exercise 9.2. (1) Let $E \subset \mathbb{R}^m$ be n-rectifiable. Show that $\mathcal{H}^n|_E$ is σ-finite.

(2) Show that Theorem 9.5 may not be true if E does not satisfy $\mathcal{H}^n(E) < \infty$.

10. Purely unrectifiable sets

Definition 10.1. A $\mathcal{H}^n$-measurable set $S \subset \mathbb{R}^m$ is *purely n-unrectifiable* if, for all n-rectifiable $E \subset \mathbb{R}^n$, $\mathcal{H}^n(S \cap E) = 0$.

Lemma 10.2. *The four corner Cantor set $K \subset \mathbb{R}^2$ is purely 1-unrectifiable.*

Proof. Observe that the coordinate projections of K have $\mathcal{L}^1$ measure zero (see Exercise 10.1). If there existed a rectifiable $\gamma \subset \mathbb{R}^2$ with $\mathcal{H}^1(\gamma \cap K) > 0$, then $\gamma \cap K$ is a 1-rectifiable set of positive measure and hence, by Corollary 9.7, one of the coordinate projections must have positive measure, a contradiction.

For a second proof see Exercise 10.2. □

Lemma 10.3. *Let $A \subset \mathbb{R}^m$ be $\mathcal{H}^m$-measurable with $\mathcal{H}^m(A) < \infty$. There exists a decomposition $A = E \cup S$ with E n-rectifiable and S purely n-unrectifiable.*

Proof. Let

$$t = \sup\{\mathcal{H}^n(E) : E \subset A, n\text{-rectifiable}\}.$$

Since $\mathcal{H}^n(A) < \infty$, $t < \infty$. Let $E_i \subset A$ be n-rectifiable with $\mathcal{H}^n(E_i) \to t$. Then $E = \bigcup_{i\in\mathbb{N}} E_i$ is n-rectifiable and is contained in A. Therefore

$$t \geq \mathcal{H}^n(E) \geq \mathcal{H}^n(E_i) \to t$$

and so $\mathcal{H}^n(E) = t$. Then $S = A \setminus E$ is purely n-unrectifiable. Indeed, if $f : \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz, $E' = (A \setminus E) \cap f(\mathbb{R}^n)$, is $\mathcal{H}^m$-measurable and so

$$t \geq \mathcal{H}^n(E \cup E') = \mathcal{H}^n(E) + \mathcal{H}^n(E') = t + \mathcal{H}^n(E').$$

□

We now state a very important theorem on the structure of purely unrectifiable sets. It requires the notion of a natural measure $\gamma_{n,m}$ on $G(n,m)$ that is invariant under the action of $SO(m)$. There are several ways to construct this measure. The simplest is to consider $G(n,m)$ as a (compact) metric space equipped with the metric

$$d(V,W) = \|\pi_V - \pi_W\|.$$

Then $\gamma_{n,m}$ is given by (a scalar multiple of) $\mathcal{H}^{n(m-n)}$. We will not discuss the specific details of this measure. When $n = 1$, we may identify $G(1,m)$ with $\mathbb{S}^{m-1}$. In this case, $\gamma_{1,m}$ is simply $\mathcal{H}^{m-1}$.

Theorem 10.4 (Besicovitch–Federer projection theorem)**.** *Let $S \subset \mathbb{R}^m$ be purely n-unrectifiable with $\mathcal{H}^n(S) < \infty$. Then, for $\gamma_{n,m}$-a.e. $V \in G(n,m)$,*

$$\mathcal{L}^n(\pi_V(S)) = 0.$$

Conversely, if $E \subset \mathbb{R}^m$ is purely n-unrectifiable with $\mathcal{H}^n(E) > 0$, for $\gamma_{n,m}$-a.e. $V \in G(n,m)$,

$$\mathcal{L}^n(\pi_V(E)) > 0.$$

Remark 10.5. The converse statement is given by Corollary 9.7.

We will prove the projection theorem for $n = 1$ and $m = 2$, which was proved by Besicovitch. First we prove some preliminary geometric properties of purely unrectifiable sets.

Lemma 10.6. *Let $E \subset \mathbb{R}^m$, $V \in G(m-n,m)$ and $0 < s < 1$. Suppose that, for every $x \in E$,*

$$E \cap C(x,V,s) \cap B(x,r) = \emptyset.$$

Then E is n-rectifiable.

Proof. Since E may be divided into countably many sets of diameter at most r, we may suppose $\operatorname{diam} E \leq r$. In this case, $\pi_{V^\perp}$ has Lipschitz inverse on E. Indeed, if $x, y \in E$ then $y \notin C(x,V,s)$ and so

$$\|\pi_{V^\perp}(x-y)\| = \operatorname{dist}(x-y,V) \geq s\|x-y\|.$$

Therefore, E is covered by a Lipschitz image of $\mathbb{R}^n$. □

Lemma 10.7. *Let $S \subset \mathbb{R}^m$ be purely n-unrectifiable, $V \in G(m-n,m)$, $0 < s < 1$ and $0 < \delta, \lambda < \infty$. If*

$$\mathcal{H}^n(S \cap C(x,V,s) \cap B(x,r)) \leq \lambda r^n s^n \tag{10.1}$$

for every $x \in S$ and $0 < r < \delta$ then

$$\mathcal{H}^n(S \cap B(a,\delta/6)) \leq 2 \cdot 20^n \lambda \delta^n$$

for every $a \in S$.

Remark 10.8. Note that this is certainly not true for a rectifiable set E; the first cone may be empty for every $x \in E$.

Proof. For a fixed $a \in S$, we may suppose $S \subset B(a,\delta/6)$. By Lemma 10.6, we may suppose that

$$S \cap C(x,V,s/4) \neq \emptyset$$

for every $x \in S$. For every $x \in S$ let

$$h(x) = \sup\{|x-y| : y \in S \cap C(x,V,s/4)\},$$

so that $0 < h(x) < \delta/3$ for all $x \in S$. Pick $x^* \in S \cap C(x,V,s/4)$ with $|x - x^*| \geq 3h(x)/4$ and let C_x be the cylinder

$$C_x = \pi_{V^\perp}^{-1}(\pi_{V^\perp}(B(x,sh(x)/4))).$$

We claim that

$$C_x \cap S \subset C(x,V,s) \cap B(x,2h(x)) \cup C(x^*,V,s) \cap B(x^*,2h(x)). \tag{10.2}$$

(Draw a picture!) Suppose $z \in C_x \cap S$ does not belong to the first set. Then

$$\begin{aligned} s\|x^* - z\| &\le \|\pi_{V^\perp}(x^* - z)| \\ &\le \|\pi_{V^\perp}(x^* - x)\| + \|\pi_{V^\perp}(x - z)\| \\ &\le s|x^* - x|/4 + sh(x)/4 \\ &\le sh(x)/2, \end{aligned}$$

where the penultimate inequality follows because $x^* \in C(x, V, s/4)$ and $z \in C_x$. Therefore

$$\begin{aligned} \|x - z\| &\ge \|x - x^*\| - \|x^* - z\| \\ &> 3h(x)/4 - h(x)/2 \\ &\ge \|\pi_{V^\perp}(x - z)\|/s. \end{aligned}$$

That is, z belongs to the first set in (10.2).

By (10.2) and (10.1),

$$\mathcal{H}^1(S \cap C_x) \le 2\lambda(2h(x)s)^n.$$

We apply Lemma 6.1 to the balls

$$\pi_{V^\perp}(B(x, sh(x)/20))$$

with $x \in S$. This gives countably many $x_i \in S$, for which these balls are disjoint, and

$$S \subset \bigcup_{i \in \mathbb{N}} C_{x_i}.$$

Therefore

$$\begin{aligned} \mathcal{H}^n(S) &\le \sum_{i \in \mathbb{N}} \mathcal{H}^n(C_{x_i}) \\ &\le 2\lambda 2^n \sum_{i \in \mathbb{N}} (sh(x_i))^n \\ &= 2\lambda 2^n 20^n \sum_{i \in \mathbb{N}} \left(\frac{sh(x_i)}{20} \right)^n. \end{aligned}$$

But, the $\pi_{V^\perp}(B(x_i, sh(x_i)/20))$ are disjoint subsets of $B(\pi_{V^\perp}(a), \delta/2) \subset V^\perp$ and so the final sum is bounded above by $(\delta/2)^n$. $\square$

Corollary 10.9. *If $S \subset \mathbb{R}^m$ is purely n-unrectifiable with $\mathcal{H}^n(S) < \infty$ then for every $V \in G(m - n, m)$, every $0 < s < 1$ and $\mathcal{H}^n$-a.e. $x \in S$,*

$$\Theta^{*,n}(S \cap C(a, V, s), a) \ge 240^{-n-1} s^n.$$

Proof. For a fixed V, s, this is immediate from the fact that $\Theta^{*,n}(S, a) \ge 2^{-n}$ almost everywhere. To obtain the conclusion for all V, s, note that the conclusion is determined by a countable dense set of V, s. $\square$

Theorem 10.10. *Let $E \subset \mathbb{R}^m$ satisfy $\mathcal{H}^n(E) < \infty$ and suppose that, for $\mathcal{H}^n$-a.e. $x \in E$, E has a unique approximate tangent plane at x. Then E is n-rectifiable.*

Proof. By Lemma 10.3, there exists a decomposition $E = E' \cup S$, where E' is n-rectifiable and S is purely n-unrectifable. We must show that $\mathcal{H}^n(S) = 0$. Note that, by applying Lemma 8.3 to E', we see that the approximate tangent plane to E at $x \in S$ is also an approximate tangent plane to S for $\mathcal{H}^n$-a.e. x.

It suffices to show, for a fixed $W \in G(n,m)$, that the set S' of $x \in S$ whose approximate tangent plane V_x lies in $B(W,\delta)$ has measure zero. Suppose not. Then, for any $\lambda > 0$, there exists an $R > 0$ such that the set S'' of those $x \in S'$ with

$$\sup_{0<r<R} \frac{\mathcal{H}^n(S \cap B(a,r) \setminus C(a, V_a, 1/3))}{r^n} < \lambda 3^{-n}$$

has positive measure. Fix an $x \in S''$. Since $\|\pi_{V_x} - \pi_W\| \leq 1/3$, for every $0 < r < R$ we have

$$C(x, W^\perp, 1/3) \cap B(x,r) \subset B(x,r) \setminus C(x, V_x, 1/3).$$

Thus, for $x \in S''$ and $0 < r < R$,

$$\mathcal{H}^n(S' \cap C(x, W^\perp, 1/3) \cap B(x,r)) < \lambda 3^{-n} r^n.$$

If $\lambda < 240^{-m-1}$, Corollary 10.9 implies $\mathcal{H}^n(S'') = 0$, a contradiction. $\square$

We now prove the Besicovitch projection theorem [1]

Theorem 10.11. *Let $S \subset \mathbb{R}^2$ be purely 1-unrectifiable with $\mathcal{H}^1(S) < \infty$. Then for $\mathcal{H}^1$-a.e. $e \in \mathbb{S}^1$,*

$$\mathcal{L}^1(\pi_e(S)) = 0.$$

We follow the presentation of Orponen [2].

From Corollary 10.9, we see that a purely unrectifiable set has many radiating out of almost every point in all directions at almost every point. We now precisely describe two ways in which this can occur.

Notation 10.12. Let $S \subset \mathbb{R}^2$ and $x \in S$. For $e \in \mathbb{S}^1$ let $l_e(x)$ be the half line $x + [0,\infty)e$ and for $I \subset \mathbb{S}^1$, let $C(I,x)$ be the cone $\bigcup_{e\in I} l_e(x)$. For $r > 0$ let $H_x(r)$ be those $e \in \mathbb{S}^1$ for which

$$|K \cap l_e(x) \cap B(x,r)| \geq 2.$$

That is, $S \cap l_e(x) \cap B(x,r)$ contains *another* point of K. Also let $H_x = \bigcap_{r>0} H_x(r)$, the directions that contain other points of S arbitrarily close to x. For $e \in \mathbb{S}^1$, we let H_e be those $x \in S$ for which $e \in H_x$.

For $R, M, \epsilon > 0$ let $D_x(R,M,\epsilon)$ be those $e \in \mathbb{S}^1$ for which there exists $0 < r < R$ and an interval $I \subset \mathbb{S}^1$ with $e \in I$ and $0 < \mathcal{H}^1(I) < \epsilon$ such that

$$\frac{\mathcal{H}^1(S \cap C(x,I) \cap B(x,r))}{r} \geq M\mathcal{H}^1(I).$$

That is, the *density* of S in the cone $C(x,I)$ at scale r is very high, compared to the length of I. Also let $D_x = \bigcap_{R,M,\epsilon>0} D_x(R,M,\epsilon)$. For $e \in \mathbb{S}^1$, we let D_e be those $x \in S$ for which $e \in D_x$.

The main step in proving Theorem 10.11 is the following.

Proposition 10.13. *Let $S \subset \mathbb{R}^2$ be purely 1-unrectifiable with $\mathcal{H}^1(S) < \infty$. For $\mathcal{H}^1$-a.e. $x \in S$, $\mathcal{H}^1(\mathbb{S} \setminus H_x \cup D_x) = 0$.*

Before proving Proposition 10.13, we will demonstrate how it is used to prove Theorem 10.11.

Lemma 10.14 (Special case of the *coarea formula*)**.** *For any $e \in \mathbb{S}^1$ and any compact $K \subset \mathbb{R}^2$,*

$$\int_{\mathbb{R}} \operatorname{card}(K \cap l_e(t))\, \mathrm{d}t \leq \mathcal{H}^1(K).$$

In particular, if $\mathcal{H}^1(K) < \infty$ then for any $e \in \mathbb{S}^1$, $\mathcal{L}^1(\pi_{e^\perp}(H_e)) = 0$.

Proof. Since K is compact,

$$f(t) = \mathrm{card}(K \cap l_e(t))$$

is a Borel function. Indeed, if, for $\delta > 0$,

$$f_\delta(t) = \max\{n \in \mathbb{N} : \exists x_1, \ldots, x_n \in K \cap l_e(t) \text{ with } \|x_i - x_j\| \geq \delta \ \forall 1 \leq i \neq j \leq n\}$$

then f_δ monotonically increases to f as $\delta \to 0$. Since K is compact, the f_δ are lower semi-continuous. Therefore, by the monotone convergence theorem, it suffices to bound the integral of each f_δ.

Fix $\delta > 0$ and cover K by sets $E_1, E_2, \ldots$ with $\operatorname{diam} E_i < \delta$ such that

$$\sum_{i \in \mathbb{N}} \operatorname{diam} E_i \leq \mathcal{H}^1(K) + \delta.$$

Note that

$$f_\delta(t) \leq \mathrm{card}(\{i : E_i \cap l_e(t) \neq \emptyset\}).$$

Therefore

$$\begin{aligned}
\int_{\mathbb{R}} f_\delta \, \mathrm{d}\mathcal{L}^1 &\leq \int_{\mathbb{R}} \sum_{i \in \mathbb{N}} \chi_{\{(i,t) : E_i \cap l_e(t) \neq \emptyset\}} \\
&= \sum_{i \in \mathbb{N}} \int_{\mathbb{R}} \chi_{\{(i,t) : E_i \cap l_e(t) \neq \emptyset\}} \\
&\leq \sum_{i \in \mathbb{N}} \operatorname{diam} E_i \\
&\leq \mathcal{H}^1(K) + \delta,
\end{aligned}$$

as required. □

Lemma 10.15. *Let $S \subset \mathbb{R}^2$ be $\mathcal{H}^1$-measurable with $\mathcal{H}^1(S) < \infty$. Then for any $e \in \mathbb{S}^1$, $\mathcal{L}^1(\pi_{e^\perp}(D_e)) = 0$.*

Proof. Fix $e \in \mathbb{S}^1$. For any $M \in \mathbb{N}$ and $t \in \pi_{e^\perp}(D_e)$ there exists an $x \in D_e$, $r_x > 0$ and an interval $e \in I_x \subset \mathbb{S}^1$ with $\operatorname{diam} I < 1/10$ such that

$$\mathcal{H}^1(S \cap C(x, I) \cap B(x, r)) \geq M \mathcal{H}^1(rI).$$

Apply Lemma 6.1 to the intervals $J_x = \pi_{e^\perp}(C(x, I_x) \cap B(x, r_x))$ to obtain a disjoint collection $J_{x_1}, J_{x_2}, \ldots \subset \mathbb{R}$ such that

$$D_e \subset \bigcup_{i \in \mathbb{N}} 5 J_{x_i}.$$

Therefore

$$\begin{aligned}
\mathcal{L}^1(D_e) &\leq \sum_{i \in \mathbb{N}} 5 \mathcal{L}^1(J_{x_i}) \\
&\leq 5 \sum_{i \in \mathbb{N}} \mathcal{H}^1(r I_{x_i}) \\
&\leq \frac{5}{M} \sum_{i \in \mathbb{N}} \mathcal{H}^1(S \cap C(x_i, I_{x_i}) \cap B(x_i, r_{x_i})) \\
&\leq \frac{5}{M} \mathcal{H}^1(S),
\end{aligned}$$

where the final inequality follows from the disjointness of the sets $C(x_i, I_{x_i}) \cap B(x_i, r_i)$, $i \in \mathbb{N}$. Since this is true for all $M \in \mathbb{N}$, $\mathcal{L}^1(D_e) = 0$. □

Proof of Theorem 10.11 using Proposition 10.13. By the inner regularity of $\mathcal{H}^1$, it suffices to prove the result for compact S. By definition, we have

$$\{(x,e) \in S \times \mathbb{S}^1 : e \notin H_x \cup D_x\} = \{(x,e) : x \notin H_e \cup D_e\}.$$

Proposition 10.13 implies that the left hand expression has $\mathcal{H}^1 \times \mathcal{H}^1$-measure zero and so Fubini's theorem implies that, for $\mathcal{H}^1$-a.e. $e \in \mathbb{S}^1$, $\mathcal{H}^1(S \setminus H_e \cup D_e) = 0$. Therefore, by Lemmas 10.14 and 10.15, $\pi_{e^\perp}(S) = 0$ for $\mathcal{H}^1$-a.e. $e \in \mathbb{S}^1$. □

Proof of Proposition 10.13. Fix $R, M, \epsilon > 0$ and $x \in S$ which satisfies the conclusion of Corollary 10.9. That is,

$$\Theta^{*,1}(S \cap C(x,I), x) \geq c_0 \mathcal{H}^1(I) \tag{10.3}$$

for every interval $I \subset \mathbb{S}^1$. It suffices to show that $\mathcal{H}^1(\mathbb{S}^1 \setminus H_x(R) \cup D_x(R,M,\epsilon)) = 0$. In fact, we will show that, for *any* $e \in \mathbb{S}^1$,

$$\Theta^{*,1}(H_x(R), e) > 0 \quad \text{or} \quad \Theta^1_*(D_x(R,M,\epsilon), e) > 0,$$

from which the result follows by the Lebesgue density theorem.

To this end, fix $e \in \mathbb{S}^1$ with $\Theta^{*,1}(H_x(R), e) = 0$. Then for all sufficiently small intervals I with $e \in I \subset \mathbb{S}^1$,

$$\mathcal{H}^1(H_x(R) \cap I) < c_0 \mathcal{H}^1(I)/4M. \tag{10.4}$$

Fix such an I. By Equation (10.3), there exists $r < R$ with

$$\mathcal{H}^1(S \cap C(x,I) \cap B(x,r)) \geq c_0 r \mathcal{H}^1(I). \tag{10.5}$$

Note that (10.4) implies

$$\mathcal{H}^1(H_x(r) \cap I) < c_0 \mathcal{H}^1(I)/4M. \tag{10.6}$$

We will prove that

$$\mathcal{H}^1(D_x(R,M,\epsilon) \cap I) \geq c_0 \mathcal{H}^1(I)/4M. \tag{10.7}$$

Since I is any sufficiently small interval containing e, this implies

$$\Theta^1_*(D_x(R,M,\epsilon), e) \geq c_0/4M > 0$$

as required.

By (10.6) we may cover $H_x(r) \cap I$ by disjoint intervals $I_1, I_2, \ldots$ with

$$\sum_{i \in \mathbb{N}} \mathcal{H}^1(I_i) < c_0 \mathcal{H}^1(I)/4M$$

(indeed, the disjointness of the intervals is shown in Exercise 10.6). By the definition of $H_x(r)$, we know that

$$S \cap C(x,I) \cap B(x,r) \subset \bigcup_{i \in \mathbb{N}} S \cap C(x,I_i) \cap B(x,r). \tag{10.8}$$

Let $\mathcal{G}$ be those $i \in \mathbb{N}$ with

$$\frac{\mathcal{H}^1(S \cap C(x,I_i) \cap B(x,r))}{r} \geq M \mathcal{H}^1(I_i). \tag{10.9}$$

Note that by (10.5) and (10.8),

$$\begin{aligned}\sum_{i \in \mathcal{G}} \frac{\mathcal{H}^1(S \cap C(x,I_i) \cap B(x,r))}{r} &\geq \frac{\mathcal{H}^1(S \cap C(x,I) \cap B(x,r))}{r} - \frac{c_0 \mathcal{H}^1(I)}{4} \\ &\geq \frac{3c_0 \mathcal{H}^1(I)}{4}.\end{aligned}$$

That is, the cones associated to $G := \bigcup_{i\in\mathcal{G}} I_i$ cover a large proportion of $S\cap C(x,I)\cap B(x,r)$. Moreover, $G \subset D_x(R,M,\epsilon)$, because if $\xi \in G$ then $\xi \in I_i$ for some $i \in \mathcal{G}$ which satisfies the definition of $D_x(R,M,\epsilon)$.

Thus, to show (10.7), it would be enough to bound $\mathcal{H}^1(G)$ from below by a multiple of $\mathcal{H}^1(I)$. But this is not necessarily true: the intervals that form G could be extremely thin compared to I. To accommodate this, we enlarge the intervals I_i with $i \in \mathcal{G}$ as follows. For each $i \in \mathcal{G}$ enlarge I_i until (10.9) becomes an equality or until I_i intersects another I_j. If the first possibility occurs then we still have $I_i \subset D_x(R,M,\epsilon)$. If the second possibility occurs then we merge the two intervals; since both sides of (10.9) are linear in I, and the boundary of each $C(x,I_i)$ contains no points of S, (10.9) remains true after the merge. By the same reasoning, both sides of (10.9) are continuous under expanding I and consequently one of the two possibilities must occur.

This results in a disjoint collection of intervals $\tilde{I}_i$ for which (10.9) is an equality. Moreover,

$$G = \bigcup_{i\in\mathcal{G}} I_i \subset \bigcup \tilde{I}_i$$

and, by construction, each $\tilde{I}_i \in D_x(R,M,\epsilon)$. Therefore

$$\begin{aligned}
\mathcal{H}^1(D_x(r_0,\epsilon,M)\cap I) &\geq \sum \mathcal{H}^1(\tilde{I}_i) \\
&= \sum \frac{\mathcal{H}^1(S\cap B(x,r)\cap C(x,\tilde{I}_i))}{rM} \\
&\geq \sum_{i\in\mathcal{G}} \frac{\mathcal{H}^1(S\cap B(x,r)\cap C(x,I_i))}{rM} \\
&\geq \frac{3c_0\mathcal{H}^1(I)}{4M}.
\end{aligned}$$

□

10.1. **Exercises.**

Exercise 10.1. Prove that the coordinate projections of the four corner Cantor set $K \subset \mathbb{R}^2$ have Lebesgue measure zero.

Exercise 10.2. A second proof that the four corner Cantor set is purely 1 unrectifiable.

Suppose that $f\colon \mathbb{R} \to \mathbb{R}^2$ satisfies $\mathcal{H}^1(f(\mathbb{R})\cap K) > 0$.

(1) Prove that there exists $x \in \mathbb{R}$ that is a density point of $f^{-1}(K)$ such that $f'(x) \neq 0$.
(2) Therefore, for sufficiently small r, $f(x-r,x+r)$ is approximated by a line segment of length $2rf'(x)$ that is mostly contained in K. Derive a contradiction.

Exercise 10.3. Prove that the decomposition given in Lemma 10.3 is unique up to $\mathcal{H}^n$-null sets.

Exercise 10.4. Think about Corollary 10.9 and proposition 10.13 in regard to the four corner Cantor set.

Exercise 10.5. Let $K \subset \mathbb{R}^2$ be compact. Show that

$$\{(x,e) : e \in H_x\} \quad \text{and} \quad \{(x,e) : e \in D_x\}$$

are Borel subsets of $K \times \mathbb{S}^1$.

Exercise 10.6. Show that in the definitions of $\mathcal{L}^n$ and $\mathcal{H}^n$ we may suppose the covering intervals, respectively sets, may be chosen to be disjoint.

Exercise 10.7 (Open problem)**.** For $k \geq 2$, does there exist a compact purely 1-unrectifiable $S \subset \mathbb{R}^2$ with $\mathcal{H}^1(S) > 0$ that intersects every line in at most k points?

References

[1] A. S. Besicovitch. On the fundamental geometrical properties of linearly measurable plane sets of points (III). *Math. Ann.*, 116(1):349–357, 1939. doi:10.1007/BF01597361.

[2] T. Orponen. Geometric measure theory. *Online lecture notes*, 2018.

Email address: david.bate@warwick.ac.uk

Zeeman Building, University of Warwick, Coventry CV4 7AL