Lecture Notes in Real Analysis

Lewis Bowen*
University of Texas at Austin

December 8, 2014

Contents

*supported in part by NSF grant DMS-0968762, NSF CAREER Award DMS-0954606 and BSF grant 2008274

1 Outer measure and measurable sets

Definition 1. A subset $O \subset \mathbb{R}$ is *open* if for every $x \in O$ there is an $\epsilon > 0$ such that the interval $(x - \epsilon, x + \epsilon) \subset O$. A set $F \subset \mathbb{R}$ is *closed* if its complement is open.

Exercise 1. For every open subset $O \subset \mathbb{R}$ there is a finite or countable collection $\{I_i\}$ of pairwise disjoint open intervals such that $O = \cup_i I_i$.

Definition 2. The *outer measure* of a subset $E \subset \mathbb{R}$ is defined by $m^*(E) = \inf_{\mathcal{C}} \sum_{(x,y)\in\mathcal{C}} |x - y|$ where the infimum is over all subsets $\mathcal{C} \subset \mathbb{R} \times \mathbb{R}$ satisfying

- $(x, y) \in \mathcal{C} \Rightarrow x < y$
- $E \subset \cup_{(x,y)\in\mathcal{C}}(x, y)$.

We are using (x, y) to mean two different things: (x, y) is either an element of $\mathbb{R}^2$ or is an open interval of $\mathbb{R}$. The context should make clear which meaning is meant.

Exercise 2. The outer measure of an interval is its length.

Exercise 3. For any subsets $E_1, E_2, \ldots, \subset \mathbb{R}$, $m^*(\cup_i E_i) \leq \sum_i m^*(E_i)$. This means that outer measure is *countably sub-additive.*

Observation 1. If $E_1 \subset E_2$ then $m^*(E_1) \leq m^*(E_2)$.

Exercise 4. For any $E \subset \mathbb{R}$ and any $\epsilon > 0$ there exists an open set $O \supset E$ such that $m^*(O) < m^*(E) + \epsilon$.

Definition 3. We say that a subset $E \subset \mathbb{R}$ is *measurable* if for every $\epsilon > 0$ there exists an open set $O \supset E$ with $m^*(O \setminus E) < \epsilon$.

Observation 2. Open sets are measurable. If $m^*(E) = 0$ then E is measurable.

Exercise 5. If $E_1, E_2, \ldots$ are measurable then $\cup_i E_i$ is also measurable.

Exercise 6. Compact sets are measurable.

Exercise 7. Closed sets are measurable.

Exercise 8. E is measurable if and only if $\mathbb{R} \setminus E$ is measurable.

Exercise 9. If $E_1, E_2, \ldots$ are measurable then $\cap_i E_i$ is also measurable.

Definition 4. Let X be a set and $\mathcal{C}$ a collection of subsets of X. We say $\mathcal{C}$ is a *σ-algebra* if it is nonempty and for every $E_1, E_2, \ldots \in \mathcal{C}$, $X \setminus E_i \in \mathcal{C}$ (for all i, $\cup_i E_i \in \mathcal{C}$ and $\cap_i E_i \in \mathcal{C}$.

Observation 3. The measurable subsets of $\mathbb{R}$ form a σ-algebra.

Definition 5. The *Borel σ-algebra* is the smallest sigma-algebra containing all of the open sets. A set is *Borel* if it is in the Borel sigma-algebra. Note that all Borel sets are measurable.

2 Measures and measurable sets

Exercise 10. If $E \subset \mathbb{R}$ is measurable then for every $\epsilon > 0$ there exists a closed set $F \subset E$ with $m^*(E \setminus F) < \epsilon$.

Proof. Because E is meas., its complement E^c is also meas. So if $\epsilon > 0$ then there exists an open set $O \supset E^c$ with $m^*(O \setminus E^c) < \epsilon$. Now O^c is a closed set, $O^c \subset E$ and $E \setminus O^c = E \cap O = O \setminus E^c$. So $m^*(E \setminus O^c) < \epsilon$. □

Exercise 11. If $E_1, E_2, \ldots \subset \mathbb{R}$ are measurable and pairwise disjoint then

$$m^*(\cup_i E_i) = \sum_i m^*(E_i).$$

Definition 6. Let X be a set and $\mathcal{C}$ a σ-algebra on X. We say that $(X, \mathcal{C})$ is a *measurable space* (or *Borel space*). Also let $\mu : \mathcal{C} \to [0, \infty]$ be a function satisfying: if $E_1, E_2, \ldots \in \mathcal{C}$ are pairwise disjoint then

$$\mu(\cup_i E_i) = \sum_i \mu(E_i).$$

Then μ is a *measure* on $(X, \mathcal{C})$ and $(X, \mathcal{C}, \mu)$ is a *measure space*. Often we omit $\mathcal{C}$ from the notation and just say "μ is a measure on X".

Observation 4. m^* is a measure on $(\mathbb{R}, \mathcal{M})$ where $\mathcal{M}$ denotes the collection of measurable subsets. From now on, we let m denote the restriction of m^* to $\mathcal{M}$. This is called *Lebesgue measure on* $\mathbb{R}$.

Exercise 12. Let $(X, \mathcal{C}, \mu)$ be a measure space. Suppose $E_1 \subset E_2 \subset \cdots \in \mathcal{C}$ and $F_1 \supset F_2 \supset \cdots \in \mathcal{C}$. Then

$$\lim_i \mu(E_i) = \mu(\cup_i E_i).$$

If $\mu(F_1) < \infty$ then $\lim_i \mu(F_i) = \mu(\cap_i F_i)$.

Definition 7. A countable intersection of open subsets is called a set of type G_δ. A countable union of closed subsets is called a set of type F_σ.

Example 1. The irrational numbers are a dense G_δ subset of the real line.

Exercise 13. Let $E \subset \mathbb{R}$. Prove that the following are equivalent.

1. E is measurable.
2. for every $\epsilon > 0$ there exists an open set $O \supset E$ with $m^*(O \setminus E) < \epsilon$;
3. there exists a set $G \supset E$ of type G_δ such that $m^*(G \setminus E) = 0$;
4. for every $\epsilon > 0$ there exists a closed set $F \supset E$ with $m^*(E \setminus F) < \epsilon$;
5. there exists a set $F \supset E$ of type F_σ such that $m^*(E \setminus F) = 0$;

6. there exists a Borel set B such that $m(E \triangle B) = 0$ where $E \triangle B = (E \setminus B) \cup (B \setminus E)$ is the *symmetric difference* of B and E;

7. for every set $A \subset \mathbb{R}$,
$$m^*(A) = m^*(A \cap E) + m^*(A \setminus E).$$

3 Cantor sets and the Cantor-Lebesgue function

...

4 Measurable functions

Definition 8. Let X, Y be topological spaces. We will assume X is endowed with a sigma-algebra so that we can meaningfully discuss measurable subsets of X. A function $f : X \to Y$ is

- *continuous* if for every open $O \subset Y$, $f^{-1}(O)$ is open;
- *measurable* if for every open $O \subset Y$, $f^{-1}(O)$ is measurable.

Observation 5. Every continuous function is measurable.

We will concern ourselves with measurable functions into the *extended reals* $\mathbb{R} \cup \{-\infty, +\infty\}$.

Exercise 14. Let $f : X \to \mathbb{R} \cup \{\pm\infty\}$ be a function. TFAE

1. f is measurable;
2. for every $a \in \mathbb{R}$, $f^{-1}(a, +\infty]$ is measurable;
3. for every $a \in \mathbb{R}$, $f^{-1}[a, +\infty]$ is measurable;
4. for every Borel subset $B \subset \mathbb{R} \cup \{\pm\infty\}$, $f^{-1}(B)$ is measurable.

Proof. Clearly (1) $\Rightarrow$ (2). So assume (2). By taking complements, we see $f^{-1}[-\infty, a]$ is measurable $\forall a$. Since
$$f^{-1}([-\infty, a)) = \cup_{r \in \mathbb{Q}, r<a} f^{-1}[-\infty, r]$$
it follows that $f^{-1}([-\infty, a))$ is measurable. By taking complements again we see that $f^{-1}[a, +\infty]$ is measurable. So (2) $\Rightarrow$ (3).

Now assume (3). We will prove (1). Since every open subset is a countable union of open intervals, it suffices to show that $f^{-1}(I)$ is measurable whenever I is an open interval. By taking complements we see that this is true if $I = [-\infty, a)$ for some a. Because $f^{-1}(a, +\infty] = \cup_{r \in Q, r>a} f^{-1}[r, +\infty]$, it is also true if $I = f^{-1}(a, +\infty]$ for some a. So it's true whenever I is an infinite interval. If I is finite then $I = (a, b)$ for some $a, b \in \mathbb{R}$ in which case $f^{-1}(I) = f^{-1}(a, +\infty] \cap f^{-1}[-\infty, b)$. So it's true in this case too.

We have now shown that 1,2,3 are equivalent. Since open sets are Borel, (4) implies (1). To see that (1) implies (4), it suffices to observe that the collection $\mathcal{C}$ of all subsets $A \subset \mathbb{R} \cup \{\pm\infty\}$ such that $f^{-1}(A)$ is measurable forms a sigma-algebra. This is because for any sets $E_1, E_2, \ldots,$

$$f^{-1}(\cup_i E_I) = \cup_i f^{-1}(E_i), f^{-1}(\cap_i E_I) = \cap_i f^{-1}(E_i), f^{-1}(E^c) = f^{-1}(E)^c.$$

□

Exercise 15. Let $f, g, b : X \to \mathbb{R}$ be functions. Suppose f, g are measurable and b is continuous. Then

- $f + g$ is measurable
- fg is measurable
- f/g is measurable if g is never 0.
- $b \circ f$ is measurable

Proof. Let $a \in \mathbb{R}$. It suffices to show $\{x \in \mathbb{R} : f(x) + g(x) > a\}$ is measurable. This follows from

$$\{x \in X : f(x) + g(x) > a\} = \bigcup_{q \in \mathbb{Q}} \{x \in X : f(x) > q\} \cap \{x \in X : g(x) > a - q\}.$$

Similarly, if $a \geq 0$ then

$$\begin{aligned} \{x \in X : f(x)g(x) > a\} &= \bigcup_{q \in \mathbb{Q}, q > 0} \{x \in X : f(x) > q\} \cap \{x \in X : g(x) > a/q\} \\ &\cup \bigcup_{q \in \mathbb{Q}, q < 0} \{x \in X : f(x) < q\} \cap \{x \in X : g(x) < a/q\}. \end{aligned}$$

The case $a < 0$ is similar. This shows the first two items. The third one is similar.

To show $b \circ f$ is measurable, suppose $O \subset \mathbb{R}$ is open. Then $b^{-1}(O)$ is open. So $f^{-1}(O)$ is measurable. □

Exercise 16. Suppose $f_i : X \to \mathbb{R}$ for $i = 1, 2, \ldots$ are measurable functions. Then $\sup_i f_i, \inf_i f_i, \limsup_i f_i$ and $\liminf_i f_i$ are measurable.

Proof. For any $a \in \mathbb{R}$,

$$\{x \in X : \sup_i f_i(x) > a\} = \cap_n \cup_i \{x \in X : f_i(x) > a - 1/n\}.$$

This shows $\sup_i f_i$ is measurable. $\inf_i f_i$ is similar.

$$\{x \in X : \limsup_i f_i(x) \geq a\} = \cap_i \{x \in X : \sup_{j > i} f_j(x) \geq a\}.$$

This shows $\limsup_i f_i$ is measurable. $\liminf_i f_i$ similar. □

Definition 9. Let $f, g : X \to \mathbb{R}$ be functions. We say that $f = g$ almost everywhere (a.e.) if $\{x \in X : \; f(x) \neq g(x)\}$ has measure zero.

Exercise 17. Suppose $f = g$ a.e. If f is measurable then so is g.

Proof. Let $Z = \{x \in X : \; f(x) \neq g(x)\}$ and let $O \subset \mathbb{R} \cup \{\pm\infty\}$ be open. Then $g^{-1}(O) \setminus Z = f^{-1}(O) \setminus Z$. So there exists a subset $Z' \subset Z$ such that $g^{-1}(O) = Z' \cup (f^{-1}(O) \setminus Z)$. Because Z, Z' have measure zero, they are both measurable. Thus $g^{-1}(O)$ is measurable. Since O is arbitrary, this proves g is measurable. □

Definition 10. The *characteristic function* or *indicator function* of a set $E \subset X$ is the function $\chi_E : X \to \mathbb{R}$ given by $\chi_E(x) = 1$ if $x \in E$ and $\chi_E(x) = 0$ otherwise. A *simple function* is a finite linear combination of characteristic functions of measurable subsets.

Exercise 18. Let $f : X \to \mathbb{R}$ be a function. TFAE

1. f is simple
2. f is measurable and the range of f is finite
3. there exist disjoint measurable sets $E_1, \ldots, E_k$ and real numbers $r_1, \ldots, r_k$ such that

$$f(x) = \sum_i r_i \chi_{E_i}.$$

Later on, we will use simple functions to develop integration theory. It will be useful to have the following approximation result:

Exercise 19. Let f be a measurable function on X. Then there exist simple functions $\{f_i\}_i$ such that $f = \lim_i f_i$ pointwise. Moreover, if $f \geq 0$ then we can choose f_i so that $0 \leq f_1 \leq f_2 \leq \cdots \leq f$.

Proof. Define f_n by

$$f(x) = \begin{cases} -n & f(x) < -n \\ k2^{-n} & k2^{-n} \leq f(x) < (k+1)2^{-n} \\ n+1 & f(x) \geq n+1 \end{cases}$$

□

5 Borel functions (tangential and optional)

Definition 11. Let X, Y be topological spaces. A function $f : X \to Y$ is *Borel* if for every open $O \subset Y$, $f^{-1}(O)$ is Borel.

Observation 6. Every continuous function is Borel and every Borel function is measurable (as long as Borel sets are measurable which is usually the case).

Exercise 20. A function $f : X \to Y$ is Borel if and only if for every Borel set $B \subset Y$, $f^{-1}(Y)$ is Borel. Hence compositions of Borel functions are Borel.

Exercise 21. There is a Borel function $f : [0,1] \to [0,1]$ and a measurable set $X \subset [0,1]$ such that $f^{-1}(X)$ is not measurable.

Exercise 22. There are measurable functions $f, g : \mathbb{R} \to \mathbb{R}$ whose composition is not measurable.

6 Semi-continuity (tangential)

We won't use this much but it is good to have in your vocabulary:

Definition 12. $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is *upper semi-continuous* if $f^{-1}[-\infty, a)$ is open for every $a \in \mathbb{R}$ (where X is a topological space). f is *lower semi-continuous* if $f^{-1}(a, +\infty]$ is open for every $a \in \mathbb{R}$.

Exercise 23. Let X be a compact metric space. Then $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is upper semi-continuous if and only if there exist continuous functions $f_1, f_2, \ldots$ such that $f = \inf_i f_i$. Similarly, f is lower semi-continuous if and only if there exist continuous functions $f_1, f_2, \ldots$ such that $f = \sup_i f_i$.

7 Littlewood's 3 principles

Littlewood's three principles are:

1. Every subset of the real line of finite measure is nearly a finite union of intervals.

2. Every measurable function is nearly continuous.

3. Every convergent sequence of functions is nearly uniformly convergent.

Let us make this rigorous:

Exercise 24 (First principle). Suppose $E \subset \mathbb{R}$ is measurable and has finite measure. Prove: for every $\epsilon > 0$ there exists a finite union of open intervals O such that $m(O \triangle E) < \epsilon$.

Proof. There exists an open set $O \supset E$ with $m(O \setminus E) < \epsilon$. Since O is a countable union of intervals, this means there is a finite union of intervals $O' \subset O$ with $m(O \setminus O') < \epsilon$. Therefore $m(O' \triangle E) < 2\epsilon$. □

We will prove:

Theorem 7.1 (Second principle: Lusin's Theorem)**.** *Suppose $f : \mathbb{R} \to \mathbb{R}$ measurable function. For every $\epsilon > 0$ there exists a continuous function g such that*

$$m(\{x \in \mathbb{R} : \ f(x) \neq g(x)\}) < \epsilon.$$

Theorem 7.2 (Third principle: Egorov's Theorem)**.** *Suppose $\{f_n\}$ is a sequence of measurable functions defined on a set $E \subset \mathbb{R}$ of finite measure. Suppose also that $f_n \to f$ pointwise a.e. For every $\epsilon > 0$ there exists a subset $B \subset E$ such that $m(E \setminus B) < \epsilon$ and f_n converges to f* uniformly *on $E \setminus B$.*

We will prove Egorov's Theorem first. So suppose $\{f_n\}, f, E$ are as in Egorov's Theorem.

Exercise 25. For every $\delta > 0$ there exists a measurable set $B \subset E$ and an integer $N > 0$ such that for every $n > N$,

$$m(\{x \in E : \ |f_n(x) - f(x)| \geq \delta\}) \leq \delta.$$

Hint: let

$$G_n = \{x \in E : \ |f_n(x) - f(x)| \geq \delta\}$$

and set $E_N = \cup_{n=N}^{\infty} G_n$. Prove $\lim_{N\to\infty} m(E_N) = 0$ and derive the result from this.

Proof. Note that $E_N \supset E_{N+1} \supset \cdots$ is a decreasing sequence. So

$$m(\cap_N E_N) = \lim_{N\to\infty} m(E_N).$$

Suppose $x \in \cap_N E_N$. Then $\limsup_{N\to\infty} |f_n(x) - f(x)| \geq \delta$. However, the latter is satisfied only on a measure zero set. So

$$0 = m(\cap_N E_N) = \lim_{N\to\infty} m(E_N).$$

Choose N large enough so that $m(E_N) \leq \delta$. Because

$$\{x \in E : \ |f_n(x) - f(x)| \geq \delta\} \subset E_N$$

(for $n \geq N$) we're done. □

Exercise 26. Prove Egorov's Theorem. Hint: apply the previous exercise repeatedly with $\delta_n = 2^{-n}\epsilon$.

Proof. By the previous exercise, there exist measurable sets B_n and integers N_n such that

- $m(B_n) < 2^{-n}\epsilon$
- for every $k > N_n$ and $x \in E \setminus B_n$,

$$|f_k(x) - f(x)| < 2^{-n}\epsilon.$$

Let $B_\infty = \cup_n B_n$. So $m(B_\infty) < \epsilon$ and if $x \in E \setminus B_\infty$ and $k > N_n$ then

$$|f_k(x) - f(x)| < 2^{-n}\epsilon.$$

This proves $\{f_k\}$ converges uniformly to f on $E \setminus B_\infty$. □

We now start with the proof of Lusin's theorem.

Exercise 27. If $f : E \to \mathbb{R}$ is simple and $\epsilon > 0$ then there exists a closed set $F \subset E$ such that $m(E \setminus F) < \epsilon$ and $f \upharpoonright F$ is continuous.

Proof. Because f is simple $f = \sum_{i=1}^n c_i \chi_{E_i}$ for some coefficients c_i and some pairwise disjoint measurable sets E_i. Let $F_i \subset E_i$ be closed sets with $m(E_i \setminus F_i) < \epsilon/n$. Let $F = \cup F_i$. Note $m(E \setminus F) < \epsilon$ and f restricted to F is continuous. $\square$

Exercise 28. If $f : \mathbb{R} \to \mathbb{R}$ is measurable and $\epsilon > 0$ then there exists a closed set $F \subset \mathbb{R}$ such that $m(\mathbb{R} \setminus F) < \epsilon$ and $f \upharpoonright F$ is continuous.

Proof. Let us first assume that $m(E) < \infty$.

Let $\{f_k\}$ be a sequence of simple functions that converge pointwise to f. For each k, let $F_k \subset E$ be a closed set with $m(\mathbb{R} \setminus F_k) < \epsilon/2^k$ such that $f_k \upharpoonright F_k$ is continuous. By Egorov's Theorem there exists $B \subset \mathbb{R}$ such that $m(B) < \epsilon$ and $\{f_k\}$ converges uniformly to f on $\mathbb{R} \setminus B$. Wlog, we may require B to be open so that $F_0 := \mathbb{R} \setminus B$ is closed. (if B is not open we may replace it with an open set $O \supset B$ such that $m(O) < \epsilon$).

Because a uniformly convergent sequence of continuous functions is continuous, $f \upharpoonright F_\infty := F_0 \cap \bigcap_k F_k$ is continuous. Note

$$m(E \setminus F_\infty) \le \sum_{k=0}^{\infty} m(E \setminus F_k) < 2\epsilon.$$

So this handles the case $m(E) < \infty$. To obtain the general case, we find finite measure sets $E_0 \subset E_1 \subset E_2 \subset \cdots \subset \mathbb{R}$ such that $\mathbb{R} = \cup E_i$. Then for each i there exists a closed set $F_i \subset E_i$ with $m(E_i \setminus F_i) < \epsilon/2^i$ such that $f \upharpoonright F_i$ is continuous. Note that $m(\mathbb{R} \setminus \cup_i F_i) < \epsilon$. So there exists finite I such that $m(\mathbb{R} \setminus \cup_{i=0}^I F_i) < \epsilon$. Observe that f is continuous on $\cup_{i=0}^I F_i$. $\square$

Exercise 29. Prove Tietze's Extension Theorem: if $F \subset \mathbb{R}$ is any closed set and $f : F \to \mathbb{R}$ any continuous function then there exists a continuous function $g : \mathbb{R} \to \mathbb{R}$ whose restriction to F equals f.

Proof. There exist pairwise disjoint open intervals $O_1, O_2, \ldots$ such that the complement of F is the union $\cup_i O_i$. Write $O_i = (a_i, b_i)$. Now define g by $g(x) = f(x)$ if $x \in F$ and

$$g(x) = \left(\frac{b_i - x}{b_i - a_i}\right) f(a_i) + \left(\frac{x - a_i}{b_i - a_i}\right) f(b_i)$$

if $x \in (a_i, b_i)$. $\square$

Remark 1. Tietze's Extension Theorem holds in much greater generality: you can allow F to be a closed subset of any normal topological space. (Normal means that every two disjoint closed subsets have disjoint open neighborhoods. For example, metric spaces are normal).

Lusin's Theorem is an immediate consequence of the previous two exercises.

Lusin's Theorem can be generalized a great deal. To explain we need:

Definition 13. A measure m on a topological space X is *regular* if for every measurable set $E \subset X$, every $\epsilon > 0$ there exists an open set $O \supset E$ and a compact set $F \subset E$ such that

$$m(O \setminus E) < \epsilon, \quad m(E \setminus F) < \epsilon.$$

For example, Lebesgue measure is regular.

Regular measures are common:

Theorem 7.3. *Every Borel measure on a locally compact secound countable space is regular.*

(We won't prove this here).

Exercise 30 (Lusin's Theorem)*.* Suppose X is a normal Hausdorff topological space and m is a σ-finite regular measure on X. Then for any measurable function $f : X \to \mathbb{R}$ and $\epsilon > 0$ there exists a continuous $g : X \to \mathbb{R}$ such that

$$m(\{x \in X : \ f(x) \neq g(x)\}) < \epsilon.$$

7.1 An aside

It can be shown that a measurable function $f : [a, b] \to \mathbb{R}$ is Riemann-integrable if and only if the set of all $x \in [a, b]$ such that f is discontinuous at x has measure zero.

Example: Let $f = \chi_{\mathbb{Q}}$. Claim: f is discontinuous everywhere.

Proof: For any $x \in \mathbb{R}$ and any $\delta > 0$ we have that $(x - \delta, x + \delta) \cap \mathbb{Q} \neq \emptyset$ and $(x - \delta, x + \delta) \cap \mathbb{Q}^c \neq \emptyset$. So there y, z with $|x - y| < \delta$ such that $y \in \mathbb{Q}$ and $z \notin \mathbb{Q}$. Then either $|f(x) - f(y)| = 1$ or $|f(x) - f(z)| = 1$. In any case, f is discontinuous at x.

8 Convergence in measure

Definition 14. Let $f_1, f_2, \dots, f_\infty$ be measurable functions all defined on some measurable subset $E \subset \mathbb{R}$. We say

- $\{f_n\}$ *converges pointwise* to f_∞ if $f_\infty(x) = \lim_n f_n(x)$ for every $x \in E$;
- $\{f_n\}$ *converges pointwise a.e.* to f_∞ if $f_\infty(x) = \lim_n f_n(x)$ for a.e. $x \in E$;
- $\{f_n\}$ *converges in measure* to f_∞ if for every $\epsilon > 0$,

$$\lim_n m(\{x \in X : \ |f_n(x) - f_\infty(x)| > \epsilon\}) = 0.$$

Exercise 31*.* Suppose E has finite measure and $f_n \to f_\infty$ on E pointwise a.e. Show that $f_n \to f_\infty$ in measure.

Proof. This follows from Egorov's Theorem. Note: it is necessary that E have finite measure. □

Exercise 32. Construct measurable functions $f_1, f_2, \dots, f_\infty : [0,1] \to \mathbb{R}$ such that $f_n \to f_\infty$ in measure but f_n does not converge pointwise a.e.

Proof. Define $f_i = \chi_{E_i}$ where $E_i \subset [0,1]$ is measurable, $m(E_i) \to 0$ and $\limsup_i E_i = [0,1]$. □

Exercise 33. Suppose $f_n \to f_\infty$ in measure. Show there exists a subsequence $\{f_{n_i}\}$ such that $f_{n_i} \to f_\infty$ pointwise a.e.

Proof. Let n_i be large enough so that if $E_i = \{x \in X : \ |f_{n_i}(x) - f(x)| > 1/i\}$ then

$$m(E_i) \le 2^{-i}.$$

Let $Z = \{x \in X : \ \limsup_i |f_{n_i}(x) - f(x)| > 0\} = \limsup_i E_i = \cap_{j=1}^\infty \cup_{i>j} E_i$. By continuity $m(Z) = 0$. So $f_{n_i} \to f_\infty$ pointwise a.e. □

9 Integration for bounded functions

We will assume throughout that $(X, \mathcal{C}, m)$ is a measure space.

Definition 15. Let $f = \sum_{i=1}^n c_i \chi_{E_i}$ be a simple function with finite measure support. Assume that the sets $E_1, \dots, E_n$ are pairwise disjoint and $c_i \ne c_j$ if $i \ne j$. (This property determines the E_i's and c_i's from f uniquely up to pemuting the indices). Then we define

$$\int f \ dm = \sum_{i=1}^n c_i m(E_i).$$

In other words

$$\int f \ dm = \sum_{c \in \mathbb{R}} c m(f^{-1}(c)).$$

Exercise 34. Suppose f, g are simple functions (each with finite measure support) and $a, b \in \mathbb{R}$. Then $af + bg$ is simple and

$$\int af + bg \ dm = a \int f \ dm + b \int g \ dm.$$

Proof. It's clear that $\int af \ dm = a \int f \ dm$ so it suffices to prove $\int f + g \ dm = \int f \ dm + \int g \ dm$.

We observe

$$
\begin{aligned}
\int f+g\ dm &= \sum_{c\in\mathbb{R}} cm(\{x\in X:\ f(x)+g(x)=c\}) \\
&= \sum_{c\in\mathbb{R}} c\sum_{b\in\mathbb{R}} m(\{x\in X:\ f(x)=b, g(x)=c-b\}) \\
&= \sum_{b\in\mathbb{R}}\sum_{c\in\mathbb{R}} (b+c)m(\{x\in X:\ f(x)=b, g(x)=c\}) \\
&= \sum_{b\in\mathbb{R}}\sum_{c\in\mathbb{R}} bm(\{x\in X:\ f(x)=b, g(x)=c\}) + \sum_{b\in\mathbb{R}}\sum_{c\in\mathbb{R}} cm(\{x\in X:\ f(x)=b, g(x)=c\}) \\
&= \sum_{b\in\mathbb{R}} bm(\{x\in X:\ f(x)=b\}) + \sum_{c\in\mathbb{R}} cm(\{x\in X:\ g(x)=c\}) \\
&= \int f\ dm + \int g\ dm.
\end{aligned}
$$

□

Exercise 35. Suppose $E\subset X$ has finite measure and $f: E\to\mathbb{R}$ is a bounded measurable function.

$$\sup_{\phi\le f}\int \phi\ dm = \inf_{f\le\psi}\int\psi\ dm$$

where the sup and inf are over simple functions from E to $\mathbb{R}$.

Proof. It's easy to see that $\le$ must occur. Indeed, if $\phi\le f\le\psi$ are as above and

$$\phi=\sum_{i=1}^n c_i\chi_{E_i},\quad \psi=\sum_{i=1}^m d_i\chi_{F_i}$$

then after refining the partition $\{E_1,\dots,E_n\}$, $\{F_1,\dots,F_m\}$ we may assume that they are equal. That is, we may assume that

$$\phi=\sum_{i=1}^k c_i\chi_{G_i},\quad \psi=\sum_{i=1}^k d_i\chi_{G_i}$$

for some measuable sets $G_1,\dots,G_k$. Since $c_i\le d_i$, the inequality $\le$ follows.

Because f is bounded there is some $M>0$ such that $|f(x)|\le M$ for all x. For $n>0$ and $k\in\mathbb{Z}, |k|\le 2^n$ let

$$E_{n,k}=\{x\in X:\ k2^{-n}\le f(x)/M<(k+1)2^{-n}\}.$$

Let $\phi_n(x)=Mk2^{-n}$ and $\psi_n(x)=M(k+1)2^{-n}$ on $E_{n,k}$. These are both simple functions and by design $|\phi_n-\psi_n|\le M2^{-n}$. So

$$\left|\int\phi_n\ dm-\int\psi_n dm\right|\le M2^{-n}\sum_k m(E_{n,k})=M2^{-n}m(E).$$

Since this tends to 0 as $n\to\infty$, it proves the exercise. □

Definition 16. With f, X as above we define

$$\int_E f\ dm = \sup_{\phi \le f} \int \phi\ dm = \inf_{f \le \psi} \int \psi\ dm.$$

If $Y \subset X$ is measurable then

$$\int_Y f\ dm = \int f\chi_Y\ dm.$$

If $[a,b] \subset X \subset \mathbb{R}$ is an interval and m is Lebesgue measure then

$$\int_{[a,b]} f\ dm = \int_a^b f(x)\ dx.$$

If $X = E$ or if E is understood from the context then we usually drop the subscript and simply write $\int f\ dm$.

Exercise 36. Suppose $E \subset X$ has finite measure and $f, g : E \to \mathbb{R}$ are bounded measurable functions and $a, b \in \mathbb{R}$. Then

1. (linearity) $\int_E af + bg\ dm = a\int f\ dm + b\int g\ dm$,
2. (monotonicity) if $f \le g$ a.e. then $\int f\ dm \le \int g\ dm$. So $f = g$ a.e. then $\int f\ dm = \int g\ dm$.
3. If $A \le f \le B$ then $Am(E) \le \int f\ dm \le Bm(E)$,
4. (finitely additive) if $A_1, A_2, \dots, A_n \subset X$ are disjoint, measurable and $\cup_i A_i = E$ then $\int f\ dm = \sum_i \int_{A_i} f\ dm$.

Proof. Suppose that $a, b > 0$. If $\phi_1 \le f \le \psi_1$ and $\phi_2 \le g \le \psi_2$ are simple functions then

$$a\phi_1 + b\phi_2 \le af + bg \le a\psi_1 + b\psi_2$$

are simple functions and

$$\int_E a\phi_1 + b\phi_2\ dm = a\int \phi_1\ dm + b\int \phi_2\ dm$$

and a similar formula holds with ψ_i in place of ϕ_i. Using the definition of the integral, this implies (1).

To prove (2) we simply notice that if $\phi \le f$ is a simple function since $\phi \le g$ we also have $\int \phi\ dm \le \int g\ dm$ by definition. So taking the sup over all $\phi \le f$ we obtain (2).

Item (3) follows from item (4) by letting f (or g) be a constant.

To see item (4), observe that $f = \sum_i f\chi_{A_i}$. So (4) follows from (1). $\square$

10 Integration for nonnegative functions

Definition 17. Let $f : X \to [0, \infty]$ be measurable (where $(X, \mathcal{C}, m)$ is a measure space). We define

$$\int f\ dm = \sup_{g \le f} \int g\ dm$$

where the sup is over all bounded measurable functions $g \le f$ with finite measure support.

Exercise 37. Suppose $f, g : X \to [0, \infty]$ are measurable and $c > 0$ is a constant. Then

1. $\int cf\ dm = c \int f\ dm$
2. $\int f + g\ dm = \int f\ dm + \int g\ dm$.
3. $f \le g \Rightarrow \int f\ dm \le \int g\ dm$.
4. if $X = A \cup B$ where A, B are disjoint measurable sets, then
$$\int f\ dm = \int_A f\ dm + \int_B f\ dm.$$

Proof. The last two inequalities are obvious. To prove the others, let $f' \le f, g' \le g$ be bounded measurable functions with finite measure supports. Because $cf' \le cf$ and $f' + g' \le f + g$ are bounded measurable functions with finite supports we have

$$c \int f'\ dm = \int cf'\ dm \le \int cf\ dm$$

$$\int f'\ dm + \int g'\ dm = \int f' + g'\ dm \le \int f + g\ dm.$$

Taking the sup over all such f', g' yields

$$\int cf\ dm \ge c \int f\ dm$$

$$\int f + g\ dm \ge \int f\ dm + \int g\ dm.$$

It follows that

$$\int f\ dm = \int (1/c)cf\ dm \ge (1/c) \int cf\ dm \ge (1/c)c \int f\ dm = \int f\ dm.$$

Since equality holds throughout, we must have that $\int cf\ dm = c \int f\ dm$ as claimed.

Let $k \le f + g$ be a bounded measurable function with finite measure support. Let $k' = \min(f, k)$ and $k'' = k - k'$. Then $k' \le f$ and $k'' \le g$ (since $k \le f + g$ implies $k - f \le g$ implies $k - k' \le g$). So

$$\int k'\ dm + \int k''\ dm = \int k\ dm \le \int f + g\ dm.$$

Taking the sup over all k proves the exercise. $\square$

Exercise 38. Let $f : \mathbb{R} \to \mathbb{R}$ be a nonnegative function and $R = \{(x, y) : \ 0 \le y \le f(x)\}$ be the region under the graph of f. Then

$$\int f \ dm = m(R).$$

Proof. By linearity, it suffices to prove this in the special case in which f has finite support.

This is clear if f is the characteristic function of an interval. Therefore, it is also true if f is a finite linear combination of such characteristic functions. Since any measurable set with finite measure is nearly a finite union of intervals (Littlewood's first principle), it is also true if $f = \chi_E$ where E is measurable. Therefore it is true if f is simple.

Using Egorov's Theorem and approximation of f by simple functions, we obtain the result for arbitrary f.

□

11 Integrable functions

Definition 18. A measurable function $f : X \to \mathbb{R}$ is *integrable* if

$$\int |f| \ dm < \infty.$$

Because $|f|$ is a nonnegative measurable function, $\int |f| \ dm$ is well-defined. In this case, let

$$f^+ = \max(f, 0)$$

$$f^- = \max(-f, 0).$$

Note: f^+, f^- are measurable functions, $f = f^+ - f^-$ and $|f| = f^+ + f^-$. Define

$$\int f \ dm = \int f^+ \ dm - \int f^- \ dm.$$

Exercise 39. Suppose $f, g : X \to \mathbb{R}$ are integrable and $c \in \mathbb{R}$ is a constant. Then

1. cf is integrable and $\int cf \ dm = c \int f \ dm$,
2. $f + g$ is integrable and $\int f + g \ dm = \int f \ dm + \int g \ dm$,
3. if $f \le g$ a.e. then $\int f \ dm \le \int g \ dm$. In particular if $f = g$ a.e. then $\int f \ dm = \int g \ dm$.
4. if $X = A \cup B$ where A, B are disjoint measurable sets, then

$$\int f \ dm = \int_A f \ dm + \int_B f \ dm.$$

Proof. Suppose that $c > 0$. Then $(cf)^+ = cf^+$ and $(cf)^- = cf^-$. This implies (1) when $c > 0$. The case $c \le 0$ is similar.

To prove (2), suppose $f = f_1 - f_2$ where f_1, f_2 are nonnegative measurable functions. Then

$$f = f^+ - f^- = f_1 - f_2$$

implies $f^+ + f_2 = f^- + f_1$. By the previous section on integrals of nonnegative functions,

$$\int f^+ \, dm + \int f_2 \, dm = \int f^+ f_2 \, dm = \int f^- + f_1 \, dm = \int f^- \, dm + \int f_1 \, dm.$$

Therefore,

$$\int f \, dm = \int f^+ \, dm - \int f^- \, dm = \int f_1 \, dm - \int f_2 \, dm.$$

Now let us apply this to the situation at hand:

$$\begin{aligned}
\int f + g \, dm &= \int f^+ - f^- + g^+ - g^- \, dm \\
&= \int (f^+ + g^+) - (f^- + g^-) \, dm \\
&= \int (f^+ + g^+) \, dm - \int (f^- + g^-) \, dm \\
&= \int f^+ \, dm + \int g^+ \, dm - \left(\int f^- \, dm + \int g^- \, dm \right) \\
&= \int f \, dm + \int g \, dm.
\end{aligned}$$

This proves (2). (3) follows from (2) since

$$0 \le \int g - f \, dm = \int g \, dm - \int f \, dm.$$

(4) follows from (2) since $f = \chi_A f + \chi_B f$. □

12 Convergence Theorems

There are four convergence theorems which state that if f_n converges in some sense to a function f then $\int f_n \, dm \to \int f \, dm$ (or at least we have an inequality if not an equality).

Exercise 40 (Bounded Convergence Theorem). Suppose $m(X) < \infty$. Let $\{f_n\}$ be a uniformly bounded sequence of measurable functions on X that converge pointwise a.e. to f. (uniformly bounded means there is a number $M > 0$ such that $|f_n(x)| \le M$ for a.e. x and every n). Then

$$\int f_n \, dm \to \int f \, dm$$

as $n \to \infty$.

Proof. This follows from Egorov's Theorem. Indeed, if $\epsilon > 0$ then, by Egorov's Theorem there is a set $Y \subset X$ such that $m(X \setminus Y) < \epsilon/(4M)$ and an N such that $n > N$ implies

$$|f_n(x) - f(x)| \leq \epsilon/(2\mu(X))$$

for all $x \in Y$. Therefore

$$\begin{aligned} |\int f_n(x)\ dm - \int f(x)\ dm| &\leq \int |f_n(x) - f(x)|\ dm \\ &= \int_Y |f_n(x) - f(x)|\ dm + \int_{X\setminus Y} |f_n(x) - f(x)|\ dm \\ &\leq (\epsilon/(2\mu(X)))\mu(X) + (\epsilon/(4M))(2M) < \epsilon. \end{aligned}$$

□

Exercise 41 (Fatou's Lemma). Let $\{f_n\}$ be an sequence of nonnegative functions. Then

$$\liminf_n \int f_n\ dm \geq \int \liminf_n f_n\ dm$$

as $n \to \infty$. Hint: use the definition of $\int \liminf_n f_n\ dm$ directly.

Proof. Let $h \leq \liminf_n f_n$ be a bounded measurable function with finite measure support. Let

$$h_n = \min(h, f_n).$$

Note $h_n \to h$ pointwise. So the bounded convergence theorem implies

$$\int h\ dm = \lim_n \int h_n\ dm \leq \liminf_n \int f_n\ dm$$

Take the sup over all h to obtain the exercise. □

Exercise 42. Find an example of a sequence of functions as in Fatou's Lemma such that the inequality is strict.

Proof. For example, we could have $f_n = (+\infty)\chi_{[0,1/n]}$ or $f_n = \chi_{[n,+\infty)}$. □

Exercise 43 (Monotone Convergence Theorem). Let $\{f_n\}$ be an increasing sequence of nonnegative functions. (This means $0 \leq f_1 \leq f_2 \leq \cdots$). Let $f(x) = \lim_n f_n(x) \in [0, +\infty]$. Then

$$\int f_n\ dm \to \int f\ dm$$

as $n \to \infty$.

Proof. By Fatou's Lemma,

$$\liminf_n \int f_n\ dm \geq \int f\ dm.$$

But for each n we have $f_n \leq f$ which implies $\int f_n\ dm \leq \int f\ dm$. Taking limits implies the opposite inequality. □

Exercise 44. If $X_1, X_2, \ldots$ are pairwise disjoint measurable sets whose union is X and $f \geq 0$ is measurable then

$$\int f\ dm = \sum_{i=1}^{\infty} \int_{X_i} f\ dm.$$

(therefore if $\mu(E) = \int_E f\ dm$ then μ is a measure on X). More generally, if $\{f_n\}$ are nonnegative functions and $f = \sum_n f_n$ then

$$\int f\ dm = \sum_n \int f_n\ dm.$$

Exercise 45 (Lebesgue's Dominated Convergence Theorem). Let $\{f_n\}$ be a sequence of measurable functions on X that converge pointwise a.e. to f. Suppose there is an integrable function g such that $0 \leq f_n \leq g$ for all n. Then

$$\int f_n\ dm \to \int f\ dm$$

Proof. Proof #1:One way to do this is to realize that $g\ dm$ is a measure. That is: we define a new measure μ on X by

$$\mu(E) = \int_E g\ dm.$$

We've already proven that this really defines a measure. Moreover, if k is any bounded measurable function then

$$\int k\ d\mu = \int kg\ dm.$$

(To see this, note that it's true for characteristic functions and therefore true for simple functions and therefore true for arbitrary bounded measurable functions).

After replacing X with the support of g if necessary we may assume that $g > 0$ a.e. By the Bounded Convergence Theorem,

$$\int f_n/g\ d\mu \to \int f/g\ d\mu.$$

(This is because $\mu(X) = \int g\ dm < \infty$ and $0 \leq f_n/g \leq 1$ is bounded). Since $\int k\ d\mu = \int kg\ dm$ for any bounded measurable function k, this implies the exercise.

Proof #2: By Fatou's Lemma,

$$\liminf_n \int f_n\ dm \geq \int f\ dm.$$

On the other hand, $g - f_n \geq 0$ converges pointwise a.e. to $g - f$. So Fatou's Lemma implies

$$\liminf_n \int g - f_n\ dm \geq \int g - f\ dm$$

In other words,

$$\limsup_n \int f_n \leq \int f\ dm.$$

□

Exercise 46. The hypotheses of Lebesgue's Dominated Convergence Theorem can be weakened by replacing $0 \le f_n \le g$ with $h \le f_n \le g$ where h, g are integrable functions.

Proof. Observe that $0 \le f_n - h \le g - h$. By the previous exercise, $\int f_n - h\, dm \to \int f - h\, dm$. Cancelling $\int -h\, dm$ from both sides proves the result. $\square$

Definition 19. If (X, m) is a measure space and $p > 0$, we let $L^p(X, m)$ be the set of all equivalence classes classes of measurable functions $f : X \to \mathbb{C}$ such that

$$\int |f|^p\, dm < \infty.$$

Here: two functions f, g are equivalent if they agree a.e. By abuse of notation, we may write $f \in L^1(X, m)$ to mean that f is integrable (for example). Later on, we will show that if $p \ge 1$ and $f, g \in L^p(X, m)$ then

$$\|f - g\|_p = \left(\int |f - g|^p\, dm \right)^{1/p}$$

is a metric on $L^p(X, m)$. We will study the topological aspects of these spaces later.

Also, we let $L^\infty(X, m)$ denote the set of equivalence classes of bounded measurable functions. We write

$$\|f\|_\infty = \sup\{a \ge 0 : \ m(f^{-1}[a, \infty]) > 0\}.$$

(This is called the essential supremum of f). This also gives a metric on $L^\infty(X, m)$: the distance between f and g is $\|f - g\|_\infty$.

13 Riemannian integration

Let f be a function on an interval $[a, b] \subset \mathbb{R}$. Let $\Sigma = \{x_0, x_1, \ldots, x_n\} \subset [a, b]$ with $x_0 = a, x_n = b$. Define

$$I_\Sigma^+(f) = \sum_{i=1}^{n} (\sup f \restriction [x_{i-1}, x_i]) |x_i - x_{i-1}|$$

and

$$I_\Sigma^-(f) = \sum_{i=1}^{n} (\inf f \restriction [x_{i-1}, x_i]) |x_i - x_{i-1}|.$$

Let $\operatorname{diam}(\Sigma) = \max_i |x_i - x_{i-1}|$. We define the Riemannian integral of f (if it exists) by

$$R \int f\, dx = \lim_{\operatorname{diam}(\Sigma) \to 0} I_\Sigma^+(f) = \lim_{\operatorname{diam}(\Sigma) \to 0} I_\Sigma^-(f).$$

We say f is *Riemann-integrable* if its Riemann integral exists and is finite.

We will prove:

Theorem 13.1. *Suppose $f : [a, b] \to \mathbb{R}$ is Riemann-integrable. Then f is integrable and*

$$R \int f \ dx = \int f \ dm.$$

Exercise 47. If f is measurable, $f \geq 0$ a.e. and $\int f \ dm = 0$ then $f = 0$ a.e.

Exercise 48. Suppose $f : X \to \mathbb{R}$ is a function satisfying

$$\sup_{\phi \leq f} \int \phi \ dm = \inf_{\psi \geq f} \int \psi \ dm$$

where the sup is over all simple functions $\phi \leq f$ and the inf is over all simple functions $\psi \leq f$. Then f is measurable.

Hint: there exist simple functions ϕ_n, ψ_n such that $\phi_n \leq f \leq \psi_n$ and $\int \psi_n - \phi_n \ dm \leq 1/n$. Let $\phi^* = \sup_n \phi_n$ and $\psi^* = \inf_n \psi_n$.

Exercise 49. Finish the proof of Theorem 13.1.

Proof. If R is Riemann-integrable then

$$\sup_{\phi \leq f} \int \phi \ dm = \inf_{\psi \geq f} \int \psi \ dm$$

where the sup is over all step functions $\phi \leq f$ and the inf is over all step functions $\psi \leq f$. So the previous exercise implies f is measurable.

Now

$$\sup_{\phi \leq f} \int \phi \ dm \leq \sup_{\phi \leq f} \int \phi \ dm \leq \int f \ dm \leq \inf_{\psi \geq f} \int \psi \ dm \leq \inf_{\psi \geq f} \int \psi \ dm$$

where the first sup and the last inf are over step functions and the second sup and first inf are over simple functions. Since the first and last quantity are equal to $R \int f \ dm$ this proves

$$R \int f \ dx = \int f \ dm.$$

$\square$

Exercise 50. A bounded function $f : [a, b] \to \mathbb{R}$ is Riemann-integrable if and only if the set of discontinuities of f has measure zero.

14 Differentiation

Motivation:

1. when does the derivative f' exist?

2. when does $\int_a^b f'\ dx = f(b) - f(a)$?

In this subsection we will show that $\frac{d}{dx}(\int_a^x f\ dm)$ exists a.e. under very general conditions. First,

Exercise 51. Let $f : X \to \mathbb{R}$ be an integrable function where $(X, \mathcal{C}, m)$ is a measure space. For every $\epsilon > 0$ there exists $\delta > 0$ such that if $E \subset X$ satisfies $m(E) < \delta$ then $\int_E |f|\ dm \le \epsilon$.

(In Wheedon-Zygmund this is referred to as "absolute continuity" of the "set function" $E \mapsto \int_E f\ dm$.)

Our next goal is to prove:

Theorem 14.1 (Lebesgue's Differentiation Theorem)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be integrable. Then for a.e. $x \in \mathbb{R}^n$,*

$$f(x) = \lim_{Q \searrow x} \frac{1}{m(Q)} \int_Q f\ dm$$

where the limit is over cubes Q centered at x.

We will need the following approximation result:

Exercise 52. Let $f \in L^1(\mathbb{R}^n)$. Then for every $\epsilon > 0$ there exists a continuous function g with compact support such that $\|f - g\|_1 \le \epsilon$. In other words, $C_c(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$ where $C_c(\mathbb{R}^n)$ denotes the subspace of compactly-supported continuous functions.

Definition 20. The Hardy-Littlewood maximal function of f is

$$f^*(x) = \sup_Q \frac{1}{m(Q)} \int_Q |f|\ dm$$

where the sup is over all cubes centered at x.

Exercise 53. Let $E \subset \mathbb{R}^n$ be bounded with positive measure. Let $f(x) = \chi_E(x)$. Prove that f^* is not integrable.

Definition 21. We say a measurable function f is in weak L^1 if there is a constant $C > 0$ such that

$$m(\{x \in \mathbb{R}^n :\ |f(x)| \ge t\}) \le C \frac{\|f\|_1}{t}$$

for all $t > 0$.

Exercise 54. Prove that all integrable functions are in weak L^1 (with constant $C = 1$). This is known as Chebyshev's inequality.

We will prove:

Theorem 14.2 (Hardy-Littlewood maximal inequality)**.** *For any $f \in L^1(\mathbb{R}^n)$ and any $t > 0$,*

$$m(\{x \in \mathbb{R}^n :\ f^*(x) \ge t\}) \le C_n \frac{\|f\|_1}{t}$$

where $C_n > 0$ is a contant depending only on the dimension n. In particular, f^* is in weak L^1.

First we need:

Exercise 55 (The Simple Vitali Covering Lemma). Let $E \subset \mathbb{R}^n$ satisfy $m^*(E) < \infty$. Let $\mathcal{C}$ be a collection of cubes covering E. Then there exists a constant $C_n > 0$ (depending only on n) and a subcollection $\mathcal{C}' \subset \mathcal{C}$ such that

- the cubes in $\mathcal{C}'$ are pairwise disjoint
- $m(\cup\{Q : Q \in \mathcal{C}'\}) \geq C_n m^*(E)$.

Hint: use a greedy algorithm.

Exercise 56. Prove the Hardy-Littlewood maximal inequality in the special case. Hint: for each x with $f^*(x) > t$ choose a cube Q_x such that $\frac{1}{m(Q_x)} \int_{Q_x} |f| \, dm \geq t$. These cubes cover $\{x \in \mathbb{R}^n : f^*(x) \geq t\}$. Use Vitali's Covering Lemma. What can you say about $\int_E |f| \, dm$ where $E = \cup\{Q : Q \in \mathcal{C}'\}$?

Proof. Following the hint, we observe that,

$$\int_E |f| \, dm \leq \int |f| \, dm = \|f\|_1.$$

On the other hand, because the cubes in $\mathcal{C}'$ are pairwise disjoint,

$$\int_E |f| \, dm = \sum_{Q \in \mathcal{C}'} \int_Q |f| \, dm \geq \sum_{Q \in \mathcal{C}} tm(Q) = tm(E).$$

Thus $m(E) \leq \frac{\|f\|_1}{t}$. By the way we chose E, $m(E) \leq (1/C_n) m(\{x : f^*(x) \geq t\})$. So

$$m(\{x : f^*(x) \geq t\}) \leq C_n \frac{\|f\|_1}{t}.$$

□

Definition 22. A point $x \in \mathbb{R}^n$ is a *Lebesgue point* of a locally integrable function f if

$$\lim_{Q \searrow 0} \frac{1}{m(Q)} \int_Q |f(y) - f(x)| \, dm(y) = 0.$$

Exercise 57. For any Lebesgue measurable function f, almost every point is a Lebesgue point. Note: this is formally stronger than Lebesgue's Differentation Theorem. Hint: First consider the case in which f is continuous with compact support. Then approximation arbitrary f by compact supported continuous functions. Use the maximal inequality to control the error term.

Proof. The special case in which f is continuous with compact support is easy. So let us assume f is arbitrary.

Let $\epsilon > 0$ and let g be a continuous function with compact support. Let

$$E_\epsilon = \{x \in X : \limsup_{Q \searrow 0} \frac{1}{m(Q)} \int_Q |f(x) - f(y)| \, dm(y) > \epsilon\}.$$

By the triangle inequality,

$$|f(x) - f(y)| \leq |f(x) - g(x)| + |g(x) - g(y)| + |g(y) - f(y)|.$$

Because

$$\limsup_{Q \searrow 0} \frac{1}{m(Q)} \int_Q |g(x) - g(y)| \, dm = 0,$$

$E_\epsilon \subset E'_\epsilon \cup E''_\epsilon$ where

$$E'_\epsilon = \{x \in X : |f(x) - g(x)| \geq \epsilon/2\}$$

$$E''_\epsilon = \{x \in X : \limsup_{Q \searrow x} |\frac{1}{m(Q)} \int_Q |g - f| \, dm| \geq \epsilon/2\}.$$

Note

$$m(E'_\epsilon) \leq \frac{2}{\epsilon} \int_E |f(x) - g(x)| \, dm \leq \frac{2}{\epsilon} \|f - g\|_1.$$

Note that

$$E''_\epsilon \subset \{x \in X : (g - f)^*(x) \geq \epsilon/2\}.$$

So the maximal inequality implies $m(E''_\epsilon) \leq \frac{2C_n}{\epsilon} \|f - g\|_1$ for some constant $C_n > 0$. Thus

$$m(E_\epsilon) \leq m(E'_\epsilon) + m(E''_\epsilon) \leq \frac{2C_n + 2}{\epsilon} \|f - g\|_1.$$

By a previous exercise, we can make $\|f - g\|_1$ as small as we wish. Therefore $m(E_\epsilon) = 0$. Since ϵ is arbitrary, this implies the theorem. □

Remark 2. The proof above is typical of a general strategy involving maximal inequalities: suppose there is a subset $\mathcal{P} \subset L^1(X)$. We wish to prove that $\mathcal{P} = L^1(X)$. (In the Differentiation Theorem, $\mathcal{P}$ is the set of all functions f satisfying the Differentiation Theorem). First we prove that $\mathcal{P}$ contains a dense subset (in the example, the dense subset was the set of compactly-supported continuous functions). Second we use a maximal inequality to prove that $\mathcal{P}$ is closed. This same strategy is employed to prove Birkhoff's pointwise ergodic theorem.

Exercise 58. A measurable function f on $\mathbb{R}^n$ is *locally integrable* if $\int_K |f| \, dm < \infty$ for every compact $K \subset \mathbb{R}^n$. The conclusion to Lebesgue's Differentiation Theorem holds for locally integrable functions.

Exercise 59. For any measurable subset $E \subset \mathbb{R}^n$,

$$\chi_E(x) = \lim_{Q \searrow x} \frac{m(Q \cap E)}{m(Q)}$$

for a.e. $x \in E$ where the limit is over all cubes Q centered at x.

Definition 23. A family $\{S\}$ of measurable sets *shrinks regularly to* x if

- the diameters of the sets S tend to 0
- if Q is the smallest cube with center x containing S then there is a constant k independent of S such that $m(Q) \leq km(S)$.

Exercise 60. Suppose x is a Lebesgue point of f. Then

$$\frac{1}{m(S)} \int_S |f(y) - f(x)| \, dm(y) \to 0$$

for any family of sets $\{S\}$ that shrink regularly to x.

15 Differentiation of monotone functions

Motivation:

1. when does f' exist?
2. when does $\int_a^b f' \, dx = f(b) - f(a)$?

Here we will study the second question. We begin with monotone functions.

Definition 24. $f : [a, b] \to \mathbb{R}$ is *monotone increasing* if $f(x) \leq f(y)$ whenever $x \leq y$. f is *monotone decreasing* if $f(x) \geq f(y)$ whenever $x \leq y$.

We will prove:

Theorem 15.1. *If $f : (a, b) \to \mathbb{R}$ is monotone increasing and measurable then f' exists a.e., f' is measurable and*

$$\int_a^b f' \, dx \leq f(b^-) - f(a^+).$$

Exercise 61. Use the Cantor-Lebesgue function to prove that the inequality above can be strict.

Definition 25. We say that a collection $\mathcal{I}$ of intervals covers a set $E \subset \mathbb{R}$ *in the sense of Vitali* if for every $\epsilon > 0$ and $x \in E$ there exists $I \in \mathcal{I}$ with $x \in I$ and $m(I) < \epsilon$. The intervals may be closed, open or half-open but we do not allow an interval to degenerate to a point.

Exercise 62 (The not-as-simple Vitali's Covering Lemma). Let $E \subset \mathbb{R}$ have finite outer measure. Suppose $\mathcal{I}$ covers E in the sense of Vitali. Then for any $\epsilon > 0$ there exists a countable subset $\mathcal{I}' \subset \mathcal{I}$ such that

- the intervals in $\mathcal{I}'$ are pairwise disjoint
- $m^*(E \setminus \cup_{I \in \mathcal{I}'} I) = 0$.

- $m^*(\cup_{I\in\mathcal{J}'}) \leq (1+\epsilon)m^*(E)$.

Proof. Let $G \supset E$ be an open set with $m(G) \leq (1+\epsilon)m^*(E)$. Without loss of generality, we may assume that every cube of $\mathcal{J}$ lies in G. So the (3) is automatic.

By the Simple Vitali Covering Lemma, there exist a finite collection $\mathcal{J}_1 \subset \mathcal{J}$ such that (a) $\mathcal{J}_1$ is pairwise disjoint and (b)

$$m(\cup_{I\in\mathcal{J}_1} I) \geq C_1 m^*(E)$$

for some $C_1 > 0$. Thus

$$\begin{aligned} m^*(E \setminus \cup_{I\in\mathcal{J}_1} I) &\leq m(G \setminus \cup_{I\in\mathcal{J}_1} I) = m(G) - m(\cup_{I\in\mathcal{J}_1} I) \\ &\leq (1+\epsilon - C_1)m^*(E). \end{aligned}$$

By choosing $\epsilon < C_1/2$ we obtain

$$m^*(E \setminus \cup_{I\in\mathcal{J}_1} I) \leq (1-\epsilon/2)m^*(E).$$

Repeating the above argument we find a finite pairwise disjoint collection $\mathcal{J}_2 \subset \mathcal{J} \setminus \mathcal{J}_1$ such that

$$m^*(E \setminus \cup_{I\in\mathcal{J}_1\cup\mathcal{J}_2} I) \leq (1-\epsilon/2)m^*(E \setminus \cup_{I\in\mathcal{J}_1} I) \leq (1-\epsilon/2)^2 m^*(E).$$

Repeating in this way we obtain the result with $\mathcal{J}' = \cup_n \mathcal{J}_n$.

□

Exercise 63. Show that in Vitali's not-so-simple Covering Lemma, we can choose $\mathcal{J}'$ to be finite if we relax the second condition to

$$m^*(E \setminus \cup_{I\in\mathcal{J}'} I) < \epsilon.$$

Definition 26. Let $f : [a, b] \to \mathbb{R}$. For $x \in [a, b]$ define the *Dini derivatives* of f by:

$$\begin{aligned} D_{ne}f(x) &:= \limsup_{h\searrow 0} \frac{f(x+h)-f(x)}{h} \\ D_{se}f(x) &:= \liminf_{h\searrow 0} \frac{f(x+h)-f(x)}{h} \\ D_{nw}f(x) &:= \limsup_{h\searrow 0} \frac{f(x)-f(x-h)}{h} \\ D_{sw}f(x) &:= \liminf_{h\searrow 0} \frac{f(x)-f(x-h)}{h}. \end{aligned}$$

(nw stands for northwest, etc). If all of these are equal at a point x then we let $f'(x)$ denote the common value. This is the *derivative* of f.

Exercise 64. Prove Theorem 15.1. Hint: For rational number $u < v$, let

$$E_{u,v} = \{x : \ D_{sw}f(x) < u < v < D_{ne}f(x)\}.$$

In order to prove that $m^*(E_{u,v}) = 0$, let $\mathcal{I}$ be the collection of all intervals of the form $[x-h, x]$ such that

$$f(x) - f(x-h) < uh.$$

Apply the previous exercise to obtain a nice finite subcollection $\mathcal{I}' \subset \mathcal{I}$. Get an upper bound on $\sum_{[c,d]\in\mathcal{I}'} f(d) - f(c)$.

Then let $F_{u,v} = E_{u,v} \cap (\cup_{I\in\mathcal{I}'} I)$. Let $\mathcal{J}$ be the collection of all intervals of the form $[x, x+h]$ with $[x, x+h] \subset I \in \mathcal{I}'$ for some I and such that such that $f(x+h) - f(x) > vh$. Apply the previous exercise again to get a nice finite subcollection $\mathcal{J}' \subset \mathcal{J}$. This time get a lower bound on $\sum_{[c,d]\in\mathcal{J}'} f(d) - f(c)$. Then use the fact that each interval in $\mathcal{J}'$ is a subinterval of an interval in $\mathcal{I}'$.

Proof. Let $\epsilon > 0$. Following the hint, we obtain a finite collection $\mathcal{I}' \subset \mathcal{I}$ of intervals such that

- the intervals in $\mathcal{I}'$ are pairwise disjoint
- $m^*(E_{u,v} \setminus \cup_{I\in\mathcal{I}'} I) < \epsilon$.
- $m(\cup_{I\in\mathcal{I}'}) \le (1+\epsilon) m^*(E_{u,v})$.

If $I = [c, d] \in \mathcal{I}'$ then $f(d) - f(c) < u(d-c) = um(I)$. Therefore,

$$\sum_{[c,d]\in\mathcal{I}'} f(d) - f(c) < um(\cup_{I\in\mathcal{I}'}) \le u(1+\epsilon) m^*(E_{u,v}).$$

Let $F_{u,v} = E_{u,v} \cap (\cup_{I\in\mathcal{I}'} I)$. Note that $m^*(F_{u,v}) \ge m^*(E_{u,v}) - \epsilon$.

By the previous exercise, we obtain a finite collection $\mathcal{J}' \subset \mathcal{J}$ of intervals such that

- the intervals in $\mathcal{J}'$ are pairwise disjoint
- $m^*(F_{u,v} \setminus \cup_{J\in\mathcal{J}'} J) < \epsilon$.
- $m(\cup_{J\in\mathcal{J}'}) \le (1+\epsilon) m^*(F_{u,v})$.

If $J = [c, d] \in \mathcal{J}'$ then $f(d) - f(c) > v(d-c) = vm(J)$. Therefore,

$$\sum_{[c,d]\in\mathcal{J}'} f(d) - f(c) > vm(\cup_{J\in\mathcal{J}'} J) \ge v(m^*(F_{u,v} - \epsilon) \ge v(m^*(E_{u,v} - 2\epsilon).$$

Because each interval in $\mathcal{J}'$ is a subinterval of an interval in $\mathcal{I}'$ and f is monotone increasing,

$$\sum_{[c,d]\in\mathcal{J}'} f(d) - f(c) \le \sum_{[c,d]\in\mathcal{I}'} f(d) - f(c).$$

So

$$u(1+\epsilon) m^*(E_{u,v}) \ge v(m^*(E_{u,v} - 2\epsilon).$$

Because $\epsilon > 0$ is arbitrary and $u < v$, this implies $m^*(E_{u,v}) = 0$. Since this holds for all rational u, v it follows that $D_{sw}f(x) = D_{ne}f(x)$ for a.e. x. In a similar way, it can be shown that all of the derivatives are equal. Therefore, f' exists a.e.

Now let

$$f_k(x) = \frac{f(x + 1/k) - f(x)}{1/k}$$

where we define $f(x) = f(b^-)$ for $x \geq b$. Observe that $f_k \to f'$ a.e. Moreover $f_k \geq 0$ since f is monotone increasing. Fatou's Lemma implies

$$\int_a^b f' \, dm \leq \liminf_k \int_a^b f_k \, dm.$$

However, (setting $h = 1/k$),

$$\int_a^b f_k \, dm = (1/h) \int_{a+h}^{b+h} f \, dm - \int_a^b f \, dm = f(b^-) - (1/h) \int_a^{a+h} f \, dm.$$

Since $f(a^+) \leq (1/h) \int_a^{a+h} f \, dm \leq f(a^+ + 1/h)$, we obtain $\int_a^b f_k \, dm \to f(a^+)$. This implies the theorem.

□

15.1 Functions of Bounded Variation

Let $f : [a, b] \to \mathbb{R}$ be a function. Let $\Gamma = \{x_1, x_2, \ldots, x_n\} \subset [a, b]$ be a finite set with $x_1 < x_2 < \cdots < x_n$. Define

$$s_\Gamma = S_\Gamma[f; a, b] = \sum_{i=1}^{n-1} |f(x_{i+1}) - f(x_i)|.$$

Let

$$V = V[f; a, b] = \sup_\Gamma s_\Gamma$$

be the *variation of f over* $[a, b]$. We say that f has *bounded variation on* $[a, b]$ if $V[f; a, b] < \infty$.

Exercise 65.

1. If f is monotone on $[a, b]$ then it has bounded variation.
2. If f, g have bounded variation and $c, d \in \mathbb{R}$ are scalars then $cf + dg$ has bounded variation.
3. If $[a', b'] \subset [a, b]$ then $V[f; a', b'] \leq V[f; a, b]$.
4. If $a < c < b$ then $V[f; a, b] = V[f; a, c] + V[f; c, b]$.

Exercise 66. Construct a continuous function on $[0, 1]$ that is of unbounded variation on every subinterval. Hint: modify the Cantor-Lebesgue construction.

Exercise 67. Suppose f is continuously differentiable on $[a,b]$. Prove that $V[f;a,b] = \int_a^b |f'|\, dx$.

For Γ as above, let

$$P_\Gamma = \sum_{i=1}^{n-1} [f(x_{i+1}) - f(x_i)]^+$$

$$N_\Gamma = \sum_{i=1}^{n-1} [f(x_{i+1}) - f(x_i)]^-.$$

Observe that $S_\Gamma = P_\Gamma + N_\Gamma$ and $f(b) - f(a) = P_\Gamma - N_\Gamma$ (assuming $a, b \in \Gamma$). Let $P = P[f;a,b] = \sup_\Gamma P_\Gamma$, $N[f;a,b] = \sup_\Gamma N_\Gamma$.

Exercise 68. $P + N = V$, $P - N = f(b) - f(a)$. Thus

$$P = \frac{1}{2}[V + f(b) - f(a)], \quad N = \frac{1}{2}[V - f(b) + f(a)].$$

Proof. Because $V \geq S_\Gamma = P_\Gamma + N_\Gamma$, $V \geq P + N$. On the other hand, $S_\Gamma \leq P + N$. So taking the sup over Γ yields $V \leq P + N$.

Since $N_\Gamma + f(b) - f(a) = P_\Gamma$, we have $N + f(b) - f(a) = P$. □

Exercise 69. Prove that f as bounded variation on $[a,b]$ if and only if f is the difference of two monotone functions.

Proof. By a previous exercise we only need to show that if f has bounded variation then it is the difference of two monotone functions. Let $P(x) = P[f;a,x]$ and $N(x) = N[f;a,x]$. These are monotone functions and the previous exercise shows $f(x) = P(x) - N(x) + f(a)$. □

Exercise 70. If f has bounded variation then the set of discontinuities of f is countable. Moreover, every discontinuity is a jump discontinuity.

Exercise 71. If f is of bounded variation on $[a,b]$ and $V(x) = V[f;a,x]$ for $a \leq x \leq b$ then

$$V'(x) = |f'(x)|$$

for a.e. $x \in [a,b]$.

hints: ??

16 Absolutely continuous functions

Next, we investigate: when does $\int_a^b f'\, dx = f(b) - f(a)$?

Definition 27. We say that f is *absolutely continuous* on an interval $[a,b]$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $\mathfrak{I}$ is any collection of pairwise disjoint intervals in $[a,b]$ with $\sum_{I \in \mathfrak{I}} m(I) < \delta$ then

$$\sum_{[c,d] \in \mathfrak{I}} |f(d) - f(c)| < \epsilon.$$

Exercise 72. If f is absolutely continuous then f is continuous.

Exercise 73. If f is Lipschitz then f is absolutely continuous. For example, if f is continuously differentiable on $[a, b]$ then f is absolutely continuous.

Exercise 74. If f is absolutely continuous on $[a, b]$ then f has bounded variation on $[a, b]$. So f' exists a.e.

Proof. Let $\epsilon > 0$ and let $\delta > 0$ be as in the definition of absolute continuity. Let $\Gamma = \{a = x_0, x_1, \ldots, x_n, b = x_{n+1}\}$ be any partition of the interval into subintervals of length $\leq \delta$. Then by partitioning Γ into collections of subintervals whose total less is $\leq \delta$ we see that

$$s_\Gamma \leq \epsilon(1 + \lceil 1/\delta \rceil).$$

(The partition of Γ can be chosen to have at most $(1 + \lceil 1/\delta \rceil)$ parts). Since this is true for every Γ, we are done. □

Definition 28. A function f for which $f' = 0$ a.e. is said to be *singular*. For example, the Cantor-Lebesgue function is singular.

Exercise 75. If f is absolutely continuous and singular on $[a, b]$ then f is constant. Hint: use Vitali's Covering Lemma.

Proof. Let E be the set of all $x \in [a, b]$ such that $f'(x) = 0$. So $m(E) = b - a$. Let $\epsilon > 0$. Because f is absolutely continuous there exists $\delta > 0$ such that if $\mathcal{J}$ is any pairwise disjoint collection of subintervals of $[a, b]$ then $\sum_{[c,d] \in \mathcal{J}} |f(d) - f(c)| < \epsilon$.

Let $\mathcal{I}$ be the collection of all intervals of the form $[x, x + h]$ with $x \in E$ and

$$|f(x + h) - f(x)| < \epsilon h.$$

By the exercise after Vitali's Covering Lemma, there exists a pairwise disjoint finite subcollection $\mathcal{I}' \subset \mathcal{I}$ such that

$$\sum_{I \in \mathcal{I}'} m(I \cap [a, b]) \geq b - a - \delta.$$

Let $\mathcal{J}$ be the collection of intervals in the complement of $\cup_{I \in \mathcal{I}'} I$ in $[a, b]$. We now have

$$\sum_{[c,d] \in \mathcal{I}' \cup \mathcal{J}} |f(d) - f(c)| < \epsilon + (b - a)\epsilon.$$

Since

$$f(b) - f(a) = \sum_{[c,d] \in \mathcal{I}' \cup \mathcal{J}} f(d) - f(c)$$

we have

$$|f(b) - f(a)| \leq \sum_{[c,d] \in \mathcal{I}' \cup \mathcal{J}} |f(d) - f(c)| < \epsilon + (b - a)\epsilon.$$

Since ϵ is arbitrary, this implies $f(b) = f(a)$.

So we have shown that if $f' = 0$ a.e. and f is absolutely continuous on $[a, b]$ then $f(b) = f(a)$. We can apply this again to any subinterval of $[a, b]$ to see that f must be constant throughout.

□

Exercise 76. f is absolutely continuous on $[a, b]$ if and only if f' exists a.e. on $[a, b]$, f' is integrable on $[a, b]$ and

$$f(x) - f(a) = \int_a^x f' dm.$$

Hint: use the previous exercise.

Proof. We have already shown that integrals are absolutely continuous. So one direction is clear.

Now assume f is absolutely continuous on $[a, b]$. We have already shown that f' exists a.e. and is integrable. Let $F(x) = \int_a^x f' dm$. By Lebesgue's Differentiation Theorem, $F' = f'$ a.e. and F is absolutely continuous. So $F - f$ is absolutely continuous and singular on $[a, b]$. By the previous exercise, $F - f$ is constant. Since $F(a) - f(a) = -f(a)$ this implies $F(x) = f(x) - f(a) = \int_a^x f' \, dx$. □

Exercise 77. If f has bounded variation on $[a, b]$ then $f = g+h$ for some absolutely continuous function g and some singular function h on $[a, b]$. Moreover, g and h are unique up to additive constants.

Proof. For the first part, let $g(x) = \int_a^x f' \, dm$ and $h = f - g$. For the second part, note that if $f = g_1 + h_1 = g_2 + h_2$ then $g_1 - g_2 = h_2 - h_1$ is both absolutely continuous and singular. □

Exercise 78. If f, g are absolutely continuous on $[a, b]$ then

$$\int_a^b gf' \, d = g(b)f(b) - g(a)f(a) - \int_a^b g'f \, dx.$$

Proof. it is easy to check that gf is absolutely continuous on $[a, b]$ and the product rule

$$(gf)' = g'f + gf'$$

holds. The formula above follows by integrating the product rule. □

17 Convex functions

Definition 29. A function $\phi : (a, b) \to \mathbb{R}$ is *convex* if for every pair a', b' with $a < a' < b' < b$ the graph of $\phi \restriction [a', b']$ lies under (or on) the graph of the line from $(a', \phi(a'))$ to $(b', \phi(b'))$. We allow $a = -\infty, b = +\infty$.

Exercise 79. ϕ is convex on (a, b) if and only if for $0 < t < 1$ and $x < y$ with $x, y \in (a, b)$ we have

$$\phi(tx + (1 - t)y) \le t\phi(x) + (1 - t)\phi(y).$$

Exercise 80. If ϕ is twice continuously differentiable and $\phi'' \geq 0$ then ϕ is convex.

Exercise 81. If ϕ is convex then it is continuous.

Exercise 82. If ϕ is convex then $\phi'(x)$ exists for all but countably many x's. Moreover, ϕ' is monotone increasing.

Exercise 83. If ϕ is convex then ϕ is locally Lipschitz. In other words, for every pair a', b' with $a < a' < b' < b$ there is a constant $C > 0$ such that if $x, y \in [a', b']$ then

$$|f(y) - f(x)| \leq C|x - y|.$$

The last exercise implies that ϕ is absolutely continuous.

Exercise 84 (Jensen's inequality). Let (X, μ) be a probability space and $f \in L^1(X, \mu)$. Suppose that ϕ is convex on the essential range of f. Then

$$\phi(\int f \, d\mu) \leq \int \phi(f) \, d\mu.$$

Hint: to get an intuition for why this result is true, consider the special case in which μ is the probability measure on $\mathbb{R}$ satisfying $\mu(\{a\}) = t, \mu(\{b\}) = 1 - t$ for some $0 < t < 1$.

Proof. Let $\gamma = \int f \, d\mu$ and let m_0 be the slope of a supporting line L at γ. In other words $m_0 = \phi'(\gamma)$ if this exists. If it does not exist, then we could, for example, set m_0 equal to any of the Dini derivatives of ϕ at γ.

Because ϕ is convex, for a.e. $x \in X$,

$$\phi(f(x)) \geq m_0(f(x) - \gamma) + \phi(\gamma).$$

(This is because $m_0(f(x) - \gamma) + \phi(\gamma)$ is "y-coordinate" of the point on L that has "x-coordinate" equal to $f(x)$). Integrating over x, we obtain the exercise. $\square$

For example, take $\phi(x) = e^x$. Then Jensen's inequality becomes,

$$\exp(\int f \, d\mu) \leq \int e^f \, d\mu.$$

If $X = \{p_1, \ldots, p_n\}$, $\mu(\{p_i\}) = 1/n$ and $f(p_i) = x_i$ then we obtain

$$\exp((x_1 + \cdots + x_n)/n) \leq (1/n)(e^{x_1} + \cdots + e^{x_n}).$$

If we put $y = e^{x_i}$ then we obtain

$$(y_1 \cdots y_n)^{1/n} \leq (1/n)(y_1 + \cdots + y_n)$$

in other words: the arithmetic mean dominates the geometric mean (for a sequence of positive numbers).

More generally, if $\mu(\{p_i\}) = w_i$ then we have

$$y_1^{w_1} \cdots y_n^{w_n} \leq \sum_i w_i y_i.$$

Exercise 85. Suppose $a, b \geq 0$, $p, q > 1$ and $1/p + 1/q = 1$. Then $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.

Proof. Wlog $a, b > 0$.

Let $\phi(x) = e^x$ and let μ be the probability measure on $\{\log(a), \log(b)\}$ defined by $\mu(\{\log(a)\}) = 1/p, \mu(\{\log(b)\}) = 1/q$. Define f on $\{\log(a), \log(b)\}$ by $f(\log(a)) = p\log(a)$ and $f(\log(b)) = q\log(b)$. By Jensen's inequality,

$$e^{p\log(a)/p + q\log(b)/q} \leq (1/p)e^{p\log(a)} + (1/q)e^{q\log(b)}.$$

□

18 L^p spaces

Recall that if (X, μ) is a measure space and $0 < p < \infty$ then $L^p(X, \mu)$ is the set of all mod 0 equivalence classes of measurable functions f on X satisfying

$$\|f\|_p L = \left(\int |f|^p \, d\mu \right)^{1/p} < \infty.$$

We also define

$$\|f\|_\infty = \sup\{t \geq 0 : \ \mu(\{x \in X : \ |f(x)| \geq t\}) > 0\}.$$

By the way, we allow f to be complex-valued by defining

$$\int f \, d\mu = \int \mathrm{Re}(f) \, d\mu + i \int \mathrm{Im}(f) \, d\mu.$$

Exercise 86 (Young's Inequality). Let $\phi : [0, \infty) \to [0, \infty)$ be continuous, strictly increasing with $\phi(0) = 0$. Let ϕ^{-1} denote its composition inverse (so $\phi \circ \phi^{-1}(x) = x$). Then for $a, b > 0$,

$$ab \leq \int_0^a \phi(x) \, dx + \int_0^b \phi^{-1}(x) \, dx.$$

Equality holds if and only if $b = \phi(a)$.

Hint: there is an almost immediate geometric proof. Simply draw the graph of ϕ and realize that it is also the graph of ϕ^{-1} (if we switch the axes) and interpret the integrals as areas under the curves.

Definition 30. We say numbers $p, q \in [1, \infty]$ are *conjugate* if $\frac{1}{p} + \frac{1}{q} = 1$. For examples, 2 is conjugate with itself and 1 is conjugate with $+\infty$.

Exercise 87 (Hölder's inequality). If p, q are conjugate then $\|fg\|_1 \leq \|f\|_p \|g\|_q$. Hints: note that if the result holds for f, g and a, b are scalars then the result holds for af and bg. So we may assume without loss of generality, that $\|f\|_p = \|g\|_q = 1$. Now use Young's inequality with $\phi(x) = x^{p-1}$.

Proof. The case when p or $q = \infty$ is obvious. So assume $1 < p, q < \infty$. Then Young's inequality implies

$$ab \le \int_0^a x^{p-1}\ dx + \int_0^b x^{1/(p-1)}\ dx = \frac{a^p}{p} + \frac{b^{1/(p-1)+1}}{1 + 1/(p-1)}$$

Notice that $1/(p-1) + 1 = \frac{p}{p-1} = \frac{1}{1-1/p} = q$. So

$$ab \le a^p/p + b^q/q.$$

Applying this to $a = |f(x)|, b = |g(x)|$ and integrating over x, we obtain

$$\|fg\|_1 \le \|f\|_p^p/p + \|g\|_q^q/q.$$

In the special case in which $\|f\|_p = \|g\|_q = 1$, this implies the result. □

Schwarz's inequality is the special case $p = q = 2$.

Exercise 88 (Minkowski's inequality). If $1 \le p \le \infty$ and $f, g \in L^p(X)$ then $\|f + g\|_p \le \|f\|_p + \|g\|_p$. Hint: the cases $p \in \{1, \infty\}$ are straightforward so assume $1 < p < \infty$. Note that

$$\int |f+g|^p\ d\mu = \int |f+g|^{p-1}|f+g|\ d\mu \le \int |f+g|^{p-1}|f|\ d\mu + \int |f+g|^{p-1}|g|\ d\mu.$$

Now think about how to apply Hölder's inequality.

Proof. Hölder's inequality implies

$$\int |f+g|^{p-1}|f|\ d\mu \le \|f\|_p \||f+g|^{p-1}\|_q.$$

Since $(1/p) + (1/q) = 1$, multiplying out the denominators gives $p + q = pq$. So $(p-1)q = p$. So

$$\||f+g|^{p-1}\|_q = \|f+g\|_p^{p/q}.$$

Similarly,

$$\int |f+g|^{p-1}|f|\ d\mu \le \|g\|_p \|f+g\|_p^{p/q}.$$

So we obtain

$$\|f+g\|_p^p \int |f+g|^p\ d\mu \le (\|f\|_p + \|g\|_p)\|f+g\|_p^{p/q}.$$

Observe that $p/q = p - 1$. So dividing both sides by $\|f+g\|_p^{p-1}$ finishes it. □

Minkowski's inequality implies that $L^p(X)$ is a metric space with metric $d(f, g) = \|f - g\|_p$ if $1 \le p \le \infty$. By contrast,

Exercise 89. Show using explicit examples that if $0 < p < 1$ then there exist $f, g \in L^p(X)$ such that $\|f - g\|_p > \|f\|_p + \|g\|_p$.

Exercise 90. Show that, if $p \geq 1$, then $L^p(X)$ is a complete metric space. Hint: Suppose $p < \infty$. If $\{f_n\}$ is Cauchy in L^p then Chebyshev's inequality implies $\{f_n\}$ is Cauchy in measure. This implies (by a previous exercise?) the existence of f such that $f_n \to f$ is measure. Now apply Fatou's Lemma.

Proof. If $p = \infty$ then for a.e. x, $\{f_n(x)\}$ is Cauchy in $\mathbb{R}$. So there exists a limit function $f \in L^\infty$ and f_n converges uniformly a.e. to f.

We may assume $p < \infty$. Assuming the hint, we have $f_n \to f$ in measure. So Fatou's Lemma implies

$$\liminf_m \int \liminf_{n_j} |f_{n_j} - f_m|^p \, d\mu \leq \liminf_{n_j, m} \int |f_{n_j} - f_m|^p \, d\mu.$$

The RHS is 0. So the LHS is also 0. We can choose the subsequence $\{n_j\}$ so that $f_{n_j} \to f$ pointwise a.e.

So $\|f_m - f\|_p \to 0$ as $m \to \infty$. This implies, in particular, that $f \in L^p$ since $\|f\|_p \leq \|f - f_n\|_p + \|f_n\|_p$. □

Exercise 91. If $1 \leq p < \infty$ then $L^p(\mathbb{R}^n)$ is separable. However, $L^\infty(\mathbb{R}^n)$ is not.

Proof. Consider the set S of simple functions of the form $\sum_{i=1}^n q_i \chi_{E_i}$ where $q_i \in \mathbb{Q}$ and E_i is a rectangle with vertices at rational coordinates. This set is countable and with a bit of effort, you can show that it is dense in $L^p(\mathbb{R}^n)$. □

18.1 Banach spaces

Definition 31. A vector space V with a function $\|\cdot\| : V \to [0, \infty)$ (called the *norm*) is a *Banach space* if

- $\|\cdot\|$ really is a norm. This means that $\|cv\| = |c|\|v\|$ for every scalar c and vector v and $\|v + w\| \leq \|v\| + \|w\|$ and $\|v\| = 0$ if and only if $v = 0$.
- V is complete with respect to the metric $d(v, w) = \|v - w\|$.

We have already proven that each $L^p(X, \mu)$ is a Banach space if $1 \leq p \leq \infty$.

Of course, $\mathbb{R}^n, \mathbb{C}^n, \ell^p, C(X), C_0(X)$ are also Banach spaces (where e.g. X is a topological space and $C(X)$ is the space of continuous functions with norm $\|f\| = \sup_x |f(x)|$).

Exercise 92. If X is a topological space, the $C(X)$ is a Banach space.

Proof. If $\{f_n\} \subset C(X)$ is Cauchy, then for each $x \in X$, $\{f_n(x)\}$ is Cauchy in $\mathbb{R}$ (or $\mathbb{C}$). So $f(x) = \lim_n f_n(x)$ exists for every x. Standard results imply f is continuous. □

Definition 32. If V is a Banach space, then V^* is the set of all linear maps $f : V \to \mathbb{C}$ (or $\mathbb{R}$ whichever is the scalar field), such that

$$\sup_{\|v\|=1} |f(v)| < \infty.$$

V^* is called the *dual space*. Its elements are *bounded linear functionals* of V. For $f \in V^*$, let

$$\|f\| := \sup_{\|v\|=1} |f(v)|.$$

Definition 33. A linear map $f : V \to \mathbb{C}$ is in V^* if and only if it is continuous.

Exercise 93. If V is a Banach space then V^* is also a Banach space.

Proof. It is elementary to check that $\|\cdot\|$ is a norm. Suppose $\{f_n\}$ is a Cauchy sequence. By the previous exercise, we may regard f_n as a continuous function on V_1, the unit norm ball of V. That is, $V^* \subset C(V_1)$. Note that the norm on V^* coincides with the norm it inherits from $C(V_1)$. So $\{f_n\}$ is Cauchy as a sequence in $C(V_1)$. Since $C(V_1)$ is complete, $\{f_n\}$ converges to $f \in C(V_1)$. It is easy to check that f must be linear. □

Exercise 94. Suppose $1 \le p \le q \le \infty$ are conjugate. Show that $L^q(X)$ naturally sits inside the dual $L^p(X)^*$. Moreover, the map from $L^q(X)$ into $L^p(X)^*$ is an isometric embedding. In case $q = \infty$, assume the measure is semi-finite (this means every subset $Y \subset X$ with $\mu(Y) > 0$ contains a subset $Y' \subset Y$ with $0 < \mu(Y') < \infty$).

Proof. **Case 1**. Suppose $1 < p, q < \infty$.

Given $\phi \in L^q(X)$, the map $f \mapsto \int f\phi \, d\mu$ is a bounded linear function on L^p (by Hölder's inequality). This shows $L^q(X)$ embeds into $L^p(X)^*$. However, you may wonder: why is it an isometric embedding? In other words, we have to prove:

$$\|\phi\|_q = \sup_{f \in L^p, f \neq 0} \frac{|\int f\phi \, d\mu|}{\|f\|_p}.$$

First assume that $\phi \ge 0$ and $\|\phi\|_q = 1$. Define $f(x) = \phi^{q-1}$. Observe that $f \in L^p$,

$$\|f\|_p^p = \int \phi^{pq-p} \, d\mu = \int \phi^q \, d\mu = 1$$

and $\int \phi f \, d\mu = \int \phi^q \, d\mu = 1$. This proves it.

Now suppose $\|\phi\|_q = 1$ but don't assume $\phi \ge 0$. Set $f(x) = |\phi^{q-1}(x)|u(x)$ where u is chosen so that $u(x) = 0$ if $\phi(x) = 0$ and $u(x) = \frac{\overline{\phi(x)}}{\phi(x)}$ otherwise. The same argument as above works.

For the general case, we observe that the statement is scale-invariant. So we are done with the case $1 < p, q < \infty$.

Case 2. Suppose $q = 1$ and $p = \infty$.

Define $f \in L^\infty$ by $f(x) = 0$ of $\phi(x) = 0$ and $f(x) = \frac{\overline{\phi(x)}}{|\phi(x)|}$ otherwise. Note that $f(x)\phi(x) = |\phi(x)|$. Therefore $\int f(x)\phi(x) \, d\mu(x) = \|\phi\|_1$ which proves this case.

Case 3. Suppose $q = \infty$ and $p = 1$.

We assume wlog that $\|\phi\|_\infty = 1$. For $n > 0$ let $E_n = \{x \in X : |\phi(x)| > 1 - 1/n\}$. Note $\mu(E_n) > 0$. Because μ is semifinite, there exists a subset $E_n' \subset E_n$ with $0 < \mu(E_n') < \infty$.

Define $f_n(x) = 0$ if $x \notin E'_n$ and $f_n(x) = \frac{\overline{\phi(x)}}{|\phi(x)|\mu(E'_n)}$ if $x \in E'_n$. Then

$$\int f_n(x)\phi(x)\ d\mu(x) = \int_{E'_n} |\phi(x)|/\mu(E'_n)\ d\mu(x) \geq (1 - 1/n).$$

Note that

$$\|f_n\|_1 = \int_{E'_n} \frac{1}{\mu(E'_n)}\ d\mu = 1.$$

Therefore,

$$1 = \sup_{f \in L^1, f \neq 0} \frac{|\int f\phi\ d\mu|}{\|f\|_1}$$

as required.

□

Later we will show that $L^q = (L^p)^*$ (unless $p = \infty$).

19 Hilbert spaces

For $f, g \in L^2(X)$, define the inner product:

$$\langle f, g \rangle = \int f\bar{g}\ d\mu.$$

By Schwarz' inequality,

$$|\langle f, g \rangle| \leq \|f\|_2 \|g\|_2.$$

In particular, it is finite.

Exercise 95.

- the inner product is sesquilinear: for any $f, g, h \in L^2(X)$,

$$\langle f, g + h \rangle = \langle f, g \rangle + \langle f, h \rangle,$$

$$\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle.$$

  Also if c is a constant then

$$\langle cf, g \rangle = c\langle f, g \rangle,$$
$$\langle f, cg \rangle = \bar{c}\langle f, g \rangle.$$

- $\langle f, f \rangle = \|f\|_2^2$.
- $\langle v, w \rangle = \overline{\langle w, v \rangle}$.

More generally,

Definition 34. A *Hilbert space* is a Banach space $\mathcal{H}$ with an inner product satisfying the conditions of the previous exercise. So $L^2(X)$ is a Hilbert space. Also ℓ^2 is also a Hilbert space.

Most of the results we will consider are true for an arbitrary Hilbert space. So from now on, fix a Hilbert space $\mathcal{H}$.

Definition 35. We say $v, w \in \mathcal{H}$ are *orthogonal* if $\langle v, w\rangle = 0$. A subset $B \subset \mathcal{H}$ is *orthonormal* if $\|v\| = 1$ for every $v \in B$ and every pair of distinct vectors in B are orthogonal. If $B \subset \mathcal{H}$ is any set then its *span* is the set of all finite linear combinations of the form $\sum_{b\in B} c_b b$ (with $c_b \in \mathbb{C}$). We say B is a *basis* if its span is dense in $\mathcal{H}$.

Exercise 96. If $E \subset \mathbb{R}^n$ has positive measure then $L^2(E)$ has a countable orthonormal basis. Moreover, every orthonormal basis of $L^2(E)$ is countable.

Proof. By a previous exercise, $L^2(E)$ is separable. Let $S = \{s_1, s_2, \dots\} \subset L^2(E)$ be a countable dense subset. Let $S' = \{s'_1, s'_2, \dots\}$ be defined inductively by: $s'_1 = s_1$ and $s'_i = s_j$ where j is the smallest number so that s_j is not contained in the span of $\{s'_1, \dots, s'_{j-1}\}$.

Define vectors v_i inductively by the Gram-Schmidt process:

$$v_1 = \frac{s_1}{\|s_1\|},$$

$$v_i = \frac{s_i - \sum_{j<i}\langle s_i, v_j\rangle v_j}{\|s_i - \sum_{j<i}\langle s_i, v_j\rangle v_j\|}.$$

Then the v_i's are orthonormal. They form a basis because they have the same span as S. □

Exercise 97. Let $\{v_i\}_{i=1}^n$ be a finite ON set and $x \in \mathcal{H}$. Let $y = \sum_{i=1}^n \langle x, v_i\rangle v_i$. Then

- $(x - y) \perp v_i$ for all i.
- $(x - y) \perp z$ for every $z \in \mathrm{Span}(v_1, \dots, v_n)$.
- y is the unique closest point to x with $y \in \mathrm{Span}(v_1, \dots, v_n)$

Proof. It's easy to check that $(x - y) \perp v_i$ for all i which implies $(x - y) \perp z$ for every $z \in \mathrm{Span}(v_1, \dots, v_n)$. Note that, in general, if $z_1 \perp z_2$ then $\|z_1 - z_2\|^2 = \|z_1\|^2 + \|z_2\|^2$. So if $w \in \mathrm{Span}(v_1, \dots, v_n)$ then

$$\begin{aligned}\|x - w\|^2 &= \langle x - y + y - w, x - y + y - w\rangle \\ &= \|x - y\|_2^2 + \|y - w\|_2^2.\end{aligned}$$

Thus $\|x - w\|^2 \geq \|x - y\|^2$ with equality iff $y = w$.

□

Exercise 98. $\{v_i\}$ is an ON basis for $\mathcal{H}$ if and only if every $x \in \mathcal{H}$ can be uniquely expressed in the form

$$x = \sum_i \langle x, v_i\rangle v_i$$

for some coefficients $c_i := \langle x, v_i\rangle \in \mathbb{C}$, called the *Fourier coefficients*. Moreover,

$$\|x\|^2 = \sum_i |c_i|^2.$$

Proof. Let $c_i = \langle x, v_i \rangle$. Let $y_n = \sum_{i=1}^n c_i v_i$. The previous exercise implies that y_n is the closest point to x on $\mathrm{Span}(\mathrm{v}_1, \ldots, \mathrm{v}_\mathrm{n})$. Because $\{v_i\}$ is a basis, we must have $\lim_n y_n = x$. Thus $\sum_{i=1}^\infty c_i v_i = x$ as required. Observe that $\|y_n\|^2 = \sum_{i=1}^n |c_i|^2$ implies $\|x\|^2 = \sum_i |c_i|^2$.

Why is this representation unique? Well, if there were two such representations $v = \sum_i c_i v_i = \sum_i d_i v_i$ then we could subtract them to obtain a representation of 0 of the form

$$0 = \sum_i (c_i - d_i) v_i.$$

However this implies

$$0 = \|0\|^2 = \| \sum_i (c_i - d_i) v_i \|^2 = \sum_i |c_i - d_i|^2.$$

So $c_i = d_i$ for all i. □

Exercise 99. If $E \subset \mathbb{R}$ has positive measure then $L^2(E)$ is isomorphic as a Hilbert space to ℓ^2.

Exercise 100. Show that $\{(2\pi)^{-1/2} e^{inx}\}_{n \in \mathbb{Z}}$ is an orthonormal set in $L^2([0, 2\pi), m)$.

Theorem 19.1 (Stone-Weierstrauss)**.** *Let X be a locally compact Hausdorff space and $\mathcal{A} \subset C_0(X)$ an algebra. Suppose $\mathcal{A}$ separates points and vanishes nowhere (i.e. for any $x, y \in X$ with $x \neq y$ there exists $f, g \in \mathcal{A}$ with $f(x) \neq f(y)$ and $g(x) \neq 0$). Then $\mathcal{A}$ is dense in $C_0(X)$.*

Exercise 101. Using the above theorem prove that $\{(2\pi)^{-1/2} e^{inx}\}_{n \in \mathbb{Z}}$ is a basis in $L^2([0, 2\pi), m)$.

20 Signed measures

Definition 36. A signed measure on a measurable space $(X, \mathcal{B})$ is a function $\nu : \mathcal{B} \to \mathbb{R} \cup \{\pm\infty\}$ such that

- ν assumes at most one of the values $-\infty, +\infty$
- $\nu(\cup_{i=1}^\infty E_i) = \sum_{i=1}^\infty \nu(E_i)$ for any sequence $\{E_i\}$ of pairwise disjoint measurable sets.

Definition 37. If ν is a signed measure then $E \in \mathcal{B}$ is

- *positive* if $\nu(E') \geq 0$ for every measurable subset $E' \subset E$
- *negative* if $\nu(E') \leq 0$ for every measurable subset $E' \subset E$
- *null* if $\nu(E') = 0$ for every measurable subset $E' \subset E$.

Exercise 102. Let E be a measurable subset with $0 < \nu(E) < \infty$. Then there is a positive set $A \subset E$ with $\nu(A) > 0$.

Hint: If E is not positive, then let n_1 be the smallest positive integer such that there is a set $E_1 \subset E$ with $\nu(E_1) < -1/n_1$. Inductively define n_k to be the smallest positive integer such that there is a set $E_k \subset E \setminus \cup_{i<k} E_i$ with $\nu(E_k) < -1/n_k$ (assuming $E \setminus \cup_{i<k} E_i$ is not already positive). Show that $A = E \setminus \cup_i E_i$ satisfies the exercise.

Proof. Following the hint, we observe that $\nu(\cup_i E_i) < \sum_i \frac{-1}{n_i}$. Since $\nu(E) = \nu(\cup_i E_i) + \nu(A)$ is finite we must have that $\nu(\cup_i E_i)$ is also finite, $\nu(A) > 0$ and $n_i \to +\infty$ as $i \to \infty$.

To show A is positive, let $B \subset A$ be measurable. So B is in the complement of $\cup_i E_i$. By definition of E_i this means that $\nu(B \cup E_i) = \nu(B) + \nu(E_i) \geq -1/(n_i - 1)$. So $\nu(B) \geq -1/(n_i - 1) - \nu(E_i) \geq -1/(n_i - 1)$. Since $n_i \to +\infty$, this implies $\nu(B) \geq 0$.

Since B is arbitrary, A is positive. □

Exercise 103 (Hahn Decomposition Theorem). Let ν be a signed measure on $(X, \mathcal{B})$. Then there are a positive set P and a negative set N such that $X = P \cup N$, $P \cap N = \emptyset$.

Hint: let λ be the sup of $\nu(A)$ over all positive sets A. Observe that a countable union of positive sets is positive.

Proof. We assume wlog that $+\infty \neq \nu(A)$ for any A. Following the hint, we obtain positive set A_i with $\nu(A_i) \geq \lambda - 1/i$ and set $P = \cup_i A_i$ and $N = X \setminus P$. □

Definition 38 (Jordan decomposition)**.** Let ν be a signed measure and P, N be sets as above. Let ν^+, ν^- be the measures defined by

$$\nu^+(E) = \nu(E \cap P), \quad \nu^-(E) = -\nu(E \cap N).$$

Observe that these are measures on $(X, \mathcal{B})$ in the usual sense (non-signed). Moreover,

$$\nu = \nu^+ - \nu^-.$$

This is called the *Jordan decomposition* of ν. The *total variation of* ν is

$$|\nu| = \nu^+ + \nu^-.$$

Definition 39. Let ν, μ be two measures on X. We say

- ν is *absolutely continuous* to μ, denoted $\nu << \mu$ if for every measurable E, $\mu(E) = 0 \Rightarrow \nu(E) = 0$;
- ν is *singular* to μ, denoted $\nu \perp \mu$, if there exists a measurable set A such that $\nu(A) = 0, \mu(A^c) = 0$.

Exercise 104. Prove that the Jordan decomposition of ν is unique in the sense that if $\nu = \mu_1 - \mu_2$ for two positive measures μ_1, μ_2 that are mutually singular then $\nu^+ = \mu_1, \nu^- = \mu_2$.

Proof. Suppose $\nu = \mu_1 - \mu_2$. Because μ_1, μ_2 are singular, there exists a decomposition $X = S_1 \cup S_2$ such that $\mu_i(S_{i+1}) = 0$. If $\mu_1(N) > 0$ then $\mu_1(N \cap S_1) > 0$ which implies

$$\nu(N \cap S_1) = \mu_1(N \cap S_1) > 0$$

a contradiction. So $\mu_1(N) = 0$. Similarly $\mu_2(P) = 0$. Similarly, $\nu^+(S_2) = 0 = \nu^-(S_1)$. If $E \subset X$ is measurable then

$$\nu(E \cap P) = \nu^+(E \cap P) = \nu^+(E) = \mu_1(E \cap P) = \mu_1(E)$$
$$\nu(E \cap N) = \nu^-(E \cap N) = \nu^-(E) = \mu_2(E \cap N) = \mu_2(E).$$

□

21 Radon-Nikodym Theorem

Before proving the Radon-Nikodym Theorem (below) we will need two exercises.

Exercise 105. For motivation, suppose $f : X \to \mathbb{R}$ is a measurable real-valued function. For each rational number r we may consider the sets $f^{-1}((-\infty, r])$. Do these sets determine f? In other words, if $g : X \to \mathbb{R}$ is another measurable function such that $\mu(f^{-1}((-\infty, r]) \triangle g^{-1}((-\infty, r])) = 0$ for every rational r, then does $g = f$ a.e. ?

Proof. Yes. To obtain a contradiction suppose there is an $\epsilon > 0$ such that the set $E = \{x \in X : \ f(x) - g(x) > \epsilon\}$ has positive measure. There must exist rational r_0 such that $f^{-1}((r_0, r_0+\epsilon]) \cap E$ has positive measure. However $g^{-1}((r_0-\epsilon, r_0]) \cap E \supset f^{-1}((r_0, r_0+\epsilon]) \cap E$ which implies that $g^{-1}(-\infty, r_0]$ is not the same as $f^{-1}((-\infty, r_0])$ (even up to measure zero). $\square$

Exercise 106. Can we reverse the previous exercise? Suppose we are given a collection of measurable sets $\{X_r\}_{r\in\mathbb{Q}}$ satisfying $r < s \Rightarrow X_r \subset X_s$ and $X = \cup_r X_r$. Is there a measurable function f such that $f \leq r$ on X_r and $f \geq r$ on X_r^c? If there is one, is it unique?

Proof. Define $f(x)$ to be the infimum of r over all r such that $x \in X_r$. To see that f is measurable, note that for any $r > 0$

$$f^{-1}(-\infty, r] = (\cup_{s \leq r} X_s) \cup (\cap_{t>r} X_t).$$

Clearly, f satisfies the exercise. Suppose g is another such function. If $g \neq f$ a.e. then there is a positive measure set $E \subset X$ and disjoint closed intervals I_f, I_g such that $f(x) \in I_f$ for $x \in E$ and $g(x) \in I_g$ for $x \in I_g$. Let r be strictly between I_f and I_g. Wlog $I_f < r < I_g$. Then $E \subset X_r$ (because $I_f < r$) but E is not a subset of X_r (because $r < I_g$). This contradiction proves it. $\square$

Exercise 107 (The Radon-Nikodym Theorem). Suppose $\nu << \mu$ and μ is σ-finite. Then there exists a measurable function f on X with $f \geq 0$ and

$$\nu(E) = \int_E f \, d\mu$$

for all measurable E. This function is unique up to measure zero.

Proof. Wlog we may assume $\mu(X) < \infty$.

For each rational r, consider the signed measure $\nu - r\mu$. By the Hahn Decomposition Theorem, we may write $X = P_r \cup N_r$ where $P_r \cap N_r = \emptyset$ are measurable sets such that P_r is positive for $\nu - r\mu$ and N_r is negative.

We don't yet know that these sets are nested so we can't yet use the previous exercise. However, note that if $s < r$ then

$$0 \geq \nu - s\mu(N_s \setminus N_r) \geq \nu - r\mu(N_s \setminus N_r) \geq 0.$$

So after ignoring a measure zero set, we can in fact assume that these are nested. So $s < r \Rightarrow N_s \subset N_r$.

So the previous exercise implies there is a function f such that $f \le r$ on N_r and $f \ge r$ on $N_r^c = P_r$.

Without loss of generality, we may assume $f \ge 0$ since $\nu \ge 0$ implies we can choose $P_0 = X$.

Let $k > 0$ be an integer and let $E_n = P_{k/n} \cap N_{(k+1)/n}$. Let E_∞ be the complement of $\cup_n E_n$ in X. So $X = E_\infty \cup \bigcup_n E_n$.

Let $S \subset X$ be measurable with $\mu(S) < \infty$. Because $E_n \subset P_{k/n}$ we have

$$\nu(S \cap E_n) - (k/n)\mu(S \cap E_n) \ge 0$$

which implies $\nu(S \cap E_n) \ge (k/n)\mu(S \cap E_n)$. Because $E_n \subset N_{(k+1)/n}$ we have

$$(k/n)\mu(S \cap E_n) \le \nu(S \cap E_n) \le ((k+1)/n)\mu(S \cap E_n).$$

Because

$$k/n \le f(x) \le (k+1)/n$$

on $P_{k/n} \cap N_{(k+1)/n}$ we have

$$k/n\mu(S \cap E_n) \le \int_{S \cap E_n} f \, d\mu \le (k+1)/n\mu(S \cap E_n).$$

Thus

$$|\nu(S \cap E_n) - \int_{S \cap E_n} f \, d\mu| \le (1/n)\mu(S \cap E_n).$$

Adding up over n we have

$$|\nu(S \cap (X \setminus E_\infty)) - \int_{S \cap (X \setminus E_\infty)} f \, d\mu| \le (1/n)\mu(S).$$

Since n is arbitrary and $\mu(S) < \infty$ this shows

$$|\nu(S \cap (X \setminus E_\infty)) - \int_{S \cap (X \setminus E_\infty)} f \, d\mu.$$

Now observe that $f = +\infty$ on E_∞. If $\mu(E_\infty) = 0$ then we're done of course since $\nu << \mu$.

If $\mu(E_\infty \cap S) > 0$ then $\nu(E_\infty \cap S) = +\infty$ since $\nu(E_\infty \cap S) - r\mu(E_\infty \cap S) \ge 0$ for every r. Also

$$\int_{E_\infty \cap S} f \, d\mu = +\infty = \nu(E_\infty \cap S).$$

So we're done for $\mu(E) < \infty$. The general case follows from this case by countable additivity.

Suppose g is another function satisfying the exercise. Then $0 = \int_E f - g \, d\mu$ for all measurable sets E. By a previous exercise, this implies $f = g$ a.e.

$\square$

Definition 40. The function f appearing in the previous exercise is called the Radon-Nikodym derivative and sometimes it is denoted by

$$f(x) = \frac{d\nu}{d\mu}(x).$$

Exercise 108 (Lebesgue Decomposition Theorem). Let μ, ν be σ-finite measures on $(X, \mathcal{B})$. Then there exist unique measures ν_{ac}, ν_{sing} such that

$$\nu = \nu_{ac} + \nu_{sing}$$

$\nu_{ac} << \mu$ and $\nu_{sing} \perp \mu$.

Proof. Let $\lambda = \mu + \nu$. Observe that $\mu << \lambda$ and $\nu << \lambda$. So there exist measurable functions f, g such that

$$\mu(E) = \int_E f \; d\lambda, \quad \nu(E) = \int_E g \; d\lambda$$

for all measurable E. Let $A = \{x \in X : \; f(x) > 0\}$ and $B = \{x \in X : \; f(x) = 0\}$. Then $X = A \cup B, A \cap B = \emptyset$. Define

$$\nu_{ac}(E) = \nu(E \cap A), \quad \nu_{sing}(E) = \nu(E \cap B).$$

The rest is clear. □

22 The Dual of L^p

Exercise 109. Let $1 \leq p \leq q \leq \infty$ be conjugate exponents. Recall from a previous exercise that $L^q(X)$ isometrically embeds into $L^p(X)^*$. Prove that this is onto. That is $L^q(X)$ is canonically isomorphic to $L^p(X)^*$.

Hint: suppose $\Phi \in L^p(X)^*$. We may assume wlog that the L^p spaces are real (as opposed to complex). Handle the case $\mu(X) < \infty$ first. Define a set function ν by

$$\nu(E) = \Phi(\chi_E).$$

Prove that ν is a signed measure. Apply the Radon-Nikodym Theorem to the measures ν^+ and ν^- from its Jordan Decomposition.

Proof. First we show ν is a measure. Let $\{E_i\}$ be pairwise disjoint measurable sets. Define f on X by

$$f(x) = \sum_i \mathrm{sign}(\Phi(\chi_{\mathrm{E_i}}))\chi_{\mathrm{E_i}}(\mathrm{x})$$

and f_n by

$$f_n(x) = \sum_{i=1}^n \mathrm{sign}(\Phi(\chi_{\mathrm{E_i}}))\chi_{\mathrm{E_i}}(\mathrm{x}).$$

Because $\mu(X) < \infty$, $f \in L^p(X)$ and $f_n \to f$ in $L^p(X)$ (by the bounded convergence theorem). Because Φ is continuous, $\Phi(f_n) \to \Phi(f)$. Thus

$$\Phi(f) = \lim_n \Phi(f_n) = \lim_n \sum_{i=1}^{n} |\nu(E_i)| = \sum_i |\nu(E_i)| < \infty.$$

Thus

$$\nu(E) = \sum_i \nu(E_i)$$

(again using continuity of Φ and the absolute summability of $\nu(E_i)$). This shows ν is a signed measure.

By the Radon-Nikodym Theorem (and the Jordan Decomposition), there is a measurable function g such that

$$\Phi(\chi_E) = \nu(E) = \int_E g \; d\mu.$$

if f is any simple function then we must have

$$\Phi(f) = \int fg \; d\mu.$$

If g is bounded then $g \in L^q$. In this case, we have that $\Phi = \Phi_g$ on the dense set of simple functions (where $\Phi_g \in (L^p)^*$ is the linear functional given by $f \mapsto \int fg \; d\mu$). Because both Φ and Φ_g are continuous, this implies $\Phi = \Phi_g$ and $\|\Phi\| = \|g\|_q$ (by a previous exercise).

Now suppose g is not necessarily bounded. Let $E_n = \{x \in X : \; |g(x)| \leq n\}$. We observe

$$\|\Phi\| \geq \|\Phi \restriction L^p(E_n)\| = \|g \restriction E_n\|_q.$$

Here $\restriction$ means "restricted to". For example, $\|\Phi \restriction L^p(E_n)\|$ is the operator norm of the restriction of Φ to $L^p(E_n)$ which we may regard as a subspace of $L^p(X)$. So $\|\Phi \restriction L^p(E_n)\| = \sup\{|\Phi(f)| : \; f \in L^p(E_n), \|f\|_p \leq 1\}$. The reason that $\|\Phi \restriction L^p(E_n)\| = \|g \restriction E_n\|_q$ is that g is bounded on E_n. So the results of the previous paragraph apply.

As $n \to \infty$, $\|g \restriction E_n\|_q \to \|g\|_q$ (since $\cup_n E_n = X$). Therefore, $\|\Phi\| \geq \|g\|_q$, and in particular $g \in L^q$. Once again, $\Phi = \Phi_g$ on the dense set of simple functions implies $\Phi = \Phi_g$ by continuity.

It is clear that g is unique (up to measure zero).

Now extend to the σ-finite case: suppose $X = \cup_i X_i$ where $\mu(X_i) < \infty$. Then for each i there is a unique $g_i \in L^q(E_i)$ such that for any $f \in L^p$, $\Phi(\chi_{X_i} f) = \int_{E_i} fg_i \; d\mu$. Note that $g_{i+1} = g_i$ a.e. on E_i. So we can define g on X by $g = g_i$ on E_i. Note that $\|g_i\|_q \leq \|\Phi\|$. So taking the limit as $i \to \infty$ we obtain $\|g\|_q \leq \|\Phi\| < \infty$ (this uses the monotone convergence theorem). In particular, $g \in L^q$. Also using Lebesgue's Dominated Convergence Theorem, we obtain $\Phi(f) = \int fg \; d\mu$ for any $f \in L^p$, finishing it.

$\square$

Exercise 110. Show that $L^1(\mathbb{R})$ is not the dual of $L^\infty(\mathbb{R})$.

23 Outer measure

Definition 41. An **outer measure** on a set X is a function μ^* defined on *all subsets of* X satisfying

- $\mu^*(\emptyset) = 0$
- $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$
- $E \subset \cup_i E_i \Rightarrow \mu * (A) \leq \sum_i \mu^*(E_i)$.

Definition 42. A subset $E \subset X$ is *measurable* with respect to an outer measure μ^* if for every $A \subset X$,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \setminus E).$$

Exercise 111. The class $\mathcal{B}$ is μ^*-measurable sets is a sigma-algebra. Hints:

1. First show that if $E_1, E_2 \in \mathcal{B}$ then $E_1 \cup E_2 \in \mathcal{B}$.
2. Next, show that if $E_1, \ldots, E_n$ are pairwise disjoint and measurable and A is arbitrary then $\mu^*(A \cap \cup_{i=1}^n E_i) = \sum_i \mu^*(A \cap E_i)$.
3. Next show that if $\{E_i\}$ are pairwise disjoint measurable sets then $\cup_i E_i$ is measurable.

Proof. Clearly, $\emptyset \in \mathcal{B}$ and $E \in \mathcal{B} \Rightarrow E^c \in \mathcal{B}$. So it suffices to show $\mathcal{B}$ is closed under countable unions. First we will show it is closed under finite unions.

Let $E_1, E_2 \in \mathcal{B}$. By subadditivity, we have

$$\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \setminus E)$$

for any A, E. So it suffices to prove the opposite inequality when $E = E_1 \cup E_2$ and A is arbitrary.

Because E_2 is measurable,

$$\mu^*(A) = \mu^*(A \cap E_2) + \mu^*(A \cap E_2^c).$$

Because E_1 is measurable

$$\mu^*(A \cap E_2^c) = \mu^*(A \cap E_2^c \cap E_1) + \mu^*(A \cap E_2^c \cap E_1^c).$$

So we have

$$\mu^*(A) = \mu^*(A \cap E_2) + \mu^*(A \cap E_2^c \cap E_1) + \mu^*(A \cap E_2^c \cap E_1^c).$$

Note

$$A \cap (E_1 \cup E_2) = (A \cap E_2) \cup (A \cap E_1 \cap E_2^c).$$

By subadditivity,

$$\mu^*(A \cap (E_1 \cup E_2)) \leq \mu^*(A \cap E_2) + \mu^*(A \cap E_1 \cap E_2^c).$$

So

$$\mu^*(A) \geq \mu^*(A \cap (E_1 \cup E_2)) + \mu^*(A \cap E_2^c \cap E_1^c).$$

Since $(E_1 \cup E_2)^c = E_1^c \cap E_2^c$ this proves $E_1 \cup E_2$ is measurable.

Note that $E_1 \setminus E_2 = E_1 \cap E_2^c = (E_1^c \cup E_2)^c$. So relative complements of measurable sets are measurable. For this reason, it suffices now to show that if $\{E_i\}$ be a sequence of *disjoint* measurable sets then the union $\cup_i E_i$ is measurable.

Let $G_n = \cup_{i=1}^n E_i$. Then G_n is measurable and

$$\mu^*(A) = \mu^*(A \cap G_n) + \mu^*(A \cap G_n^c) \geq \mu^*(A \cap G_n) + \mu^*(A \cap G_\infty^c).$$

Unfortunately, we don't know that μ^* is continuous, so we can't just take a limit as $n \to \infty$.

However, because E_n is measurable,

$$\begin{aligned} \mu^*(A \cap G_n) &= \mu^*(A \cap G_n \cap E_n) + \mu^*(A \cap G_n \cap E_n^c) \\ &= \mu^*(A \cap E_n) + \mu^*(A \cap G_{n-1}). \end{aligned}$$

By induction,

$$\mu^*(A \cap G_n) = \sum_{i=1}^n \mu^*(A \cap E_i).$$

So

$$\mu^*(A) \geq \mu^*(A \cap G_\infty^c) + \sum_{i=1}^n \mu^*(A \cap E_i).$$

Now we take a limit as $n \to \infty$ to obtain

$$\mu^*(A) \geq \mu^*(A \cap G_\infty^c) + \sum_{i=1}^\infty \mu^*(A \cap E_i) \geq \mu^*(A \cap G_\infty^c) + \mu^*(A \cap G_\infty)$$

where the last inequality uses subadditivity. □

Exercise 112. Show $\mathcal{B}$ is complete in the sense that if $E \in \mathcal{B}$ has $\mu^*(E) = 0$ then every subset of E is measurable.

Proof. Let $F \subset E$ and $A \subset X$ be arbitrary. Then

$$\mu^*(A) \geq \mu^*(A \cap F^c) = \mu^*(A \cap F) + \mu^*(A \cap F^c).$$

The first inequality is from monotonicity and the second is also from monotonicity (since $A \cap F \subset F \subset E$). Subadditivity gives the oppositive inequality and thereby proves F is measurable. □

Exercise 113. If $\mu = \mu^* \restriction \mathcal{B}$ then μ is a measure. Hint: first show finite additivity.

Proof. If $E_1, E_2 \in \mathcal{B}$ are disjoint then, because E_2 is measurable,

$$\mu(E_1 \cup E_2) = \mu((E_1 \cup E_2) \cap E_2) + \mu((E_1 \cup E_2) \setminus E_2) = \mu(E_2) + \mu(E_1).$$

So we have finite addiitivity.

Suppose $\{E_i\} \subset \mathcal{B}$ are pairwise disjoint. Then

$$\mu(\cup_{i=1}^{\infty} E_i) \geq \mu(\cup_{i=1}^{n} E_i) = \sum_{i=1}^{n} \mu(E_i).$$

By taking limits we have

$$\mu(\cup_{i=1}^{\infty} E_i) \geq \sum_{i=1}^{\infty} \mu(E_i).$$

The oppositive inequality follows from subadditivity.

□

24 Carathéodory's Extension Theorem

Suppose $\mathcal{A}$ is algebra on a set X and μ_0 is like a measure on $\mathcal{A}$ in the sense that $\mu_0 : \mathcal{A} \to [0, \infty)$ is countably additive: if $E_1, E_2, \ldots \in \mathcal{A}$ are pairwise disjoint *and if* $\cup E_i \in \mathcal{A}$ then $\mu_0(\cup_i E_i) = \sum_i \mu_0(E_i)$. Can we promote μ_0 to a measure on the sigma-algebra generated by $\mathcal{A}$? The extension theorem says 'yes'. We will use the extension theorem to construct product measures and (if there's time in the semester) to prove the Riesz-Markov Theorem characterizing positive linear functionals on $C_c(X)$ for locally compact spaces X in terms of measures. It can also be used to construct Hausdorff measure and Haar measure but we won't do that (not enough time in the semester).

To begin, we use μ_0 to construct an outer measure. Define μ^* by:

$$\mu^*(A) = \inf \sum_i \mu_0(E_i)$$

where the infimum is taken over all collections $\{E_i\}$ of sets $E_i \in \mathcal{A}$ with $A \subset \cup_i E_i$. We need to show that μ^* is an outer measure. But first:

Exercise 114. If $A \in \mathcal{A}$ then $\mu^*(A) = \mu_0(A)$.

Proof. if $\{E_i\} \subset \mathcal{A}$ satisfies $A \subset \cup_i E_i$ then we'd like to say $\mu_0(A) \leq \sum_i \mu_0(E_i)$. Unfortunately, we don't know $\cup_i E_i \in \mathcal{A}$ so we can't do this directly. So let $B_n = (A \cap E_n) \setminus (\cup_{i<n} E_i)$. Then $B_n \in \mathcal{A}$ and $A = \cup_n B_n$ is a disjoint union. So

$$\mu_0(A) = \sum_i \mu_0(B_i) \leq \sum_i \mu_0(E_i)$$

(the last inequality occurs because $B_i \subset E_i$. Since $\{E_i\}$ is arbitrary, this implies $\mu_0(A) \leq \mu^*(A)$. The opposite inequality is also true since we can take $E_1 = A$ and $E_i = \emptyset$ for all $i > 1$. □

Exercise 115. μ^* is an outer measure.

Proof. It's clear that $\mu^*(\emptyset) = 0$ and μ^* is monotone. To show countable subadditivity, let $E \subset X$ and let $\{E_i\}$ be subsets with $E \subset \cup_i E_i$. We must show $\mu^*(E) \le \sum_i \mu^*(E_i)$. If $\mu(E_i) = +\infty$ for some i then we are done. So we can asume $\mu^*(E_i) < \infty$ for each i. Let $\epsilon > 0$. By definition of μ^* there exist sets $A_{ij} \in \mathcal{A}$ such that $E_i \subset \cup_j A_{ij}$ and

$$\mu^*(E_i) \ge -\epsilon/2^i + \sum_j \mu_0(A_{ij}).$$

Since the sets $\{A_{ij}\}$ cover E, we have

$$\mu(E) \le \sum_{ij} \mu_0(A_{ij}) \le \sum_i \mu^*(E_i) + \epsilon/2^i \le \epsilon + \sum_i \mu^*(E_i).$$

Since ϵ is arbitrary, this implies the exercise. $\square$

Exercise 116. Every set in $\mathcal{A}$ is measurable.

Proof. Let $E \subset X$ be arbitrary and $A \in \mathcal{A}$. Because of subadditivity, it suffices to show

$$\mu^*(E) \ge \mu^*(E \cap A) + \mu^*(E \setminus A).$$

We can assume wlog that $\mu^*(E) < \infty$. Let $\epsilon > 0$. By definition of μ^* there exist sets $\{A_i\} \subset \mathcal{A}$ such that $E \subset \cup_i A_i$ and $\mu^*(E) > -\epsilon + \sum_i \mu(A_i)$.

Because μ restricted to $\mathcal{A}$ is additive, we have

$$\mu(A_i) = \mu(A_i \cap A) + \mu(A_i \setminus A)$$

for each i. Adding up over i gives

$$\mu^*(E) > -\epsilon + \sum_i \mu(A_i \cap A) + \mu(A_i \cap A^c) \ge -\epsilon + \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

where in the last line we use subadditivity and monotonicity. Since this is true for every ϵ, we are done. $\square$

We now know that we can extend μ_0 originally defined on $\mathcal{A}$ to a measure μ defined on a sigma-algebra containing $\mathcal{A}$. Is this extension unique?

Exercise 117. Let $\mathcal{A}_\sigma$ be the collection of all countable unions of sets in $\mathcal{A}$ and $\mathcal{A}_{\sigma\delta}$ the collection of all countable intersections of sets in $\mathcal{A}_\sigma$. Suppose μ is σ-finite. Then for every μ^*-meaurable set E there exists $A \in \mathcal{A}_{\sigma\delta}$ with $E \subset A$ and $\mu(A \setminus E) = 0$.

Proof. Since μ is sigma-finite, $X = \cup_i X_i$ where $\mu(X_i) < \infty$.

By definition of μ, for every n there exist sets $A_{n,i,j} \in \mathcal{A}$ such that

$$-2^{-i}/n + \sum_{j=1}^{\infty} \mu(A_{n,i,j}) \leq \mu(E \cap X_i)$$

and $E \cap X_i \subset \cup_j A_{n,i,j}$.

So if $A_{n,i} = \cup_j A_{n,i,j}$ then $A_{n,i} \in \mathcal{A}_\sigma$ and $-2^{-i}/n + \mu(A_{n_i}) \leq \mu(E \cap X_i)$. Now let $A_n = \cup_i A_{n,i}$. Then $A_n \in \mathcal{A}_\sigma$, $E \subset A_n$ and

$$\mu(A_n \setminus E) \leq \sum_{i=1}^{\infty} \mu(A_{n,i} \setminus E \cap X_i) \leq 1/n.$$

Finally let $A = \cap_n A_n$. Then $A \in \mathcal{A}_{\sigma\delta}$, $E \subset A$ and we have $\mu(A \setminus E) \leq 1/n$ for all n (by monotonicity) which implies $\mu(A \setminus E) = 0$. □

Exercise 118. If μ_0 is σ-finite and $\mathcal{B}$ is the smallest sigma-algebra containing $\mathcal{A}$ then $\mu \restriction \mathcal{B}$ is the unique measure on $\mathcal{B}$ that agrees with μ_0 on $\mathcal{A}$.

Proof. Let ν be a measure on $\mathcal{B}$ that agrees with μ_0 on $\mathcal{A}$. By the previous exercise it is enough to show that for every $E \in \mathcal{A}_{\sigma\delta}$, $\nu(E) = \mu(E)$.

If $\{A_i\}_{i=1}^{\infty} \subset \mathcal{A}$ then $\cup_i A_i$ can be expressed as a disjoint union of sets in $\mathcal{A}$. Therefore $\nu(\cup_i A_i) = \mu(\cup_i A_i)$. So $\nu = \mu$ on $\mathcal{A}_\sigma$.

Consider sets $A_i, B_j \in \mathcal{A}$ $(i, j \in \mathbb{N})$. Note that $(\cup_i A_i) \cap (\cup_j B_j) = \cup_{i,j} A_i \cap B_j \in \mathcal{A}_\sigma$. So finite intersections of sets in $\mathcal{A}_\sigma$ are in $\mathcal{A}_\sigma$.

Now if $A_1, A_2, \ldots \in \mathcal{A}_\sigma$ and $B_n = \cap_{i=1}^{n} A_i$ then $B_n \in \mathcal{A}_\sigma$ and $\cap_i A_i = \cap_n B_n$. Moreover $\cap_n B_n$ is a decreasing intersection. That is $B_1 \supset B_2 \supset \cdots$. Assuming $A_1 = B_1$ has finite measure, then

$$\mu(\cap_i A_i) = \mu(\cap_n B_n) = \lim_n \mu(B_n) = \lim_n \nu(B_n) = \nu(\cap_n B_n) = \nu(\cap_i A_i).$$

In the general case, write $X = \cup_i X_i$ where $\mu(X_i) < \infty$ for all i. Then we have shown $\mu(X_i \cap \cap_j A_j) = \nu(X_i \cap \cap_j A_j)$ for all i. As we can assume the X_i's are pairwise disjoint, this proves that $\mu(\cap_j A_j) = \nu(\cap_j A_j)$ and therefore $\mu = \nu$ on $\mathcal{A}_{\sigma\delta}$ and so $\mu = \nu$ on $\mathcal{B}$.

□

If we only assume that μ_0 is finitely additive (instead of countably additive) then it might not extend to a measure on the sigma-algebra generated by $\mathcal{A}$. Indeed there are finitely additive probability measures on $\mathbb{Z}$ (defined on all subsets) that are shift-invariant. No such measure extends to a countably-additive measure.

25 Product measures

Given measure spaces $(X, \mathcal{X}, \mu)$ and $(Y, \mathcal{Y}, \nu)$ we'd like to construct a canonical measure space on the product $X \times Y$. If $A \in \mathcal{X}, B \in \mathcal{Y}$ then $A \times B$ is a **rectangle**. Let $\mathcal{R}$ denote the collection of all rectangles. Define $\lambda : \mathcal{R} \to [0, \infty]$ by $\lambda(A \times B) = \mu(A)\nu(B)$ (with the convention that $0\infty = 0$).

We would like to use Caratheodory's Theorem to extend λ to a measure on $X \times Y$. First we need to extend it to the algebra generated by $\mathcal{R}$. For that we need:

Exercise 119. If $A \times B = \sqcup_i A_i \times B_i$ then

$$\lambda(A \times B) = \sum_i \lambda(A_i \times B_i).$$

Proof. Observe that

$$\sum_i \nu(B_i)\chi_{A_i}(x) = \nu(B)\chi_A(x).$$

By the monotone convergence theorem,

$$\sum \int \nu(B_i)\chi_{A_i}(x)\ d\mu(x) = \int \sum \nu(B_i)\chi_{A_i}(x)\ d\mu(x) = \int \nu(B)\chi_A(x)\ d\mu(x) = \nu(B)\mu(A).$$

By direct computation the LHS is $\sum_i \lambda(A_i \times B_i)$. □

Let $\mathcal{A}$ be the algebra generated by $\mathcal{R}$.

Exercise 120. Every set $R \in \mathcal{A}$ is a finite union of rectangles.

Proof. It suffices to show that if $\mathcal{A}'$ the set of all finite unions of rectangles, then $\mathcal{A}'$ is an algebra. Note that

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \times B_2)$$

so $\mathcal{R}$ (and thus $\mathcal{A}'$) is closed under finite intersections. Also

$$(A \times B)^c = A^c \times B \cup A \times B^c \cup A^c \times B^c$$

So $\mathcal{A}'$ is closed under complementation. Obviously, $\mathcal{A}'$ is closed under finite unions; so this finishes it. □

Exercise 121. Any finite union of rectangles can be expressed as a finite *disjoint* union of rectangles.

Proof.

$$\begin{gathered}(A_1 \times B_1) \cup (A_2 \times B_2) \\ = (A_1 \cap A_2^c) \times (B_1 \cap B_2^c) \cup (A_1 \cap A_2^c) \times (B_1 \cap B_2) \\ \cup (A_1^c \cap A_2) \times (B_1^c \cap B_2) \cup (A_1^c \cap A_2) \times (B_1 \cap B_2) \\ \cup (A_1 \cap A_2) \times (B_1 \cap B_2^c) \cup (A_1 \cap A_2) \times (B_1^c \cap B_2) \cup (A_1 \cap A_2) \times (B_1 \cap B_2).\end{gathered}$$

□

Exercise 122. If $A \times B \in \mathcal{R}$ and $A_i \times B_i \in \mathcal{R}$ are countably many pairwise disjoint with $A \times B = \cup_i A_i \times B_i$ then $\lambda(A \times B) = \sum_i \lambda(A_i \times B_i)$.

Proof. For any $x \in X$,

$$\sum_i \nu(B_i)\chi_{A_i}(x) = \nu(B)\chi_A(x).$$

This is because $\{x\} \times B$ is the disjoint union of $\{x\} \times B_i$ over those i with $x \in A_i$. Now integrate over x,

$$\int \sum_i \nu(B_i)\chi_{A_i}(x) = \int \nu(B)\chi_A(x)\ d\mu(x) = \lambda(A \times B).$$

The Monotone Convergence Theorem allows us to switch the sum and integral in the LHS. So we obtain

$$\lambda(A \times B) = \sum_i \int \nu(B_i)\chi_{A_i}(x)\ d\mu(x) = \sum_i \lambda(A_i \times B_i).$$

□

Exercise 123. If $R_1, \dots, R_n$ are pairwise disjoint rectangles and $S_1, \dots, S_m$ are also pairwise disjoint rectangles and $\cup_i R_i = \cup_j S_j$ then

$$\sum_i \lambda(R_i) = \sum_j \lambda(S_j).$$

Proof. Observe that $R_i = \cup_j R_i \cap S_j$. So $\lambda(R_i) = \sum_j \lambda(R_i \cap S_j)$ and

$$\sum_i \lambda(R_i) = \sum_{i,j} \lambda(R_i \cap S_j).$$

Similarly, $\sum_j \lambda(S_j) = \sum_{i,j} \lambda(R_i \cap S_j)$. □

If $R_1, \dots, R_n$ are disjoint rectangles then we define $\lambda(\cup_i R_i) = \sum_i \lambda(R_i)$. The previous exercise show that this depends only on $\cup_i R_i$. So we have defined λ on the algebra generated by $\mathcal{R}$. The previous exercise also implies λ is countably additive. So Caratheodory's extension Theorem implies that λ extends to a measure on the sigma-algebra containing $\mathcal{R}$. If μ and ν are sigma-finite then λ is also sigma-finite and in this case λ is uniquely defined. We call it the **product measure** and denote it by $\mu \times \nu$.

Example: note that the direct product of Lebesgue measure on the real line with itself is Lebesgue measure on $\mathbb{R}^2$.

26 Fubini's Theorem

Theorem 26.1 (Fubini's Theorem)**.** *Let $(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$ be two complete measure spaces and f be an integrable function on $X \times Y$. Then*

1. *for a.e. $x \in X$, $f_x : Y \to \mathbb{C}$, $f_x(y) = f(x, y)$ is integrable on Y*
2. *for a.e. $y \in Y$, $f^y : X \to \mathbb{C}$, $f^y(x) = f(x, y)$ is integrable on X*
3. *$\int_Y f(x, y)\ d\nu(y)$ is an integrable function of x*
4. *$\int_X f(x, y)\ d\mu(x)$ is an integrable function of y*
5. *$\int(\int_Y f(x, y)\ d\nu(y))d\mu(x) = \int_{X\times Y} f\ d\mu \times \nu = \int(\int_X f(x, y)\ d\mu(x))d\nu(y)$.*

We will prove the theorem first for characteristic functions of sets. So if $E \subset X \times Y$, then we set $E_x = \{y \in Y : (x, y) \in E\}$, $E^y = \{(x, y) : (x, y) \in E\}$. These are called *cross sections.*

Recall that the collection of measurable sets of $X \times Y$ is the smallest σ-algebra containing all of the rectangles which is complete wrt $\mu \times \nu$.

Exercise 124. If $E \subset X \times Y$ is in the sigma-algebra generated by Ω then E_x and E^y are measurable for a.e. x. Hint: first show this is true for rectangles then the collection of sets for which it is true is a sigma-algebra.

Proof. Let Ω be the collection of all sets $E \subset X \times Y$ such that E_x is measurable for a.e. x. Observe that Ω contains all measurable rectangles $\Omega \supset \mathcal{R}$. Observe that Ω is a sigma-algebra because

- $X \times Y \in \Omega$
- $E \in \Omega \Rightarrow (E^c)_x = (E_x)^c \Rightarrow E^c \in \Omega$
- if $E_i \in \Omega$ and $E = \cup_i E_i$ then $E_x = \cup_i (E_i)_x$ hence $E \in \Omega$

□

Exercise 125. Suppose μ, ν are sigma-finite. Then for any measurable $E \subset X \times Y$, E_x and E^y are measurable for a.e. x. Moreover,

$$\mu \times \nu(E) = \int \mu(E^y)\ d\nu(y) = \int \nu(E_x)\ d\mu(x).$$

Hint: Let Ω be the collection of all measurable sets E satisfying the above. Show Ω contains the rectangles and finite unions of rectangles. Conclude that Ω contains the algebra $\mathcal{A}$ generated by rectangles. Then show Ω contains $\mathcal{A}_\sigma$ and $\mathcal{A}_{\sigma\delta}$. Use a previous approximation result to conclude the exercise. It may be easier to handle the case that $\mu(X)$ and $\nu(Y)$ are finite first.

Proof. We follow the hint. Clearly Ω contains all rectangles.

If $E_1, E_2, \ldots \in \Omega$ are pairwise disjoint and $E = \cup_i E_i$ then

$$\begin{aligned} \mu \times \nu(E) &= \sum_i \mu \times \nu(E_i) = \sum_i \int \mu(E_i^y)\ d\nu(y) \\ &= \int \sum_i \mu(E_i^y)\ d\nu(y) = \int \mu(E^y)\ d\nu(y) \end{aligned}$$

where the third equality follows from the Monotone Convergence Theorem. So $E \in \Omega$. Since every set in the algebra generated by the rectangles is a finite disjoint union of rectangles, this proves that the algebra $\mathcal{A}$ generated by the rectangles is contained in Ω.

If $E_1 \subset E_2 \subset \cdots$ is an increasing sequence of subsets of Ω then by the Monotone Convergence Theorem, $E := \cup_i E_i$ is also in Ω. So $\mathcal{A}_\sigma \subset \Omega$.

Now assume that $\mu(X)$ and $\nu(Y)$ are finite.

If $E_1 \supset E_2 \supset \cdots \in \Omega$ then $\cap_i E_i \in \Omega$ by Lebesgue's Dominated Convergence Theorem. So $\mathcal{A}_{\sigma\delta} \subset \Omega$.

If $\mu \times \nu(E) = 0$ then by a previous exercise there exists $F \in \mathcal{A}_{\sigma\delta}$ with $E \subset F$ and $\mu \times \nu(F) = 0$. Since $E^y \subset F^y$ for every y, this implies $E \in \Omega$.

Now let E be an arbitrary measurable set. By a previous exercise, there exist $F \in \mathcal{A}_{\sigma\delta}$ such that $E \subset F$ and $\mu \times \nu(F \setminus E) = 0$. So

$$\begin{aligned} \mu \times \nu(E) &= \mu \times \nu(F) = \int \mu(F^y)\ d\nu(y) = \int \mu(F^y) - \mu((F \setminus E)^y)\ d\nu(y) \\ &= \int \mu(E^y)\ d\nu(y). \end{aligned}$$

So $E \in \Omega$. This proves the exercise when $\mu \times \nu(X \times Y) < \infty$.

The general case follows from the finite measure case because Ω is closed under countable disjoint unions. □

Exercise 126 (Tonelli's Theorem). If $f \geq 0$ is measurable on $X \times Y$ then

$$\iint f(x,y)\ d\mu(x)d\nu(y) = \iint f(x,y)\ d\nu(y)d\mu(x) = \int f\ d\mu \times \nu.$$

In particular $x \mapsto f(x,y)$ is measurable for a.e. y.

Proof. The previous exercise states that this result is true when f is a characteristic function. By linearity, it holds for simple functions. By the Monotone Convergence Theorem, it holds for nonnegative functions. □

Exercise 127. Prove Fubini's Theorem.

Proof. This follows from Tonelli's Theorem by splitting f into real and imaginary parts and then splitting each of these parts into positive and negative parts. □

27 Convolution

Given measurable functions f, g on $\mathbb{R}^n$ we define the function $f * g$ by

$$f * g(x) = \int f(t)g(x-t)\ dt$$

whenever this integral exists. This is called the **convolution**. Before we go further, there is one minor detail to contend with: how do we know $f * g$ is measurable? Well suppose we knew that the two variables function $F(x,t) = f(t)g(x-t)$ is measurable. Then $f * g(x) = \int F(x,t)\ dt$ is measurable by Fubini's Theorem (assuming as we do, that F is integrable). But how do we know that F is measurable? First of all, the function $G(x,t) = f(t)g(x)$ is measurable. In fact, since f and g are pointwise limits of simple functions, G is also a pointwise limit of simple functions. Now $F = G \circ H$ where $H(x,t) = (t, x-t)$. H is continuous but we know that *pre*-compositions of continuous functions with measurable functions do not have to be measurable. (Post-compositions are OK). What to do? Well, it is an exercise to show that any measurable function agrees with a Borel function almost everywhere. (To prove this, use the lemma before the proof of the Radon-Nikodym Theorem). Compositions of Borel functions are Borel. So we can find a Borel function G' that agrees with G almost everywhere and $G' \circ H = G \circ H$ almost everywhere because H^{-1} takes measure zero sets to measure zero sets.

Observe that if $s = x - t$ then

$$\int f(t)g(x-t)\ dt = \int f(x-s)g(s)\ ds.$$

So $f * g = g * f$.

Example: suppose $f = \frac{1}{2\epsilon}\chi_{[-\epsilon,\epsilon]}$. Then

$$f * g(x) = \frac{1}{2\epsilon}\int_{-\epsilon}^{\epsilon} g(x-t)\ dt$$

is the average value of g in the interval $[x-\epsilon, x+\epsilon]$. Observe:

1. if $g \in L^1$ then $f * g \in L^1$ and $\|f * g\|_1 \le \|g\|_1$.
2. if $g \in L^1$ then $f * g$ is continuous!
3. $\lim_{\epsilon \searrow 0} f * g(x) = g(x)$ a.e. Why? (there is a one line answer).

The main technical goal of this section is to generalize this example.

Some motivation:

1. Probability theory: suppose f and g are probability distributions, $X, Y \in \mathbb{R}^n$ are chosen independently at random with laws f, g respectively. What's the distribution of $X+Y$? Of course, it is $f*g$. In particular, if $X_1, X_2, \ldots, X_n$ are iid random variables with distribution f then $X_1 + \cdots + X_n$ has distribution equal to the n-fold convolution of f with itself. (These facts are especially easy to verify if we replace $\mathbb{R}^n$ with $\mathbb{Z}^n$).

2. Image processing: it is common to 'clean up' photographs by convolving with a Gaussian. For example, suppose a satellite image gives the height $h(x, y)$ above sea level of a point (x, y) on the Earth. This function can be very 'noisy'. So we replace it with $h * g$ where $g(x, y) = Ce^{-x^2-y^2}$ is a Gaussian. This has the effect of replacing the value $h(x, y)$ with the weighted average $\int h(x-t, y-w)g(t, w)\, dtdw$ allowing features of the landscape to be more easily identifiable.

3. Fourier analysis: the Fourier transform of $f \in L^1(\mathbb{R})$ is the function

$$\hat{f}(t) = \int f(x)e^{-ixt}\, dx.$$

   It can be shown that $\widehat{f * g} = \hat{f}\hat{g}$. So convolution is transformed into ordinary pointwise multiplication.

4. PDE's: for example, solutions to the heat equation can be expressed in terms of convolution. For example, let

$$\Phi_t(x) = \frac{1}{\sqrt{4\pi t}} \exp(-x^2/4t).$$

   Let $f(x)$ denote an initial temperature distribution on $\mathbb{R}$. Let $u(x, t)$ denote the temperature distribution at time $t \geq 0$ (so $u(x, 0) = f(x)$). Then

$$u(x, t) = f * \Phi_t(x) = \int f(y)\Phi_t(x - y)\, dy.$$

   On physical considerations you might expect that $u(x, t)$ is smooth in the spatial variable for $t > 0$. This is correct if, for example, $f \in L^1$. (BTW Φ is the solution to the heat equation $u_t = u_{xx}$ with initial distribution δ_0).

27.1 Norm inequalities

Exercise 128.

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1$$

with equality if $f \geq 0$ and $g \geq 0$.

Proof. Since $|f * g| \leq |f| * |g|$, we may assume that $f \geq 0$ and $g \geq 0$.

$$\begin{aligned} \|f * g\|_1 &= \iint f(t)g(x-t)\, dtdx = \iint f(t)g(x-t)\, dxdt = \int f(t)\left(\int g(x-t)\, dx\right) dt \\ &= \left(\int f\, dt\right)\left(\int g\, dx\right) = \|f\|_1\|g\|_1. \end{aligned}$$

□

One nice consequence of this formula is that convolution defines a product structure on L^1. Indeed, $L^1(\mathbb{R}^n)$ is a **Banach algebra**: it is a Banach space with a product satisfying associativity, distributivity and $\|xy\| \leq \|x\|\|y\|$. Associativity is nontrivial: you should check this:

Exercise 129. Prove that $(f * g) * h = f * (g * h)$ assuming $f, g, h \in L^1(\mathbb{R}^n)$.

Exercise 130. If $f \in L^p$ and $g \in L^1$ then

$$\|f * g\|_p \leq \|f\|_p \|g\|_1.$$

More generally, if $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$, $f \in L^p, g \in L^q$ then

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Proof. Let us just prove the first statement. The second is a homework exercise. We assume wlog that $f, g \geq 0$. The cases $p = 1, p = \infty$ are easy and left to the reader. Let q be a conjugate exponent to p. Then

$$\begin{aligned} f * g(x) &= \int f(t)g(x-t)\ dt = \int f(t)g(x-t)^{1/p} g(x-t)^{1/q}\ dt \\ &\leq \left(\int f^p(t)g(x-t)\ dt\right)^{1/p} \left(\int g(x-t)\ dt\right)^{1/q} \\ &= (f^p * g(x))^{1/p} \|g\|_1^{1/q}. \end{aligned}$$

Raising to the p-power and integrating over x and using the previous exercise we obtain

$$\|f * g\|_p^p \leq \|f\|_p^p \|g\|_1 \|g\|_1^{p/q}.$$

Taking p-th roots finishes it. $\square$

27.2 Smoothness

Let $K \in L^1(\mathbb{R}^n)$. Then convolution with K defines an operator $L^1(\mathbb{R}^n) \to L^1(\mathbb{R}^n)$, $f \mapsto f * K$. This operator has norm $= \|K\|_1$ by the previous theorems.

One very practical purpose of convolution is as a tool for increasing the regularity of a function. That is: the convolution of an arbitrary L^1 function with a smooth function will be smooth. So if K is smooth enough then convolution with K defines an operator from $L^1(\mathbb{R}^n)$ to smooth functions on $\mathbb{R}^n$.

To make this precise, recall that $C^m(\mathbb{R}^n)$ is the space of functions all of whose derivatives of order $\leq m$ exist and are continuous.

Exercise 131. Let $1 \leq p < \infty$. For any $f \in L^p(\mathbb{R}^n)$ and $t \in \mathbb{R}^n$ let $\tau_t(f) \in L^p(\mathbb{R}^n)$ be the function $\tau_t(f)(x) = f(x - t)$. Show that

$$\lim_{t \to 0} \|\tau_t(f) - f\|_p = 0.$$

Proof. This is obvious if f is continuous with compact support. Such functions are dense in $L^p(\mathbb{R}^n)$. Alternatively, it is easy to check that the result is true if f is a characteristic function (using absolute continuity of the integral). By taking linear combinations, it is true for simple functions and because simple functions are dense, it is true. $\square$

Exercise 132. Let $f \in L^p(\mathbb{R}^n), K \in C^m(\mathbb{R}^n)$ and suppose K has compact support. Then $f * K \in C^m$ and

$$D^\alpha(f * K) = f * D^\alpha K$$

if $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $|\alpha| = \sum_i \alpha_i \leq m$.

Hint: work out the case $m = 0$ first. This means that K is continuous and you have to show that $f * K$ is continuous.

Proof. Let's first handle the case $m = 0$. By changing variables,

$$\begin{aligned} |f * K(x) - f * K(y)| &= |\int f(t)[K(x-t) - K(y-t)]\, dt| = |\int f(x-t)[K(t) - K(t-x+y)]\, dt| \\ &\leq \|f\|_p \|K - \tau_{x-y}(K)\|_q. \end{aligned}$$

Because K is continuous and compactly supported, $\|K - \tau_{x-y}(K)\|_q \to 0$ as $x - y \to 0$. This proves continuity (in fact, uniform continuity).

Now suppose $m > 0$. Let $e_i \in \mathbb{R}^n$ denote the i-th basis vector. Then

$$\begin{aligned} \frac{f * K(x + he_i) - f * K(x)}{h} &= \int f(t) \frac{K(x - t + he_i) - K(x-t)}{h}\, dt \\ &= \int f(t) \frac{\partial K}{\partial x_i}(x - t + h'e_i)\, dt \end{aligned}$$

where $h' \in [0, h]$ is determined by the Mean Value Theorem (h' depends on t and x). As $h \to 0$, $\frac{\partial K}{\partial x_i}(x - t + h'e_i) \to \frac{\partial K}{\partial x_i}(x - t)$ uniformly in t (because K has compact support). By the Uniform Convergence Theorem, $\partial(f * K)/\partial x_i = f * (\partial K/\partial x_i)$. The general case follows by induction. $\square$

27.3 Approximate identities

As mentioned above, convolution defines a product structure on L^1. Is there a multiplicative identity element? This would have to be a function $g \in L^1$ such that for any $f \in L^1$, $f * g = f$.

Such a thing exists only if you allow g to be a measure instead of a function. To be precise, if $f \in L^1$ and μ is a finite measure on $\mathbb{R}^n$ then their convolution $\mu * f$ is the function defined by

$$\mu * f(x) = \int f(x - t)\, d\mu(t).$$

Let δ_0 be the Dirac measure concentrated on 0. Then $\delta_0 * f = f$. So δ_0 is the multiplicative identity. Of course, its not a function. However, we will see that there exist sequences $\{K_j\}$ of functions such that $f * K_j \to f$ as $j \to \infty$. Convergence will either mean in L^p (for appropriate p) or pointwise at continuity points.

Remark 3. BTW, we can also speak of convolution of measures:

$$\mu * \nu(E) = \iint \chi_E(x+y)\ d\mu(x)d\nu(y)$$

on $\mathbb{R}^n$ (or on any group with a measurable structure). If $f \in L^1(\mathbb{R}^n)$ then we can identify f with the measure given by $f(E) = \int_E f\ dm$ where m is Lebesgue measure. You can check that these notions are compatible. To summarize the paragraph above: the convolution identity is δ_0, a measure. However, we can find functions $K_j \in L^1$ such that *if we interpret these functions as measures* then they converge (in a weak* sense to be explained in later lectures) to δ_0. Since convolution is continuous in L^1 (and in the weak* topology), it shouldn't surprise us that $f * K_j \to f$.

Exercise 133. Suppose $\{K_j\} \subset L^1(\mathbb{R}^n)$ is a sequence satisfying

1. $\int K_j\ dm = 1$
2. $\sup_j \|K_j\|_1 < \infty$
3. for every $\delta > 0$,
$$\lim_j \int_{|x|>\delta} |K_j(x)|\ dm(x) = 0.$$

Then for every $f \in L^p$, $1 \le p < \infty$

$$\lim_{\epsilon \searrow 0} f * K_j = f$$

where the limit is in $L^p(\mathbb{R}^n)$. In other words, $\|f * K_j - f\|_p \to 0$ as $\epsilon \searrow 0$.

Proof.

$$\begin{aligned}
|f * K_j(x) - f(x)| &= \left| \int [f(x-t) - f(x)] K_j(t)\ dm(t) \right| \\
&\le \int |(f(x-t) - f(x)) K_j(t)|\ dm(t) \\
&= \int |f(x-t) - f(x)| |K_j(t)|^{\frac{1}{p}} |K_j(t)|^{\frac{1}{q}}\ dm(t) \\
&\le \left(\int |f(x-t) - f(x)|^p |K_j(t)|\ dm(t) \right)^{1/p} \left(\int |K_j(t)|\ dm(t) \right)^{1/q} \\
&= \left(\int |f(x-t) - f(x)|^p |K_j(t)|\ dm(t) \right)^{1/p} \|K_j\|_1^{1/q}.
\end{aligned}$$

So for any $\delta > 0$,

$$\begin{aligned}
\|f * K_j - f\|_p^p &= \int |f * K_j(x) - f(x)|^p \, dm(x) \\
&\le \|K_j\|_1^{p/q} \int \left(\int |f(x-t) - f(x)|^p |K_j(t)| \, dm(t) \right) \, dm(x) \\
&= \|K_j\|_1^{p/q} \iint |f(x-t) - f(x)|^p |K_j(t)| \, dm(x) dm(t) \\
&= \|K_j\|_1^{p/q} \int \|\tau_t(f) - f\|_p^p |K_j(t)| \, dm(t) \\
&\le \|K_j\|_1^{p/q} \int_{|t|<\delta} \|\tau_t(f) - f\|_p^p |K_j(t)| \, dm(t) \\
&\quad + \|K_j\|_1^{p/q} \int_{|t|\ge\delta} 2^p \|f\|_p^p |K_j(t)| \, dm(t).
\end{aligned}$$

Let $\eta > 0$. Then there exists $\delta > 0$ such that $\|\tau_t(f) - f\|_p^p < \eta$ for all $|t| < \delta$. So for this particular δ we have

$$\begin{aligned}
\limsup_{j\to\infty} \|f * K_j - f\|_p^p &\le \limsup_{j\to\infty} \|K_j\|_1^{p/q+1}\eta + \limsup_{j\to\infty} \|K_j\|_1^{p/q} \int_{|t|\ge\delta} 2^p \|f\|_p^p |K_j(t)| \, dm(t) \\
&\le \limsup_{j\to\infty} \|K_j\|_1^{p/q}\eta.
\end{aligned}$$

Since η is arbitrary, this does it. □

Exercise 134. If $1 \le p < \infty$ then $C_c^\infty(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$.

Proof. The functions with compact support are dense in L^p; so it suffices to prove that if $f \in L^p$ has compact support then there exists a sequence of compact supported smooth functions converging to f (in L^p norm).

Let K be a compactly support smooth function with $\int K = 1$. For $\epsilon > 0$, let $K_\epsilon(x) = \epsilon^{-n}K(x/\epsilon)$. Observe that $\int K_\epsilon = 1$.

Then for any $f \in L^p$ with compact support $f * K_\epsilon \to f$ as $\epsilon \searrow 0$ and $f * K_\epsilon$ are compactly supported and smooth. □

28 The space of measures

Let X be a topological space and let $C_0(X)$ denote the space of continuous functions on X that "vanish at infinity". This means that for every $f \in C_0(X)$ and $\epsilon > 0$ there exists a compact set $K \subset X$ such that for all $x \notin K$, $|f(x)| < \epsilon$. The space $C_0(X)$ is a Banach space with the norm $\|f\| = \sup_{x\in X} |f(x)|$.

Let $C_c(X) \subset C_0(X)$ denote the subspace of continuous functions that are compactly supported. This space is dense in $C_0(X)$.

Suppose μ is a Borel measure on X and f is a continuous function on X. Then we define

$$\mu(f) := \int f \, d\mu$$

whenever this integral exists. This defines a linear functional on $C_c(X)$ (assuming that μ is finite on compact sets). Moreover if $f \geq 0$ then $\mu(f) \geq 0$. The Riesz-Markov Theorem is a converse to this:

Theorem 28.1 (Riesz-Markov Theorem)**.** *Let X be a locally compact Hausdorff space. Let Λ be a positive linear functional on $C_c(X)$. Being positive means that $\Lambda(f) \geq 0$ for every nonnegative $f \in C_c(X)$. Then there is a unique Radon measure μ on X such that $\Lambda(f) = \mu(f)$ for every $f \in C_c(X)$. A Radon measure is a Borel measure which is finite on compact sets and is inner regular as explained below.*

From this result we can obtain a complete picture of $C_0(X)^*$:

Theorem 28.2 (Riesz Representation Theorem)**.** *For every $\rho \in C_0(X)^*$ there is a finite signed Borel measure μ such that*

$$\rho(f) = \mu(f).$$

One of the applications of these results (that we will not prove) is the existence and uniqueness of Haar measure. To be precise, let G be a locally compact group (e.g. $GL(n, \mathbb{R})$, $O(n)$, the isometry group of hyperbolic n-space, the absolute Galois group of a number field, etc). Then there exists a Radon measure μ on G that is invariant under left multiplication ($\mu(gE) = \mu(E)$ for Borel $E \subset G$ and $g \in G$). It is unique up to scalar multiplication. For example, Lebesgue measure is the Haar measure on $\mathbb{R}^n$ (as an additive group).

The first issue related to the Riesz-Markov Theorem we will deal with is the uniqueness. Why should it be true that if $\mu(f) = \nu(f)$ for every $f \in C_c(X)$ that $\mu = \nu$? This question is connected with the issue of regularity of a measure explained next.

28.1 Regularity

Definition 43. A measure μ on a topological space X is *inner regular* if for every Borel set E,

$$\mu(E) = \sup\{\mu(K) : \ K \subset E\}$$

where the sup is over all compact sets in E. The measure is *outer regular* if for every Borel E,

$$\mu(E) = \inf\{\mu(O) : \ O \supset E\}$$

where the inf is over all open sets containing E. The measure is *regular* if it is both inner and outer regular. For example, Lebesgue measure on $\mathbb{R}^n$ is regular.

A *Polish space* is a separable complete metric space. A locally compact space is a topological space in which every point has an open neighborhood whose closure is compact. In this section we will study Borel measures on a locally compact Polish space (X, d). While many of the results we obtain generalize beyond this setting we will restrict our attention to these spaces in order to avoid technicalities.

Exercise 135. Suppose X is a compact metric space. Let μ denote a finite Borel measure on X. Then μ is regular.

Proof. Let Ω denote the collection of all Borel subsets E of X such that

$$\mu(E) = \sup\{\mu(K) : \ K \subset E\}$$

and

$$\mu(E) = \inf\{\mu(O) : \ O \supset E\}.$$

Let $E \subset X$ be closed (and therefore compact) and let $O_n = \{x \in X : \ d(x, E) < 1/n\}$. Then O_n is open and $E = \cap_n O_n$. This shows that $E \in \Omega$.

Observe that $E \in \Omega$ if and only if $E^c \in \Omega$.

Now suppose $E_1, E_2, \ldots \in \Omega$. Let $\epsilon > 0$ and $K_i \subset E_i$ be compact such that $\mu(E_i \setminus K_i) < \epsilon/2^i, \mu(O_i \setminus E_i) < \epsilon/2^i$. Then $\mu(\cup_i E_i \setminus \cup_i K_i) < \epsilon$. So there is some n such that $\mu(\cup_{i=1}^{\infty} E_i \setminus \cup_{i=1}^{n} K_i) < \epsilon$. Since $\cup_{i=1}^{n} K_i$ is compact and $\epsilon > 0$ is arbitrary, this shows $\mu(E) = \sup\{\mu(K) : \ K \subset E\}$.

Similarly, $E_i \subset O_i$ be open such that $\mu(E_i \setminus O_i) < \epsilon/2^i$. Then $\mu(\cup_i O_i \setminus \cup_i E_i) < \epsilon$. Since $\cup_i O_i$ is open and $\epsilon > 0$ is arbitrary, this shows $\mu(E) = \inf\{\mu(O) : \ O \supset E\}$.

So we have shown Ω contains all closed sets and is closed under complementation and countable unions. So it contains all Borel sets. This proves μ is regular. □

Exercise 136. If $K \subset X$ is compact and X is locally compact and Polish then there exist precompact open sets $O_1 \supset O_2 \supset \cdots$ such that $\cap_i O_i = K$.

Proof. Because K is compact and X is locally compact, K can be covered by a finite number of open sets whose closure is compact. Taking the union of the sets, we see $K \subset U$ for some open set U whose closure is compact. Let $W_n = \{x \in U : \ d(x, K) < 1/n\}$. This is an open set with compact closure and $\cap_n W_n = K$. □

Exercise 137. Suppose X is a locally compact Polish space. Let μ denote a Borel measure on X such that $\mu(K) < \infty$ for every compact $K \subset X$. Then μ is regular.

Proof. Let $E \subset X$ be Borel. Let $K_1 \subset K_2 \subset ..$ be increasing compact sets such that $X = \cup_i K_i$. For each i there exists a compact set $L_i \subset E_i \cap K_i$ with $\mu(L_i) > -\epsilon/2^i + \mu(E \cap K_i)$. Without loss of generality we may assume $L_1 \subset L_2 \subset ...$ So

$$\mu(\cup_i L_i) = \lim_i \mu(L_i) \geq \lim_i -\epsilon/2^i + \mu(E_i \cap K_i) = \mu(E).$$

So $\mu(E) = \sup\{\mu(K) : \ K \subset E\}$.

Let $O_i \supset E \cap K_i$ be an open subset of K_i with $\mu(O_i) \leq \epsilon/2^i + \mu(E \cap K_i)$. Now O_i need not be open in X. However, because it is open there exists an open set $O_i' \subset X$ with $O_i' \cap K_i = O_i$.

By the previous exercise there exists a precompact open set $V_i \subset X$ such that $K_i \subset V_i$ and $\mu(V_i \setminus K_i) < \epsilon/2^i$.

After replacing O_i' with $O_i' \cap V_i$ if necessary, we may assume that $\mu(O_i') < \mu(E \cap K_i) + \epsilon/2^{i-1}$. So $\cup_i O_i'$ is an open set containing E with $\mu(\cup_i O_i' \setminus E) < 2\epsilon$. This proves that

$$\mu(E) = \inf\{\mu(O) : O \supset E\}.$$

Since E is arbitrary, μ is regular.

□

Definition 44. A *Radon measure* on a space X is an inner regular Borel measure μ such that $\mu(K) < \infty$ for every compact K. We have just shown that every Borel measure on a locally compact Polish space which is finite on compact sets is a Radon measure.

We will need:

Theorem 28.3 (Tietze's Extension Theorem)**.** *Suppose X is a normal space (this means that X is Hausdorff and if C_1, C_2 are disjoint closed subsets of X then there exist disjoint open sets O_1, O_2 with $C_i \subset O_i$ for $i = 1, 2$). Then if $K \subset O \subset X$ where K is closed and O is open then there exists a continuous function $f : X \to [0, 1]$ such that $f(x) = 1$ for $x \in K$ and $f(x) = 0$ for $x \notin O$.*

Exercise 138. Suppose μ, ν are two Radon measures on a locally compact Polish space X and $\mu(f) = \nu(f)$ for every $f \in C_c(X)$. Then $\mu = \nu$. Hint: first prove that $\mu(K) = \nu(K)$ for compact sets K. Use Tietze's extension theorem.

Proof. Let $K \subset X$ be compact and $O_1 \supset O_2 \supset \cdots$ precompact open sets with $\cap_n O_n = K$. By Tietze's extension theorem there exist continuous functions $f_n : X \to [0, 1]$ with $f_n(x) = 1$ for $x \in K$ and $f_n(x) = 0$ for $x \notin O_n$.

$$\mu(K) = \lim_n \mu(f_n) = \lim_n \nu(f_n) = \nu(K)$$

by the bounded convergence theorem. Since μ and ν are regular, this implies $\mu = \nu$. □

28.2 Some examples

Now that we know that Radon measures are determined by their values of compactly supported continuous functions, we may consider some topological considerations. Let $M(X)$ denote the space of all Radon measures on X with the following topology. We say that $\{\mu_n\}_{n \in \mathbb{N}}$ converges to μ_∞ if

$$\mu_n(f) \to \mu_\infty(f)$$

for every $f \in C_c(X)$.

Exercise 139. Suppose $X = \mathbb{R}^n$. Construct a sequence of probability measures μ_n that are absolutely continuous to Lebesgue measure and converge to the Dirac measure δ_0 (this is the measure on $\mathbb{R}^n$ defined by $\delta_0(E) = 1$ if $0 \in E$ and $\delta_0(E) = 0$ otherwise.

Exercise 140. Construct a Borel probability measure on the middle thirds Cantor set as a limit of measures on $\mathbb{R}$ each of which is absolutely continuous to Lebesgue measure. This example (and the previous one) show that a limit of absolutely continuous measures can be singular.

Proof. Let $K_1 = [0,1], K_2 = [0,1/3] \cup [2/3,1], K_3 = [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1]$ etc be the sets in the constructions of the Cantor set. So if C denotes the Cantor set then $C = \cap_n K_n$ is a decreasing intersection. Observe that $m(K_n) = (2/3)^{n-1}$. Define measures μ_n on $\mathbb{R}$ by

$$\mu_n(E) = (3/2)^{n-1} m(E \cap K_n).$$

Each μ_n is a probability measure. We claim that the limit $\mu_\infty = \lim_n \mu_n$ exists and is supported on C. Observe that if f is a characteristic function of an interval then $\mu_n(f)$ eventually stabilizes. In particular, $\lim_n \mu_n(f)$ exists. By linearity, the same holds for simple functions. Also $\lim_n |\mu_n(f)| \le \|f\|_\infty$ since this holds for each n individually. We can approximate any $f \in C_c(\mathbb{R})$ by step functions in the $\|\cdot\|_\infty$ norm. So we obtain that $\lim_n \mu_n(f)$ exists for every $f \in C_c(\mathbb{R})$. This explains that $\mu_\infty = \lim_n \mu_n$ exists. The fact that it is supported on C follows because: if $O \subset C$ is any nonempty open set then it contains one of the intervals I in the construction. If I is an interval of K_n then $\mu_n(O) \ge \mu_n(I) = (1/2)^{n-1}$. Moreover, if $k > n$ then $\mu_k(I) = \mu_n(I)$. So we must have $\mu_\infty(O) \ge (1/2)^{n-1} > 0$. This shows that the support of μ_∞ contains C. It is obvious that the support is contained in C. $\square$

If X is compact then it can be proven that the space of probability measures $P(X) \subset M(X)$ is also compact (note: $P(X)$ embeds into $[-1,1]^{C(X)}$ which is compact).

Exercise 141. Suppose $T : X \to X$ is a homeomorphism and $\mu \in P(X)$. Let $T_*\mu \in P(X)$ be the measure $T_*\mu(E) = \mu(T^{-1}E)$. Show that T_* is a homeomorphism of $P(X)$. Also show that if $\mu \in P(X)$ then any limit point of

$$\frac{1}{n}\sum_{i=1}^{n} T_*^i \mu$$

is a fixed point for T. Therefore, there exists a T-invariant measure on X.

The latter exercise is a fundamental result in dynamics...

(Stuff I might add in later: weak* convergence, the Cantor set example, the space of invariant measures, the existence of invariant measures, horseshoe example?, Gauss' transformation?, circle rotations)

29 Fourier series

29.1 Definition and convolution

Let $\mathbb{T} \subset \mathbb{C}$ denote the unit circle. We can consider $\mathbb{T}$ as a group under multiplication.

Exercise 142. Classify the homomorphisms from $\mathbb{T}$ to itself.

We consider $\mathbb{T}$ with Lebesgue measure normalized to have mass 1. In other words, consider the map $\pi : \mathbb{R} \to \mathbb{T}$ given by $\pi(x) = e^{ix}$. This maps $[0, 2\pi)$ onto $\mathbb{T}$ and we consider $\mathbb{T}$ with the push-forward measure divided by 2π.

Exercise 143. If $h : \mathbb{T} \to \mathbb{T}$ is a homomorphism and $f \in L^1(\mathbb{T})$. Let

$$\hat{f}(h) := \int_{\mathbb{T}} f(x)h(x^{-1})\ dx.$$

Show that for any $f, g \in L^1(\mathbb{T})$,

$$\widehat{f * g}(h) = \hat{f}(h)\hat{g}(h).$$

Proof.

$$\begin{aligned}
\widehat{f * g}(h) &= \int_{\mathbb{T}} \widehat{f * g}(x)h(x^{-1})\ dx = \int_{\mathbb{T}}\int_{\mathbb{T}} f(xy^{-1})g(y)h(x^{-1})\ dydx \\
&= \int_{\mathbb{T}}\int_{\mathbb{T}} f(xy^{-1})g(y)h(x^{-1}y)h(y^{-1})\ dydx \\
&= \int_{\mathbb{T}}\int_{\mathbb{T}} f(xy^{-1})g(y)h(x^{-1}y)h(y^{-1})\ dxdy \\
&= \hat{f}(h)\hat{g}(h).
\end{aligned}$$

□

To put this another way, we identify $\mathbb{T}$ with $[0, 2\pi)$. Then the exercise states that if $f \in L^1(0, 2\pi)$, $n \in \mathbb{Z}$ and

$$\hat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\ dx$$

then $\widehat{f * g}(n) = \hat{f}(n)\hat{g}(n)$ where the convolution is taken mod 2π.

Remark 4. We can consider Fourier analysis on an arbitrary locally compact abelian group G. Let $\hat{G} = \mathrm{Hom}(G, \mathbb{T})$. This is also an abelian group under pointwise addition. (We will write the group law of both G and $\hat{G}$ additively). It is a theorem that $\hat{G}$ is also locally compact. Because these groups are locally compact, they each admit a Haar measure. Now if $h \in \mathrm{Hom}(G, \mathbb{T})$ and $f \in L^1(G)$ then we write

$$\hat{f}(h) = \int f(x)h(-x)\ dx.$$

This defines a homomorphism from the Banach algebra $L^1(G)$ to $\mathbb{C}$ in the sense that it is linear and $\widehat{f * g}(h) = \hat{f}(h)\hat{g}(h)$. In other words, we have a natural map form $\mathrm{Hom}(G, \mathbb{T})$ to $\mathrm{Hom}(L^1(G), \mathbb{C})$ given by $h \mapsto (f \mapsto \hat{f}(h))$. It can be shown that this map is bijective.

29.2 L^2

Let e_n denote the function on $\mathbb{T}$ given by $e_n(x) = e^{inx}$ where we have identified $\mathbb{T}$ with $[0, 2\pi)$. (We could instead write $e_n(z) = z^n$ where we consider $\mathbb{T} \subset \mathbb{C}$). A **trigonometric polynomial** is a finite linear combination of the e_n's. That is to say, a trigonometric polynomial is any function in the span of the e_n's.

Exercise 144. The functions e_n ($n \in \mathbb{Z}$) are orthonormal in $L^2(\mathbb{T})$.

Note that $\hat{f}(n) = \langle f, e_n \rangle$.

We'd like to prove that if $f \in L^2(\mathbb{T})$ then

$$f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle e_n = \sum_{n \in \mathbb{Z}} \hat{f}(n) e_n$$

in the sense that this sum converges in L^2. However, we would need to know that the e_n's form a basis. This is not obvious.

To put things another way, we would like to show that

$$f = \lim_{N \to \infty} \sum_{n=-N}^{N} \langle f, e_n \rangle e_n$$

(in L^2). This implies that the trigonometric polynomials are dense in L^2.

Consider

$$D_N(x) = \sum_{n=-N}^{N} e^{inx}.$$

These are called **Dirichlet kernels**.

Exercise 145. If $f \in L^2(\mathbb{T})$ then

$$\sum_{n=-N}^{N} \langle f, e_n \rangle e^{inx} = D_N * f(x).$$

Proof. $f * e_n(x) = \int_{\mathbb{T}} f(t) e^{in(x-t)}\, dt = e^{inx} \int_{\mathbb{T}} f(t) e^{-int}\, dt = \langle f, e_n \rangle e^{inx}$. So the result follows by summing n from $-N$ to N. $\square$

Unfortunately it's not true that the D_N's form an approximation to the identity. So we let

$$F_N(x) = \frac{1}{N+1} \sum_{k=0}^{N} D_k.$$

These are called **Fejér kernels**. (They are Cesaró sums of Dirichlet kernels).

Exercise 146. By the way, why should the Fejer kernels behave better than the Dirichlet kernels? Well, if $\{x_i\}$ is sequence we could consider the partial sums $s_n = \sum_{i=1}^{n} x_i$. Sometimes the partial sums do not converge. For example, this occurs if $x_i = (-1)^i$. The **Cesaro sums** of are defined by $c_n := \frac{1}{n}\sum_{i=1}^n s_i$. It can happen that the Cesaro sums converge even when the partial sums do not. For example, this is true when $x_i = (-1)^n$. (Check!). However if the partial sums converge then the Cesaro sums also converge. And they converge to the same limit.

Exercise 147. If $f \in L^2(\mathbb{T})$ then

$$\frac{1}{N+1}\sum_{k=0}^{N}\sum_{n=-k}^{k}\langle f, e_n\rangle e^{inx} = F_N * f(x).$$

We will show that F_N does form an approximation to the identity. From this it follows that $F_N * f \to f$ as $N \to \infty$ in $L^2(\mathbb{T})$. This $F_N * f$ is a trigonometric polynomial (that is, it is in the span of the e_n's), it follows that the e_n's form an ON basis.

First we study the Dirichlet kernels in a bit more detail. Observe that

$$\begin{aligned} D_N(x) &= e^{-iNx}\sum_{n=0}^{2N} e^{inx} = e^{-iNx}\frac{1-e^{i(2N+1)x}}{1-e^{ix}} \\ &= \frac{e^{-iNx}-e^{i(N+1)x}}{1-e^{ix}} = \frac{\exp(-i(N+1/2)x)-\exp(i(N+1/2)x)}{\exp(-ix/2)-\exp(ix/2)} \\ &= \frac{-2\sin((N+1/2)x)}{-2\sin(x/2)} = \frac{\sin((N+1/2)x)}{\sin(x/2)}. \end{aligned}$$

Exercise 148. Show that D_N is an even periodic function that oscillates rapidly when N is large, $D_N(0) = 2N+1$ and $D_N(\pi) = (-1)^N$. Also

$$\int_{\mathbb{T}} D_N \, dx = 1.$$

Now

$$D_k = \frac{\sin((k+1/2)x)}{\sin(x/2)} = \frac{\exp(i(k+1/2)x)-\exp(-i(k+1/2)x)}{(2i)\sin(x/2)}.$$

So

$$(2i)\sin(x/2)D_k = \exp(i(k+1/2)x) - \exp(-i(k+1/2)x).$$

So

$$\begin{aligned} (2i\sin(x/2))^2 D_k(x) &= [\exp(ix/2)-\exp(-ix/2)][\exp(i(k+1/2)x)-\exp(-i(k+1/2)x)] \\ &= \exp(i(k+1)x) - \exp(ikx/2) - \exp(-ikx/2) + \exp(-i(k+1)x). \end{aligned}$$

So

$$
\begin{aligned}
(2i\sin(x/2))^2 F_N(x) &= \frac{1}{N+1}(2i\sin(x/2))^2 \sum_{k=0}^{N} D_k(x) \\
&= \frac{1}{N+1}\sum_{k=0}^{N} \exp(i(k+1)x) - \exp(ikx/2) - \exp(-ikx/2) + \exp(-i(k+1)x) \\
&= \frac{1}{N+1}\left(\exp(i(N+1)x) + \exp(-i(N+1)x) - 2\right) \\
&= \frac{1}{N+1}\left(\exp(i(N+1)x/2) - \exp(-i(N+1)x/2)\right) \\
&= \frac{1}{N+1}(2i\sin((N+1)x/2))^2.
\end{aligned}
$$

Thus

$$F_N(x) = \frac{\sin((N+1)x/2)^2}{(N+1)\sin(x/2)^2}.$$

Exercise 149. Show $F_N(x) \geq 0$, $F_N(0) = N+1$, $\int_{\mathbb{T}} F_N(x)\ dx = 1$ and for any $0 < \delta < 2\pi$,

$$\int_{|x|>\delta, x\in\mathbb{T}} F_N(x)\ dx \to 0$$

as $N \to \infty$.

Exercise 150. For any $f \in L^p(\mathbb{T})$ $(1 \leq p < \infty)$, $f * F_N \to f$ as $N \to \infty$ where convergence is in L^p-norm. Similarly, if $f \in C(\mathbb{T})$ then $f * F_N \to f$ uniformly.

Exercise 151. Trigonometric polynomials are dense in $L^p(\mathbb{T})$ $(1 \leq p < \infty)$. Of course, they are not dense in $L^\infty(\mathbb{T})$.

Exercise 152 (Weierstrauss Approximation Theorem)*.* Trigonometric polynomials are dense in $C(\mathbb{T})$.

Exercise 153 (Weierstrauss Approximation Theorem)*.* Polynomials are dense in $C([0,1])$.

Exercise 154 (Weyl's Equidistribution Theorem)*.* Let α be an irrational number and $f \in C(\mathbb{T})$. Then for every $x \in \mathbb{T}$,

$$\lim_{n\to\infty} \frac{1}{n+1}\sum_{k=0}^{n} f(x+k\alpha) = \int_{\mathbb{T}} f(t)\ dt.$$

Moreover the same holds true if f is the characteristic function of an interval.

Exercise 155 (Parseval's Identity)*.* For any $f, g \in L^2(\mathbb{T})$,

$$\langle f, g\rangle = \int_{\mathbb{T}} f\bar{g}\ dx = \langle \hat{f}, \hat{g}\rangle = \sum_{n=-\infty}^{+\infty} \hat{f}(n)\overline{\hat{g}(n)}.$$

Exercise 156. Show that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \pi^2/6$. Hint: consider the function $f(x) = \frac{\pi - x}{2}$ on $[0, 2\pi)$. Compute the Fourier coefficients of this function and use $\frac{1}{2\pi}\int_0^{2\pi} f(x)^2 \, dx = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2$.

Exercise 157. The map $f \mapsto \hat{f}$ gives an isomorphism from $L^2(\mathbb{T})$ to $\ell^2(\mathbb{Z})$.

Proof. Parseval's identity shows that $f \mapsto \hat{f}$ is an isometric embedding of $L^2(\mathbb{T})$ into $\ell^2(\mathbb{Z})$. But how do we know it's surjective? Let $c = (c_k) \in \ell^2(\mathbb{Z})$. Define $f_n \in L^2(\mathbb{T})$ by $f_n = \sum_{|k|\le n} c_k e_k$. Note $\hat{f}_n(k) = c_k$ if $|k| \le n$ and $\hat{f}_n(k) = 0$ otherwise. Note that $\{f_n\}$ is Cauchy with limit $f = \sum_{k\in\mathbb{Z}} c_k e_k$. Since $f \mapsto \hat{f}$ is continuous, we have $\hat{f} = \lim_n \hat{f}_n$. Of course, this could be computed directly now. □

Exercise 158 (Fourier Inversion Theorem). Let $f \in L^1(\mathbb{T})$ and assume that the Fourier series of f converges absolutely: $\sum_{n\in\mathbb{Z}} |\hat{f}(n)| < \infty$. Then there exists a continuous function $g \in C(\mathbb{T})$ such that $f = g$ a.e. Moreover, the Fourier series converges uniformly to g.

Proof. Let $g = \sum_{n\in\mathbb{Z}} \hat{f}(n) e_n$. This converges absolutely, so $g \in C(\mathbb{T})$. Of course, $\hat{g} = \hat{f}$, so $g = f$ a.e. (this is because $\widehat{g - f} = 0$ so $g - f = 0$). We have shown that the Cesaro means of the Fourier series of g converge to g uniformly. This implies the same statement about the Fourier series. □

By contrast, there exist continuous functions $f \in C(\mathbb{T})$ such that the Fourier series of f does not converge to f pointwise on an uncountable set. A deep Theorem of Carleson and Hunt shows that if $f \in L^p(\mathbb{T})$, $1 < p < \infty$ then the Fourier series of f converges almost everywhere. Kolmogorov provided a counterexample with $f \in L^1(\mathbb{T})$.

29.3 Derivatives and absolute convergence

Under what conditions on f can we guarantee that its Fourier transform $\hat{f}$ converges absolutely? We will see that if f is continuously differentiable then this is true. Moreover, we can express $\hat{f'}$ in terms of $\hat{f}$. Actually, f need not be continuously differentiable for some of these results to hold; it need only be absolutely continuous. We will also see that the regularity of f is reflected in the rate of decay of $\hat{f}$.

As usual, we will identify $\mathbb{T}$ with the interval $[0, 2\pi)$. We will say that a function f is *periodic* if it is defined on $[0, 2\pi]$ and $f(0) = f(2\pi)$. Of course, we could extend f to all of $\mathbb{R}$ by the formula $f(x) = f(x + 2\pi k)$ for $k \in \mathbb{Z}$.

Exercise 159. Suppose f is an absolutely continuous function on $[0, 2\pi]$ and $f(0) = f(2\pi)$. Then

$$\widehat{f'}(n) = \hat{f}(n)(in).$$

(In other words, if $f \sim \sum c_n e^{inx}$ then $f' \sim \sum c_n (in) e^{inx}$). Hint: integration by parts.

Proof. Following the hint,

$$\begin{aligned}
\int_{\mathbb{T}} f' e^{-inx}\, dx &= \frac{1}{2\pi}\int_0^{2\pi} f'(x)e^{-inx}\, dx \\
&= \frac{1}{2\pi}[f(x)e^{-inx}]_0^{2\pi} - \frac{1}{2\pi}\int_0^{2\pi}(-in)f(x)e^{-inx}\, dx \\
&= \hat{f}(n)(in).
\end{aligned}$$

□

Recalling that the indefinite integral of any integrable function is absolutely continuous, we obtain the first part of:

Exercise 160. If $F(x) = \int_0^x f(t)\, dt$ where $f \in L^2([0,2\pi))$ satisfies $\int_0^{2\pi} f\, dt = 0$ then

$$\widehat{f}(n) = \hat{F}(n)(in).$$

Moreover, if $f \in L^2([0,2\pi))$ then the Fourier series $\sum_{n\in\mathbb{Z}} \hat{F}(n)e^{inx}$ converges absolutely and uniformly to F. In particular, this is true whenever F is continuously differentiable.

Proof. We have $f = \sum_n \hat{f}(n)e_n$ in the sense of L^2. Also $F \sim \sum_n \frac{\hat{f}(n)}{in}e_n$. So Cauchy's inequality implies

$$\sum_n \left|\frac{\hat{f}(n)}{in}e_n\right| \leq \left(\sum_n |\hat{f}(n)|^2\right)^{1/2}\left(\sum_n |1/n|^2\right)^{1/2} < \infty.$$

So the Fourier series of F converges absolutely and uniformly. □

Exercise 161 (Riemann-Lebesgue). If $f \in L^1(\mathbb{T})$ then $\hat{f}(n) \to 0$ as $n \to \pm\infty$. In fact,

$$|\hat{f}(n)| \leq (1/2)\|f - \tau_{\pi/n}f\|_1$$

where $\tau_{\pi/n}f(x) = f(x+\pi/n)$. Hint: compute $\hat{f}(n)$ and use the change of variables $x \mapsto x + \pi/n$.

Proof.

$$\begin{aligned}
\hat{f}(n) &= \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}\, dx \\
&= \frac{-1}{2\pi}\int_0^{2\pi} f(x+\pi/n)e^{-inx}\, dx.
\end{aligned}$$

So

$$\hat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}\left(f(x) - f(x+\pi/n)\right)e^{-inx}\, dx.$$

□

Definition 45. f is *Hölder* of exponent $\alpha > 0$ if

$$|f(x+h) - f(x)| \leq C|h|^\alpha$$

for some constant $C > 0$ and for all $x, h \in \mathbb{T}$.

Exercise 162. If f is Hölder of exponent $0 < \alpha < 1$ then $\hat{f}(n) = O(|n|^{-\alpha})$. If f is Lipschitz then $\hat{f}(n) = o(|n|^{-1})$. If $f \in C^2(\mathbb{T})$ then $\hat{f}(n) = O(|n|^{-2})$.

Proof. The first statement follows immediately from the previous exercise. If f is Lipschitz then it is absolutely continuous. So this statement follows from previous exercises. □

29.4 Localization

Theorem 29.1. *Suppose $f \in L^1(\mathbb{T})$, $x_0 \in \mathbb{T}$, $f(x_0^+)$ and $f(x_0^-)$ both exist and there is a constant $C > 0$ such that*

$$|f(y) - f(x_0^+)| \leq C(y - x_0), \quad y > x_0$$
$$|f(x_0^-) - f(y)| \leq C(x_0 - y), \quad y < x_0$$

for all y in some neighborhood of x_0. (E.g. f could be Lipschitz in a neighborhood of x_0). Then

$$\lim_{N\to\infty} \sum_{n=-N}^{N} \hat{f}(n)\exp(inx_0) = \frac{f(x_0^-) + f(x_0^+)}{2}.$$

Proof. Recall that

$$f * D_N(x_0) = \sum_{n=-N}^{N} \hat{f}(n)\exp(inx_0) = D_N * f(x_0)$$

where $D_N(x_0) = \frac{\sin((N+1/2)x_0)}{\sin(x_0/2)}$. So

$$D_N * f(x_0) = \frac{1}{2\pi}\int_0^{2\pi} D_N(t) f(x_0 - t)\, dt = \frac{1}{2\pi}\int_0^{2\pi} f(x_0 - t)\frac{\sin((N+1/2)t)}{\sin(t/2)}\, dt.$$

Let $L = \frac{f(x_0^-)+f(x_0^+)}{2}$. Since $\int_{\mathbb{T}} D_n(t)\, dt = 1$, we have

$$\begin{aligned}
D_N * f(x_0) - L &= \frac{1}{2\pi}\int_0^{2\pi} [f(x_0 - t) - L]\frac{\sin((N+1/2)t)}{\sin(t/2)}\, dt \\
&= \frac{1}{2\pi}\int_0^{\pi} [f(x_0 - t) - L]\frac{\sin((N+1/2)t)}{\sin(t/2)}\, dt + \frac{1}{2\pi}\int_{-\pi}^{0} [f(x_0 - t) - L]\frac{\sin((N+1/2)t)}{\sin(t/2)}\, dt \\
&= \frac{1}{2\pi}\int_0^{\pi} [f(x_0 - t) - L]\frac{\sin((N+1/2)t)}{\sin(t/2)}\, dt + \frac{1}{2\pi}\int_0^{\pi} [f(x_0 + t) - L]\frac{\sin((N+1/2)t)}{\sin(x_0/2)t}\, dt \\
&= \frac{1}{2\pi}\int_0^{\pi} [f(x_0 - t) + f(x_0 + t) - 2L]\frac{\sin((N+1/2)t)}{\sin(t/2)}\, dt \\
&= \frac{1}{2\pi}\int_0^{\pi} \phi(t)\sin((N+1/2)t)\, dt
\end{aligned}$$

where

$$\phi(t) = \frac{[f(x_0 - t) + f(x_0 + t) - 2L]}{\sin(t/2)} = \frac{[f(x_0 - t) - f(x_0^-)] + [f(x_0 + t) - f(x_0^+)]}{\sin(t/2)}.$$

Note $\phi(t)$ is integrable for $t \in [\epsilon, \pi]$ and bounded for $t \in [0, \epsilon]$ (for some $\epsilon > 0$). Therefore $\phi \in L^1([0, \pi])$. So the Riemann-Lebesgue Lemma implies

$$\lim_{N\to\infty} \frac{1}{2\pi} \int_0^\pi \phi(t) \sin((N + 1/2)t)\ dt = 0.$$

This part might not seem so clear because we proved the Riemann-Lebesgue Lemma in a slightly different form. Let's go through the steps again (it's good practice anyway). We consider ϕ to be a π-periodic function (by extending its domain of definition). We make the change of variables $s = t + (\pi/(N + 1/2))$. Then

$$\int_0^\pi \phi(t) \sin((N + 1/2)t)\ dt = -\int_{(\pi/(N+1/2))}^{\pi+(\pi/(N+1/2))} \phi(s - (\pi/(N + 1/2))) \sin((N + 1/2)s)\ ds.$$

So

$$\begin{aligned}&\int_0^\pi \phi(t) \sin((N + 1/2)t)\ dt \\ &= \frac{1}{2}\left[\int_0^\pi \phi(t) \sin((N + 1/2)t)\ dt - \int_{(\pi/(N+1/2))}^{\pi+(\pi/(N+1/2))} \phi(s - (\pi/(N + 1/2))) \sin((N + 1/2)s)\ ds\right].\end{aligned}$$

We bound the above in absolute value by

$$\leq \frac{1}{2}\int_{(\pi/(N+1/2))}^{\pi} |\phi(t) - \phi(t - (\pi/(N+1/2)))|\ dt + \int_0^{(\pi/(N+1/2))} |\phi(t)|\ dt + \int_{\pi-(\pi/(N+1/2))}^{\pi} |\phi(t)|\ dt.$$

Because $\phi \in L^1$, the later tends to zero as $N \to \infty$. $\square$

For example, we let f denote the sign function on $[-\pi, \pi)$ which we identify with the circle. Then

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{sign(x)e^{-ixn}\ dx} = \frac{1 - (-1)^{\mathrm{n}}}{\mathrm{in}\pi}.$$

So $\hat{f}(n) = 0$ if n is even and $\hat{f}(n) = \frac{-2i}{n\pi}$ otherwise. Taking advantage of the fact that sin is odd and cos is even and $e^{ix} = \cos(x) + i\sin(x)$ we have

$$\sum_{n=-N}^{N} \hat{f}(n)e^{inx} = \frac{4}{\pi}\sum_{n=1,odd}^{N} \frac{\sin(nx)}{n}.$$

So if $0 < x < \pi$, we obtain

$$\lim_{N\to\infty} \frac{4}{\pi}\sum_{n=1,odd}^{N} \frac{\sin(nx)}{n} = 1.$$

With $x = \pi/2$, we obtain

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)} = \pi/4.$$

Theorem 29.2. *If $f \in L^1(\mathbb{T})$ and there is an open interval $I \subset \mathbb{T}$ such that $f = 0$ in I then $s_N(f)$ converges uniformly to f on every compact subset of I where $s_N(f) = f * D_N = \sum_{n=-N}^{N} \hat{f}(n) e_n$.*

Proof. The proof of this theorem is similar to the proof of the previous result. To be precise, we assume $x_0 \in I$ so that $L = 0$. Then

$$\begin{aligned} D_N * f(x_0) &= \frac{1}{2\pi} \int_0^{2\pi} f(x_0 - t) \frac{\sin((N+1/2)t)}{\sin(t/2)}\, dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} \phi_{x_0}(t) \sin((N+1/2)t)\, dt \end{aligned}$$

where

$$\phi_{x_0}(t) = \frac{f(x_0 - t)}{\sin(t/2)}.$$

Let $\epsilon > 0$ be small enough so that $(-\epsilon + x_0, \epsilon + x_0) \in I$. For $|t| < \epsilon$ we have $\phi_{x_0}(t) = 0$. So we can write $\phi_{x_0}(t) = f(x_0 - t)\chi(t)$ where χ is a continuous function satisfying $\chi(t) = \sin(t/2)$ whenever $|t| \geq \epsilon$. By the change of variables $s = t + (\pi/(N+1/2))$, it follows that

$$2\pi D_N * f(x_0) = \int_0^{2\pi} f(x_0 - t)\chi(t) \sin((N+1/2)t)\, dt$$

$$= -\int_0^{2\pi} f(x_0 - s + 2\pi/(N+1/2))\chi(s - 2\pi/(N+1/2)) \sin((N+1/2)s)\, ds.$$

So

$$\begin{aligned} |4\pi D_N f * (x_0)| &\leq \int_0^{2\pi} |f(x_0 - t)\chi(t) - f(x_0 - t + 2\pi/(N+1/2))\chi(t - 2\pi/(N+1/2))| |\sin((N+1/2)t)|\, dt \\ &\leq \int_0^{2\pi} |f(x_0 - t)\chi(t) - f(x_0 - t + 2\pi/(N+1/2))\chi(t - 2\pi/(N+1/2))|\, dt. \\ &\leq \int_0^{2\pi} |f(x_0 - t) - f(x_0 - t + 2\pi/(N+1/2))| |\chi(t)|\, dt \\ &\quad + \int_0^{2\pi} |f(x_0 - t + 2\pi/(N+1/2))| |\chi(t) - \chi(t - 2\pi/(N+1/2))|\, dt. \end{aligned}$$

The last integral tends to zero as $N \to \infty$ uniformly in x_0 because χ is continuous. We let $M = \|\chi\|$. We bound the first integral by

$$M \int_0^{2\pi} |f(x_0 - t) - f(x_0 - t + 2\pi/(N+1/2))|\, dt = M \int_0^{2\pi} |f(t) - f(t + 2\pi/(N+1/2))|\, dt.$$

Since this is independent of x_0 and tends to zero as $N \to \infty$ we are done.

□