Lecture notes on measure theory

Lawrence Reeves

lreeeves@unimelb.edu.au

2021 semester 1

Preface

These notes draw on material from:

Measure Theory, D. Cohn, Birkhauser, 1980.

Real Analysis, G. Folland, Wiley, 1999.

Measure Theory, P. Halmos, Springer-Verlag, 1974.

Measure Theory, G. Hjorth, unpublished notes, 2010. (available on the LMS)

Lectures

Lecture 1: Examples and definition

Motivated by applications to probability, volume and integration, we want to assign a value from $[0, \infty]$ to subsets of X. Some basic properties we'd like to insist upon are:

1. $\mu(\emptyset) = 0$
2. If $A_1, A_2, \ldots$ are disjoint, then $\mu(\bigcup_i A_i) = \sum_i \mu(A_i)$ (**countable additivity**)

Before we give the full definition of a measure let's note some examples.

Examples 1. (a) $X = \{0,1\}^n$, $\mu : P(X) \to [0, \infty)$, $\mu(A) = |A|/2^n$. (We'll later extend to $X = \{0,1\}^{\mathbb{N}}$.)

(b) (**counting measure**) $\mu(A) = \begin{cases} |A| & A \text{ finite} \\ \infty & \text{otherwise} \end{cases}$

(c) (**Vitali set**) Suppose that $X = \mathbb{R}$ and that (in addition to the above) μ also satisfies:

$$\mu(\{a\}) = 0 \qquad \mu((a,b)) = b - a \qquad \mu(A + x) = \mu(A) \quad (\textbf{translation invariance})$$

We show that it is impossible to have such a function that is defined on the whole of $P(X)$, by showing that there exists $V \subset [0,1]$ whose measure can be neither zero nor non-zero.

Define an equivalence relation on $[0,1]$ by $x \sim y$ iff $(x - y) \in \mathbb{Q}$. Let V be such that it contains exactly one element from each equivalence class (this needs the axiom of choice). If μ were defined on all subsets of $\mathbb{R}$, then what should $\mu(V)$ be equal to? Note that

$$\begin{aligned}
[0,1] \subset \bigcup_{q \in \mathbb{Q} \cap [-1,1]} (V + q) \subset [-1, 2] &\implies \mu([0,1]) \leqslant \mu\Big(\bigcup_{q \in \mathbb{Q} \cap [-1,1]} (V + q)\Big) \leqslant \mu([-1,2]) \\
&\implies 1 \leqslant \sum_{q \in \mathbb{Q} \cap [-1,1]} \mu(V + q) \leqslant 3 \\
&\implies 1 \leqslant \sum_{q \in \mathbb{Q} \cap [-1,1]} \mu(V) \leqslant 3
\end{aligned}$$

We get a contradiction to the first inequality if $\mu(V) = 0$ and a contradiction to the second if $\mu(V) > 0$.

(d) Another notable example showing that not all subsets of $\mathbb{R}^3$ are (Lebesgue) measurable is provided by the Banach-Tarski Theorem: Let S denote the unit sphere in $\mathbb{R}^3$. There exists a partition $S = A_1 \sqcup A_2 \cdots \sqcup A_n$ and elements $g_i \in SO(3)$ (i.e., rotations) and $k < n$ such that:

$$S = g_1 A_1 \sqcup \cdots \sqcup g_k A_k \quad \text{and} \quad S = g_{k+1} A_{k+1} \sqcup \cdots \sqcup g_n A_n$$

(e) (**Cantor Ternary Set**) Define sets $C_i \subset [0,1]$ recursively as follows:

$$\begin{aligned}
C_1 &= [0, \tfrac{1}{3}] \cup [\tfrac{1}{3}, 1] \\
C_2 &= [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{1}{3}] \cup [\tfrac{2}{3}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1] \\
C_{i+1} &= \tfrac{1}{3} C_i \cup (\tfrac{1}{3} C_i + \tfrac{2}{3})
\end{aligned}$$

Then define $C = \bigcap_i C_i$.

The Lebesgue measure of C_i satisfies $m(C_{i+1}) = \frac{2}{3} m(C_i)$ and it follows that $m(C) = 0$. (We will shortly be defining Lebesgue measure, but for the moment this hopefully seems reasonable.) Any subset $B \subset C$ must also have measure zero. Since C has the same cardinality as $\mathbb{R}$, there are many such subsets. They are all *Lebesgue measurable*.

We've seen that we can't expect to have a measure that is defined on all subsets. What classes of sets is it reasonable to define a measure on?

Definition 2. Let X be a non-empty set. An **algebra** on X is a non-empty subset $\mathcal{A} \subset P(X)$ that is closed under taking complements and finite unions:

$$A \in \mathcal{A} \implies X \setminus A \in A, \qquad A, B \in \mathcal{A} \implies A \cup B \in A$$

A σ-**algebra** is an algebra that is closed under countable unions:

$$\{A_i\}_{i \in \mathbb{N}} \subset \mathcal{A} \implies \bigcup_{i \in \mathbb{N}} A_i \in \mathcal{A}$$

Examples 3. 1) $P(X)$ is a σ-algebra. $\{\emptyset, X\}$ is a σ-algebra.

2) $\mathcal{A} = \{A \subset X \mid A \text{ is countable or } X \setminus A \text{ is countable}\}$ is a σ-algebra.

3) $\mathcal{A}_0 = \{A \subset \mathbb{R} \mid A = A_1 \sqcup A_2 \sqcup \cdots \sqcup A_n, n \in \mathbb{N}, A_i = (a_i, b_i] \text{ or } A_i = (a_i, b_i]^c\}$ is an algebra (on $\mathbb{R}$). $\mathcal{A}_0$ is the algebra generated by $\{(a, b] \mid a < b\}$. It is not a σ-algebra since $\cup_i(-1, -1/i] = (-1, 0) \notin \mathcal{A}_0$.

4) Let X be a topological space. The **Borel** σ-**algebra** is the σ-algebra generated by the open sets. This includes all open sets, closed sets, F_σ-sets (e.g., $\mathbb{Q} \subset \mathbb{R}$), G_δ-sets, etc.

Definition 4. A **measure space** is a triple $(X, \mathcal{A}, \mu)$ of a (non-empty) set X, a σ-algebra on X and a function $\mu : \mathcal{A} \to [0, \infty]$ satisfying

1) $\mu(\emptyset) = 0$

2) If $\{A_i\}_{i \in \mathbb{N}} \subset \mathcal{A}$ is a disjoint family, then $\mu(\bigcup_i A_i) = \sum_i \mu(A_i)$

The measure μ is called **finite** if $\mu(X) < \infty$. It is called σ-**finite** if there exist $A_i \in A$ such that $\mu(A_i) < \infty$ and $X = \sum_{i \in \mathbb{N}} A_i$. A set $A \in \mathcal{A}$ is called **null** if $\mu(A) = 0$. The measure μ is called **complete** if every subset of a null set is in $\mathcal{A}$.

Example 5. X an uncountable set, $\mathcal{A}$ the countable or co-countable subsets, $\mu(A) = 0$ if A countable and $\mu(A) = 1$ otherwise.

Exercises

Exercise 1. A family of sets $\mathcal{R} \subset \mathcal{P}(X)$ is called a **ring** if it is closed under finite unions and differences (i.e., if $A, B \in \mathcal{R}$, then $A \cup B \in \mathcal{R}$ and $A \setminus B \in \mathcal{R}$). A ring that is closed under countable unions is called a σ-**ring**.

Show that:

a) Rings (resp. σ-rings) are closed under finite (resp. countable) intersections.

b) A ring (resp. σ-ring) $\mathcal{R}$ is an algebra (resp. σ-algebra) iff $X \in \mathcal{R}$.

c) If $\mathcal{R}$ is a σ-ring, then $\{A \subset X \mid A \in \mathcal{R} \text{ or } A^c \in \mathcal{R}\}$ is a σ-algebra.

d) If $\mathcal{R}$ is a σ-ring, then $\{A \subset X \mid A \cap B \in \mathcal{R} \text{ for all } B \in \mathcal{R}\}$ is a σ-algebra.

Exercise 2. Let $\mathcal{A}$ be an infinite σ-algebra. Show that:

a) $\mathcal{A}$ contains an infinite sequence of disjoint sets.

b) $|\mathcal{A}| \geqslant 2^{\aleph_0}$

Exercise 3. Given $K \subset \mathcal{P}(X)$, the $\boldsymbol{\sigma}$-**algebra generated** by K is defined to be the intersection of all σ-algebras on X that contain K. Show that the σ-algebra generated by K, is the union of the σ-algebras generated by L as L ranges over all countable subsets of K.

Exercise 4. Let μ and ν be measures on $(X, \mathcal{A})$ and $a, b \in [0, \infty)$. Show that $a\mu + b\nu$ is a measure on $(X, \mathcal{A})$.

Exercise 5. If $(X, \mathcal{A}, \mu)$ is a measure space and $A, B \in \mathcal{A}$. Show that $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(A \cap B)$.

Exercise 6. Let $(X, \mathcal{A}, \mu)$ be a measure space. Show that:

1. If $A, B \in \mathcal{A}$ and $\mu(A \Delta B) = 0$, then $\mu(A) = \mu(B)$. ($A \Delta B$ denotes the symmetric difference of A and B)
2. Show that $A \sim B$ iff $\mu(A \Delta B) = 0$ defines an equivalence relation on $\mathcal{A}$
3. For $A, B \in \mathcal{A}$ define $d(A, B) = \mu(A \Delta B)$. Show that d is a metric on $\mathcal{A}/\sim$

Lecture 2: Premeasures and outer measures

Lemma 6. *Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $A, B \in \mathcal{A}$ and $\{A_i\}_{i \in \mathbb{N}} \subset \mathcal{A}$.*

1) $A \subset B \implies \mu(A) \leqslant \mu(B)$ **(monotonicity)**
2) $\mu(\bigcup_i A_i) \leqslant \sum_i \mu(A_i)$ **(subadditivity)**
3) *If $A_i \subset A_{i+1}$ for all i, then* $\mu(\bigcup A_i) = \lim_i \mu(A_i)$ **(continuity from below)**
4) *If $A_i \supset A_{i+1}$ for all i and $\mu(A_i) < \infty$ for some i, then* $\mu(\bigcap A_i) = \lim_i \mu(A_i)$ **(continuity from above)**

Proof. The first two parts are left as an exercise. For the third (setting $A_0 = \emptyset$),

$$\begin{aligned}\mu(\bigcup_i A_i) = \mu(\bigcup_i (A_i \setminus A_{i-1})) &= \sum_{i \in \mathbb{N}} \mu(A_i \setminus A_{i-1}) && \text{(countable additivity)} \\ &= \lim_{n \to \infty} \sum_{i=1}^{n} \mu(A_i \setminus A_{i-1}) = \lim_{n \to \infty} \mu(A_n)\end{aligned}$$

For the fourth part, we can assume that $\mu(A_1) < \infty$. Define $B_i = A_1 \setminus A_i$. Then $B_i \subset B_{i+1}$ and

$$\begin{aligned}\mu(A_1) - \mu(\bigcap_i A_i) = \mu(A_1 \setminus \bigcap_i A_i) = \mu(\bigcup_i B_i) &= \lim_i \mu(B_i) && \text{(by 3)} \\ &= \lim_i \mu(A_1 \setminus A_i) = \lim_i (\mu(A_1) - \mu(A_i)) = \mu(A_1) - \lim_i (\mu(A_i))\end{aligned}$$

□

Lemma 7 (Completion Lemma)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $\bar{\mathcal{A}} = \{A \cup B \mid A \in \mathcal{A}, B \subset N$ for some null $N \in \mathcal{A}\}$. Then $\bar{\mathcal{A}}$ is a σ-algebra and there is a unique extension of μ to a complete measure on $\bar{\mathcal{A}}$.*

Proof. Exercise. □

We want to mimic the way in which areas in $\mathbb{R}^2$ can be estimated/defined using grids to construct measures on an arbitrary set. More precisely, given a premeasure we contruct an outer measure and then a measure. After giving a general construction, we will use it to define Lebesgue measure on $\mathbb{R}$.

Definition 8. Let $\mathcal{A}_0$ be an algebra on X. A **premeasure** is a function $\mu_0 : \mathcal{A}_0 \to [0, \infty]$ that satisfies

1) $\mu_0(\emptyset) = 0$
2) If $\{A_i\}_{i \in \mathbb{N}}$ is a disjoint collection of elements of $\mathcal{A}_0$ and $\cup_i A_i \in \mathcal{A}_0$, then $\mu_0(\cup_i A_i) = \sum_i \mu_0(A_i)$ **(countably additive on its domain)**

The second condition says, vaguely, that there is no immediate obstruction to extending μ_0 (and $\mathcal{A}_0$) to a measure.

Example 9. Consider the algebra of Example 3.3. The function $\mu_0 : \mathcal{A}_0 \to [0, \infty]$ given by

$$\mu_0(\cup_i (a_i, b_i]) = \sum_i b_i - a_i \qquad \mu_0((-\infty, b]) = \infty \qquad \mu_0((a, \infty)) = \infty \qquad \mu_0(\emptyset) = 0$$

is a premeasure.

Definition 10. An **outer measure** on a (non-empty) set X is a function $\lambda : P(X) \to [0, \infty]$ that satisfies

1) $\lambda(\emptyset) = 0$
2) $A \subset B \implies \lambda(A) \leqslant \lambda(B)$ **(monotonicity)**
3) $\lambda(\cup_i A_i) \leqslant \sum_i \lambda(A_i)$ **(countable subadditivity)**

Exercises

Exercise 7. Let λ be an outer measure on X and $(A_n)_{n\in\mathbb{N}}$ a disjoint sequence of λ-measurable sets. Show that $\lambda(B \cap (\cup_{n\in\mathbb{N}} A_n)) = \sum_{n\in\mathbb{N}} \lambda(B \cap A_n)$ for any $B \subset X$.

Exercise 8. Let μ be a finite measure on $(X, \mathcal{A})$, and let λ be the outer measure on X induced by μ. Suppose that $A \subset X$ satisfies $\lambda(A) = \lambda(X)$. Show that:

1. If $B, C \in \mathcal{A}$ and $A \cap B = A \cap C$, then $\mu(B) = \mu(C)$.
2. Let $\mathcal{A}_A = \{A \cap B \mid B \in \mathcal{A}\}$ and define a function ν on $\mathcal{A}_A$ by $\nu(A \cap B) = \mu(B)$. Show that $\mathcal{A}_A$ is a σ-algebra on A and ν is a measure on $\mathcal{A}_A$.

Lecture 3: Constructing measures

To obtain an outer measure we can start with a class of sets on which some notion of size/measure has been fixed (e.g., intervals in $\mathbb{R}$) and then approximate arbitrary subsets by countable unions. The following lemma makes this precise.

Lemma 11. *Let $\mathcal{K} \subset P(X)$ and $\rho : \mathcal{K} \to [0, \infty]$ be such that $\emptyset \in \mathcal{K}$, $X \in \mathcal{K}$ and $\rho(\emptyset) = 0$. Define $\lambda : P(X) \to [0, \infty]$ by*

$$\lambda(A) = \inf\{\sum_{i \in \mathbb{N}} \rho(K_i) \mid K_i \in \mathcal{K}, A \subset \cup_i K_i\}$$

Then λ is an outer measure on X.

Proof. It's clear that $\lambda(\emptyset) = 0$. Monotonicity is also immediate from the definition of λ. To establish countable subadditivity, let $\{A_i\}_i \subset P(X)$ and let $\epsilon > 0$. For each i there is a sequence $\{A_{ij}\} \subset \mathcal{K}$ such that $A_i \subset \cup_j A_{ij}$ and $\sum_j \rho(A_{ij}) < \lambda(A_i) + 2^{-i}\epsilon$. Then

$$\cup_i A_i \subset \cup_i \cup_j A_{ij} \implies \lambda(\cup_i A_i) \leqslant \sum_{i,j} \rho(A_{i,j}) \leqslant \sum_i (\lambda(A_i) + 2^{-i}\epsilon) = \epsilon + \sum_i \lambda(A_i)$$

Since this holds for any $\epsilon > 0$ we must have $\lambda(\cup_i A_i) \leqslant \sum_i \lambda(A_i)$.

□

Definition 12. Let λ be an outer measure on X. A subset $A \subset X$ is called λ-**measurable** if the following holds for all $B \subset X$:

$$\lambda(B) = \lambda(B \cap A) + \lambda(B \cap A^c)$$

Note that we always have $\lambda(B) \leqslant \lambda(B \cap A) + \lambda(B \cap A^c)$ by subadditivity. If $\lambda(B) = \infty$, then the above equality holds (for any A).

Lemma 13. *Let μ_0 be a premeasure on an algebra $\mathcal{A}_0$. Let $\lambda : P(X) \to [0, \infty]$ be the outer measure defined in Lemma 11 (with $\mathcal{K} = \mathcal{A}_0$). Then $\lambda|_{\mathcal{A}_0} = \mu_0$ and every element of $\mathcal{A}_0$ is λ-measurable.*

Proof. It's immediate from the construction of λ that $\lambda(A) \leqslant \mu_0(A)$ for all $A \in \mathcal{A}_0$. To establish the reverse inequality, suppose that $A \subset \cup_i A_i$ with $A_i \in \mathcal{A}_0$. We want to show that $\mu_0(A) \leqslant \sum_i \mu_0(A_i)$. Let $B_i = A_i \setminus \cup_{j<i} A_j$. Then the B_i are disjoint, $\cup_i B_i = \cup_i A_i$ and

$$\begin{aligned} \mu_0(A) = \mu_0(A \cap \cup_i B_i) = \mu_0(\cup_i (A \cap B_i)) &= \sum_i \mu_0(A \cap B_i) && \text{(since } \mu_0 \text{ is a premeasure)} \\ &\leqslant \sum_i \mu_0(A_i) && \text{(since } A \cap B_i \subset B_i \subset A_i) \end{aligned}$$

To establish the second claim fix $A \in \mathcal{A}_0$ and $B \subset X$. We need to show that $\lambda(B) \geqslant \lambda(B \cap A) + \lambda(B \cap A^c)$. Suppose $B \subset \cup_i B_i$ for some $B_i \in \mathcal{A}_0$.Then

$$\begin{aligned} \lambda(B \cap A) + \lambda(B \cap A^c) &\leqslant \lambda(\cup_i (B_i \cap A)) + \lambda(\cup_i (B_i \cap A^c)) && \text{(monotonicity)} \\ &\leqslant \sum_i \lambda(B_i \cap A) + \sum_i \lambda(B_i \cap A^c) && \text{(subadditivity)} \\ &= \sum_i \mu_0(B_i \cap A) + \sum_i \mu_0(B_i \cap A^c) && \text{(first part of current result)} \\ &= \sum_i \mu_0(B_i \cap A) + \mu_0(B_i \cap A^c) \\ &= \sum_i \mu_0(B_i) && (\mu_0 \text{ is additive on its domain}) \end{aligned}$$

Since this inequality holds for any cover $B \subset \cup_i B_i$, we conclude that $\lambda(B \cap A) + \lambda(B \cap A^c) \leqslant \lambda(B)$. □

Lecture 4: Carathéodory's Extension Theorem

Proposition 14. *Let λ be an outer measure on X and $\mathcal{A} \subset P(X)$ the collection of all λ-measurable sets. Then $\mathcal{A}$ is a σ-algebra and λ restricted to $\mathcal{A}$ is a complete measure.*

Proof. That $\mathcal{A}$ is closed under complementation is clear from the definition of λ-measurable (it's symmetric in A and A^c). To show that $\mathcal{A}$ is closed under finite unions, let $A_1, A_2 \in \mathcal{A}$ and $B \subset X$.

$$\begin{aligned}
\lambda(B) &= \lambda(B \cap A_1) + \lambda(B \cap A_1^c) \\
&= \lambda((B \cap A_1) \cap A_2) + \lambda((B \cap A_1) \cap A_2^c) + \lambda((B \cap A_1^c) \cap A_2) + \lambda((B \cap A_1^c) \cap A_2^c) \\
&\geqslant \lambda(B \cap (A_1 \cup A_2)) + \lambda(B \cap (A_1 \cup A_2)^c) && \text{(subadditivity)}
\end{aligned}$$

since

$$B \cap (A_1 \cup A_2) = (B \cap A_1 \cap A_2) \cup (B \cap A_1 \cap A_2^c) \cup (B \cap A_1^c \cap A_2)$$

Therefore $A_1 \cup A_2 \in \mathcal{A}$ and $\mathcal{A}$ is closed under finite unions ($\mathcal{A}$ is an algebra).
Also, for disjoint $A_1, A_2 \in \mathcal{A}$ we have

$$\lambda(A_1 \cup A_2) = \lambda((A_1 \cup A_2) \cap A_1) + \lambda((A_1 \cup A_2) \cap A_1^c) = \lambda(A_1) + \lambda(A_2)$$

That is, λ is finitely additive on $\mathcal{A}$.
Now to establish that $\mathcal{A}$ is closed under countable disjoint unions. Let $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{A}$ be a disjoint family of sets. Define $A = \cup_i A_i$ and let $B \subset X$ and $n \in \mathbb{N}$. Then

$$\begin{aligned}
\lambda(B \cap A) &\geqslant \lambda(B \cap (\cup_{i\leqslant n} A_i)) && \text{(monotonicity)} \\
&= \lambda(\cup_{i\leqslant n}(B \cap A_i)) \\
&= \sum_{i\leqslant n} \lambda(B \cap A_i) && (A_i \in \mathcal{A} \text{ and are disjoint})
\end{aligned}$$

On the other hand

$$\begin{aligned}
\lambda(B \cap A) &= \lambda(\cup_i (B \cap A_i)) \\
&\leqslant \sum_i \lambda(B \cap A_i) && \text{(subadditivity)}
\end{aligned}$$

Therefore

$$\lambda(B \cap A) = \sum_i \lambda(B \cap A_i) \tag{$*$}$$

Since $\cup_{i\leqslant n} A_i \in \mathcal{A}$ we have

$$\begin{aligned}
\lambda(B) &= \lambda(B \cap (\cup_{i\leqslant n} A_i)) + \lambda(B \cap (\cup_{i\leqslant n} A_i)^c) \\
&\geqslant \lambda(B \cap (\cup_{i\leqslant n} A_i)) + \lambda(B \cap A^c) && \text{(monotonicity)} \\
&= \sum_{i\leqslant n} \lambda(B \cap A_i) + \lambda(B \cap A^c) \\
&\xrightarrow[n\to\infty]{} \sum_{i\in\mathbb{N}} \lambda(B \cap A_i) + \lambda(B \cap A^c) \\
&= \lambda(B \cap A) + \lambda(B \cap A^c) && \text{(by ($*$) above)}
\end{aligned}$$

Therefore $\mathcal{A}$ is closed under countable disjoint unions.
Putting $B = X$ in ($*$), we get that λ is countably additive on $\mathcal{A}$: $\lambda(\cup_i A_i) = \sum_i \lambda(A_i)$.
If the sets $\{A_i\}_i$ are not necessarily disjoint, we still have

$$\cup_i A_i = \cup_i (A_i \setminus \cup_{j<i} A_j) \in \mathcal{A}$$

All that remains is to show that $\lambda|_{\mathcal{A}}$ is complete. Suppose $A \in \mathcal{A}$ is null, that is, $\lambda(A) = 0$ and let $C \subset A$. We need to show that $C \in \mathcal{A}$. For any $B \subset X$ we have

$$\begin{aligned}
\lambda(B) &\leqslant \lambda(B \cap C) + \lambda(B \cap C^c) && \text{(subadditivity of } \lambda\text{)} \\
&= 0 + \lambda(B \cap C^c) && \text{(monotonicity of } \lambda\text{, noting that } B \cap C \subset A\text{)} \\
&\leqslant \lambda(B) && \text{(monotonicity of } \lambda\text{)}
\end{aligned}$$

□

Theorem 15 (Carathéodory's Extension Theorem)**.** *Let μ_0 be a premeasure on an algebra $\mathcal{A}_0$. Let $\mathcal{A}$ be the σ-algebra generated by $\mathcal{A}_0$. Then there is a measure μ on $\mathcal{A}$ such that*

1) μ extends μ_0;

2) If ν is any measure on $\mathcal{A}$ that extends μ_0, then $\nu(A) \leqslant \mu(A)$ for all $A \in \mathcal{A}$ with equality if $\mu(A) < \infty$;

3) If μ_0 is σ-finite, then μ is the unique *extension of μ_0 to $\mathcal{A}$.*

Proof. Let λ be the outer measure obtained from μ_0 as in Lemma 11 and let $\mathcal{M}$ be the collection of λ-measureable sets. From Proposition 14 we know that $\mathcal{M}$ is a σ-algebra and that $\lambda|_{\mathcal{M}}$ is a complete measure. From Lemma 13 we have that $\mathcal{A}_0 \subset \mathcal{M}$ and $\lambda|_{\mathcal{A}_0} = \mu_0$. Since $\mathcal{A}_0 \subset \mathcal{M}$ and $\mathcal{M}$ is a σ-algabra, we have $\mathcal{A}_0 \subset \mathcal{A} \subset \mathcal{M}$. Defining $\mu = \lambda|_{\mathcal{A}}$ we have a measure on $\mathcal{A}$ with $\mu|_{\mathcal{A}_0} = \lambda|_{\mathcal{A}_0} = \mu_0$.

To establish the second part, suppose that ν is any measure on $\mathcal{A}$ with $\nu|_{\mathcal{A}_0} = \mu_0$. Let $A \in \mathcal{A}$ and $A_i \in \mathcal{A}_0$ such that $A \subset \cup_i A_i$. Then

$$\nu(A) \leqslant \nu(\cup_i A_i) \leqslant \sum_i \nu(A_i) = \sum_i \mu_0(A_i)$$

and it follows that $\nu(A) \leqslant \lambda(A) = \mu(A)$. Also,

$$\begin{aligned}
\nu(\cup_i A_i) &= \lim_n \nu(\cup_{i=1}^n A_i) && \text{(continuity from below)} \\
&= \lim_n \mu(\cup_{i=1}^n A_i) && \text{(since } \cup_{i=1}^n A_i \in \mathcal{A}_0 \text{)} \\
&= \mu(\cup_i A_i)
\end{aligned}$$

Suppose that $\mu(A) < \infty$. Fix $\epsilon > 0$ and choose the $A_i \in \mathcal{A}_0$ such that $\mu(\cup_i A_i) < \mu(A) + \epsilon$. Then

$$\mu(A) \leqslant \mu(\cup_i A_i) = \nu(\cup_i A_i) = \nu(A) + \nu((\cup_i A_i) \setminus A) \leqslant \nu(A) + \mu((\cup_i A_i) \setminus A) \leqslant \nu(A) + \epsilon$$

Therefore, $\mu(A) < \infty$ implies that $\mu(A) \leqslant \nu(A)$.

Suppose, for the third claim, that μ_0 is σ-finite. That is, that there exist disjoint $A_i \in \mathcal{A}_0$ with $X = \cup_i A_i$ and $\mu_0(A_i) < \infty$. Then, for any $A \in \mathcal{A}$ we have

$$\nu(A) = \nu(A \cap \cup_i A_i) = \nu(\cup_i (A \cap A_i)) = \sum_i \nu(A \cap A_i) = \sum_i \mu(A \cap A_i) = \mu(A)$$

since $\mu(A \cap A_i) \leqslant \mu(A_i) = \mu_0(A_i) < \infty$. □

Lecture 5: Borel measures on $\mathbb{R}$

Before looking at measures on $\mathbb{R}$, let's note the following example as an application of the Extension Theorem.

Example 16. Let $X = \{0,1\}^{\mathbb{N}}$ and consider the elements of X as infinite words on the alphabet $\{0,1\}$. For each $w \in \{0,1\}^{<\mathbb{N}}$ let $A_w = \{x \in X \mid w \text{ is a prefix of } u\}$. Define $\mu_0(A_w) = 2^{-\ell(w)}$, where $\ell(w)$ is the length of the word w. For example, $\mu_0(X) = 1$ and $\mu_0(A_0) = \mu_0(A_1) = 1/2$, Define $\mathcal{A}_0$ to be the set of all finite unions of sets of the form A_w. Then $\mathcal{A}_0$ is an algebra and μ_0 extends to a σ-finite premeasure on $\mathcal{A}_0$. By the above theorem, this extends (uniquely) to a measure on the σ-algebra generated by $\mathcal{A}_0$. For example, $\mu(\{x\}) = 0$ and $\mu(\cup_i A_{10^i1}) = 1/4$.

We now apply the Carathéodory Extension Theorem to obtain Lebesgue measure on $\mathbb{R}$. Let $\mathcal{B}_{\mathbb{R}} \subset P(\mathbb{R})$ denote the Borel σ-algebra, that is, $\mathcal{B}_{\mathbb{R}}$ is generated by the open subset of $\mathbb{R}$. We want to consider the possible measure spaces $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \mu)$. A slight generalisation of the usual construction of Lebesgue measure will give all such measures (having the property that bounded intervals have finite measure).

Let $\mathcal{A}_0 \subset P(\mathbb{R})$ be the algebra generated by the collection of all 'fingernail' intervals: $S = \{(a,b] \mid a, b \in \mathbb{R}, a < b\}$.

Exercise 9. Every element of $\mathcal{A}_0$ can be written as a finite *disjoint* union of the form $A_1 \sqcup \cdots \sqcup A_n$, where $A_i \in S$ for $i < n$ and either $A_n \in S$ or $A_n^c \in S$.

Exercise 10. The σ-algebra generated by $\mathcal{A}_0$ is exactly $\mathcal{B}_{\mathbb{R}}$.

Lemma 17. *Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing, right-continuous function. Define $\mu_0 : \mathcal{A}_0 \to [0, \infty]$ by*

$$\mu_0(\cup_{i=1}^n (a_i, b_i]) = \sum_i F(b_i) - F(a_i)$$

where the intervals $(a_i, b_i]$ are disjoint. Then μ_0 is a premeasure on $\mathcal{A}_0$.

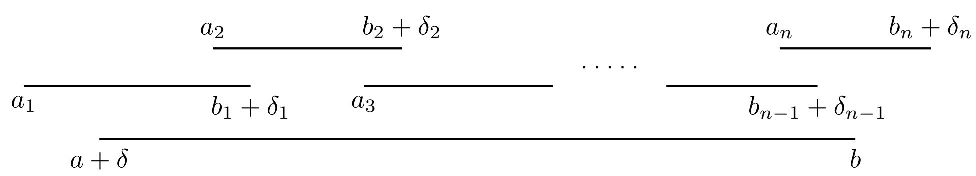
Proof. It's an exercise to check that μ_0 is well-defined and finitely additive. It remains to show that if $\{A_i\}_i \subset \mathcal{A}_0$ is a disjoint family and $\cup_i A_i \in \mathcal{A}_0$, then $\mu_0(\cup_i A_i) = \sum_i \mu(A_i)$. There is no lose in generality in assuming that $A_i = (a_i, b_i]$ and $A \in S$ (see exercise before this result). Suppose that $A = (a,b]$ for some $a, b \in \mathbb{R}$. We need to show that $\sum_i \mu_0((a_i, b_i]) = \mu_0((a,b])$. We have

$$\begin{aligned}
\mu_0((a,b]) &= \mu_0(\cup_i (a_i, b_i])) \\
&= \mu_0(\cup_{i \leqslant n}(a_i, b_i]) + \mu_0((a,b] \setminus \cup_{i \leqslant n}(a_i, b_i]) \\
&\geqslant \sum_{i \leqslant n} \mu_0((a_i, b_i])
\end{aligned}$$

Since this holds for all n, we conclude that

$$\mu_0((a,b]) \geqslant \sum_i \mu_0((a_i, b_i])$$

For the reverse inequality we will use a compactness argument. Fix $\epsilon > 0$. Since F is right continuous, for all i there is a $\delta_i > 0$ such that $F(b_i + \delta_i) - F(b_i) < \epsilon 2^{-i}$ and a $\delta > 0$ such that $F(a + \delta) - F(a) < \epsilon$. Noting that $[a + \delta, b]$ is compact and contained in $\cup_i (a_i, b_i + \delta_i)$, there is a finite subcover. Relabelling if necessary, we can assume that $a_{i+1} < b_i + \delta_i < a_{i+2}$.



Then

$$\begin{aligned}
\sum_i \mu_0((a_i,b_i]) &\geqslant \sum_{i\leqslant n} \mu_0((a_i,b_i]) = \sum_{i\leqslant n} F(b_i) - F(a_i) \\
&\geqslant \sum_{i\leqslant n} F(b_i+\delta_i) - \epsilon 2^{-i} - F(a_i) \\
&\geqslant (\sum_{i\leqslant n} F(b_i+\delta_i) - F(a_i)) - \epsilon \\
&= F(b_1+\delta_1) - F(a_1) + (\sum_{2\leqslant i\leqslant n-1} F(b_i+\delta_i) - F(a_i)) + F(b_n+\delta_n) - F(a_n) - \epsilon \\
&\geqslant F(b_1+\delta_1) - F(a+\delta) + (\sum_{2\leqslant i\leqslant n-1} F(b_i+\delta_i) - F(a_i)) + F(b) - F(a_n) - \epsilon \\
&\geqslant F(b_1+\delta_1) - F(a) - \epsilon + (\sum_{2\leqslant i\leqslant n-1} F(b_i+\delta_i) - F(a_i)) + F(b) - F(a_n) - \epsilon \\
&= F(b) - F(a) - 2\epsilon + F(b_1+\delta_1) + (\sum_{2\leqslant i\leqslant n-1} -F(a_i) + F(b_i+\delta_i)) - F(a_n) \\
&= F(b) - F(a) - 2\epsilon + (\sum_{1\leqslant i\leqslant n-1} F(b_i+\delta_i) - F(a_{i+1})) \\
&\geqslant F(b) - F(a) - 2\epsilon
\end{aligned}$$

Since this holds for all ϵ, we have $\sum_i \mu_0((a_i,b_i]) \geqslant F(b) - F(a)$.

Exercise 11. Finish the proof by considering the case in which $A = (a,b]^c$.

□

We now show that every Borel measure on $\mathbb{R}$ (such that intervals have finite measure) can be constructed using an appropriate function F.

Theorem 18. *Let F be as above.*

1) *There exists a unique Borel measure $\mu_F : \mathcal{B}_{\mathbb{R}} \to [0,\infty]$ satisfying $\mu_F(a,b] = F(b) - F(a)$.*
2) *For two such functions F and G, $\mu_F = \mu_G$ iff $F - G$ is a constant.*
3) *Suppose that $\mu : \mathcal{B}_{\mathbb{R}} \to [0,\infty]$ is a measure satisfying $\mu(a,b] < \infty$ for all $a < b \in \mathbb{R}$. Then $\mu = \mu_F$ for some (increasing, right continuous) function $F : \mathbb{R} \to \mathbb{R}$.*

Proof. The preceding result gives a premeasure μ_0 on $\mathcal{A}_0$, which then, by the Extension Theorem, extends uniquely to a measure $\mu_F : \mathcal{B}_{\mathbb{R}} \to [0,\infty]$. The second part is left as an exercise.
For the third part, define $F : \mathbb{R} \to \mathbb{R}$ by

$$F(x) = \begin{cases} \mu(0,x] & x \geqslant 0 \\ -\mu(x,0] & x < 0 \end{cases}$$

Then F is increasing since μ is monotone. That F is right continuous follows from the fact that μ is continuous from below (for the case $x < 0$) and continuous from above (for $x \geqslant 0$). □

Given such a measure μ_F we can consider its extension to a complete measure $\mu : \mathcal{M}_\mu \to [0,\infty]$. Such measure are called **Lebesgue-Stieltjes measures**.

Lecture 6: Properties of Lebesgue measure. Measurable functions

Exercise 12. Show that for any $A \in \mathcal{M}_\mu$ we have $\mu(A) = \inf\{\sum_i \mu(a_i, b_i) \mid A \subset \cup_i (a_i, b_i)\}$.

Proposition 19. *For all $A \in \mathcal{M}_\mu$ the following hold:*

1) $\mu(A) = \inf\{\mu(V) \mid V \supset A, V \text{ open}\}$ **(outer regularity)**
2) $\mu(A) = \sup\{\mu(K) \mid K \subset A, K \text{ compact}\}$ **(inner regularity)**

Proof. The first follows from the exercise above. For the second part, suppose first that A is bounded. Let $\epsilon > 0$. There is an open V such that $V \supset \bar{A} \setminus A$ and $\mu(V) \leqslant \mu(\bar{A} \setminus A) + \epsilon$. Let $K = \bar{A} \setminus V$. Then K is compact (being closed and bounded), $K \subset A$ and

$$\begin{aligned}
\mu(K) &= \mu(A) - \mu(A \cap V) && (\text{since } A = (A \cap K) \cup (A \cap K^c) = K \cup (A \cap V)) \\
&= \mu(A) - (\mu(V) - \mu(V \setminus A)) \\
&\geqslant \mu(A) - \mu(V) + \mu(\bar{A} \setminus A) \\
&\geqslant \mu(A) - \epsilon
\end{aligned}$$

It remains to consider the case in which A is unbounded. For $i \in \mathbb{N}$ let $A_i = A \cap [-i, i]$ and let $K_i \subset A_i$ be compact with $\mu(K_i) \geqslant \mu(A_i) - 2^{-i}$. Then

$$\lim_i \mu(K_i) \leqslant \mu(A) = \lim_i \mu(A_i) \leqslant \lim_i (\mu(K_i) + 2^{-i}) = \lim_i \mu(K_i)$$

□

Exercise 13. Let $\mu : \mathcal{M}_\mu \to [0, \infty]$ be a Lebesgue-Stieljes measure. Show that the following are equivalent for $A \subset \mathbb{R}$.

a) $A \in \mathcal{M}_\mu$
b) $A = V \setminus N$ for some G_δ set V and null set N
c) $A = C \cup N$ for some F_σ set C and null set N

In the case in which F is the identity function, the resulting complete measure is called **Lebesgue measure**. We'll denote it by $m : \mathcal{L} \to [0, \infty]$. Note that $\mathcal{B}_{\mathbb{R}} \subsetneq \mathcal{L}$.

Proposition 20. *For all $A \in \mathcal{L}$ and $x \in \mathbb{R}$ the following hold:*

1) $m(A + x) = m(A)$
2) $m(xA) = |x| m(A)$

Exercise 14. Prove the above proposition.

Measurable functions

We now turn to integration. Our first step is to define the class of functions with which we will be able to work.

Definition 21. Let $\mathcal{A}$ be a σ-algebra on a set X and $\mathcal{B}$ a σ-algebra on a set Y. A function $f : X \to Y$ is called $(\mathcal{A}, \mathcal{B})$-**measurable** if $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$.

In the case in which $(Y, \mathcal{B}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ we will say $\mathcal{A}$-**measurable** in place of $(\mathcal{A}, \mathcal{B})$-measurable. We will often be concerned with the case $(X, \mathcal{A}) = (\mathbb{R}, \mathcal{L})$ and $(Y, \mathcal{B}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ in which case we'll usually say **Lebesgue measurable** in place of $(\mathcal{A}, \mathcal{B})$-measurable.

Example 22. Let B be a subset of X. Then $\mathbb{1}_B$ is $\mathcal{A}$-measurable if and only if $B \in \mathcal{A}$.

Exercise 15. Let $f : X \to \mathbb{R}$ be a function.

a) Show that if f is $\mathcal{A}$-measurable and $A \in \mathcal{A}$, then the restriction $f|_A$ is $\mathcal{A}$-measurable.

b) Let $\{A_i\}_i \subset \mathcal{A}$ with $\cup_i A_i = X$. Show that if the restrictions $f|_{A_i}$ are all $\mathcal{A}$-measurable, then f is $\mathcal{A}$-measurable.

Exercise 16. Let X, Y be topological spaces and $\mathcal{B}_X$, $\mathcal{B}_Y$ the respective Borel σ-algebras. Show that any continuous function $f : X \to Y$ is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.

Exercise 17. a) Let $\psi : \mathbb{R}^2 \to \mathbb{R}$ be given by $\psi(x, y) = xy$. Show that ψ is $(\mathcal{B}_{\mathbb{R}^2}, \mathcal{B}_{\mathbb{R}})$-measurable.

b) Let $\xi : \mathbb{R}^2 \to \mathbb{R}$ be given by $\xi(x, y) = x + y$. Show that ξ is $(\mathcal{B}_{\mathbb{R}^2}, \mathcal{B}_{\mathbb{R}})$-measurable.

c) Let $\mathcal{A}$ be a σ-algebra on a set X. Suppose the $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ are $\mathcal{A}$-measurable. Show that $f + g$ and fg are $\mathcal{A}$-measurable. Hint: Define $\varphi : X \to \mathbb{R}^2$ by $\varphi(x) = (f(x), g(x))$ and show that φ is $(\mathcal{A}, \mathcal{B}_{\mathbb{R}^2})$-measurable.

Lemma 23. *Let $\mathcal{A}$ be a σ-algebra on a set X. Let $(f_i)_{i\in\mathbb{N}}$ be a sequence of $\mathcal{A}$-measurable functions $f_i : X \to \bar{\mathbb{R}}$. Then the following functions are $\mathcal{A}$-measurable:*

$$\min\{f_i, f_2\} \qquad \max\{f_1, f_2\} \qquad \sup f_i \qquad \inf f_i \qquad \limsup f_i \qquad \liminf f_i$$

Proof. Let $s : X \to \bar{\mathbb{R}}$ be given by $s(x) = \sup_i f_i(x)$. Then

$$s^{-1}(a, \infty] = \cup_1^\infty f_i^{-1}(a, \infty] \in \mathcal{A} \qquad \text{since} \qquad f^{-1}(a, \infty] \in \mathcal{A}$$

Since $\mathcal{B}_{\bar{\mathbb{R}}}$ is generated by $\{(a, \infty] \mid a \in \mathbb{R}\}$, it follows that s is $\mathcal{A}$-measurable.
Let $l : X \to \bar{\mathbb{R}}$ be given by $l(x) = \limsup_i f_i(x)$. Then

$$l^{-1}(a, \infty] = \cap_{n\in\mathbb{N}} \cup_{i\geqslant n} f_i^{-1}(a, \infty] \in \mathcal{A}$$

and l is $\mathcal{A}$-measurable.

Exercise 18. Finish of the remaining cases.

□

Exercise 19. Show that the supremum of an uncountable family of Borel measurable functions $\{f_i : \mathbb{R} \to \mathbb{R} \mid i \in I\}$ can fail to be Borel measurable.

Lecture 7: Measurable functions and integration

Definition 24. A function $h : X \to \mathbb{R}$ is **simple** if it is a linear combination of characteristic functions of elements of $\mathcal{A}$. That is,

$$h = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}$$

for some $n \in \mathbb{N}$, $a_i \in \mathbb{R}$, $A_i \in \mathcal{A}$.

Exercise 20. A function $h : X \to \mathbb{R}$ is simple iff h is $\mathcal{A}$-measurable and finite-valued (i.e., $|h(X)| < \infty$).

Note that if h is simple, we can assume that the sets A_i are disjoint since $h = \sum_{a \in h(X)} a \mathbb{1}_{h^{-1}(a)}$.

Lemma 25. *Let $f : X \to \bar{\mathbb{R}}$ be $\mathcal{A}$-measurable and suppose $f \geqslant 0$. There exists a sequence of simple functions $(h_i)_{i\in\mathbb{N}}$ such that*

1) $0 \leqslant h_i \leqslant h_{i+1} \leqslant f$
2) $h_i(x) \to f(x)$ *for all* $x \in X$
3) $h_i \to f$ *uniformly on any set on which f is bounded.*

Proof. For $i \in \mathbb{N}$ subdivide the interval $[0, 2^i]$ into disjoint fingernail intervals of length 2^{-i}.
Let the subintervals be $I_{ij} = (a_{ij}, b_{ij}]$ and define $A_{ij} = f^{-1}(I_{ij}) \in \mathcal{A}$ and $A_{i,\infty} = f^{-1}(2^i, \infty) \in \mathcal{A}$. Setting

$$h_i = \sum_{j=1}^{2^{2i}} a_{ij} \mathbb{1}_{A_{ij}} + 2^i \mathbb{1}_{A_{i,\infty}}$$

gives the desired sequence.

□

Let $(X, \mathcal{A}, \mu)$ be a measure space. A property (of points in X) is said to hold **almost everywhere** if there exists $N \in \mathcal{A}$ with $\mu(N) = 0$ and such that the property holds on $X \setminus N$.

Exercise 21. Give an example of two function $f, g : \mathbb{R} \to \mathbb{R}$ that agree on a dense subset of $\mathbb{R}$, but for which $f(x) \neq g(x)$ almost everywhere on X (with respect to Lebesgue measure on $\mathbb{R}$.)

Proposition 26. *Let $(X, \mathcal{A}, \mu)$ be a measure space and let $\bar{\mathcal{A}}$ be the completion of $\mathcal{A}$ (with respect to μ). Then a function $f : X \to \bar{\mathbb{R}}$ is $\bar{\mathcal{A}}$-measurable if and only if there are $\mathcal{A}$-measurable functions $f_0, f_1 : X \to \bar{\mathbb{R}}$ such that*

1) $f_0 \leqslant f \leqslant f_1$ *holds everywhere on X, and*
2) $f_0 = f_1$ *holds μ-almost everywhere on X.*

Proof. Suppose first that such f_0 and f_1 exist and let $N \in \mathcal{A}$ be such that $f_0 = f_1 = f$ on $X \setminus N$. Then for any $B \in \mathcal{B}_\mathbb{R}$

$$f^{-1}(B) = (f_0^{-1}(B) \cap N^c) \cup (f^{-1}(B) \cap N) \in \bar{\mathcal{A}}$$

For the converse, suppose first that f is simple and that $f \geqslant 0$, that is, $f = \sum_{i=1}^{k} a_i \mathbb{1}_{A_i}$ for some $a_i \geqslant 0$ and $A_i \in \bar{\mathcal{A}}$. Since $A_i \in \bar{\mathcal{A}}$, there exist $A_{i,0}, A_{i,1} \in \mathcal{A}$ such that $A_{i,0} \subset A_i \subset A_{i,1}$ and $\mu(D_i \setminus C_i) = 0$. The functions $f_0 = \sum_{i=1}^{k} a_i \mathbb{1}_{A_{i,0}}$ and $f_1 = \sum_{i=1}^{k} a_i \mathbb{1}_{A_{i,1}}$ satisfy the two conditions above.
Suppose now that $f : X \to \mathbb{R}$ is $\bar{\mathcal{A}}$-measurable and that $f \geqslant 0$. By proceeding lemma, there exists a sequence $(h_i)_i$ of positive simple functions such that $f(x) = \lim_i h_i(x)$ for all $x \in X$. We have already seen that for each h_i there exist $\mathcal{A}$-measurable functions $h_{i,0}$ and $h_{i,1}$ such that $h_{i,0} \leqslant h_i \leqslant h_{i,1}$ for all $x \in X$ and $h_{i,0} = h_{i,1}$ μ-almost everywhere. Take $f_0 = \limsup_i h_{i,0}$ and $f_1 = \liminf_i h_{i,1}$.
If f is measurable but not necessarily positive, we can apply the above argument to the two functions $f^+ = \max(f, 0)$ and $f^- = \min(f, 0)$.

□

Integration

We define the integral first for simple positive functions and then extend to measurable positive functions and then to arbitrary measurable functions.

Given a measure space $(X, \mathcal{A}, \mu)$ and a simple function $h : X \to \mathbb{R}$ with $h = \sum_{i=1}^{k} a_i \mathbb{1}_{A_i}$ for some $a_i \geqslant 0$ and disjoint $A_i \in \mathcal{A}$, the **integral** of h with respect to μ is defined to be

$$\int h \, d\mu = \sum_{i=1}^{k} a_i \mu(A_i) \in [0, \infty]$$

For $A \in \mathcal{A}$ define $\int_A h \, d\mu = \int h \mathbb{1}_A \, d\mu$.

Exercise 22. Let $h, g : X \to \mathbb{R}$ be positive simple functions and let $c \in [0, \infty]$. Show that

1) $\int ch \, d\mu = c \int h \, d\mu$
2) $\int h + g \, d\mu = \int h \, d\mu + \int g \, d\mu$
3) $h \leqslant g \implies \int h \, d\mu \leqslant \int g \, d\mu$
4) $A \mapsto \int_A h \, d\mu$ determines a measure on $\mathcal{A}$

Now for a positive measurable function $f : X \to \bar{\mathbb{R}}$ define

$$\int f \, d\mu = \sup \left\{ \int h \, d\mu \mid h \text{ simple, } 0 \leqslant h \leqslant f \right\}$$

Exercise 23. Let f, g be positive measurable functions and $c \geqslant 0$. Show that

1) $f \leqslant g \implies \int f \, d\mu \leqslant \int g \, d\mu$
2) $\int cf \, d\mu = c \int f \, d\mu$

Lemma 27. *Let $f : X \to \bar{\mathbb{R}}$ be positive and measurable. Then*

$$\int f \, d\mu = 0 \iff f = 0 \ \mu\text{-almost everywhere}$$

Proof. Suppose that there exists a μ-null set $N \in \mathcal{A}$ such that $f \mid_{N^c} = 0$. If h is a simple function such that $0 \leqslant h \leqslant f$, then we can write h as $h = \sum_{i=1}^{k} a_i A_i$ where the A_i are disjoint and $\cup_i A_i = N$. Therefore $\int h \, d\mu = 0$. Since this holds for all such h, we have that $\int f \, d\mu = 0$.

For the converse we have $f^{-1}(0, \infty] = \cup_{i \in \mathbb{N}} f^{-1}(1/i, \infty]$ and therefore

$$\begin{aligned} \mu(f^{-1}(0, \infty]) \neq 0 &\implies \mu(f^{-1}(1/i, \infty]) \neq 0 \text{ for some } i \\ &\implies f > \frac{1}{i} \mathbb{1}_A \text{ for some } A \in \mathcal{A} \text{ with } \mu(A) > 0 \\ &\implies \int f \, d\mu \geqslant \frac{1}{i} \int \mathbb{1}_{A_i} \, d\mu = \frac{1}{i} \mu(A_i) > 0 \end{aligned}$$

□

Lecture 8: Montotone convergence theorem

Lemma 28. *Let $f : X \to \bar{\mathbb{R}}$ be positive and measurable. If $(h_i)_i$ is a sequence of simple functions such that $0 \leqslant h_i \leqslant h_{i+1} \leqslant f$ and $h_i(x) \to f(x)$ for all $x \in X$ (such exist by Lemma 25), then $\int f\,d\mu = \lim \int h_i\,d\mu$.*

Proof. Note first that $\int h_i\,d\mu \leqslant \int f\,d\mu$ since $h_i \leqslant f$. Therefore $\int f\,d\mu \geqslant \lim \int h_i\,d\mu$. For the reverse inequality suppose that h is any simple function such that $0 \leqslant h \leqslant f$ and let $\epsilon \in (0,1)$. Define

$$A_i = \{x \in X \mid h_i(x) \geqslant \epsilon h(x)\}$$

Then note that

- $A_i \in \mathcal{A}$
- $A_{i+1} \supset A_i$
- $\cup_i A_i = X$ since $h_i \to f > \epsilon h$
- $\int h_i\,d\mu \geqslant \int_{A_i} h_i\,d\mu \geqslant \int_{A_i} \epsilon h\,d\mu = \epsilon \int_{A_i} h\,d\mu$
- $\int_{A_i} h\,d\mu \to \int h\,d\mu$ since $A \mapsto \int_A h\,d\mu$ is a measure and hence continuous from below

Therefore $\lim \int h_i\,d\mu \geqslant \epsilon \int h\,d\mu$ and since this holds for all ϵ we conclude that $\lim \int h_i\,d\mu \geqslant \int h\,d\mu$. Taking the supremum over all h gives $\lim \int h_i\,d\mu \geqslant \int f\,d\mu$ as desired. □

Exercise 24. Use Lemmas 25 and 28 to show that if $f, g : X \to \bar{\mathbb{R}}$ are $\mathcal{A}$-measurable and positive, then

$$\int f + g\,d\mu = \int f\,d\mu + \int g\,d\mu$$

Monotone convergence theorem

Theorem 29 (Montone Convergence Theorem). *Let $(f_i)_{i\in\mathbb{N}}$ be a sequence of measurable positive functions $f_i : X \to \bar{\mathbb{R}}$ such that $f_i \leqslant f_{i+1}$. Then*

$$\int \lim_i f_i\,d\mu = \lim_i \int f_i\,d\mu$$

Proof. The proof is similar to that of Lemma 28. Let $f = \lim f_i$ (which is equal to $\sup f_i$). Since $f_i \leqslant f$ we have $\lim \int f_i\,d\mu \leqslant \int f\,d\mu$. For the reverse inequality, let $\epsilon \in (0,1)$. Suppose that h is a simple function with $0 \leqslant h \leqslant f$. Let $A_i = \{x \in X \mid f_i(x) \geqslant \epsilon h(x)\}$. Then $A_i \subset A_{i+1}$ and $\cup_i A_i = X$. Also, $A_i \in \mathcal{A}$ since $A_i = (f_i - \epsilon h)^{-1}[0,\infty]$. Then

$$\int f_i\,d\mu \geqslant \int_{A_i} f_i\,d\mu \geqslant \int_{A_i} \epsilon h\,d\mu = \epsilon \int_{A_i} h\,d\mu \to \epsilon \int h\,d\mu$$

So,

$$\lim \int f_i\,d\mu \geqslant \epsilon \int h\,d\mu \qquad \text{for all } \epsilon \in (0,1)$$

therefore,

$$\lim \int f_i\,d\mu \geqslant \int h\,d\mu \qquad \text{for all } h$$

therefore,

$$\lim \int f_i\,d\mu \geqslant \int f\,d\mu$$

□

Corollary 30. *Let $(f_i)_{i\in\mathbb{N}}$ be a sequence of measurable positive functions $f_i : X \to \bar{\mathbb{R}}$. Then*

$$\int \sum_i f_i\,d\mu = \sum_i \int f_i\,d\mu$$

Proof. Let $g_n = \sum_{i=1}^n f_i$. From Exercise 24 we have that $\int g_n \, d\mu = \sum_1^n \int f_i \, d\mu$. Applying the above theorem to the g_n we get

$$\begin{aligned} \int \sum_{i\in\mathbb{N}} f_i \, d\mu &= \int \lim_n g_n \, d\mu \\ &= \lim_n \int g_n \, d\mu \qquad \text{(by the MCT)} \\ &= \lim_n \sum_1^n \int f_i \, d\mu \qquad \text{(finite version)} \\ &= \sum_{i\in\mathbb{N}} \int f_i \, d\mu \end{aligned}$$

□

Exercise 25. Show that in the statement of the MCT it is enough to insist that for each i, $f_i \leqslant f_{i+1}$ μ-almost everywhere.

Example 31. To see that the hypothesis that the sequence $(f_i)_i$ be increasing (almost everywhere) is needed, consider $f_i = \mathbb{1}_{(i,i+1)}$. For this sequence, we have

$$\int \lim_i f_i \, d\mu = \int 0 \, d\mu = 0 \quad \text{whereas} \quad \lim_i \int f_i \, d\mu = \lim_i 1 = 1$$

Lecture 9: Fatou's lemma and the dominated convergence theorem

This result is sometimes useful to show that a function is integrable and to provide an upper bound on the value of the integral.

Lemma 32 (Fatou's Lemma)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space and $(f_i)_i$ a sequence of measurable positive functions on X. Then*

$$\int \liminf f_i \, d\mu \leqslant \liminf \int f_i \, d\mu$$

Proof.

$$\begin{aligned} & \inf_{i \geqslant n} f_i \leqslant f_j \qquad \forall\, j \geqslant n \\ \implies & \int \inf_{i \geqslant n} f_i \, d\mu \leqslant \int f_j \, d\mu \qquad \forall\, j \geqslant n \\ \implies & \int \inf_{i \geqslant n} f_i \, d\mu \leqslant \inf_{j \geqslant n} \int f_j \, d\mu \end{aligned} \tag{$*$}$$

Letting $n \to \infty$ and applying the MCT we get

$$\begin{aligned} \int \liminf f_i \, d\mu &= \lim_{n \to \infty} \int \inf_{i \geqslant n} f_i \, d\mu \qquad \text{(MCT)} \\ &\leqslant \liminf \int f_j \, d\mu \qquad \text{(by } (*)\text{)} \end{aligned}$$

□

Corollary 33. *With f_i as above, suppose that f is a positive measurable function such that $f_i \to f$ μ-almost everywhere. Then*

$$\int f \, d\mu \leqslant \liminf \int f_i \, d\mu$$

□

So far we've defined the integral for positive functions. Extending the definition to cover measurable functions that aren't positive is straightforward. Given a measurable function $f : X \to \bar{\mathbb{R}}$ we have $f = f^+ - f^-$, where $f^+ = \max(0, f)$ and $f^- = \max(0, -f)$ are both positive and measurable. If $\int f^+ \, d\mu < \infty$ or $\int f^- \, d\mu < \infty$, then we say that the integral **exists** and define

$$\int f \, d\mu = \int f^+ \, d\mu \; - \; \int f^- \, d\mu$$

We say that f is **integrable** if both $\int f^+ \, d\mu < \infty$ and $\int f^- \, d\mu < \infty$. In the case in which $(X, \mathcal{A}, \mu)$ is $(\mathbb{R}, \mathcal{L}, m)$ we sometimes say **Lebesgue integrable**.

Exercise 26. Define $\mathscr{L}^1(X, \mathcal{A}, \mu, \mathbb{R})$ to be the set of all integrable functions $f : X \to \mathbb{R}$. Show that $\mathscr{L}^1(X, \mathcal{A}, \mu, \mathbb{R})$ forms a vector space and that the integral is linear functional.

Proposition 34. *Let $f : X \to \bar{\mathbb{R}}$ be a measurable function. If f is integrable then*

$$\left| \int f \, d\mu \right| \leqslant \int |f| \, d\mu$$

Proof. Note that f is measurable iff $|f| = f^+ + f^-$ is measurable.

$$f \text{ integrable} \iff f^+ \text{ and } f^- \text{ are both integrable} \iff |f| = f^+ + f^- \text{ is integrable}$$

If f and $|f|$ are integrable then we have,

$$\left| \int f \, d\mu \right| = \left| \int f^+ \, d\mu - \int f^- \, d\mu \right| \leqslant \int f^+ \, d\mu + \int f^- \, d\mu = \int |f| \, d\mu$$

□

Exercise 27. Find an example of a function $f : \mathbb{R} \to \mathbb{R}$ such that f is not Lebesgue integrable, but $|f|$ is Lebesgue integrable.

Theorem 35 (Dominated Convergence Theorem)**.** *Let g be a positive integrable function and let f and $\{f_i\}_i$ be measurable functions $X \to \bar{\mathbb{R}}$. Suppose that*

1) $f = \lim f_i$ $\quad\mu$-almost everywhere

2) $|f_i| \leqslant g$ $\quad\mu$-almost everywhere for all i

Then f and all f_i are integrable and

$$\int f\, d\mu = \lim_i \int f_i\, d\mu$$

Proof. The integrability of f and the f_i follows from that of g.
For each i, both $g + f_i$ and $g - f_i$ are positive measurable function. Applying (the corollary to) Fatou's lemma to the sequence $(g + f_i)_i$ we get

$$\int g + f\, d\mu \leqslant \liminf \int g + f_i\, d\mu = \int g\, d\mu + \liminf \int f_i\, d\mu$$

Similarly, considering the sequence $(g - f_i)_i$

$$\int g - f\, d\mu \leqslant \liminf \int g - f_i\, d\mu = \int g\, d\mu - \limsup \int f_i\, d\mu$$

Therefore

$$\liminf \int f_i\, d\mu \geqslant \int f\, d\mu \geqslant \limsup \int f_i\, d\mu$$

□

Now some consequences of the DCT. Firstly, we can extend the result of Corollary 30 to functions that are not necessarily positive.

Proposition 36. *Let $(f_i)_i$ be a sequence of integrable functions $f_i \in \mathscr{L}^1(X, \mathcal{A}, \mu, \mathbb{R})$ and suppose that $\sum_{i\in\mathbb{N}} \int |f_i|\, d\mu < \infty$. Then $\sum_i f_i$ converges (almost everywhere) to a function in $\mathscr{L}^1$ and*

$$\int \sum_i f_i\, d\mu = \sum_i \int f_i\, d\mu$$

Proof. By Corollary 30 we have that $\int \sum_i |f_i|\, d\mu = \sum_i \int |f_i|\, d\mu$. Therefore, since $\sum_{i\in\mathbb{N}} \int |f_i|\, d\mu < \infty$ there is a function $g \in \mathscr{L}^1$ such that $g = \sum_i |f_i|$ almost everywhere. Also, $\sum_i f_i(x)$ converges for almost all $x \in X$. Applying the DCT to the sequence of partial sums $\sum_1^n f_i$ (noting that $|\sum_1^n f_i| \leqslant g$ almost everywhere) we conclude that

$$\int \sum_{i\in\mathbb{N}} f_i\, d\mu = \int \lim_n \sum_1^n f_i\, d\mu = \lim \sum_1^n \int f_i\, d\mu = \sum_{i\in\mathbb{N}} \int f_i\, d\mu$$

□

Next we observe that the simple functions are (in an appropriate sense) dense in $\mathscr{L}^1$.

Proposition 37. *Let $f \in \mathscr{L}^1(X, \mathcal{A}, \mu, \mathbb{R})$ and let $\epsilon \in (0, \infty)$. There exists an integrable simple function h such that*

$$\int |f - h|\, d\mu < \epsilon$$

Proof. Fix a sequence of simple functions h_i such that $|h_i| \leqslant |h_{i+1}| \leqslant |f|$ and $h_i(x) \to f(x)$ for all $x \in X$. Apply the DCT to the sequence $|f - h_i| \leqslant 2|f|$ to get $\lim \int |f - h_i|\, d\mu = \int 0\, d\mu = 0$. □

Lecture 10: The spaces $\mathscr{L}^p$ and L^p

We've already encountered $\mathscr{L}^1$. We now define related spaces $\mathscr{L}^p$ and L^p (for $p \in [1, \infty]$) and consider some properties.

Let $(X, \mathcal{A}, \mu)$ be a measure space and let $p \in [1, \infty)$. We define

$$\mathscr{L}^p(X, \mathcal{A}, \mu, \mathbb{R}) = \{ f : X \to \mathbb{R} \mid |f|^p \text{ is integrable } \}$$

Exercise 28. Verify that $\mathscr{L}^p(X, \mathcal{A}, \mu, \mathbb{R})$ forms a vector space. (For closure under addition it's useful to observe that $|f(x) + g(x)|^p \leqslant 2^p|f(x)|^p + 2^p|g(x)|^p$.)

A function $f : X \to \mathbb{R}$ is called **essentially bounded** if there exists M such that the set $\{x \in X \mid |f(x)| > M\}$ is μ-null[1]. We now define

$$\mathscr{L}^\infty(X, \mathcal{A}, \mu, \mathbb{R}) = \{f : X \to \mathbb{R} \mid f \text{ is } \mathcal{A}\text{-measurable and essentially bounded}\}$$

As with $\mathscr{L}^p$ for $p < \infty$, $\mathscr{L}^\infty$ equipped with the usual operations forms a vector space.

We can define a seminorm on $\mathscr{L}^p$ by

$$\|f\|_p = \left(\int |f|^p \, d\mu \right)^{\frac{1}{p}} \qquad \text{for } 1 \leqslant p < \infty$$
$$\|f\|_\infty = \inf\{M \mid \text{ the set } \{x \in X \mid |f(x)| > M\} \text{ is null}\}$$

Exercise 29. Let $f \in \mathscr{L}^\infty$. Show that $\{x \in X \mid |f(x)| > \|f\|_\infty\}$ is μ-null.

Proposition 38 (Hölder's inequality)**.** *Let $f \in \mathscr{L}^p$ and $g \in \mathscr{L}^q$ where $p, q \in [1, \infty]$ satisfy $\frac{1}{p} + \frac{1}{q} = 1$. Then $fg \in \mathscr{L}^1$ and*

$$\int |fg| \, d\mu \leqslant \|f\|_p \|g\|_q$$

Proof. Outline. Suppose first that $f \in \mathscr{L}^1$ and $g \in \mathscr{L}^\infty$. Then $|f(x)g(x)| \leqslant |f(x)|\|g\|_\infty$ almost everywhere. It follows that $fg \in \mathscr{L}^1$ and that $\int |fg| \, d\mu \leqslant \|g\|_\infty \int |f| \, d\mu = \|f\|_p \|g\|_\infty$.

Now suppose that p (hence q) is in $(1, \infty)$.

Exercise 30. Show that for all $x, y \in [0, \infty)$ we have $xy \leqslant x^p/p + y^q/q$.

Then for all x we have $|f(x)g(x)| \leqslant \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q$ and so $fg \in \mathscr{L}^1$ and

$$\int |fg| \, d\mu \leqslant \frac{1}{p} \int |f|^p \, d\mu + \frac{1}{q} \int |g|^q \, d\mu$$

If $\|f\|_p = \|g\|_q = 1$ then the above gives the required inequality. Otherwise replace f by $f/\|f\|_p$ and g by $g/\|g\|_q$. (We can assume that $\|f\|_p$ and $\|g\|_q$ are non-zero, since otherwise the result is clearly true.) □

Proposition 39 (Minkowski's inequality)**.** *Let $p \in [1, \infty]$ and let $f, g \in \mathscr{L}^p$. Then*

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p$$

Proof. Outline. If $p = 1$ we have

$$\|f + g\|_1 = \int |f + g| \, d\mu \leqslant \int |f| \, d\mu + \int |g| \, d\mu = \|f\|_1 + \|g\|_1$$

If $p = \infty$ we have

$$|f(x) + g(x)| \leqslant |f(x)| + |g(x)| \leqslant \|f\|_\infty + \|g\|_\infty$$

outside a null set.

[1] In the case in which $(X, \mathcal{A}, \mu)$ is not σ-finite we should use 'locally μ-null'.

Now suppose that $p \in (1,\infty)$ and let $q \in (1,\infty)$ be such that $1/p+1/q=1$. Note that $|f+g|^{p-1} \in \mathscr{L}^q$ because $f+g \in \mathscr{L}^p$ and $(p-1)q = p$.

$$\begin{aligned}\int |f+g|^p \, d\mu &\leqslant \int (|f|+|g|)\,|f+g|^{p-1}\, d\mu = \int |f|\,|f+g|^{p-1}\, d\mu + \int |g|\,|f+g|^{p-1}\, d\mu \\ &\leqslant \|f\|_p \||f+g|^{p-1}\|_q + \|g\|_p \||f+g|^{p-1}\|_q \qquad \text{(Hölder's inequality, twice)} \\ &= (\|f\|_p + \|g\|_p)(\int |f+g|^p \, d\mu)^{1/q}\end{aligned}$$

Assuming $\int |f+g|^p \, d\mu \neq 0$, we obtain $(\int |f+g|^p \, d\mu)^{1/p} \leqslant \|f\|_p + \|g\|_p$ as required. If $\|f+g\|_p = 0$ the result is clear. ☐

Corollary 40. *The function $f \mapsto \|f\|_p$ is a seminorm on $\mathscr{L}^p$.* ☐

We don't get a norm on $\mathscr{L}^p$ because there are non-zero functions with $\|f\|_p = 0$. Let $\mathscr{N}^p = \{f \in \mathscr{L}^p \mid \|f\|_p = 0\}$ and define L^p to be the quotient $\mathscr{L}^p/\mathscr{N}^p$. The elements of L^p consist of equivalence classes of the relation on $\mathscr{L}^p$ given by $f \sim g$ iff $\|f-g\|_p = 0$. The equivalence class is sometimes denoted $\bar{f}$.

Exercise 31. Show that for $f,g \in \mathscr{L}^p$, $f \sim g \implies \|f\|_p = \|g\|_p$.

It follows that we $\|\cdot\|_p$ induces a function on L^p (also denoted $\|\cdot\|_p$) and, by the above corollary, it is a norm on L^p.

Exercise 32. Show that the following defines an inner product on L^2.

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

Theorem 41. *Let $p \in [1,\infty]$. The normed space L^p (equipped with the norm $\|\cdot\|_p$) is complete.*

Proof. A normed space is complete iff every absolutely convergent series is convergent. Let $f_i \in \mathscr{L}^p$ be such that $\sum_{i\in\mathbb{N}} \|f_i\|_p < \infty$.
Consider first the case in which $p = \infty$. Let N_i be a null set such that $|f_i(x)| \leqslant \|f_i\|_\infty$ on N_i^c and let $N = \cup_i N_i$. The series $\sum_i f_i(x)$ converges for all $x \notin N$. The function $f = \mathbb{1}_{N^c} \sum_i f_i$ is bounded and measurable and

$$\|f - \sum_{i=1}^n f_i\|_\infty = \|\sum_{i\geqslant n+1} f_i\|_\infty \leqslant \sum_{i\geqslant n+1} \|f_i\|_\infty \xrightarrow[n\to\infty]{} 0$$

Now suppose that $p \in [1,\infty)$. Minkowski's inequality gives

$$\left(\int \left(\sum_{i=1}^n |f_i|\right)^p d\mu\right)^{1/p} = \left\|\sum_{i=1}^n |f_i|\right\|_p \leqslant \sum_{i=1}^n \|f_i\|_p \qquad (*)$$

This holds for all n. Applying the MCT to the sequence of functions $(\sum_{i=1}^n |f_i|)^p$ we get

$$\begin{aligned}\int (\sum_{i=1}^\infty |f_i|)^p \, d\mu &= \lim_n \int (\sum_{i=1}^n |f_i|)^p \, d\mu \qquad \text{(MCT)} \\ &\leqslant \lim_n \left(\sum_{i=1}^n \|f_i\|_p\right)^p \qquad \text{(by } (*)\text{)} \\ &= \left(\sum_{i=1}^\infty \|f_i\|_p\right)^p \\ &< \infty\end{aligned}$$

Therefore $g = (\sum_{i=1}^\infty |f_i|)^p$ is integrable. It follows that $\sum_{i=1}^\infty |f_i(x)|$ converges for all x in a conull set $C \in \mathcal{A}$. Define $f = \mathbb{1}_C \sum_{i\in\mathbb{N}} f_i$. Then f is measurable and in $\mathscr{L}^p$ since $|f|^p \leqslant g$. Moreover, for all $x \in C$ we have

$$0 = \lim_n (f(x) - \sum_{i=1}^n f_i(x)) \quad \text{and} \quad |f(x) - \sum_{i=1}^n f_i(x)|^p \leqslant g(x)$$

Using the DCT we then conclude that

$$\begin{aligned}
\lim_n \left\| f - \sum_{i=1}^n f_i \right\|_p &= \lim_n \left(\int |f - \sum_{i=1}^n f_i|^p \, d\mu \right)^{1/p} = \left(\lim_n \int |f - \sum_{i=1}^n f_i|^p \, d\mu \right)^{1/p} \\
&= \left(\int |f - \sum_{i=1}^\infty f_i|^p \, d\mu \right)^{1/p} \qquad \text{(DCT)} \\
&= 0
\end{aligned}$$

□

We mention the following. Further details about L^p can be found in Rudin's book (for example).

Proposition 42. *The simple functions determine a dense subspace of* L^p. □

Proposition 43. *Let* $p \in [1, \infty)$. *If* μ *is* σ*-finite and* $\mathcal{A}$ *is countably generated, then* L^p *is separable.* □

Lecture 11: Signed measures

To be able to talk about the idea of differentiating one measure with respect to another and for other applications it's useful to relax the requirement that measures be positive valued.

Definition 44. Let $\mathcal{A}$ be a σ-algebra on a set X. A **signed measure** is a function $\nu : \mathcal{A} \to \bar{\mathbb{R}}$ (with at most one of $-\infty$ and $+\infty$ in its image) satisiying:

1) $\nu(\emptyset) = 0$
2) If $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{A}$ is a disjoint family, then $\nu(\bigcup_i A_i) = \sum_i \nu(A_i)$ (and the sum is absolutely convergent)

A signed measure is called **finite** if neither $+\infty$ nor $-\infty$ occur among its values.

Examples 45. a) Let $f \in \mathscr{L}^1(X, \mathcal{A}, \mu, \mathbb{R})$. Then $\nu(A) = \int_A f\, d\mu$ gives a signed measure on $(X, \mathcal{A})$. Notice that $\nu = \nu^+ - \nu^-$ where ν^+ and ν^- are (positive) measures given by $\nu^\pm(A) = \int_A f^\pm\, d\mu$.

b) Let ν^+ and ν^- be (positive) measures on $(X, \mathcal{A})$, at least one of which is finite. Then $\nu = \nu^+ - \nu^-$ is a signed measure on $(X, \mathcal{A})$.

Lemma 46. *Let ν be a signed measure on $(X, \mathcal{A})$. Let $(A_i)_i$ be a sequence of sets from $\mathcal{A}$.*

1) *If $(A_i)_i$ is increasing, then $\nu(\cup_i A_i) = \lim_i \nu(A_i)$.*
2) *If $(A_i)_i$ is decreasing and $\nu(A_1) < \infty$, then $\nu(\cap_i A_i) = \lim_i \nu(A_i)$.*

Exercise 33. Prove this lemma. (See Lemma 6.)

Suppose ν is a signed measure on $(X, \mathcal{A})$. A set $A \in \mathcal{A}$ is called **positive** if $\nu(B) \geqslant 0$ for all subsets $B \subset A$ with $B \in \mathcal{A}$. Similarly, $A \in \mathcal{A}$ is called **negative** if $\nu(B) \leqslant 0$ for all subsets $B \subset A$ with $B \in \mathcal{A}$.

Exercise 34. Show that a countable union of positive sets is positive.

Theorem 47 (Hahn decomposition theorem)**.** *Let ν be a signed measure on $(X, \mathcal{A})$. Then*

1) *There exists a positive set $P \in \mathcal{A}$ and a negative set $N \in \mathcal{A}$ such that $X = P \cup N$ and $P \cap N = \emptyset$.*
2) *If $X = P' \cup N'$ is another such partition, then $\nu(P\Delta P') = \nu(N\Delta N') = 0$.*

Proof. We can assume (by replacing ν with $-\nu$ if necessary) that ν does not take value $+\infty$. Let $\delta = \sup\{\nu(A) \mid A \in \mathcal{A} \text{ is positive }\}$. There exist positive sets $P_i \in \mathcal{A}$ such that $\nu(P_i) \to \delta$. The set $P = \cup_i P_i$ is positive because it's a countable union of positive sets. Moreover, $\nu(P) = \delta$ since $\nu(P) = \nu(P \setminus P_i) + \nu(P_i) \geqslant \nu(P_i)$, and therefore $\delta < \infty$.

Let $N = X \setminus P$. We want to show that N is negative. Note first that if $A \subset N$ is positive, then $\nu(A) = 0$, since $\nu(P) = \delta \geqslant \nu(A \cup P) = \nu(A) + \nu(P)$. If $A \subset N$ and $\nu(A) > 0$, then (since A is not positive) there exists a subset $B \subset A$ with $\nu(B) < 0$ and therefore $\nu(A \setminus B) > \nu(A)$.

Suppose, for a contradiction, that N is not negative. Then there exists $A \subset N$ with $\nu(A) > 0$. Define a partial order on the set $\Sigma = \{A \subset N \mid A \in \mathcal{A}, \nu(A) > 0\}$ by

$$A_1 \preccurlyeq A_2 \iff (A_1 = A_2 \text{ or } (A_2 \subset A_1 \text{ and } \nu(A_1) < \nu(A_2))$$

If $A_1 \preccurlyeq A_2 \preccurlyeq \cdots$, then $\cap_i A_i \in \Sigma$ and $A_i \preccurlyeq \cap_i A_i$. By Zorn's lemma[2], there is a maximal element $A \in \Sigma$. This is a contradiction since $\nu(A) > 0$ and therefore A contains a subset of strictly larger measure. Therefore N is negative.

If $X = P' \cup N'$ is another partition with P' positive and N' negative, then $P \setminus P' \subset P$ and $P \setminus P' \subset (P')^c = N'$. Therefore $\mu(P \setminus P') = 0$.

□

A decomposition as given in the above theorem is called a **Hahn decomposition**. Two signed measures μ and ν are said to be **mutually singular** (written $\mu \perp \nu$) if there exist $M, N \in \mathcal{A}$ such that $M \cap N = \emptyset$, $M \cup N = X$, M is μ-null and N is ν-null.

[2] It's possible to give a proof that does not appeal to Zorn's lemma. See the notes by Greg Hjorth.

Corollary 48 (Hahn-Jordan decomposition theorem)**.** *Let ν be a signed measure on $(X, \mathcal{A})$. There exist unique positive measures ν^+ and ν^- such that $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$.*

Exercise 35. Prove the above corollary.

Using the above result we can now define the integral of a function $f \in L^1(\nu^+) \cap L^1(\nu^-)$ with respect to a signed measure ν by

$$\int f\, d\nu = \int f\, d\nu^+ - \int f\, d\nu^-$$

The **variation** of a signed measure ν is the positive measure $|\nu|$ defined by $|\nu| = \nu^+ + \nu^-$. The **total variation** $\|\nu\|$ of ν is defined by $\|\nu\| = |\nu|(X)$. Denote by $M(X, \mathcal{A}, \mathbb{R})$ the set of all finite signed measures on $(X, \mathcal{A})$.

Exercise 36. Verify that $M(X, \mathcal{A}, \mathbb{R})$ forms a vector space (using the obvious operations) and that the total variation is a norm.

Exercise 37. Show that if $A, B \in \mathcal{A}$ are disjoint, then $|\nu(A)| + |\nu(B)| \leqslant \|\nu\|$

Lecture 12: Signed measures (continued)

Proposition 49. *The space $M(X, \mathcal{A}, \mathbb{R})$ equipped with the total variation norm is complete.*

Proof. Let $(\nu_i)_i$ be a Cauchy sequence in M. For any $A \in \mathcal{A}$, $|\nu_i(A) - \nu_j(A)| \leqslant \|\nu_i - \nu_j\|$. Therefore, for a fixed $A \in \mathcal{A}$ the sequence $(\nu_i(A))_i$ is a Cauchy sequence of real numbers and hence is convergent. Define a function $\nu : \mathcal{A} \to \mathbb{R}$ by $\nu(A) = \lim_i \nu_i(A)$. We'll check that ν is a signed measure and that $\nu_i \to \nu$.

It's clear that $\nu(\emptyset) = 0$ and that ν is finitely additive. Let $(A_i)_i$ be a decreasing sequence of sets from $\mathcal{A}$ with $\cap_i A_i = \emptyset$. Given $\epsilon > 0$ let N be such that $|\nu(A) - \nu_i(A)| < \epsilon/2$ for all $i \geqslant N$ and $A \in \mathcal{A}$. By Lemma 46, $\lim_i \nu_N(A_i) = 0$. Let K be such that $|\nu_N(A_i)| \leqslant \epsilon/2$ whenever $i \geqslant K$. Then, for $i \geqslant K$

$$|\nu(A_i)| \leqslant |\nu(A_i) - \nu_N(A_i)| + |\nu_N(A_i)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

The countable additivity of ν now follows since

$$\nu(\cup_{i \in \mathbb{N}} B_i) = \nu(\cup_{i=1}^k B_i \bigcup \cup_{i=k+1}^\infty B_i) = \sum_{i=1}^k \nu(B_i) + \nu(\cup_{i=k+1}^\infty B_i)$$

It remains to show that $\|\nu - \nu_i\| \to 0$. Let $\epsilon > 0$ and N be such that $\|\nu_i - \nu_j\| < \epsilon$ whenever $i, j \geqslant N$. Let $X = P_j \cup N_j$ be a Hahn decomposition for $\nu - \nu_j$.

$$|\nu_i(P_j) - \nu_j(P_j)| + |\nu_i(N_j) - \nu_j(N_j)| \leqslant \|\nu_i - \nu_j\|$$

therefore

$$\begin{aligned} \|\nu - \nu_j\| = |\nu(P_j) - \nu_j(P_j)| + |\nu(N_j) - \nu_j(N_j)| &= \lim_i \left(|\nu_i(P_j) - \nu_j(P_j)| + |\nu_i(N_j) - \nu_j(N_j)|\right) \\ &\leqslant \lim_i \|\nu_i - \nu_j\| \\ &\leqslant \lim_i \epsilon = \epsilon \end{aligned}$$

□

Definition 50. Let $(X, \mathcal{A})$ be a measure space, let ν be a signed measure on $(X, \mathcal{A})$, and let μ be a positive measure on $(X, \mathcal{A})$. We say that ν is **absolutely continuous** with respect to μ if $\mu(A) = 0 \implies \nu(A) = 0$ for all $A \in \mathcal{A}$. We will write this as $\nu \ll \mu$.

Exercise 38. Show that

a) $\nu \ll \mu$ iff ($\nu^+ \ll \mu$ and $\nu^- \ll \mu$)

b) ($\nu \ll \mu$ and $\nu \perp \mu$) $\implies \nu = 0$

The term 'absolutely continuous' is motivated by the following exercise.

Exercise 39. Let ν be a finite signed measure and μ a positive measure on $(X, \mathcal{A})$. Show that $\nu \ll \mu$ iff

$$\forall\ \epsilon > 0\ \exists\ \delta > 0 \text{ such that } |\nu(A)| < \epsilon \text{ whenever } \mu(A) < \delta$$

Lebesgue-Radon-Nikodym theorem

Suppose that $f \in L^1(X, \mathcal{A}, \mu, \mathbb{R})$ is positive. Then, as we've seen, $\nu(A) = \int_A f\, d\mu$ defines a positive measure on $(X, \mathcal{A})$. This measure ν is clearly absolutely continuous with respect to μ. In fact, every finite measure on $(X, \mathcal{A})$ that is absolutely continuous with respect to μ arises in this way.

Theorem 51 (Lebesgue-Radon-Nikodym theorem). *Let μ be a σ-finite measure on $(X, \mathcal{A})$ and let ν be a σ-finite signed measure on $(X, \mathcal{A})$. Then there exist unique σ-finite signed measures λ and ρ on $(X, \mathcal{A})$ such that*

1. $\nu = \lambda + \rho$
2. $\lambda \perp \mu$
3. $\rho \ll \mu$

Further, there exists $f : X \to \mathbb{R}$ such that $\rho(A) = \int_A f \, d\mu$ for each $A \in \mathcal{A}$.
The function f is unique up to μ-almost everywhere equality.

In particular, if $\nu \ll \mu$ then we have $\nu(A) = \int_A f \, d\mu$ for each $A \in \mathcal{A}$. This is sometimes written as $d\nu = f d\mu$. The function f is called the **Radon-Nikodym derivative** of ν with respect to μ and is sometimes denoted $\frac{d\nu}{d\mu}$.

Proof of Lebesgue-Radon-Nikodym Theorem. We consider first the case in which both μ and ν are positive and finite. Define $F = \{f : X \to [0, \infty] \mid \int_A f \, d\mu \leqslant \nu(A) \; \forall A \in \mathcal{A}\}$. Note that the zero function is in F and that if $f, g \in F$ then $\max\{f, g\} \in F$. Let $\alpha = \sup\{\int f \, d\mu \mid f \in F\}$. Then $\alpha \leqslant \nu(X) < \infty$. We use the MCT to show that there is a function in F that achieves this value. Let $f_i \in F$ be such that $\int f_i \, d\mu \to \alpha$ and define $g_i = \max\{f_1, \ldots, f_i\}$ and $f = \sup_{i \in \mathbb{N}} f_i$. We have $g_{i+1} \geqslant g_i$ and that $f = \lim_i g_i$. By the MCT we have $\int f \, d\mu = \lim_i \int g_i \, d\mu$ and therefore $f \in F$. Also,

$$\alpha \geqslant \int g_i \, d\mu \geqslant \int f_i \, d\mu \to \alpha$$

Therefore $\int_A f \, d\mu = \alpha$. Now define $\lambda(A) = \nu(A) - \int_A f \, d\mu$ and $\rho(A) = \int_A f \, d\mu$. Clearly, $\nu = \lambda + \rho$ and $\rho \ll \mu$. We need to show that $\lambda \perp \mu$.

Continued next lecture... $\square$

Lecture 13: The Lebesgue-Radon-Nikodym theorem

Last lecture we began the proof of the following:

Theorem (Lebesgue-Radon-Nikodym theorem)**.** *Let μ be a σ-finite measure on $(X, \mathcal{A})$ and let ν be a σ-finite signed measure on $(X, \mathcal{A})$. Then there exist unique σ-finite signed measures λ and ρ on $(X, \mathcal{A})$ such that*

1. $\nu = \lambda + \rho$ *2.* $\lambda \perp \mu$ *3.* $\rho \ll \mu$

Further, there exists $f : X \to \mathbb{R}$ such that $\rho(A) = \int_A f \, d\mu$ for each $A \in \mathcal{A}$.
The function f is unique up to μ-almost everywhere equality.

We will use the following.

Exercise 40. Let λ and μ be finite positive measures on $(X, \mathcal{A})$. Then, either $\lambda \perp \mu$ or there exist $\epsilon > 0$ and $A \in \mathcal{A}$ such that $\mu(A) > 0$ and $\lambda \geqslant \epsilon\mu$ on A (i.e., A is positive for $\lambda - \epsilon\mu$).

Proof of L-R-N continued. We want to show that λ (as defined previously) satisfies $\lambda \perp \mu$. Suppose, for a contradiction, that λ and μ are not mutually singular. From Exercise40 there exist $\epsilon > 0$ and $B \in \mathcal{A}$ such that $\mu(B) > 0$ and $\lambda \geqslant \epsilon\mu$ on B. Then, for any $A \in \mathcal{A}$ we have

$$\begin{aligned} \epsilon\mu(A \cap B) \leqslant \lambda(A \cap B) \leqslant \lambda(A) = \nu(A) - \int_A f \, d\mu \\ \implies \int_A (f + \epsilon \mathbb{1}_B) \, d\mu \leqslant \nu(A) \\ \implies f + \epsilon \mathbb{1}_B \in F \end{aligned}$$

But this contradicts the choice of α since $\int_X f + \epsilon \mathbb{1}_B \, d\mu = \alpha + \epsilon\mu(B) > \alpha$.
For the uniqueness of λ and ρ, suppose that $\nu(A) = \lambda'(A) + \int_A f' \, d\mu$ for all $A \in \mathcal{A}$ and that $\lambda' \perp \mu$. Then $(\lambda - \lambda') \perp \mu$ since $\lambda \perp \mu = 0$ and $\lambda' \perp \mu = 0$. Say $X = Y \cup Z$ where Y is $(\lambda - \lambda')$-null and Z is μ-null. Then

$$(\lambda - \lambda')(A) = (\lambda - \lambda')(A \cap Z) = \int_{A \cap Z} f' - f \, d\mu = 0$$

Therefore $\lambda = \lambda'$. Also, for any $C \subset Y$

$$(\lambda - \lambda')(C) = 0 \implies \int_C f' - f \, d\mu = 0$$

Therefore $f = f'$ μ-a.e.
Now we consider the case in which both ν and μ are σ-finite and positive. Let $A_i \in \mathcal{A}$ be disjoint and such that $X = \cup_i A_i$ and $\mu(A_i) < \infty$ and $\nu(A_i) < \infty$. Define $\mu_i(A) = \mu(A \cap A_i)$ and $\nu_i(A) = \nu(A \cap A_i)$. From the previous case there are λ_i and ρ_i such that $\nu_i = \lambda_i + \rho_i$, $\lambda_i \perp \mu_i$, $\rho_i \ll \mu_i$, and $f_i : X \to \mathbb{R}$ such that $\rho_i(A) = \int_A f_i \, d\mu_i$ (with $f_i|_{A_i^c} = 0$). Define $\lambda = \sum_i \lambda_i$ and $f = \sum_i f_i$. Then

$$\nu(A) = \sum_i \nu_i(A) = \sum_i \left(\lambda_i(A) + \int_A f_i \, d\mu_i \right) = \lambda(A) + \sum_i \int_{A \cap A_i} f_i \, d\mu = \lambda(A) + \int_A f \, d\mu$$

That $\lambda \perp \mu$ follows from Exercise 41 below.
The general case, in which ν is a signed measure follows from applying the above to ν^+ and ν^-

□

Exercise 41. Let μ be a measure and suppose that λ_i are measures satisfying $\lambda_i \perp \mu$. Show that $\sum_i \lambda_i \perp \mu$.

Exercise 42. Suppose $\nu, \nu_1, \nu_2 \ll \mu$ and $g \in L^1(\nu)$ and $\mu \ll \xi$. Establish the following

a) $\dfrac{d(\nu_1 + \nu_2)}{d\mu} = \dfrac{d\nu_1}{d\mu} + \dfrac{d\nu_2}{d\mu}$ b) $\displaystyle\int g \, d\nu = \int g \frac{d\nu}{d\mu} \, d\mu$ c) $\dfrac{d\nu}{d\xi} = \dfrac{d\nu}{d\mu} \dfrac{d\mu}{d\xi}$

Remarks. The condition that μ be σ-finite is needed. Suppose, for example, that μ is counting measure on $X = [0,1]$ and that ν is Lebesgue measure on $[0,1]$. Then $\nu \ll \mu$ but $d\nu/\,d\mu$ does not exist.

We finish this section on signed measures by noting the relationship with functions of *bounded variation.* When we considered Borel measures on $\mathbb{R}$ we saw that there is a bijection between the set of all bounded positive measures on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ and the set of all bounded non-decreasing right-continuous functions $F : \mathbb{R} \to \mathbb{R}$ that satisfy $\lim_{x\to-\infty} F(x) = 0$. (This follows from Theorem 18.)

Suppose that ν is a finite signed measure on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Define a function $F_\nu : \mathbb{R} \to \mathbb{R}$ by

$$F_\nu(x) = \nu((-\infty, x])$$

It's easy to check that $\lim_{x\to-\infty} F_\nu(x) = 0$. Moreover, writing $\nu = \nu^+ + \nu^-$ and using the result for positive measures, we can show that F_ν is right-continuous.

If $t_0 < t_1 < t_2 < \cdots < t_k$ is an increasing sequence of of real numbers then

$$\sum_{i=1}^{k} |F_\nu(t_i) - F_\nu(t_{i-1})| = \sum_{i=1}^{k} |\, \nu(t_{i-1}, t_i]\, | \leqslant \|\nu\|$$

In general, a function $F : \mathbb{R} \to \mathbb{R}$ is said to be of **bounded variation** if

$$\sup\{\sum_i |F(t_i) - F(t_{i-1})| \mid (t_i)_i \text{ is increasing finite sequence } \} < \infty$$

The function F_ν is therefore right-continuous and of bounded variation. In fact

Proposition 52. *The map $\nu \mapsto F_\nu$ is a bijection between the set of all finite signed measures on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ and the set of all right-continuous functions $F : \mathbb{R} \to \mathbb{R}$ of bounded variation such that $\lim_{x\to-\infty} F(x) = 0$.*

Proof. Sketch. We've already argued that F_ν is of the right form. Suppose that $F_\mu = F_\nu$. Then $F_{\mu^+} - F_{\mu^-} = F_{\nu^+} - F_{\nu^-}$. From Theorem 18 it follows that $\mu^+ + \nu^- = \nu^+ + \mu^-$ and hence that $\mu = \nu$. For surjectivity note that if F is a right-continuous function of bounded variation, then there exist bounded right-continuous non-decreasing functions F^+ and F^- such that $F = F^+ - F^-$. To see this, let $F^\pm = (V_F \pm F)/2$, where $V_F(x)$ is the variation of F over $(-\infty, x]$. Then apply Theorem 18.

□

Lecture 14: Product measures

Given measure spaces $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ we would like to combine μ and ν to obtain a measure on $X \times Y$. Define the **product** of $\mathcal{A}$ and $\mathcal{B}$ to be the σ-algebra on $X \times Y$ given by

$$\mathcal{A} \otimes \mathcal{B} = \langle \{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\} \rangle \subset P(X \times Y)$$

That is, $\mathcal{A} \otimes \mathcal{B}$ is the σ-algebra generated by the collection of all **rectangles**, meaning a set of the form $A \times B = \{(a, b) \mid a \in A, b \in B\}$ for some $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Let $\mathcal{R}_0$ be the collection of all subsets of $X \times Y$ that can be written as a finite disjoint union of rectangles.

Exercise 43. Check that $\mathcal{R}_0$ is an algebra of sets and that $\mathcal{A} \otimes \mathcal{B}$ is the σ-algebra generated by $\mathcal{R}_0$.

Exercise 44. Show that $\mathcal{B}_{\mathbb{R}} \otimes \mathcal{B}_{\mathbb{R}} = \mathcal{B}_{\mathbb{R}^2}$.

We will use μ and ν to define a premeasure ξ_0 on $\mathcal{R}_0$ which then extends, by Carathéodory's Extension Theorem (Theorem 15), to a measure on $\mathcal{A} \otimes \mathcal{B}$. Define $\xi_0 : \mathcal{R}_0 \to [0, \infty]$ by

$$\xi_0(\cup_i A_i \times B_i) = \sum_i \mu(A_i)\nu(B_i)$$

Exercise 45. Check that ξ_0 is well-defined and is a premeasure on $\mathcal{R}_0$.

By Carathéodory's Extension Theorem, ξ_0 extends to a measure on $(X \times Y, \mathcal{A} \otimes \mathcal{B})$. If μ and ν are each σ-finite, then there is a unique such extension which we call the **product measure** and denote by $\mu \times \nu$. Note that

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

for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$ and $\mu \times \nu$ is the unique such measure.

Example 53. Let m be Lebesgue measure on $(\mathbb{R}, \mathcal{L})$. The measure $m \times m$ on $(\mathbb{R}^2, \mathcal{L} \otimes \mathcal{L})$ is not complete. Let $N \in \mathcal{L}$ be non-empty and μ-null, and let $V \in \mathcal{P}(\mathbb{R}) \setminus \mathcal{L}$ (e.g., the Vitali set). Then $N \times \mathbb{R} \in \mathcal{L} \otimes \mathcal{L}$ and $m \times m(N \times \mathbb{R}) = 0$, however $N \times V \notin \mathcal{L} \otimes \mathcal{L}$ and $N \times V \subset N \times \mathbb{R}$.

We define **Lebesgue measure on** $\mathbb{R}^n$ to be the completion of the measure $m \times \cdots \times m$ on $\mathcal{B}_{\mathbb{R}} \otimes \cdots \otimes \mathcal{B}_{\mathbb{R}} = \mathcal{B}_{\mathbb{R}^n}$ and denote the measure space by $(\mathbb{R}^n, \mathcal{L}^n, m^n)$.

Exercise 46. Show that the completion of $\mathcal{B}_{\mathbb{R}^n}$ with respect to m^n is equal to the completion of $\mathcal{L} \otimes \cdots \otimes \mathcal{L}$ with respect to m^n.

We want to compare integration with respect to a product measure $\mu \times \nu$ with integration first with respect to μ and then with respect to ν.

Definition 54. For a subset $S \subset X \times Y$ and $x \in X$ and $y \in Y$ define sets $S_x \subset Y$ and $S^y \subset X$ (called **sections**) by

$$S_x = \{y \in Y \mid (x, y) \in S\} \qquad S^y = \{x \in X \mid (x, y) \in S\}$$

For a function f with domain $X \times Y$ define functions f_x on Y and f^y on X by

$$f_x(y) = f(x, y) \qquad f^y(x) = f(x, y)$$

Lemma 55. *Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces.*

1) If $S \in \mathcal{A} \otimes \mathcal{B}$, then $S_x \in \mathcal{B}$ and $S^y \in \mathcal{A}$

2) If $f : X \times Y \to \bar{\mathbb{R}}$ is $(\mathcal{A} \otimes \mathcal{B})$-measurable, then f_x is $\mathcal{B}$-measurable and f^y is $\mathcal{A}$-measurable.

Proof. Let $\Sigma = \{S \subset X \times Y \mid S_x \in \mathcal{B}$ for all $x \in X$, and $S^y \in \mathcal{A}$ for all $y \in Y\}$. Check that Σ is a σ-algebra, and that Σ contains all rectangles. It follows that $\mathcal{A} \otimes \mathcal{B} \subset \Sigma$.
The second part follows from the first since $(f_x)^{-1}(C) = (f^{-1}(C))_x$ and $(f^y)^{-1}(C) = (f^{-1}(C))^y$. □

In the proof of Proposition 58 below we will need the Monotone Class Theorem.

Definition 56. A collection $\mathcal{M} \subset \mathcal{P}(X)$ of subsets of X is called a **monotone class** if $X \in \mathcal{M}$ and $\mathcal{M}$ is closed under both countable increasing unions and countable decreasing intersections.

Every σ-algebra is a monotone class, but a monotone class need not be a σ-algebra. However, the monotone class generated by an algebra is always a σ-algebra by the following result.

Theorem 57 (Monotone Class Theorem)**.** *If $\mathcal{A}_0 \subset \mathcal{P}(X)$ is an algebra of sets, then the monotone class generated by $\mathcal{A}_0$ coincides with the σ-algebra generated by $\mathcal{A}_0$.*

Proof. Let $\mathcal{M}$ be the monotone class and $\mathcal{A}$ the σ-algebra generated by $\mathcal{A}_0$. The inclusion $\mathcal{M} \subset \mathcal{A}$ is immediate. For the reverse inclusion we need to show that $\mathcal{M}$ is a σ-algebra.

We first show that $A, B \in \mathcal{M}$ implies $A \cap B \in \mathcal{M}$. Given $A \in \mathcal{M}$, define $\mathcal{M}(A) = \{B \in \mathcal{M} \mid A \cap B \in \mathcal{M}\}$. It's easy to check that $\mathcal{M}(A)$ is a monotone class. Since $\mathcal{A}_0$ is an algebra and $\mathcal{A}_0 \subset \mathcal{M}$

$$A \in \mathcal{A}_0 \implies \mathcal{A}_0 \subset \mathcal{M}(A) \implies \mathcal{M} \subset \mathcal{M}(A)$$

Therefore, if $A \in \mathcal{A}_0$ and $B \in \mathcal{M}$, then $A \cap B \in \mathcal{M}$. Since $\mathcal{M}(B)$ is a monotone class containing $\mathcal{A}_0$, we have $\mathcal{M} \subset \mathcal{M}(B)$. Therefore, $\mathcal{M}$ is closed under finite intersections.

Now observe that $\mathcal{M}$ is closed under countable intersections because it is closed under finite intersections and is a monotone class.

All that remains is to check that $\mathcal{M}$ is closed under complementation. Let $\mathcal{N} = \{A \in \mathcal{M} \mid A^c \in \mathcal{M}\} \subset \mathcal{M}$. It's easy to check that $\mathcal{N}$ is a monotone class and that $\mathcal{A}_0 \subset \mathcal{N}$. Therefore $\mathcal{M} \subset \mathcal{N}$, because $\mathcal{M}$ is the monotone class generated by $\mathcal{A}_0$. $\square$

Lecture 15: Fubini's theorem

We would like to evaluate an integral with respect to a product measure as two iterated integrals.

Proposition 58. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces and let $S \in \mathcal{A} \otimes \mathcal{B}$. Then*

1) the map $x \mapsto \nu(S_x)$ is $\mathcal{A}$-measurable

2) the map $y \mapsto \mu(S^y)$ is $\mathcal{B}$-measurable

3) $\mu \times \nu(S) = \int \nu(S_x)\, d\mu = \int \mu(S^y)\, d\nu$

Proof. Suppose first that both μ and ν are both finite. Let $\Sigma \subset \mathcal{A} \otimes \mathcal{B}$ be the collection of sets S for which the proposition holds. Our strategy is to show that Σ contains the algebra $\mathcal{R}_0$ of all disjoint unions of rectangles and that Σ is a monotone class. It then follows from the Monotone Class Theorem that $\Sigma \supset \mathcal{A} \otimes \mathcal{B}$.

If $S = A \times B$ is a rectangle, then $\nu(S_x) = \mathbb{1}_A(x)\nu(B)$. The map $x \mapsto \nu(S_x)$ is simple, hence measurable. Similarly, the map $y \mapsto \mu(S^y)$ is measurable because $\mu(S^y) = \mu(A)\mathbb{1}_B(y)$. Also

$$\mu \times \nu(S) = \mu(A)\nu(B) = \int \mathbb{1}_A \nu(B)\, d\mu = \int \mu(A)\mathbb{1}_B\, d\nu$$

Therefore Σ contains all rectangles. Similarly, any finite disjoint union of rectangles is in Σ, that is, $\mathcal{R}_0 \subset \Sigma$.

Now to show that Σ is a monotone class. Suppose that $(A_i)_{i\in\mathbb{N}}$ is an increasing sequence of elements of Σ and let $A = \cup_i A_i$. Let $f_i : Y \to [0, \infty]$ be the $\mathcal{B}$-measurable function given by $f_i(y) = \mu(A_i^y)$. The functions f_i increase pointwise to the function $f : Y \to [0, \infty]$, $f(y) = \mu(A^y)$. Therefore f is $\mathcal{B}$-measurable and, by the Monotone Convergence Theorem, we have

$$\int \mu(A^y)\, d\nu = \lim \int \mu(A_i^y)\, d\nu = \lim \mu \times \nu(A_i) = \mu \times \nu(A)$$

Similarly, the map $x \mapsto \nu(A_x)$ is $\mathcal{A}$-measurable and $\int \nu(A_x)\, d\mu = \mu \times \nu(A)$. Hence $A \in \Sigma$ and Σ is closed under countable increasing unions. Suppose now that $B = \cap_{i\in\mathbb{N}} B_i$ is a countable decreasing union of elements $B_i \in \Sigma$. The map $y \mapsto \mu(B_1^y)$ is in $L^1(\nu)$ since $\mu(B_1^y) \leqslant \mu(X) < \infty$ and $\nu(Y) < \infty$. Applying the Dominated Convergence Theorem gives

$$\int \mu(B^y)\, d\nu = \lim \int \mu(B_i^y)\, d\nu = \lim \mu \times \nu(B_i) = \mu \times \nu(B)$$

Similarly, $\int \nu(B_x)\, d\mu = \mu \times \nu(B)$ and $B \in \Sigma$. The result therefore holds in the case that μ and ν are both finite.

For the general case write $X \times Y$ as an increasing union $X \times Y = \cup_i X_i \times Y_i$ where $\mu(X_i) < \infty$ and $\nu(Y_i) < \infty$. For any $S \in \mathcal{A} \otimes \mathcal{B}$ we have from the finite case above that

$$\mu \times \nu(S \cap (X_i \times Y_i)) = \int \mathbb{1}_{X_i} \nu(S_x \cap Y_i)\, d\mu = \int \mathbb{1}_{Y_i} \mu(S^y \cap X_i)\, d\nu$$

Then apply the Monotone Convergence Theorem. □

Proposition 59. *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces and let $f : X \times Y \to \bar{\mathbb{R}}$ be positive and $\mathcal{A} \otimes \mathcal{B}$-measurable. Then*

1) the function $x \mapsto \int_Y f_x\, d\nu$ is $\mathcal{A}$-measurable

2) the function $y \mapsto \int_X f^y\, d\mu$ is $\mathcal{B}$-measurable

3) $\displaystyle\int_{X\times Y} f\, d(\mu \times \nu) = \int_Y \left(\int_X f^y\, d\mu \right) d\nu = \int_X \left(\int_Y f_x\, d\nu \right) d\mu$

Proof. First note that for $f = \mathbb{1}_S$ with $S \in \mathcal{A} \otimes \mathcal{B}$ we have $f_x = \mathbb{1}_{S_x}$ and so $\int f_x\, d\nu = \nu(S_x)$. Therefore, in this case, part 1 follows from the previous proposition as does $\int_X \int_Y f_x\, d\nu\, d\mu = \mu \times \nu(S) = \int_{X\times Y} f\, d(\mu \times \nu)$. Part 2 and the remaining equality in part 3 are similar.

The result holds for positive simple $\mathcal{A} \otimes \mathcal{B}$-measurable functions by the linearity properties of the integral. For the general case apply Lemma 25 and the Monotone Convergence Theorem. □

Remark. The function f in the above result is not assumed to be integrable. The result can sometimes be used to decide whether or not $|f|$ (hence f) is integrable.

Theorem 60 (Fubini's Theorem)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces and let $f \in \mathscr{L}^1(\mu \times \nu)$. Then*

1) *$f_x \in \mathscr{L}^1(\nu)$ for μ-almost all $x \in X$ and the function $x \mapsto \int_Y f_x \, d\nu$ is in $\mathscr{L}^1(\mu)$*
2) *$f^y \in \mathscr{L}^1(\mu)$ for ν-almost all $y \in Y$ and the function $y \mapsto \int_X f^y \, d\mu$ is in $\mathscr{L}^1(\nu)$*
3) $\displaystyle\int_{X \times Y} f \, d(\mu \times \nu) = \int_Y \left(\int_X f^y \, d\mu \right) d\nu = \int_X \left(\int_Y f_x \, d\nu \right) d\mu$

Proof. Write $f = f^+ - f^-$ with $f^\pm$ positive and integrable. By the previous proposition, the functions $x \mapsto \int (f^+)_x \, d\nu$ and $x \mapsto \int (f^-)_x \, d\nu$ are $\mathcal{A}$-measurable and integrable. Since the functions are integrable, they are finite μ-almost everywhere. Therefore f_x is ν-integrable for μ-almost all x. So part 1 holds. Using the previous proposition we also have

$$\begin{aligned}\int f \, d(\mu \times \nu) &= \int f^+ \, d(\mu \times \nu) - \int f^- \, d(\mu \times \nu) \\ &= \int \left(\int (f^+)_x \, d\nu \right) d\mu - \int \left(\int (f^-)_x \, d\nu \right) d\mu \\ &= \int \left(\int f_x \, d\nu \right) d\mu\end{aligned}$$

Similar arguments apply to f^y. □

Example 61. The hypothesis that the measures be σ-finite is needed. For example, consider $X = Y = [0, 1]$, $\mathcal{A} = \mathcal{B} = \mathcal{B}_{[0,1]}$, μ is Lebesgue measure and ν is counting measure. Let $S = \{(x, x) \mid x \in [0, 1]\}$. Then $S \in \mathcal{A} \otimes \mathcal{B}$ and

$$\begin{aligned}&\int \int (\mathbb{1}_S)^y \, d\mu \, d\nu = \int 0 \, d\nu = 0 \\ &\int \int (\mathbb{1}_S)_x \, d\nu \, d\mu = \int 1 \, d\mu = 1 \\ &\int \mathbb{1}_S \, d(\mu \times \nu) = \mu \times \nu(S) = \infty\end{aligned}$$

As an application of the above results on product measures let's consider the *convolution* of two Lebesgue integrable functions.

Proposition 62. *Let $f, g \in \mathscr{L}^1(\mathbb{R}, \mathcal{B}_\mathbb{R}, m)$. The function $f * g$ defined by*

$$f * g(x) = \begin{cases} \int f(x - t) g(t) \, d\mu(t) & \text{if } t \mapsto f(x-t)g(t) \text{ is Lebesgue integrable} \\ 0 & \text{otherwise} \end{cases}$$

belongs to $\mathscr{L}^1(\mathbb{R}, \mathcal{B}_\mathbb{R}, m)$ and $\|f * g\|_1 \leqslant \|f\|_1 \, \|g\|_1$.

Proof. The function $(x, t) \mapsto f(x - t)$ is the composition of a continuous and a Borel function and is therefore Borel. Similarly, $(x, t) \mapsto g(x)$ is Borel. Hence $(x, t) \mapsto f(x - t)g(x)$.

We have

$$\begin{aligned}\int |f(x-t)g(t)| d(m \times m) &= \int \int |f(x-t)g(t)| \, dm(x) \, dm(t) && \text{(Proposition 59)} \\ &= \int \|f\|_1 |g(t)| \, dm(t) && \text{(Lebesgue measure is translation invariant)} \\ &= \|f\|_1 \|g\|_1\end{aligned}$$

Therefore the function $(x, t) \mapsto f(x - t)g(t)$ is in $\mathscr{L}^1(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2}, m \times m)$ and then by Fubini's theorem we have that $t \mapsto f(x - t)g(t)$ is integrable for almost all x. Finally

$$\begin{aligned}&|f * g(x)| \leqslant \int |f(x-t)g(t)| \, dm(t) \\ \implies &\|f * g\|_1 \leqslant \int \int |f(x-t)g(t)| \, dm(t) \, dm(x) = \|f\|_1 \|g\|_1\end{aligned}$$

□

Exercise 47. Let m^n be Lebesgue measure on $(\mathbb{R}^n, \mathcal{L}^n)$ and let $A \in \mathcal{L}^n$.

1. Show that

$$m^n(A) = \inf\{m^n(V) \mid V \supset A, V \text{ open}\} = \sup\{m^n(K) \mid K \subset A, K \text{ compact}\}$$

2. Suppose that $m^n(A) < \infty$. Let $\epsilon > 0$. Show that there exists a finite disjoint collection of rectangles $\mathcal{R}_i \in \mathcal{B}_{\mathbb{R}^n}$ (i.e., sets of the form $\mathcal{R} = A_1 \times \cdots \times A_n$ with $A_i \in \mathcal{B}_{\mathbb{R}}$) such that $m^n(A \Delta \cup_{i=1}^k \mathcal{R}_i) < \epsilon$.

Lecture 16: Lebesgue measure on $\mathbb{R}^n$

We have already defined the measure space $(\mathbb{R}^n, \mathcal{L}^n, m^n)$. We want to note some useful properties. When it is clear from the context, we will sometimes write m in place of m^n for the measure.

Proposition 63. *Let $A \in \mathcal{L}^n$. Then*

1. $m(A) = \inf\{m(V) \mid V \supset A, V \subset \mathbb{R}^n \text{ open}\} = \sup\{m(K) \mid K \subset A, K \subset \mathbb{R}^n \text{ compact}\}$
2. *If $m(A) < \infty$, then for all $\epsilon > 0$ there exists a finite collection of disjoint rectangles R_i, whose sides are intervals, such that $m(A \Delta \cup_{i=1}^k R_i) < \epsilon$*
3. *m is invariant under translations and rotation*

Outline of proof. (in lecture) □

Lemma 64. *Let $\mathcal{B}$ be a collection of open balls in $\mathbb{R}^n$ and let $A = \cup_{B \in \mathcal{B}} B$. Let $c \in \mathbb{R}$ be such that $c < m(A)$. Then there exist disjoint $B_1, \ldots, B_k \in \mathcal{B}$ such that $\sum_{i=1}^k m(B_i) > 3^{-n} c$.*

Proof. (in lecture) □

Definition 65. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called **locally integrable** if $\int_A |f| \, dm < \infty$ for all bounded $A \in \mathcal{L}^n$. Denote by L^1_{loc} the set of all such functions. For $f \in L^1_{loc}$, $r > 0$ and $x \in \mathbb{R}^n$ define

$$A_r f(x) = \frac{1}{m(B(r,x))} \int_{B(r,x)} f(y) \, dy$$

The **Hardy-Littlewood maximal function** is given by

$$Hf(x) = \sup_{r>0} A_r |f|(x)$$

Lemma 66. *The function $A_r f$ is continuous in both r and x.*

Theorem 67 (Maximal Theorem)**.** $\exists C > 0 \ \forall f \in L^1 \ \forall \alpha > 0$

$$m(\{x \mid Hf(x) > \alpha\}) \leqslant \frac{C}{\alpha} \int |f| \, dm$$

Proof. (in lecture)

□

Theorem 68. *Let $f \in L^1_{loc}$. Then*

$$\lim_{r \to 0} A_r f(x) = f(x) \qquad \text{for almost all } x \in \mathbb{R}^n$$

Proof. (in lecture) □

Lecture 17: Hausdorff measure

We would like to measure the *size* of subsets of a metric space (X, d) in a way that doesn't assume any extra structure on the subset. For example, it should work for subsets of $\mathbb{R}^n$ that are not submanifolds.

Definition 69. A **metric outer measure** on X is an outer measure λ on X such that

$$\lambda(A \cup B) = \lambda(A) + \lambda(B) \qquad \text{whenever } d(A, B) > 0$$

Lemma 70. *Let λ be a metric outer measure on X. Every Borel set in X is λ-measurable.*

Proof. First note that, since the λ-measurable sets form a σ-algebra, it's sufficient to show that closed subsets of X are λ-measurable. Let $F \subset X$ be closed. We need to show that for all $A \subset X$ with $\lambda(A) < \infty$ we have

$$\lambda(A) \geqslant \lambda(A \cap F) + \lambda(A \cap F^c)$$

Let $A_i = \{x \in A \cap F^c \mid d(x, F) \geqslant 1/i\}$. Since F is closed we have $\cup_i A_i = A \cap F^c$. Also

$$\begin{aligned} \lambda(A) &\geqslant \lambda((A \cap F) \cup A_i) && \text{(monotonicity of outer measures)} \\ &= \lambda(A \cap F) + \lambda(A_i) && \text{(metric outer measure)} \end{aligned}$$

We will be done if we show that $\lim_i \lambda(A_i) = \lambda(A \cap F^c)$. Let $C_i = A_{i+1} \setminus A_i$. Note that

$$d(C_{i+1}, A_i) \geqslant \frac{1}{i(i+1)} \qquad \text{and therefore} \qquad \lambda(A_{i+2}) \geqslant \lambda(A_i) + \lambda(C_{i+1})$$

Induction gives

$$\lambda(A) \geqslant \lambda(A_{2i+1}) \geqslant \sum_{j=1}^{i} \lambda(C_{2j}) \quad \text{and} \quad \lambda(A) \geqslant \lambda(A_{2i}) \geqslant \sum_{j=1}^{i} \lambda(C_{2j-1})$$

The two infinite series are therefore convergent and hence $\sum_{j \geqslant i} \lambda(C_j) \xrightarrow[i \to \infty]{} 0$. Since

$$\lambda(A \cap F^c) = \lambda(A_i \cup \bigcup_{j=i}^{\infty} C_j) \leqslant \lambda(A_i) + \sum_{j \geqslant i} \lambda(C_j)$$

we have

$$\lambda(A \cap F^c) \leqslant \liminf \lambda(A_i) \leqslant \limsup \lambda(A_i) \leqslant \lambda(A \cap F^c)$$

□

Definition 71. Let $n \geqslant 0$ and $\delta > 0$. For $A \subset X$ define

$$H_{n,\delta}(A) = \inf\{\sum \operatorname{diam}(A_i)^n \mid A \subset \cup_{i \in \mathbb{N}} A_i, \operatorname{diam}(A_i) \leqslant \delta\}$$

The n-**dimensional Hausdorff measure** of a set A is defined to be

$$H_n(A) = \lim_{\delta \to 0} H_{n,\delta}(A)$$

Exercise 48. Show that for $n = 0$ this is the same as counting measure.

Lecture 18: Hausdorff measure (continued)

Proposition 72. *H_n is a metric outer measure on X.*

Proof. That $H_{n,\delta}$ is an outer measure follows from Lemma 11. It follows that H_n is an outer measure. To see that it is a metric outer measure, consider $A, B \subset X$ with $d(A, B) > 0$. Let $\{C_i\}_{i\in\mathbb{N}}$ and $\delta > 0$ be such that $A \cup B \subset \cup_i C_i$ and $\operatorname{diam}(C_i) \leqslant \delta < d(A, B)$. No C_i can have non-empty intersection with both A and B. Let $I, J \subset \mathbb{N}$ be such that $C_i \cap A \neq \emptyset$ iff $i \in I$ and $C_i \cap B \neq \emptyset$ iff $i \in J$. Then $I \cap J = \emptyset$ and

$$A \cup B \subset \left(\bigcup_{i\in I} C_i\right) \cup \left(\bigcup_{i\in J} C_i\right) \qquad A \subset \bigcup_{i\in I} C_i \qquad B \subset \bigcup_{i\in J} C_i$$

Therefore $\sum_{i\in\mathbb{N}} \operatorname{diam}(C_i)^n \geqslant H_{n,\delta}(A) + H_{n,\delta}(B)$ and hence $H_{n,\delta}(A \cup B) \geqslant H_{n,\delta}(A) + H_{n,\delta}(B)$. Letting $\delta \to 0$ gives the required result. □

Remark. It follows that all Borel subsets of X are H_n-measurable and therefore H_n gives a measure on $(X, \mathcal{B}_X)$. The measure H_n on $(\mathbb{R}^n, \mathcal{B}_\mathbb{R})$ is a scalar multiple of Lebesgue measure.

Exercise 49. Show that H_n is invariant under isometries of X.

Lemma 73. *Let $A \subset X$.*

1) $H_n(A) < \infty \implies H_m(A) = 0$ *for all* $m > n$
2) $H_n(A) > 0 \implies H_m(A) = \infty$ *for all* $m < n$.

Proof. The two parts are equivalent. For the first suppose $A \subset X$ satisfies $H_n(A) < \infty$. Then for all $\delta > 0$ sufficiently small there exists a covering $A \subset \cup_i A_i$ with $\operatorname{diam}(A_i) < \delta$ and $\sum_i \operatorname{diam}(A_i)^n \leqslant H_n(A) + 1$. If $m > n$ we have

$$\sum_{i\in\mathbb{N}} \operatorname{diam}(A_i)^m \leqslant \delta^{m-n} \sum_{i\in\mathbb{N}} \operatorname{diam}(A_i)^n \leqslant \delta^{m-n}(H_n(A) + 1)$$

Therefore $H_{m,\delta}(A) \leqslant \delta^{m-n}(H_n(A) + 1)$ and letting $\delta \to 0$ gives $H_m(A) = 0$. □

The **Hausdorff dimension** of $\emptyset \neq A \subset X$ is defined to be $\inf\{m \geqslant 0 \mid H_m(A) = 0\}$. By the above lemma, this is equal to $\sup\{m \geqslant 0 \mid H_m(A) > 0\}$.

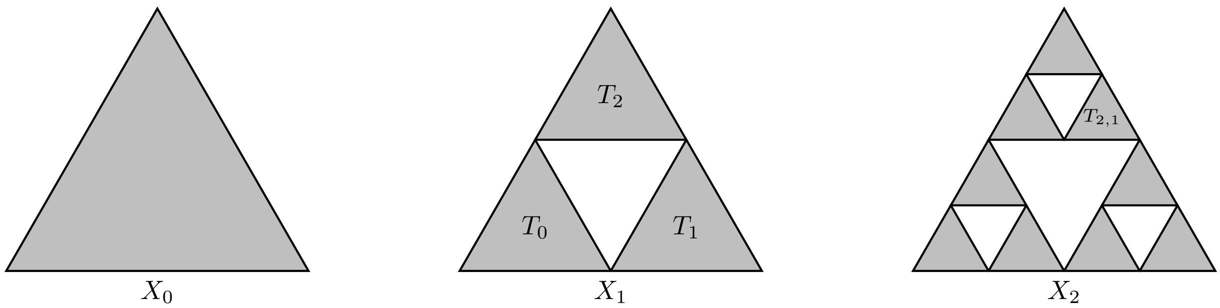
It's possible to show that if $X = \mathbb{R}^m$ and A is a C^1-submanifold, then this gives the expected dimension. But what about more complicated subsets? For example what is the Hausdorff dimension of the ternary Cantor set $C \subset \mathbb{R}$? We know that $H_1(C) = 0$ and $H_0(C) = \infty$. It turns out that C has Hausdorff dimension $\log_3 2$.

Lecture 19: Self-similarity and fractional Hausdorff dimension

The Hausdorff dimension of a submanifold of $\mathbb{R}^n$ agrees with our existing notion of dimension for such a space.[3] Interestingly, the Hausdorff dimension need not, in general, be an integer.

Sierpinski triangle

The **Sierpinski triangle** is the closed subset $X \subset \mathbb{R}^2$ defined as $X = \cap_{i\in\mathbb{N}} X_i$ where $X_i \supset X_{i-1}$ are defined recursively as indicated below.



The set X_{i+1} is made up of three scaled copies of X_i arranged by translations.

Exercise 50. Show that X is compact, has cardinality $2^{\aleph_0}$ and has Lebesgue measure zero.

We'll show that X has Hausdorff dimension $d = \log_2 3$. A similar argument applies to the ternary Cantor set and other *self-similar* sets in $\mathbb{R}^n$.

Claim 1. $H_d(X) \leqslant 1$

Proof. The set X_k is made up of 3^k triangles, each of diameter 2^{-k}. Therefore $H_{d,2^{-k}}(X) \leqslant 3^k(2^{-k})^d = 1$. □

Therefore to show that X has dimension d it will be sufficient (using Lemma 73) to show that $H_d(X) > 0$. To do this we define an outer measure λ on $\mathbb{R}^2$ by declaring that each level k triangle should have measure 3^{-k} and applying Lemma 11. That is, for $A \subset \mathbb{R}^2$, define

$$\lambda(A) = \inf\{\sum_i 3^{-\ell(w_i)} \mid w_i \in \{0,1,2\}^*, A \cap X \subset \cup_i T_{w_i}\}$$

Exercise 51. Check that

a) $\lambda(T_w) = 3^{-\ell(w)}$ for all $w \in \{0,1,2\}^*$

b) $\lambda(A) = 0$ if $A \cap X = \emptyset$

c) $\lambda(X) = 1$

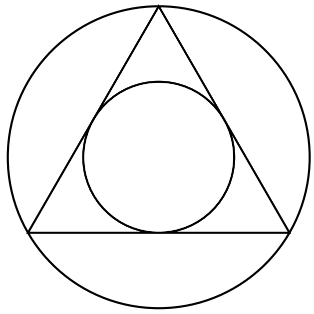
d) λ is a metric outer measure

Let μ be the Borel measure obtained by restricting λ to $\mathcal{B}_{\mathbb{R}^2}$ (Lemma 70 and Proposition 14).

Claim 2. *There exists $N > 0$ such that if $B \subset \mathbb{R}^2$ is a ball of radius $\delta \leqslant 1$, then $\mu(B) \leqslant N\delta^d$.*

Proof. If w has length k, then the triangle T_w contains a ball of radius $r(k) = 2^{-k-1}3^{-1/2}$ and is contained within a ball of radius $R(k) = 2^{-k}3^{-1/2}$.

[3] See, for example, Folland §11.2



Suppose that B intersects $M(k)$ level k triangles. Then the ball B', having the same centre as B but with radius $\delta + 2R(k)$, contains $M(k)$ *disjoint* smaller balls (each of radius $r(k)$).

Adding up the (Lebesgue) areas, we get that $M(k)r(k)^2 \leqslant (\delta + 2R(k))^2$ and hence

$$M(k) \leqslant (\delta + 2R(k))^2 r(k)^{-2} = 12(\delta 2^k + 2/\sqrt{3})^2$$

Define $N = 12(2 + 2/\sqrt{3})^2$. Now fix k such that $1/2^{k+1} \leqslant \delta \leqslant 1/2^k$. Note that $M(k) \leqslant N$ and

$$\mu(B) \leqslant 3^{-k} M(k) \leqslant \delta^d N$$

□

Claim 3. $H_d(X) > 0$

Proof. Let $\delta \leqslant 1$ and suppose that $\{A_i\}_i$ is a sequence of set $A_i \subset \mathbb{R}^2$ with $\operatorname{diam}(A_i) = \delta_i \leqslant \delta$ and $X \subset \cup_i A_i$. Each A_i is contained within a ball, B_i of diameter δ_i. Note that $\mu(B_i) \leqslant N\delta_i^d/2^d$ where N is as in the previous claim. Then

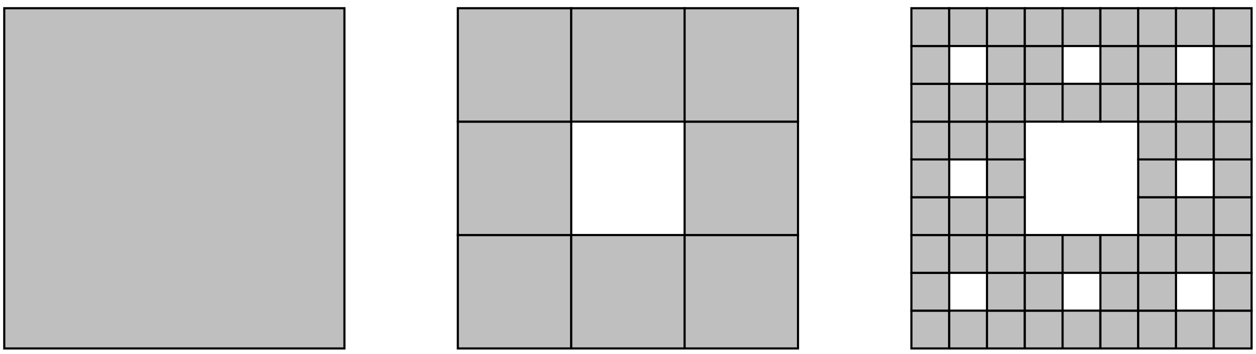
$$\sum_i \operatorname{diam}(A_i)^d = \sum_i \delta_i^d \geqslant \frac{2^d}{N}\sum \mu(B_i) \geqslant \frac{2^d}{N}\mu(\cup_i B_i) \geqslant \frac{2^d}{N}\mu(X) = \frac{2^d}{N}$$

It follows that $H_{d,\delta}(X) \geqslant 2^d/N$ and therefore $H_d(X) \geqslant 2^d/N > 0$ □

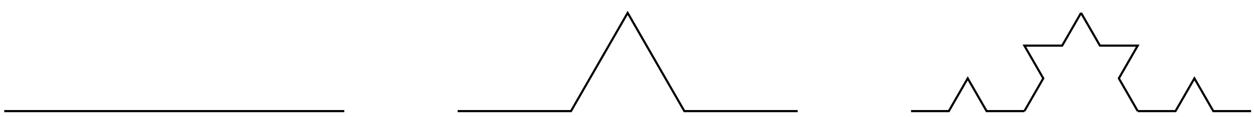
Exercise 52. Adapt the proof to show that the ternary Cantor set has Hausdorff dimension $\log_3 2$.

Examples 74.

The **Sierpinski carpet** has Hausdorff dimension $\log_3 8$.



The **snowflake** curve has Hausdorff dimension $\log_3 4$.



Lecture 20: LCH spaces

We will look at measures on locally compact Hausdorff (LCH) spaces and the Riesz Representation Theorem which relates measures and linear functionals on a certain space. Let's start by recalling some definitions.

Definition 75. A topological space (X, τ) is **locally compact** if every point has an open neighbourhood whose closure is compact. That is, $\forall x \in X \;\; \exists V \in \tau$ such that $x \in V$ and $\overline{V}$ (the closure of V) is compact.

A topological space (X, τ) is **Hausdorff** if for each distinct pair of points $x \neq y \in X$ there exist $V_x, V_y \in \tau$ such that $x \in V_x$, $y \in V_y$ and $V_x \cap V_y = \emptyset$.

The **support** of a continuous function $f : X \to \mathbb{R}$, denoted $\operatorname{supp}(f)$, is the closure of the set $\{x \in X \mid f(x) \neq 0\}$. Denote by $\mathcal{K}(X)$ the set of all compactly supported continuous functions on X.

Examples 76. Some examples of LCH spaces are: $\mathbb{R}^n$, compact metric spaces, any open subset of an LCH space. The set $\mathbb{Q} \subset \mathbb{R}$ endowed with the induced topology is *not* locally compact.

We will know establish some useful results about LCH spaces.

Exercise 53. Let X be a Hausdorff topological space and $K_1, K_2 \subset X$ disjoint compact subsets. Show that there are disjoint open subsets $V_1, V_2 \subset X$ such that $K_1 \subset V_1$ and $K_2 \subset V_2$. Hint: Consider first the case in which $K_1 = \{x\}$.

Exercise 54. Let X be a Hausdorff topological space and let $K \subset X$ be compact. Suppose that $V_1, V_2 \subset X$ are open sets and that $K \subset V_1 \cup V_2$. Show that there exist compact $K_1, K_2 \subset X$ such that $K_1 \subset V_1$, $K_2 \subset V_2$ and $K = K_1 \cup K_2$. Hint: apply previous result to $K \setminus V_i$.

Lemma 77. *Let X be an LCH space. Let $K, V \subset X$ be such that K is compact, V is open and $K \subset V$. There exists an open set $U \subset X$ such that $\overline{U}$ is compact and $K \subset U \subset \overline{U} \subset V$.*

Proof. Since X is locally compact, for all $k \in K$ there is an open set $V_k \ni k$ such that $\overline{V}_k$ is compact. Replacing V_k with $V_k \cap V$, we can assume that $V_k \subset V$. Since $\overline{V}_k \setminus V_k$ and $\{k\}$ are compact and disjoint, they can be separated by a pair of disjoint open sets, say $A_k \supset \overline{V}_k \setminus V_k$ and $B_k \supset \{k\}$. Then $W_k = B_k \cap V_k$ is an open set that contains k, $\overline{W}_k$ is compact and moreover $\overline{W_k} \subset V_k \subset V$.

Now, since K is covered by the family of open sets $\{W_k \mid k \in K\}$ and K is compact, there is a finite set $F \subset K$ such that $K \subset \cup_{k \in F} W_k$. Define $U = \cup_{k \in F} W_k$. Finally, note that $\overline{U} = \cup_{k \in F} \overline{W_k}$ is compact and contained in V. □

To state the following standard result we recall that a topological space is **normal** if it is Hausdorff and if every pair of disjoint closed subsets can be separated by a pair of disjoint open subsets.

Exercise 55. Show that every compact Hausdorff topological space is normal.

Theorem 78 (Urysohn's Lemma)**.** *Let X be a normal topological space and let A and B be disjoint closed subsets of X. There exists a continuous function $f : X \to [0, 1]$ such that $f|_A = 0$ and $f|_B = 1$.*

Proof. We describe a family of open sets $\{V_d \mid d \in D\}$, indexed by the set of dyadic rationals in the interval $(0, 1)$ and satisfying

$$A \subset V_r \subset \overline{V_r} \subset V_s \subset \overline{V_s} \subset B^c$$

whenever $r < s$. Given such a family, we define $f : X \to [0, 1]$ by

$$f(x) = \begin{cases} \inf\{r \mid x \in V_r\} & x \in \cup_{r \in D} V_r \\ 1 & \text{otherwise} \end{cases}$$

The function f is continuous since, for $a \in (0, 1)$, we have $f^{-1}([0, a)) = \cup_{r<a} V_r$ and $f^{-1}((a, 1]) = \cup_{r>a} (\overline{V}_r)^c$. That $f|_A = 0$ and $f|_B = 1$ is clear from the definition of f.

It remains to construct the sets V_r. Since X is normal there exists an open set $V_{1/2}$ such that $A \subset V_{1/2} \subset \overline{V}_{1/2} \subset B^c$. Similarly, from normality applied to the closed sets A and $V_{1/2}^c$ we have an open set $V_{1/4}$ such that $A \subset V_{1/4} \subset \overline{V}_{1/4} \subset V_{1/2}$. Considering the closed sets $\overline{V}_{1/2}$ and B, we get an open set $V_{3/4}$ such that $V_{1/2} \subset V_{3/4} \subset \overline{V}_{3/4} \subset B^c$. Continue inductively to define V_r for all r. □

We first list two consequences of Urysohn's lemma.

Proposition 79. *Let X be an LCH space. Let $K, V \subset X$ be subsets of X such that $K \subset V$, K is compact and V is open. There exists $f \in \mathcal{K}(X)$ such that $\mathbb{1}_K \leqslant f \leqslant \mathbb{1}_V$ and* $\operatorname{supp}(f) \subset V$.

Proof. By the above lemma 77, there is a open set U such that $\overline{U}$ is compact and $K \subset U \subset \overline{U} \subset V$. By Urysohn's lemma applied to the space $\overline{U}$, there is a continuous function $g : \overline{U} \to [0,1]$ such that $g|_K = 1$ and $g|_{\overline{U}\setminus U} = 0$. Define $f : X \to [0,1]$ by $f(x) = g(x)$ for $x \in \overline{U}$ and $f(x) = 0$ if $x \notin \overline{U}$. The function f is continuous because for a closed $F \subset [0,1]$ we have that $f^{-1}(F)$ is closed:

$$f^{-1}(F) = \begin{cases} g^{-1}(F) & 0 \notin F \\ g^{-1}(F) \cup (\overline{U})^c = g^{-1}(F) \cup U^c & 0 \in F \end{cases}$$

$\square$

Lecture 21: Regular measures

Proposition 80. *Let X be a LCH space and let $f \in \mathcal{K}(X)$. Suppose that $\{V_i\}_{i=1}^n$ is an open cover of* $\operatorname{supp}(f)$. *There exist $f_i \in \mathcal{K}(X)$ such that* $\operatorname{supp}(f_i) \subset V_i$ *and $f = f_1 + \cdots + f_n$.*

Proof. It is enough to establish the case in which $n = 2$: $\operatorname{supp}(f) \subset V_1 \cup V_2$. By Exercise 54 above there exist compact $K_i \subset V_i$ such that $\operatorname{supp}(f) = K_1 \cup K_2$. By Proposition 79 there are $g_i \in \mathcal{K}(X)$ satisfying $\mathbb{1}_{K_i} \leqslant g_i \leqslant \mathbb{1}_{V_i}$ and $\operatorname{supp}(g_i) \subset V_i$. Define $g_3 = g_2 - \min(g_1, g_2)$ and note that $\operatorname{supp}(g_3) \subset \operatorname{supp}(g_2) \subset V_2$. If $x \in \operatorname{supp}(f)$, then $g_1(x) + g_3(x) = g_1(x) + g_2(x) - \min(g_1(x), g_2(x)) = \max(g_1(x), g_2(x)) = 1$. The functions $f_1 = fg_1$ and $f_2 = fg_3$ have the required properties. □

Definition 81. Let X be a Hausdorff topological space. A measure μ on $(X, \mathcal{A})$ is **regular** if $\mathcal{A} \supset \mathcal{B}_X$ and

1) $\mu(K) < \infty$ for all compact K
2) μ is *outer regular*: $\mu(A) = \inf\{\mu(V) \mid A \subset V,\, V \text{ open}\}$
3) μ is *inner regular on open sets*: $\mu(V) = \sup\{\mu(K) \mid K \subset V,\, K \text{ compact}\}$

Examples 82. Lebesgue measure on $\mathbb{R}$ is a regular. Every finite Borel measure on $\mathbb{R}$ is regular. (Lecture 6)

We are going to relate regular Borel measures on X to linear functionals on $\mathcal{K}(X)$. If μ is a regular measure on X, then $f \mapsto \int f\, d\mu$ is a linear functional on $\mathcal{K}(X)$.

Definition 83. A linear functional Λ on $\mathcal{K}(X)$ is **positive** is $\Lambda(f) \geqslant 0$ whenever $f \geqslant 0$.

Exercise 56. Show that if Λ is positive and $f \leqslant g$, then $\Lambda(f) \leqslant \Lambda(g)$.

For $V \subset X$ open and $f \in \mathcal{K}(X)$, the notation $f \prec V$ indicates that $0 \leqslant f \leqslant \mathbb{1}_V$ and $\operatorname{supp}(f) \subset V$.

Lemma 84. *Let X be an LCH space and μ a regular Borel measure on X. Then, for $V \subset X$ open we have*

$$\mu(V) = \sup\{\int f\, d\mu \mid f \in \mathcal{K}(X), f \prec V\}$$

Proof. It's clear that $\mu(V) \geqslant \sup\{\int f\, d\mu \mid f \in \mathcal{K}(X), f \prec V\}$ by inner regularity. For the reverse inequality, suppose that $\alpha < \mu(V)$. By inner regularity, there is a compact $K \subset V$ such that $\mu(K) > \alpha$. By Proposition 79 there is $f \in \mathcal{K}(X)$ with $\mathbb{1}_K \leqslant f$ and $f \prec V$. We have $\alpha < \mu(K) < \int f\, d\mu$ and therefore

$$\alpha \leqslant \sup\{\int f\, d\mu \mid f \in \mathcal{K}(X), f \prec V\}$$

□

We're now ready to prove the representation theorem.

Theorem 85. *Let X be an LCH space and let Λ be a positive linear functional on $\mathcal{K}(X)$. Then there exists a unique regular Borel measure μ on X such that for all $f \in \mathcal{K}(X)$*

$$\Lambda(f) = \int f\, d\mu$$

Proof. We start with the uniqueness. Suppose that μ and ν are regular measures with $\Lambda(f) = \int f\, d\mu = \int f\, d\nu$ for all $f \in \mathcal{K}(X)$. By Lemma 84 we have $\mu(V) = \nu(V)$ for all open $V \subset X$. By outer regularity it follows that $\mu(A) = \nu(A)$ for all Borel $A \subset X$.

Now we turn to the existence part of the statement. The idea is to define a function on open sets first (motivated by Lemma 84), then extend to an outer measure, then show that Borel sets are measurable and finally that the resulting measure is regular. For an open set $V \subset X$ define

$$\rho(V) = \sup\{\Lambda(f) \mid f \in \mathcal{K}(X), f \prec V\} \tag{$*$}$$

Now define $\lambda : P(X) \to [0, \infty]$ by

$$\lambda(A) = \inf\{\rho(V) \mid V \supset A, V \text{ open}\} \tag{$\dagger$}$$

Note that $\lambda(A) = \rho(A)$ for open A, since if $A \subset V$ then $\rho(A) \leqslant \rho(V)$.

Lecture 22: Riesz representation theorem

Continuing with the proof from last lecture.

Claim 1. *λ is an outer measure.*

It's clear that $\lambda(\emptyset) = 0$ and that $\lambda(A) \leqslant \lambda(B)$ whenever $A \subset B$.

We first establish countable subadditivity for open subsets $\{V_i\}_{i\in\mathbb{N}}$. Let $V = \cup_i V_i$. If $f \in \mathcal{K}(X)$ satisfies $f \prec V$, then $\mathrm{supp}(f)$ is a compact subset of $\cup_i V_i$. Therefore there exists $n \in \mathbb{N}$ such that $\mathrm{supp}(f) \subset \cup_{i=1}^n V_i$. By Proposition 80 there exist $f_i \in \mathcal{K}(X)$ such that $\mathrm{supp}(f_i) \subset V_i$ and $f = f_1 + \cdots + f_n$. We have

$$\Lambda(f) = \sum_{i=1}^n \Lambda(f_i) \leqslant \sum_{i=1}^n \lambda(V_i) \leqslant \sum_{i\in\mathbb{N}} \lambda(V_i)$$

Since the above holds for all $f \in \mathcal{K}(X)$ with $f \prec V$, we conclude

$$\lambda(\cup_{i\in\mathbb{N}} V_i) = \rho(V) \leqslant \sum_{i\in\mathbb{N}} \lambda(V_i)$$

Now for any sequence of sets $A_i \subset X$ such that $\sum_i \lambda(A_i) < \infty$ and for any $\epsilon > 0$, let V_i be an open set such that $A_i \subset V_i$ and $\lambda(V_i) < \lambda(A_i) + \epsilon 2^{-i}$. Then

$$\lambda(\cup_i A_i) \leqslant \lambda(\cup_i V_i) \leqslant \sum_i \lambda(V_i) \leqslant \epsilon + \sum_i \lambda(A_i)$$

It follows that λ is countably subadditive and hence an outer measure.

Claim 2. *Every Borel subset of X is λ-measurable.*

Since the λ-measurable sets form a σ-algebra, it's enough to show that open sets are λ-measurable. Let $U, V \subset X$ be open. Given $\epsilon > 0$ there exists $f \in \mathcal{K}(X)$ with $f \prec U \cap V$ and $\Lambda(f) > \lambda(U \cap V) - \epsilon$. Similarly, since $U \cap (\mathrm{supp}(f))^c$ is open, there exists $g \in \mathcal{K}(X)$ with $g \prec U \cap (\mathrm{supp}(f))^c$ and $\Lambda(g) > \lambda(U \cap (\mathrm{supp}(f))^c) - \epsilon$. We have that $f + g \prec U$ and

$$\begin{aligned}
\lambda(U) &\geqslant \Lambda(f+g) \qquad \text{(from definition of } \rho) \\
&= \Lambda(f) + \Lambda(g) \\
&> \lambda(U \cap V) + \lambda(U \cap (\mathrm{supp}(f))^c) - 2\epsilon \\
&\geqslant \lambda(U \cap V) + \lambda(U \cap V^c) - 2\epsilon \qquad \text{(monotonicity of } \lambda)
\end{aligned}$$

Since this holds for all $\epsilon > 0$, we conclude that $\lambda(U) \geqslant \lambda(U \cap V) + \lambda(U \cap V^c)$. Now, to establish that V is λ-measurable. Let $A \subset X$ with $\lambda(A) < \infty$. There exists an open $U \supset A$ with $\lambda(U) < \lambda(A) + \epsilon$. Therefore,

$$\begin{aligned}
\lambda(A) &> \lambda(U) - \epsilon \\
&\geqslant \lambda(U \cap V) + \lambda(U \cap V^c) - \epsilon \qquad \text{(from above)} \\
&\geqslant \lambda(A \cap V) + \lambda(A \cap V^c) - \epsilon \qquad \text{(monotonicity)}
\end{aligned}$$

Since this holds for all $\epsilon > 0$, we have $\lambda(A) \geqslant \lambda(A \cap V) + \lambda(A \cap V^c)$ and hence V is λ-measurable. This establishes that the Borel sets are λ-measurable.

Claim 3. *Let $A \subset X$ and $f \in \mathcal{K}(X)$. If $\mathbb{1}_A \leqslant f$, then $\lambda(A) \leqslant \Lambda(f)$.*

Given $\epsilon \in (0,1)$, define $V_\epsilon = \{x \in X \mid f(x) > 1 - \epsilon\}$. Then $V_\epsilon \supset A$ and, since f is continuous, V_ϵ is open. If $g \in \mathcal{K}(X)$ and $g \leqslant \mathbb{1}_{V_x}$, then

$$\begin{aligned}
g \leqslant f/(1-\epsilon) &\implies \Lambda(g) \leqslant \frac{1}{1-\epsilon}\Lambda(f) \\
&\implies \lambda(V_\epsilon) \leqslant \frac{1}{1-\epsilon}\Lambda(f) \qquad \text{(by } (*))
\end{aligned}$$

Therefore $\lambda(A) \leqslant \frac{1}{1-\epsilon}\Lambda(f)$ for any $\epsilon \in (0,1)$, and it follows that $\lambda(A) \leqslant \Lambda(f)$.

Claim 4. *Let $K \subset X$ be compact and $f \in \mathcal{K}(X)$. If $0 \leqslant f \leqslant \mathbb{1}_K$, then $\Lambda(f) \leqslant \lambda(K)$.*

For any open set V with $V \supset K$, we have $f \prec V$ and therefore $\lambda(V) \geqslant \Lambda(f)$ by $(*)$. Since this holds for any such V, from $(\dagger)$ we have that $\lambda(K) \geqslant \Lambda(f)$.

Now let μ denote the measure obtained by restricting λ to $\mathcal{B}_X$.

Claim 5. *The measure μ is regular.*

Let $K \subset X$ be compact. By Proposition 79, there exists $f \in \mathcal{K}(X)$ with $\mathbb{1}_K \leqslant f$. By Claim 3 above, $\lambda(K) \leqslant \Lambda(f) < \infty$. Therefore μ is finite on compact sets. The outer regularity of μ follows immediately from the way in which λ was defined in $(\dagger)$. For inner regularity (on open sets), suppose that $V \subset X$ is open and $\mu(V) > 0$.
Then

$$
\begin{aligned}
\mu(V) = \lambda(V) &= \sup\{\Lambda(f) \mid f \in \mathcal{K}(X), f \prec V\} \qquad \text{(by } (*)\text{)} \\
&\leqslant \sup\{\lambda(\mathrm{supp}(f)) \mid f \in \mathcal{K}(X), f \prec V\} \qquad \text{(by Claim 4)} \\
&\leqslant \sup\{\lambda(K) \mid K \subset V,\ K \text{ compact}\}
\end{aligned}
$$

Let $0 < \alpha < \mu(V)$. By $(*)$ there exists $f \in \mathcal{K}(X)$ such that $f \prec V$ and $\Lambda(f) > \alpha$. Then $\lambda(\mathrm{supp}(f)) \geqslant \Lambda(f) > \alpha$. It follows that $\mu(V) = \sup\{\lambda(K) \mid K \subset V,\ K \text{ compact}\}$ and the claim is established.

Lecture 23: Riesz representation theorem (ctd)

The proof of the theorem will be complete once we have shown the following claim.

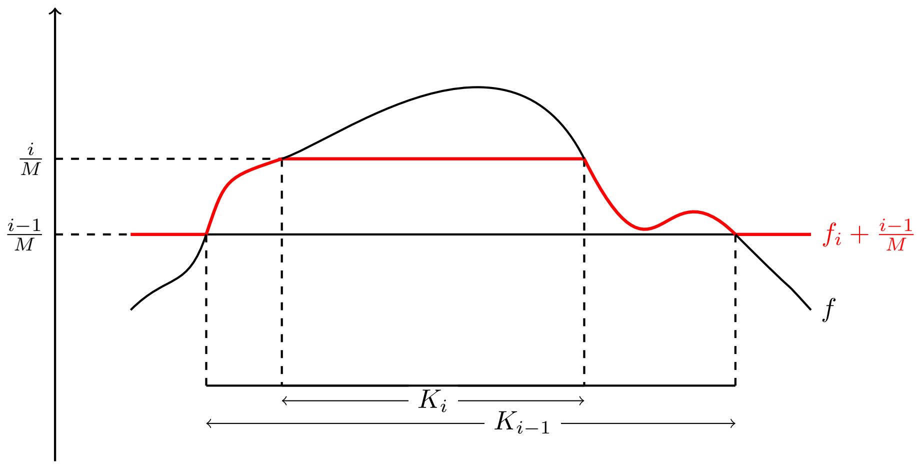
Claim 6. *For all $f \in \mathcal{K}(X)$ we have $\int f\,d\mu = \Lambda(f)$*

It's enough to establish the claim in the case in which f takes values in $[0,1]$. Fix $M \in \mathbb{N}$ and define a decreasing sequence of subsets by

$$K_0 = \operatorname{supp}(f) \qquad \text{and} \qquad K_i = \{x \in X \mid f(x) \geqslant i/M\} \qquad \text{for } i \geqslant 1$$

For $1 \leqslant i \leqslant M$ let $f_i \in \mathcal{K}(X)$ be defined by

$$f_i(x) = \begin{cases} 0 & x \notin K_{i-1} \\ f(x) - \frac{i-1}{M} & x \in K_{i-1} \setminus K_i \\ \frac{1}{M} & x \in K_i \end{cases}$$



Then we have $f = \sum_{i=1}^{M} f_i$ and

$$\frac{1}{M}\mathbb{1}_{K_i} \leqslant f_i \leqslant \frac{1}{M}\mathbb{1}_{K_{i-1}} \implies \frac{1}{M}\mu(K_i) \leqslant \int f_i\,d\mu \leqslant \frac{1}{M}\mu(K_{i-1})$$

So we have

$$\frac{1}{M}\sum_{i=1}^{M}\mu(K_i) \leqslant \int f\,d\mu \leqslant \frac{1}{M}\sum_{i=1}^{M}\mu(K_{i-1}) \tag{‡}$$

Now note that for any open set $V \supset K_{i-1}$ we have $Mf_i \prec V$ and hence $\mu(V) \geqslant M\Lambda(f_i)$ by $(*)$. Since this holds for any such V, by outer regularity we have that $\mu(K_{i-1}) \geqslant M\Lambda(f_i)$. By Claim 3 we also have $\mu(K_i) \leqslant \Lambda(Mf_i)$. Therefore

$$\frac{1}{M}\sum_{i=1}^{M}\mu(K_i) \leqslant \Lambda(f) \leqslant \frac{1}{M}\sum_{i=1}^{M}\mu(K_{i-1}) \tag{$\star$}$$

Combining (‡) and ($\star$) we get

$$\begin{aligned} |\Lambda(f) - \int f\,d\mu| &\leqslant \frac{1}{M}(\mu(K_0) - \mu(K_M)) \\ &\leqslant \frac{1}{M}(\mu(\operatorname{supp}(f))) \end{aligned}$$

Since this holds for all M and since μ is finite on compact sets, we deduce that $\int f\,d\mu = \Lambda(f)$. □

Exercise 57. Show that for all compact $K \subset X$ we have $\mu(K) = \inf\{\Lambda(f) \mid f \in \mathcal{K}(X), f \geqslant \mathbb{1}_K\}$.

Exercise 58. Let X be LCH. Let $\mathcal{A}$ be a σ-algebra on X with $\mathcal{A} \supset \mathcal{B}_X$. Suppose that μ is a regular measure on $(X, \mathcal{A})$. Show that the completion of μ is regular.

Exercise 59. Let X be an LCH space, Y a closed subset of X and ν a regular Borel measure on Y. Let Λ be the positive linear functional on $\mathcal{K}(X)$ given by $\Lambda(f) = \int f|_Y \, d\nu$. Show that the regular measure on X induced by Λ (as in the theorem) is given by $\mu(A) = \nu(A \cap Y)$.

Exercise 60. Let X be an LCH space and μ a a regular Borel measure on X. Show that μ is inner regular on all σ-finite (Borel) sets.

Exercise 61. Let μ be a σ-finite regular Borel measure on and LCH space X and let $A \in \mathcal{B}_X$. Show that $\mu_A(B) = \mu(B \cap A)$ defines a regular Borel measure on X.

Exercise 62. Let $\mathbb{R}_d$ denote $\mathbb{R}$ endowed with the discrete topology, and let $X = \mathbb{R} \times \mathbb{R}_d$.

a) Let $f : X \to \mathbb{R}$ be a function. Show that $f \in \mathcal{K}(X)$ if and only if $f^y \in \mathcal{K}(\mathbb{R})$ for all y and $f^y = 0$ for all but finitely many y.

b) Define a positive linear functional on $\mathcal{K}(X)$ by $\Lambda(f) = \sum_{y\in\mathbb{R}} \int f(x, y)\, dx$ and let μ the associated regular Borel measure on X. Show that $\mu(A) = \infty$ for any $A \subset X$ such that $A^y \neq \emptyset$ for uncountably many y.

c) Let $A = \{0\} \times \mathbb{R}_d$. Show that $\mu(A) = \infty$ (just previous part!) but $\mu(K) = 0$ for all compact $K \subset A$. (So μ is not inner regular on A.)

d) Let $A = (\mathbb{R} \setminus \{0\}) \times \mathbb{R}_d$. Show that the measure given by $\mu_A(B) = \mu(A \cap B)$ is not a regular Borel measure on X.

Dual of $C_0(X)$

We now look at regular signed measures. The class of functions that arise, $C_0(X)$ is slightly larger than $\mathcal{K}(X)$. We will show that the Banach space $M_r(X)$ of finite signed regular measures on X is isometrically isomorphic to the dual of the Banach space $C_0(X)$. Note that these results can be extended to complex valued measures.

Definition 86. A function $f : X \to \mathbb{R}$ **vanishes at infinity** if for all $\epsilon > 0$ the set $\{x \mid |f(x)| \geqslant \epsilon\}$ is compact. Denote by $C_0(X)$ the set of all such functions:

$$C_0(X) = \{f \in C(X) \mid f \text{ vanishes at infinity}\}$$

Since elements of $C_0(X)$ are bounded, $\|f\|_\infty = \sup\{|f(x)| \mid x \in X\}$ defines a norm on $C_0(X)$.

Exercise 63. Let X be an LCH space. Show that $C_0(X)$ is the closure of $\mathcal{K}(X)$ in the uniform norm. Show that $C_0(X)$ is a Banach space

It follows that if μ is a regular Borel measure on X, then the associated positive linear functional on $\mathcal{K}(X)$ extends continuously to $C_0(X)$ iff it is bounded with respect to the uniform norm. Since $\mu(X) = \sup\{\int f\, d\mu \mid f \in \mathcal{K}(X), 0 \leqslant f \leqslant 1\}$, this happens when $\mu(X) < \infty$. So positive bounded linear functionals on $C_0(X)$ are given by integration with respect to a finite regular Borel measure.

Definition 87. A finite signed measure ν on an LCH space is called **regular** if $|\nu|$ is regular. Denote by $M_r(X)$ the set of all regular finite signed Borel measures on X.

Exercise 64. Let ν be a finite signed measure on $(X, \mathcal{B}_X)$. Show that the following are equivalent:

a) ν is regular,

b) ν^+ and ν^- are both regular,

c) ν is a linear combination of finite regular Borel (positive) measures.

Exercise 65. Show that $M_r(X)$ is a closed subspace of $M(X)$ (with the total variation norm on $M(X)$).

Lemma 88. *Let X be an LCH space and $\nu \in M_r(X)$. Then*

$$\forall A \in \mathcal{B}_X\ \forall \epsilon > 0\ \exists K \subset A \text{ compact such that } |\nu(A) - \nu(B)| < \epsilon \text{ whenever } B \in \mathcal{B}_X \text{ with } K \subset B \subset A$$

Proof. Since $|\nu|$ is regular and $|\nu|(A) < \infty$, there is a compact $K \subset A$ such that $|\nu|(A \setminus K) < \epsilon$.

Then for any $B \in \mathcal{B}_X$ with $K \subset B \subset A$ we have

$$|\nu(A) - \nu(B)| = |\nu(A \setminus B)| \leqslant |\nu|(A \setminus B) \leqslant |\nu|(A \setminus K) < \epsilon$$

□

Continuous linear functionals on $C_0(X)$ admit a version of a Jordan decomposition.

Lemma 89. *Let X be an LCH space. If $\Lambda \in C_0(X)^*$, there exist positive functionals $\Lambda^+, \Lambda^- \in C_0(X)^*$ such that $\Lambda = \Lambda^+ - \Lambda^-$.*

Proof. Given $f \in C_0(X)$ with $f \geqslant 0$ define $\Lambda^+(f) = \sup\{\Lambda(g) \mid g \in C_0(X), 0 \leqslant g \leqslant f\}$. Since $|\Lambda(g)| \leqslant \|\Lambda\|\|g\|_\infty \leqslant \|\Lambda\|\|f\|_\infty$, the supremum is finite and $|\Lambda^+(f)| \leqslant \|\Lambda\|\|f\|_\infty$.

Exercise 66. Show that $0 \leqslant \Lambda^+(f)$, $\Lambda^+(tf) = t\Lambda^+(f)$ and $\Lambda^+(f_1 + f_2) = \Lambda^+(f_1) + \Lambda^+(f_2)$ (for all $t \geqslant 0$ and $f_1, f_2 \geqslant 0$).

Now extend to any $f \in C_0(X)$ by defining $\Lambda^+(f) = \Lambda^+(f^+) - \Lambda^+(f^-)$

Exercise 67. Show that Λ^+ is a positive linear functional on $C_0(X)$.

We have

$$|\Lambda^+(f)| \leqslant \max\{\Lambda^+(f^+), \Lambda^+(f^-)\} \leqslant \|\Lambda\| \max\{\|f^+\|_\infty, \|f^-\|_\infty\} = \|\Lambda\|\|f\|_\infty$$

and therefore $\|\Lambda^+\| \leqslant \|\Lambda\|$. Define $\Lambda^- = \Lambda^+ - \Lambda$. Then Λ^- is clearly linear and continuous. It is positive since if $f \geqslant 0$, then $\Lambda^+(f) \geqslant \Lambda(f)$ (by the definition of Λ^+). □

Lecture 24: Dual of $C_0(X)$

Theorem 90. *Let X be an LCH space. The map $M_r(X) \to C_0(X)^*$ that sends μ to the linear functional $f \mapsto \int f\,d\mu$ is an isometric isomorphism.*

Proof. Given $\nu \in M_r(X)$ define $\Lambda_\nu \in C_0(X)^*$ by $\Lambda_\nu(f) = \int f\,d\nu$. It's readily checked that Λ_ν is indeed linear, that $|\Lambda_\nu(f)| \leqslant \|f\|_\infty \|\nu\|$ and that the map $\Phi : \nu \mapsto \Lambda_\nu$ is linear.

So we have a linear map $\Phi : M_r(X) \to C_0(X)^*$ given by $\Phi(\nu) = \Lambda_\nu$ satisfying $\|\Phi(\nu)\| \leqslant \|\nu\|$. We want to show that Φ is norm preserving and surjective.

To show that Φ is norm preserving, let $\nu \in M_r(X)$ and $\epsilon > 0$. Let $X = P \cup N$ be a Hahn decomposition corresponding to ν. By Lemma 88 there are compact $K_P \subset P$ and $K_N \subset N$ such that

$$\|\nu\| - \epsilon < |\nu(K_P)| + |\nu(K_N)| \leqslant |\nu|(K_P) + |\nu|(K_N)$$

Let $f \in \mathcal{K}(X)$ be such that $\|f\|_\infty \leqslant 1$ and $f|_{K_P} = 1$ and $f|_{K_N} = -1$. Then (with $K = K_P \cup K_N$)

$$\int_K f\,d\nu = |\,\nu(K_P)| + |\,\nu(K_N)| > \|\nu\| - \epsilon \quad \text{and} \quad \left|\int_{K^c} f\,d\nu\right| \leqslant |\nu|(K^c) < \epsilon$$

Therefore $|\int f\,d\nu| > \|\nu\| - 2\epsilon$. Since $\|f\|_\infty \leqslant 1$ and $\epsilon > 0$ was arbitrary, it follows that $\|\Phi(\nu)\| \geqslant \|\nu\|$. So Φ is norm preserving.

Now to show that Φ is surjective. Suppose first that $\Lambda \in C_0(X)^*$ is positive. Applying the Riesz representation theorem to $\Lambda|_{\mathcal{K}(X)}$ provides a regular Borel measure μ with $\Lambda(f) = \int f\,d\mu$ for all $f \in \mathcal{K}(X)$. By Lemma 84 we have

$$\mu(X) = \sup\{\int f\,d\mu \mid f \in \mathcal{K}(X), f \prec X\} = \sup\{\Lambda(f) \mid f \in \mathcal{K}(X), 0 \leqslant f \leqslant 1\} \leqslant \|\Lambda\|$$

We have a finite measure μ such that $\Lambda(f) = \Phi(\mu)(f)$ for all $f \in \mathcal{K}(X)$. Because $\mathcal{K}(X)$ is dense in $C_0(X)$ and Λ and Φ are continuous, the equality holds for all $f \in C_0(X)$. The surjectivity of Φ follows from this and the preceding lemma.

□

Let's note the following special case.

Corollary 91. *Let X be a compact metric space. Then $M(X)$ is isometrically isomorphic to $C(X)^*$.* □

Topological groups

We're going to look at measures on topological groups. We'll see that any locally compact group admits a measure that is invariant under the action of the group in itself.

Definition 92. A **topological group** is a group G endowed with a topology such that the map $G \times G \to G$, given by $(g, h) \mapsto gh$ and the map $G \to G$ given by $g \mapsto g^{-1}$ are continuous. A **locally compact group** is a topological group that is locally compact and Hausdorff.

Examples 93.

1) $(\mathbb{R}, +)$ and $(\mathbb{R} \setminus \{0\}, \times)$ are locally compact groups.
2) $(\mathbb{Q}, +)$ is a topological group, but not locally compact.
3) $\{z \in \mathbb{C} \mid |z| = 1\}$ is a locally compact group.
4) Any topological vector space is a topological group (underlying abelian group).
5) Any group endowed with the discrete topology is a locally compact group.

Exercise 68. Show that $G = \{\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right] \in GL_2(\mathbb{R}) \mid a > 0\}$ is a locally compact group.

Lemma 94. *Let G be a topological group and $g \in G$.*

1) *The functions* $h \mapsto gh$ *and* $h \mapsto hg$ *are homeomorpishms from* G *to* G.

2) *If* $K, L \subset G$ *are compact, then so too are* gK, Kg, KL *and* K^{-1}.

Proof. The map $h \mapsto gh$ is continuos as it is the composition of two continuous maps: $G \to G \times G$ given by $h \mapsto (g, h)$ and $G \times G \to G$ given by $(k, h) \mapsto kh$. It has a continuous inverse given by the map $h \mapsto g^{-1}h$, and is therefore a homeomorphism. Similarly for the second listed map in the first part.

Since the image of a compact set under a continuous map is compact, the sets gK, Kg and K^{-1} are compact. Since $K \times L$ is a compact subset of $G \times G$, KL is compact.

□

Lecture 25: Topological groups

Lemma 95. *Let G be a topological group and $V \subset G$ on open subset such that $1_G \in V$.*

1) *There exists an open $U \subset V$ such that $1_G \in U$ and $UU \subset V$.*

2) *There exists an open $U \subset V$ such that $1_G \in U$ and $U = U^{-1}$.*

Proof. The set $W = \{(g,h) \mid gh \in V\}$ is an open neighbourhood of $(1,1)$ in $G \times G$. Therefore there are open neighbourhoods U_1, U_2 of 1 in G such that $U_1 \times U_2 \subset W$. The set $U = U_1 \cap U_2 \subset G$ is an open, contains 1_G and satisfies $UU \subset V$.

For the second part, note that U^{-1} is an open neighbourhood of 1, and define $S = U \cap U^{-1}$. □

Lemma 96. *Let G be a topological group. Every open subgroup of G is also closed.*

Proof. Let $H \leqslant G$ be open. Then the complement of H is a union of cosets of H

$$H^c = \bigcup_{g \in H^c} gH$$

Because each coset is open (Lemma 94) , H^c is open. □

Definition 97. A function $f : G \to \mathbb{R}$ is **left uniformly continuous** if for all $\epsilon > 0$ there exists an open neighbourhood $V \ni 1$ such that $|f(g) - f(h)| < \epsilon$ whenever $h \in gV$. The function f is **right uniformly continuous** is the same condition holds with Vg in place of gV.

Exercise 69. Consider the locally compact group $G = \{\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right] \in GL_2(\mathbb{R}) \mid a > 0\}$. Construct a function $f : G \to \mathbb{R}$ that is right uniformly continuous, but not left uniformly continuous.

Proposition 98. *Let G be a locally compact group. Every function in $\mathcal{K}(G)$ is both left uniformly continuous and right uniformly continuous.*

Proof. Let $f \in \mathcal{K}(G)$ and $K = \operatorname{supp}(f)$ and let $\epsilon > 0$. Since f is continuous, for all $x \in K$ there is an open $V_x \ni 1$ such that $|f(x) - f(y)| < \epsilon/2$ whenever $y \in xV_x$. By Lemma 95 there is an open $U_x \subset V_x$ such that $1 \in U_x$ and $U_x U_x \subset V_x$. The set $\{xU_x \mid x \in K\}$ is an open cover of K. Because K is compact there exists a finite subset $\{x_1, \dots, x_n\} \subset K$ such that $K \subset x_1 U_{x_1} \cup \dots \cup x_n U_{x_n}$. Since $U_{x_1} \cap \dots \cap U_{x_n}$ is an open neighbourhood of 1, by Lemma 95 there is an open $V \subset U_{x_1} \cap \dots \cap U_{x_n}$ such that $1 \in V$ and $V = V^{-1}$. We will show that for all $h, g \in G$ we have

$$h \in gV \implies |f(g) - f(h)| < \epsilon \tag{$*$}$$

Note first that $(*)$ clearly holds if $h, g \in K^c$. So suppose now that $g \in K$ and $h \in gV$. Then $g \in x_i U_{x_i} \subset x_i V_{x_i}$ for some $i \in \{1, \dots, n\}$ and $h \in gV \subset gU_{x_i} \subset x_i U_{x_i} U_{x_i} \subset x_i V_{x_i}$. That is, there is an i such that $h, g \in x_i V_{x_i}$. Therefore $|f(x_i) - f(g)| < \epsilon/2$ and $|f(x_i) - f(h)| < \epsilon/2$ and therefore $|f(g) - f(h)| < \epsilon$.

Now suppose that we have $h \in K$ and $h \in gV$. Because V is symmetric, $h \in gV$ is equivalent to $g \in hV$. We can therefore apply the argument of the previous paragraph. The left uniform continuity of f is shown.

The argument for right uniform continuity is similar. □

Corollary 99. *Let G be a locally compact group, let μ be a regular Borel measure on G, and let $f \in \mathcal{K}(G)$. The functions $g \mapsto \int f(gh)\, d\mu(h)$ and $g \mapsto \int f(hg)\, d\mu(h)$ are continuous.*

Proof. Let $g_0 \in G$ and $V \subset G$ an open neighbourhood of g_0 such that $\overline{V}$ is compact. Let $K = \operatorname{supp}(f)$. For each $g \in V$ the function $h \mapsto f(hg)$ is continuous and has support contained within the compact set $K(\overline{V})^{-1}$.

Let $\epsilon > 0$. Choose $\epsilon' > 0$ such that $\epsilon' \mu(K(\overline{V})^{-1}) < \epsilon$. By preceding proposition f is left uniformly continuous and hence there is an open neighbourhood U of 1 such that $|f(a) - f(b)| < \epsilon'$ whenever $a, b \in G$ satisfy $a \in bU$. Then for $g \in V \cap g_0 U$ and $h \in G$ we have $hg \in hg_0 U$ and therefore

$$\begin{aligned} \left| \int f(hg)\, d\mu(h) - \int f(hg_0)\, d\mu(h) \right| &\leqslant \int |f(hg) - f(hg_0)|\, d\mu(h) \\ &\leqslant \epsilon' \mu(K(\overline{V})^{-1}) \\ &\leqslant \epsilon \end{aligned}$$

Therefore the function $g \mapsto \int f(hg)\, d\mu(h)$ is continuous at the point $g_0 \in G$.
The argument for the continuity of the other function in the statement is entirely similar. $\square$

Lecture 26: Haar measure

Definition 100. Let G be a locally compact group. A Borel measure μ on G is **left-invariant** if $\mu(gA) = \mu(A)$ for all $g \in G$ and $A \in \mathcal{B}_G$. The measure μ is **right-invariant** if $\mu(Ag) = \mu(A)$ for all $g \in G$ and $A \in \mathcal{B}_G$. A **left Haar measure** on G is a non-zero regular Borel measure on G that is left-invariant. A **right Haar measure** on G is a non-zero regular Borel measure on G that is right-invariant.

Examples 101.

1) Lesbesgue measure on $\mathbb{R}^n$ is both a left (and right) Haar measure.
2) Counting measure on any group endowed with the discrete topology.
3) $G = (\mathbb{R}_{>0}, \times)$, $\mu(A) = \int_A \frac{1}{x} dx$ is a left (and right) Haar measure.
4) $G = \{\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right] \in GL_2(\mathbb{R}) \mid a > 0\}$, $\mu(A) = \int_{A'} x^{-2} dm^2$ (where $A' \subset \mathbb{R}^2$ corresponds to A, and m^2 is Lebesgue measure on $\mathbb{R}^2$) gives a left Haar measure on G.

Before considering an existence statement for Haar measures, let's note the following property.

Lemma 102. *Let G be a locally compact group and μ a left Haar measure on G. Then*

1) $\mu(V) > 0$ for every non-empty open $V \subset G$,

2) $\int f \, d\mu > 0$ for every $f \in \mathcal{K}(G)$ with $f \geqslant 0$, $f \neq 0$.

Proof. Because μ is not the zero measure and is regular , there exists a compact $K \subset G$ such that $\mu(K) > 0$. Let $V \subset G$ be non-empty and open. Then, by compactness of K, there exist $g_1, \ldots, g_n \in G$ such that $K \subset \cup_{i=1}^n g_i V$. Therefore

$$0 < \mu(K) \leqslant \mu(\cup_{i=1}^n g_i V) \leqslant \sum_{i=1}^n \mu(g_i V) = \sum_{i=1}^n \mu(V) = n\mu(V)$$

For the second part, the conditions on f imply the existence of $\epsilon > 0$ and open $V \subset G$ such that $f|_V \geqslant \epsilon$. Therefore

$$\int f \, d\mu \geqslant \int \epsilon \mathbb{1}_V \, d\mu = \epsilon\mu(V) > 0$$

□

Theorem 103. *Let G be a locally compact group. There exists a left Haar measure on G.*

Proof. Denote by $\mathcal{C}$ the collection of all compact subsets of G and denote by $\mathcal{V}$ the collection of all open neighbourhoods of the identity in G. The idea of the proof is to first construct a function $\rho : \mathcal{C} \to [0, \infty)$ that is monotonic, finitely additive and satisfies $\rho(K) = \rho(gK)$ for all $g \in G$ and $K \in \mathcal{C}$. We then use ρ to define an outer measure and then a measure.

For $K \subset G$ compact and $A \subset G$ with non-empty interior define

$$(K : A) = \min\{n \geqslant 0 \mid \exists\ g_1, \ldots, g_n \in G, K \subset \cup_{i=1}^n g_i A\}$$

Fix a compact set K_0 with non-empty interior. Given $V \in \mathcal{V}$, define $\rho_V : \mathcal{C} \to [0, \infty)$ by

$$\rho_V(K) = \frac{(K : V)}{(K_0 : V)}$$

Exercise 70. Show that

a) $\rho_V(K) \leqslant (K : K_0)$

b) $\rho_V(gK) = \rho_V(K)$

c) $K \subset L \implies \rho_V(K) \leqslant \rho_V(L)$

d) $\rho_V(K \cup L) \leqslant \rho_V(K) + \rho_V(L)$

Let $X = \prod_{K \in \mathcal{C}} [0, (K : K_0)]$ endowed with the product topology. By Tychonoff's Theorem, X is compact. Since $\rho_V(K) \leqslant (K : K_0)$, ρ_V can be identified with an element of X. For each $W \in \mathcal{V}$ define

$$D(W) = \{\rho_V \mid V \in \mathcal{V}, V \subset W\} \subset X \qquad \text{and} \qquad S = \bigcap_{W \in \mathcal{V}} \overline{D(W)}$$

Since X is compact, the set S is non empty. Fix $\rho \in S$ (so we have a function $\rho : \mathcal{C} \to [0, \infty)$).

Claim 1.

1) $\rho(K \cup L) \leqslant \rho(K) + \rho(L)$

2) $K \subset L \implies \rho(K) \leqslant \rho(L)$

3) $K \cap L = \emptyset \implies \rho(K \cup L) = \rho(K) + \rho(L)$

4) $\rho(gK) = \rho(K)$

For fixed $K, L \in \mathcal{C}$ the map $X \to \mathbb{R}$ given by $x \mapsto x(K) + x(L) - x(K \cup L)$ is continuous and non-negative at each point of $D(W)$. Therefore $\rho(K) + \rho(L) - \rho(K \cup L) \geqslant 0$. Parts 2 and 4 are similar.

Lecture 27: Existence of Haar measure

Continuing from last lecture.

Suppose that $K_1 \cap K_2 = \emptyset$. Let V_1 and V_2 be disjoint open sets with $V_i \supset K_i$.

Exercise 71. Show that there exist $U_1, U_2 \in \mathcal{V}$ such that $K_i U_i \subset V_i$.

Let $U = U_1 \cap U_2$. Then $K_1 U$ and $K_2 U$ are disjoint. Therefore, no gU^{-1} can intersect both K_1 and K_2. It follows that for any $W \in \mathcal{V}$ with $W \subset U^{-1}$ we have $\rho_W(K_1 \cup K_2) = \rho_W(K_1) + \rho_W(K_2)$. Therefore, the map $X \to \mathbb{R}$ given by $x \mapsto x(K_1) + x(K_2) - x(K_1 \cup K_2)$ vanishes at each element of $D(U^{-1})$. Since $\rho \in \overline{(D(U^{-1}))}$, part 3 of the claim is established.

We now define an outer measure λ on G first on open sets by

$$\lambda(V) = \sup\{\rho(K) \mid K \in \mathcal{C}, K \subset V\}$$

and then on all subsets by

$$\lambda(A) = \inf\{\lambda(V) \mid V \text{ open}, V \supset A\}$$

Exercise 72. Show that λ is an outer measure.

Claim 2. *All Borel subsets of G are λ-measurable*

As in the proof of the Riesz representation theorem, we show that for open sets $U, V \subset G$ we have

$$\lambda(U) \geqslant \lambda(U \cap V) + \lambda(U \cap V^c)$$

Let $\epsilon > 0$. There is a compact subset $K \subset U \cap V$ such that $\rho(K) > \lambda(U \cap V) - \epsilon$. Now choose a compact $L \subset U \cap K^c$ such that $\rho(L) > \lambda(U \cap K^c) - \epsilon$. Then K and L are disjoint and $\rho(L) > \lambda(U \cap V^c) - \epsilon$. Therefore

$$\lambda(U) \geqslant \lambda(K \cup L) \geqslant \rho(K \cup L) = \rho(K) + \rho(L) > \lambda(U \cap V) + \lambda(U \cap V^c) - 2\epsilon$$

The establishes the claim.

Let μ be the measure given by restricting λ to $\mathcal{B}_G$.

Claim 3. *The measure μ is regular.*

If V is an open set having compact closure, then

$$\lambda(V) = \sup\{\rho(K) \mid K \in \mathcal{C}, K \subset V\} \leqslant \rho(\overline{V}) < \infty$$

Given a compact K, there is an open $V \supset K$ with compact closure by Lemma 77. Therefore $\mu(K) \leqslant \mu(V) < \infty$.

That μ is outer regular follows from its definition. For inner regularity (on open sets), note that if $K \subset V$ with K compact and V open we have $\rho(K) \leqslant \mu(V)$ and therefore $\rho(K) \leqslant \mu(K)$. Therefore $\mu(V) = \sup\{\rho(K) \mid K \subset V\} \leqslant \sup\{\mu(K) \mid K \subset V\} \leqslant \mu(V)$

The proof of the theorem is complete once we note that μ is translation invariant and non-zero because ρ has those properties.

□

Lecture 28: Uniqueness of Haar measure

Next we turn to the question of uniqueness for Haar measures. If μ is a left Haar measure on a locally compact group G, then so too is $c\mu$ for for any $c > 0$. We will see that any two left Haar measures on G are related in this way.

We will need the following result about iterated integrals on LCH spaces.

Lemma 104. *Let X and Y be LCH spaces and let μ and ν be regular Borel measures on X and Y respectively. For $h \in \mathcal{K}(X \times Y)$ we have*

$$\int_X \int_Y h(x,y)\, d\nu(y)\, d\mu(x) = \int_Y \int_X h(x,y)\, d\mu(x)\, d\mu(y)$$

$\square$

Theorem 105. *Let G be a locally compact group. If μ and ν are left Haar measures on G, then there exists $c > 0$ such that $\nu = c\mu$.*

Proof. Fix a non-zero function $f \in \mathcal{K}(G)$ with $f \geqslant 0$. We will show that for all $g \in \mathcal{K}(G)$

$$\frac{\int g\, d\mu}{\int f\, d\mu} = \frac{\int g\, d\nu}{\int f\, d\nu} \tag{$*$}$$

From this it follows that $\int g\, d\nu = c \int g\, d\mu$ with $c = \int f\, d\nu / \int f\, d\mu$. Since this holds for all $g \in \mathcal{K}(G)$, the Riesz representation theorem tells us that $\nu = c\mu$.

It remains to establish $(*)$. Let $h \in \mathcal{K}(G \times G)$ be the function given by

$$h(x,y) = \frac{g(x) f(yx)}{\int f(zx)\, d\nu(z)}$$

Note that $x \mapsto \int f(zx)\, d\nu(z)$ is continuous by Corollary 99 and non-zero by Lemma 102. Also, $\mathrm{supp}(h) \subset \mathrm{supp}(g) \times \mathrm{supp}(f)\,\mathrm{supp}(g)^{-1}$.

$$\begin{aligned}
\int_X \int_Y h(x,y)\, d\nu\, d\mu &= \int_Y \int_X h(x,y)\, d\mu\, d\nu && \text{(lemma above)}\\
&= \int_Y \int_X h(y^{-1}x, y)\, d\mu\, d\nu && (\mu \text{ is translation invariant})\\
&= \int_X \int_Y h(y^{-1}x, y)\, d\nu\, d\mu && \text{(lemma above)}\\
&= \int_X \int_Y h(y^{-1}, xy)\, d\nu\, d\mu && (\nu \text{ is translation invariant})
\end{aligned}$$

For our choice of h this gives

$$\begin{aligned}
\int_X \int_Y \frac{g(x) f(yx)}{\int f(zx)\, d\nu(z)}\, d\nu\, d\mu &= \int_X \int_Y \frac{g(y^{-1}) f(x)}{\int f(zy^{-1})\, d\nu(z)}\, d\nu\, d\mu\\
\int_X \frac{g(x)}{\int f(zx)\, d\nu(z)} \int_Y f(yx)\, d\nu\, d\mu &= \int_X f(x)\, d\mu \int_Y \frac{g(y^{-1})}{\int f(zy^{-1})\, d\nu(z)}\, d\nu\\
\int_X g(x)\, d\mu &= \int_X f(x)\, d\mu \int_Y \frac{g(y^{-1})}{\int f(zy^{-1})\, d\nu(z)}\, d\nu\\
\frac{\int g\, d\mu}{\int f\, d\mu} &= \int_Y \frac{g(y^{-1})}{\int f(zy^{-1})\, d\nu(z)}\, d\nu
\end{aligned}$$

$\square$

Lecture 29: Properties of Haar measure

We now investigate some properties of Haar measure.

Exercise 73. Let μ be a left Haar measure on a locally compact group G. Show that μ is finite if and only if G is compact.

Let μ be a left Haar measure on G and $g \in G$. Since $x \mapsto xg$ is a homeomorphism of G, the formula $\mu_g(A) = \mu(Ag)$ defines a regular Borel measure on G. Moreover, for all $h \in G$ and $A \in \mathcal{B}_G$ we have

$$\mu_g(hA) = \mu(hAg) = \mu(Ag) = \mu_g(A)$$

Therefore μ_g is a left Haar measure and hence $\mu_g = \Delta(g)\mu$ for some $\Delta(g) > 0$.

Definition 106. The function $\Delta : G \to \mathbb{R}$ given by $g \mapsto \Delta(g)$ is the **modular function** of G.

If ν is another left Haar measure on G, then $\nu = c\mu$ and so

$$\nu_g = c\mu_g = c\Delta(g)\mu = \Delta(g)\nu$$

It follows that the modular function does not depend on the particular left Haar measure used.

Example 107. For the group $G = \{\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right] \in GL_2(\mathbb{R}) \mid a > 0\} \leqslant GL_2(\mathbb{R})$, the modular function is given by $\Delta(\left[\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right]) = 1/a$.

Exercise 74. Given a function $f : G \to \mathbb{R}$ and $x \in G$, define $f_x : G \to \mathbb{R}$ by $f_x(y) = f(yx^{-1})$. Show that

$$\int f_x \, d\mu = \Delta(x) \int f \, d\mu$$

Lemma 108. *Let G be a locally compact group with modular function Δ.*

1) Δ is continuous.

2) $\Delta(gh) = \Delta(g)\Delta(h)$

Proof. For the continuity statement fix $f \in \mathcal{K}(G)$ non-negative and non-zero. Then $\int f \, d\mu > 0$ by Lemma 102 and

$$x \mapsto \int f_x \, d\mu = \Delta(x) \int f \, d\mu$$

is continuous by Corollary 99.

For the second part note that

$$\Delta(gh)\mu(A) = \mu(Agh) = \Delta(h)\mu(Ag) = \Delta(h)\Delta(g)\mu(A)$$

□

Definition 109. A locally compact group G is **unimodular** if $\Delta(g) = 1$ for all $g \in G$.

Clearly, if G is abelian, then it is unimodular.

Proposition 110. *If $G/[G,G]$ is finite, then G is unimodular.*

Proof. We saw in the lemma above that Δ is a continuous homomorphism from G to the abelian group $((0,\infty), \times)$. It therefore factors through the abelianisation $G/[G,G]$. Therefore, if $G/[G,G]$ is finite, then $\Delta(G)$ is a finite subgroup of $((0,\infty), \times)$. The only finite subgroup is the trivial subgroup. □

Proposition 111. *Every compact group is unimodular.*

Proof. Since Δ is continuous, it is bounded on the compact G. If $g \in G$ satisfied $\Delta(g) > 1$, then $\Delta(g^n) = \Delta(g)^n$ would be unbounded. □

Example 112. The group $GL(n, \mathbb{R})$ is unimodular. This follows from the fact that

$$\mu(A) = \int_A \frac{1}{|\det(a)|^n} dm(a)$$

defines a left and right Haar measure on $GL(n, \mathbb{R})$. (Where m is Lebesgue measure on $\mathbb{R}^{n^2}$.)

Lecture 30: Polish spaces

Aside from LCH spaces another class on which it is fruitful to consider a measure are Polish spaces.

Definition 113. A **Polish space** is a topological space that is separable and admits a compatible complete metric.

Examples 114.

1. $\mathbb{R}^n$
2. compact metric spaces, e.g., $\{0,1\}^{\mathbb{N}}$
3. separable Banach spaces, e.g., $C([0,1])$

Before developing some properties of Polish spaces, let's note the following definitions.

Definition 115. A **standard Borel space** is a measurable space $(X, \mathcal{A})$ (i.e., $\mathcal{A}$ is a σ-algebra on X) such that there exists a Polish topology on X with $\mathcal{A}$ the Borel σ-algebra. A **Borel probability space** is a standard Borel space equipped with a probability measure.

Amazingly, any two uncountable standard Borel spaces are 'Borel isomorphic'.

Back to Polish spaces.

Proposition 116. *Every closed subspace of a Polish space is Polish. Every open subspace of a Polish space is Polish.*

Proof. Every subspace of a separable metrizable space is separable. What needs to be shown is the the open (or closed) subspace is completely metrizable. Let d be a complete metric on a Polish space X.

If $F \subset X$ is closed, then the restriction of d to F is a complete metric on F.

Suppose that $V \subsetneq X$ is open. Define a metric on V by

$$d_V(x,y) = d(x,y) + \left| \frac{1}{d(x,V^c)} - \frac{1}{d(y,V^c)} \right|$$

To see that this metric is compatible with the subspace topology on V note first that the function $x \mapsto d(x, V^c)$ is continuous. Consider a sequence of points $(x_i)_{i\in\mathbb{N}}$ with $x_i \in V$. Then $x_i \to x \in V$ with respect to d_V, if and only if $x_i \to x$ with respect to d.

Now to see that the metric d_V is complete. Suppose that $(x_i)_{i\in\mathbb{N}}$ is a Cauchy sequence in (V, d_V). Since $d(x_i,x_j) \leqslant d_V(x_i,x_j)$, the sequence is also Cauchy in (X,d). It therefore converges in (X,d) to a point $x \in X$. Note that it must be the case that $x \in V$, since otherwise we would have $d(x_i, V^c) \to 0$ and therefore for all $i \in \mathbb{N}$, $d_V(x_i,x_j) \to \infty$ as $j \to \infty$. Therefore $x \in V$ and $x_i \to x$ with respect to d_V. □

Corollary 117. *Every second countable LCH space is Polish.*

Outline of proof. The one-point compactification of X is compact, Hausdorff and second countable. Therefore X is an an open subset of a Polish space. □

Lecture 31: Borel measures on Polish spaces

Proposition 118. *The product of a sequence (finite or infinite) of Polish spaces is Polish.*

Proof. Let $(X_i)_i$ be a sequence of Polish spaces. Fix a metric d_i on X_i that is complete and satisfies $d_i(x,y) \leqslant 1$ for all $x, y \in X_i$. Define a metric on $X = \prod_i X_i$ by

$$d(x,y) = \sum_i 2^{-i} d_i(x_i, y_i)$$

Then d is compatible with the product topology on $\prod_i X_i$ and is complete (exercise!). To show that the space is separable we show that there is a countable basis for its topology. For each i let $\mathscr{V}_i$ be a countable basis for the topology on X_i. A countable basis for the product topology on X is given by the collection of all sets of the form

$$V_1 \times V_2 \times \cdots \times V_k \times X_{k+1} \times X_{k+2} \times \cdots$$

where $V_i \in \mathscr{V}_i$ □

It follows from the above proposition that $\mathbb{N}^{\mathbb{N}}$ (sometimes called **Baire space**) is Polish.

Proposition 119. *Every finite Borel measure on a Polish space is regular.*

Proof. Suppose that X is a Polish space and μ is a probability measure on $(X, \mathcal{B}_X)$. Fix a complete metric d on X. We first show that for all $A \in \mathcal{B}_X$

$$\begin{aligned} \mu(A) &= \inf\{\mu(V) \mid V \supset A, V \text{ open}\} \qquad (*) \\ &= \sup\{\mu(F) \mid F \subset A, F \text{ closed}\} \end{aligned}$$

Let $\mathcal{A} \subset \mathcal{B}_X$ be the collection of elements $A \in \mathcal{B}_X$ such that $(*)$ holds. We show that $\mathcal{A}$ contains the open sets and is a σ-algebra. From which it follows that $\mathcal{A} = \mathcal{B}_X$. Let $V \subset X$ be open and define $F_i = \{x \in V \mid d(x, V^c) \geqslant \frac{1}{i}\}$.

$$\begin{aligned} V &= \bigcup_{i \in \mathbb{N}} F_i \\ \mu(V) &= \lim \mu(F_i) \qquad \text{(continuity from below)} \end{aligned}$$

So V satisfies $(*)$.
Now suppose that $A \in \mathcal{A}$. We have

$$\begin{aligned} \mu(A^c) &= \mu(X) - \mu(A) \\ &= \mu(X) - \sup\{\mu(F) \mid F \subset A, \text{closed}\} \\ &= \inf\{\mu(X) - \mu(F) \mid F \subset A, \text{closed}\}\} \\ &= \inf\{\mu(F^c) \mid F \subset A, \text{closed}\}\} \\ &= \inf\{\mu(V) \mid V \supset A, \text{open}\} \end{aligned}$$

and similarly

$$\begin{aligned} \mu(A^c) &= \mu(X) - \inf\{\mu(V) \mid V \supset A, \text{open}\} \\ &= \sup\{\mu(X) - \mu(V) \mid V \supset A, \text{open}\}\} \\ &= \sup\{\mu(F) \mid F \subset A, \text{closed}\}\} \end{aligned}$$

Therefore $A^c \in \mathcal{A}$. Now suppose that $(A_i)_{i \in \mathbb{N}} \subset \mathcal{A}$ and let $\epsilon > 0$. For each i let $F_i \subset A_i$ be closed with $\mu(A_i \setminus F_i) < \epsilon 2^{-(i+1)}$. Let k be such that $\mu((\cup_{i \in \mathbb{N}} A_i) \setminus (\cup_{i=1}^k A_i)) < \epsilon/2$. Then we have

$$\mu((\cup_{i \in \mathbb{N}} A_i) \setminus (\cup_{i=1}^k F_i)) < \epsilon/2 + \epsilon \sum_{i=1}^k 2^{-(i+1)} < \epsilon$$

Therefore $\mu((\cup_{i \in \mathbb{N}} A_i) = \sup\{\mu(F) \mid F \subset A, \text{closed}\}$. Now choose $V_i \supset A_i$ with $\mu(V_i \setminus A_i) < \epsilon 2^{-i}$. We have

$$\mu((\cup_{i \in \mathbb{N}} V_i) \setminus (\cup_{i \in \mathbb{N}} A_i)) \leqslant \mu(\cup_{i \in \mathbb{N}} (V_i \setminus A_i)) \leqslant \sum_{i \in \mathbb{N}} \mu(V_i \setminus A_i) < \sum \epsilon 2^{-i} = \epsilon$$

Therefore $\mu((\cup_{i\in\mathbb{N}}A_i) = \inf\{\mu(V) \mid V \supset A, \text{open}\}$, and we have shown that $\mathcal{A} = \mathcal{B}_X$.

We know turn to showing inner regularity, that is that $\mu(A) = \sup\{\mu(K) \mid K \subset A, \text{compact}\}$ for all $A \in \mathcal{B}_X$. Given that the Borel sets satisfy $(*)$ above, it is enough to show that $\mu(F) = \sup\{\mu(K) \mid K \subset F, \text{compact}\}$ for all closed sets $F \subset X$.

Let $F \subset X$ be closed. Fix $\epsilon > 0$ and a dense subset $\{x_i\}_{i\in\mathbb{N}} \subset F$. For $i, j \in \mathbb{N}$ let B_i^j be the closed set given by

$$B_i^j = \{y \in F \mid d(x_i, y) \leqslant 2^{-j}\}$$

Since the x_i are dense, for each $j \in \mathbb{N}$ we have $F \subset \cup_{i\in\mathbb{N}} B_i^j$. Let $N_j \in \mathbb{N}$ be such that $\mu(F \setminus (\cup_{i=1}^{N_j} B_i^j)) < \epsilon 2^{-j}$. Define

$$K = \cap_{j\in\mathbb{N}} \cup_{i\leqslant N_j} B_i^j$$

Note that K is a closed subset of X. Also, for each j, K can be covered by finitely many balls of radius 2^{-j}. It follows that K is compact. Finally, note that

$$\begin{aligned}
\mu(F \setminus K) &= \mu(F \cap K^c) \\
&= \mu(F \cap (\cup_{j\in\mathbb{N}} \cap_{i\leqslant N_j} (B_i^j)^c)) \\
&= \mu(\cup_{j\in\mathbb{N}}(F \setminus (\cup_{i\leqslant N_j} B_i^j))) \\
&\leqslant \sum_{j\in\mathbb{N}} \epsilon 2^{-j} = \epsilon
\end{aligned}$$

□

Lecture 32: Maps between Polish spaces

Theorem 120 (Lusin's Theorem)**.** *Let X and Y be Polish spaces and μ a Borel probability measure on X. If $f : X \to Y$ is Borel, then for all $\epsilon > 0$ there is a compact $K \subset X$ such that $f\mid_K$ is continuous and $\mu(K^c) < \epsilon$.*

Proof. Let $\{V_i\}_{i\in\mathbb{N}}$ be a basis for the topology of Y. By hypothesis $f^{-1}(V_i) \in \mathcal{B}_X$. Fix an $\epsilon > 0$. Since μ is regular (by Proposition 119) there is an open $U_i \supset f^{-1}(V_i)$ with $\mu(U_i \setminus f^{-1}(V_i)) < \epsilon 2^{-(i+1)}$. Define $A \in \mathcal{B}_X$ by

$$A = X \setminus \cup_i (U_i \setminus f^{-1}(V_i))$$

Then $f|_A$ is continuous since

$$(f|_A)^{-1}(V_i) = f^{-1}(V_i) \cap A = U_i \cap A$$

Also,

$$\mu(\cup_i (U_i \setminus f^{-1}(V_i))) \leqslant \sum_i \mu(U_i \setminus f^{-1}(V_i)) \leqslant \epsilon/2$$

Since μ is regular, there exists a compact set $K \subset A$ such that $\mu(K) > \mu(A) - \epsilon/2$ and we have

$$\mu(K^c) = 1 - \mu(K) < 1 + \frac{\epsilon}{2} - \mu(A) = 1 + \frac{\epsilon}{2} - (1 - \mu(\cup_i U_i \setminus f^{-1}(V_i))) < \epsilon$$

□

Proposition 121. *Let X and Y be Polish spaces and $f : X \to Y$ a continuous map. Then $f(X)$ is measurable with respect to any Borel probability measure on Y.*

Proof. Fix compatible metrics complete d_X and d_Y on X and Y respectively and a Borel probability measure ν on Y. We need to show that there exist $E, F \in B_Y$ such that $E \subset f(X) \subset F$ and $\nu(F \setminus E) = 0$.

Exercise 75. Let $F \subset X$ be a closed subset and $\epsilon > 0$. Show that there exist closed non-empty subsets $F_i \subset F$ such that $F = \cup_i F_i$, $\operatorname{diam}(F_i) < \epsilon$ and $\operatorname{diam}(f(F_i)) < \epsilon$.

Using the above exercise we inductively define a collection of closed sets $\{F_w \subset X \mid w \in \mathbb{N}^{<\infty}\}$ with the following properties:

$$F_\emptyset = X \qquad F_w = \cup_{i\in\mathbb{N}} F_{w^\wedge i} \qquad \operatorname{diam}(F_w) < 2^{-\ell(w)} \qquad \operatorname{diam}(f(F_w)) < 2^{-\ell(w)}$$

Now, for each $w \in \mathbb{N}^{<\infty}$ let $B_w \in \mathcal{B}_Y$ be such that $f(F_w) \subset B_w \subset \overline{f(F_w)}$ with $\nu(B_w)$ minimal. Then

$$f(F_w) = \cup_{i\in\mathbb{N}} f(F_{w^\wedge i}) \subset \cup_i B_{w^\wedge i} \subset \cup_i \overline{f(F_{w^\wedge i})} \subset \overline{f(F_w)}$$

Therefore $\nu(B_w \cap \cup_{i\in\mathbb{N}} B_{w^\wedge i}) = \nu(B_w)$ since $B_w \cap \cup_{i\in\mathbb{N}} B_{w^\wedge i}$ is Borel and therefore can not have measure smaller than B_w by choice of B_w.

We have $\nu(B_w \setminus \cup_{i\in\mathbb{N}} B_{w^\wedge i}) = 0$. Define $A = \cup_{w\in\mathbb{N}^{<\infty}} (B_w \setminus \cup_{i\in\mathbb{N}} B_{w^\wedge i})$. Then $A \in \mathcal{B}_Y$ and $\nu(A) = 0$. Let $E = B_\emptyset \setminus A$ and $F = B_\emptyset$. Then we have $E \subset F$, $f(X) \subset F$ and $\nu(F \setminus E) = \nu(B_\emptyset \setminus (B_\emptyset \setminus A)) = \nu(B_\emptyset \cap A) = 0$. So we will be done if we show that $E \subset f(X)$.

Let $y \in B_\emptyset \setminus A$. Then there exists $i_1 \in \mathbb{N}$ such that $y \in B_{i_1}$. Similarly, since $y \in B_{i_1} \setminus A$, there exists $i_2 \in \mathbb{N}$ such that $y \in B_{i_1{}^\wedge i_2}$. Continuing in this way, there exists a sequence $(i_j)_{j\in\mathbb{N}}$ such that for all $n \in \mathbb{N}$ $y \in B_{i_1{}^\wedge i_2{}^\wedge \ldots {}^\wedge i_n}$. Note that it follows that $F_{i_1{}^\wedge i_2{}^\wedge \ldots {}^\wedge i_n} \neq \emptyset$. Let $x_n \in F_{i_1{}^\wedge i_2{}^\wedge \ldots {}^\wedge i_n}$. The sequence $(x_n)_n$ is Cauchy and therefore convergent to, say, $x \in X$. Since f is continuous and $d(y, f(x_n)) < 2^{-n}$, we have that $y = f(x)$.

□

Lecture 33: Brief introduction to ergodic theory

Definition 122. Let $(X, \mathcal{A}, \mu)$ be a standard Borel probability space. A measurable map $T : X \to X$ is called **measure preserving** if $\forall A \in \mathcal{A}$, $\mu(T^{-1}(A)) = \mu(A)$. A set $A \in \mathcal{A}$ is called **invariant** if $T^{-1}(A) = A$. The system $(X, \mathcal{A}, \mu, T)$ is called **ergodic** if every invariant set is either null or co-null.

Example 123. $X = \{0, 1\}^{\mathbb{N}}$ and μ as defined in Lecture 5. Let $T : X \to X$ be the 'left shift', that is, $T(x)(n) = x(n + 1)$. The map T is measure preserving and the system is ergodic.

Poincaré Recurrence Lemma. *Let $(A, \mathcal{A}, \mu)$ be a standard Borel probability space, $T : X \to X$ a measure preserving map, and $A \in \mathcal{A}$ such that $\mu(A) \neq 0$. Then for almost all $x \in A$, $\exists n \in \mathbb{N}$ such that $T^n(x) \in A$.*

Proof. For $n \geqslant 0$ define $A_n = \{x \in X \mid T^n(x) \in A \text{ and } \forall k > n\ T^k(x) \notin A\}$. Note that

$$
\begin{aligned}
&A_n \in \mathcal{A} \qquad \text{since} \qquad A_n = T^{-n}(A) \cap \left(\cup_{k>n} T^{-k}(A)\right)^c \\
&T^{-1}(A_n) = A_{n+1} \\
&A_n = T^{-n}(A_0) \\
&\mu(A_n) = \mu(A_0) \\
&A_m \cap A_n = \emptyset \qquad \text{if } m \neq n
\end{aligned}
$$

As the measure is finite, we conclude that $\mu(A_0) = 0$. $\square$

Lecture 34: Maximal ergodic theorem

Define

$$f^*(x) = \sup_{n \in \mathbb{N}} \frac{1}{n} \sum_{k=1}^{n-1} f \circ T^k(x)$$
$$E = \{x \in X \mid f^*(x) > 0\}$$

Maximal Ergodic Theorem. *Let $f \in \mathscr{L}^1(X, \mathcal{A}, \mu)$ and define f^* and E as above. Then*

$$\int_E f \, d\mu \geqslant 0$$

Proof. Done in lecture, and only with the assumption that T is injective. □

Corollary 124. *Let $\alpha \in \mathbb{R}$ and define $E_\alpha = \{x \in X \mid f^* > \alpha\}$. Then*

$$\int f \, d\mu \geqslant \alpha \mu(E_\alpha)$$

Proof. Apply the theorem to the function $f - \alpha$. □

Lecture 35: Pointwise ergodic theorem

Our main goal in this introduction to ergodic theory has been to prove the following theorem (also known as Birkhoff's Ergodic Theorem).

Pointwise Ergodic Theorem. *Let $(X, \mathcal{A}, \mu)$ be a standard Borel probability space, $T : X \to X$ a measure preserving function, and $f \in \mathscr{L}^1(X, \mathcal{A}, \mu)$. Then*

1) $\displaystyle \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x)$ *exists for μ-almost all $x \in X$.*

2) If T is ergodic, then

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) = \int f\, d\mu$$

Proof. Given $\alpha, \beta \in \mathbb{R}$ with $\alpha < \beta$ define

$$E_{\alpha,\beta} = \{x \in X \mid \liminf \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) < \alpha < \beta < \limsup \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x)\}$$

We show that $\mu(E_{\alpha,\beta}) = 0$.
Note that $T(E_{\alpha,\beta}) \subset E_{\alpha,\beta}$. Restricting T to $E_{\alpha,\beta}$ and applying Corollary 124 gives

$$\int_{E_{\alpha,\beta}} f\, d\mu \geqslant \beta\mu(E_{\alpha,\beta})$$

Applying the some reasoning to $-f$ we obtain

$$\int_{E_{\alpha,\beta}} -f\, d\mu \geqslant (-\alpha)\mu(E_{\alpha,\beta})$$

Combining the two inequalities above gives $\mu(E_{\alpha,\beta}) = 0$.
Now for the second part. It suffices to establish the result for positive f. Since T is ergodic, there exists $\alpha \in \mathbb{R}$ such that $\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) \to \alpha$ for μ-almost all $x \in X$. We want to show that $\alpha = \int f\, d\mu$. Note first that

$$0 \leqslant \int \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k\, d\mu = \int f\, d\mu \qquad (T \text{ is measure preserving})$$

Therefore

$$\begin{aligned}
\alpha &= \lim \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x) \\
&= \int \lim \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x)\, d\mu \\
&\leqslant \liminf \int \frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k(x)\, d\mu && \text{(Fatou)} \\
&= \int f\, d\mu
\end{aligned}$$

We claim that it is also the case that $\int f\, d\mu \leqslant \alpha$. Let $\epsilon > 0$. We show that $\int f\, d\mu < \alpha + \epsilon$.
Let $g : X \to \mathbb{R}$ be measurable and bounded such that $g \leqslant f$ and $\int |f - g|\, d\mu < \epsilon$. From the first part of the current theorem $\exists \gamma \in \mathbb{R}$ such that $\frac{1}{n}\sum_{k=0}^{n-1} g \circ T^k(x) \to \gamma$ for almoset all x. Since g is bounded $\exists \beta$ such that $\frac{1}{n}\sum_{k=0}^{n-1} g \circ T^k(x) < \beta$. By the Dominated Convergence Theorem

$$\lim \int \frac{1}{n}\sum_{k=0}^{n-1} g \circ T^k\, d\mu = \int \lim \frac{1}{n}\sum_{k=0}^{n-1} g \circ T^k\, d\mu = \int \gamma\, d\mu = \gamma$$

On the other hand, since T is measure preserving

$$\int \frac{1}{n}\sum_{k=0}^{n-1} g\circ T^k\,d\mu = \frac{1}{n}\sum_{k=0}^{n-1}\int g\circ T^k\,d\mu = \frac{1}{n}\sum_{k=0}^{n-1}\int g\,d\mu = \int g\,d\mu$$

Therefore $\int g\,d\mu = \gamma$. Further,

$$\int \frac{1}{n}\sum_{k=0}^{n-1} g\circ T^k\,d\mu \leqslant \int \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k\,d\mu$$

Taking limits gives $\gamma \leqslant \alpha$ and therefore

$$\int f\,d\mu < \int g\,d\mu + \epsilon = \gamma + \epsilon \leqslant \alpha + \epsilon$$

□