Math 564: Real Analysis and Measure Theory Notes

Ale Suarez Serrano

December 6, 2025

Let me know if you spot any mistakes!

Contents

CHAPTER 1

Motivation for Measure Theory

1.1 Probability

We understand well the probability theory of n coin tosses where the probability of the coin showing up 1 is some $p \in (0,1)$ and it showing up 0 is $1-p$. Then for every word $w \in 2^n := \{0,1\}^n$, the probability of coin tosses resulting in this word is

$$\mathbb{P}_p(w) = p^{\text{number of 1's in } w} \cdot (1-p)^{\text{number of 0's in } w}.$$

What if $n = \infty$? In other words, we consider the space $2^{\mathbb{N}} = \{0,1\}^{\mathbb{N}}$ of infinite binary sequences, with the same probabilities of tossing 1 or 0. Then how do we define the probability of "event" in this space?

1.2 Geometry

We would like to have a robust notion of volume in $\mathbb{R}^d$ for $d \geq 1$,i.e. we would like to determine the volume of a large class of subsets of $\mathbb{R}^d$. We know that the volume of a **box**

$$B := I_1 \times I_2 \times \cdots \times I_d \subset \mathbb{R}^d$$

(where $I_j \subset \mathbb{R}$ is an interval) should be

$$Volume(B) = ln(I_1)ln(I_2) \cdot \cdots \cdot ln(Id)$$

where $ln(I)$ is the difference of the right and left endpoints of I. We want to extend this to a class of sets which are closed under countable operations: compliments, countable unions, countable intersections.

1.3 Analysis

The class of Riemann integrable functions is not closed under pointwise limits; indeed, even a pointwise limit of continuous functions on $[0, 1]$ is tipically not Riemann integrable. But the whole subject of Analysis is about approximations/limits, so we would like to extend the class of integrable functions so it becomes closed under countable pointwise limits. Clearly, for a subset $B \subset \mathbb{R}^d$, the integral of its indicator function $\mathbb{1}_B$ will simply be $volume(B)$, so this task subsumes the previous task about volume.

CHAPTER 2

Measures, Their Construction, and Properties

2.1 Polish Spaces

We now define a very robust class of metric spaces that we will be working with throughout and that arise naturally in analysis and related fields.

Definition 2.1.1. A metric space (X, d) is called **Polish** if d is a **complete metric** (every d-Cauchy sequences converges) and X is **separable** (i.e. there is a countable dense set).

Definition 2.1.2. In a metric space X, a **topological basis** is a collection $\mathcal{U}$ of open subsets of X such that every open set is a union (maybe uncountable) of sets in $\mathcal{U}$.

Proposition 1. A metric space X is separable $\iff$ it is **2nd countable**, i.e. it has a countable basis of open sets.

Proof. Homework. □

Example 1 (Examples of Polish Spaces)**.**

(a) $\mathbb{R}$ or more generally, $\mathbb{R}^d$ with the metric

$$d_\infty(\vec{x}, \vec{y}) := \max_{i=1}^{d} |x_i - y_i| \,.$$

We know from undergrad analysis that this is a complete metric. Also, rationals are dense and countable, so $\mathbb{Q}^d \subset \mathbb{R}^d$ is dense and countable. Note that open intervals with rational endpoints form a countable basis for $\mathbb{R}$ and thus open boxes form a countable basis for $\mathbb{R}^d$.

We can also equop $\mathbb{R}^d$ with other equivalent complete metrics (two metrics are **equivalent** if they induce the sample open sets), namely, for $1 \leq p < \infty$:

$$d_p(\vec{x}, \vec{y}) := \left(\sum_{i=1}^{d} |x_i - y_i|^p \right)^{1/p} .$$

One can show that d_p is **bi-Lipschitz equivalent** to d_∞,i.e. There is a constant $C_p \leq 0$ such that

$$\frac{1}{C_p} \cdot d_\infty \leq d_p \leq C_p \cdot d_\infty.$$

In particular, the spaces $(\mathbb{R}^d, d_p)$ are Polish, for $1 \leq p \leq \infty$. It's also easy to see that $\lim_{p\to\infty} d_p = d_\infty$ (homework).

(b) If (X, d) is a polish metric space, then any closed subset is still Polish with the same metric (indeed, closedness ensures completeness of d and any subspace of a 2nd countable space is 2nd countable). What about open subsets, say $(0, 1)$ in $\mathbb{R}$? The same metric won't work because it won't be complete, but maybe we can take a different equivalent metric that is complete. Indeed, d_∞ is complete in $\mathbb{R}$ and $\mathbb{R}$ "looks like " $(0, 1)$,i.e. they are **homeomorphic** (i.e. there is a bijective continuous function with a continuous inverse between the two spaces).Thus, we can "copy" the complete metric from $(0, 1)$ via any homeomorphism. More concretely,

$$d(x, y) = d_\infty(x, y) + \left| \frac{1}{d_\infty(x, \{0, 1\})} - \frac{1}{d_\infty(y, \{0, 1\})} \right|$$

is a complete metric on $(0, 1)$ equivalent to d_∞. Such sets are called **Polishable**, and it is a theorem of Descriptive Set Theory that a subset of a Polish space is Polishable if and only if it is C_δ (countable intersection of open sets).

(c) The space $C([0, 1])$ of continuous functions on $[0, 1]$ with the **uniform metric** :

$$d_u(f, g) = \max \sup_{x\in[0,1]} |f(x) - g(x)| ,$$

is Polish. Indeed, we know from undergrad analysis that a uniformly Cauchy sequence of continuous functions converges to a continuous function, so d_u is complete.

As for separability, polynomials with rational coefficients form a countable dense set (by the Weistrass theorem), or more precisely, piece-wise linear functions (with finitely many pieces) with rational breakpoints form a countable dense set.

(d) The tree spaces: Cantor space $2^{\mathbb{N}}$ and Baire space $\mathbb{N}^{\mathbb{N}}$.

Let A be a nonempty countable set (e.g. $A = 2 := \{0,1\}$ or $A = \mathbb{N}$). Let $X = A^{\mathbb{N}}$ be the set of infinite sequences of elements of A. We depict $A^{\mathbb{N}}$ as the infinite branches through the tree $A^{<\mathbb{N}}$:= the set of finite sequences in A.

INCLUDE TREE DRAWING.

We equip $A^{\mathbb{N}}$ wit the following metric: If $x \neq y \in A^{\mathbb{N}}$, then:

$$d(x,y) = 2^{-\Delta(x,y)}$$

where $\Delta(x,y) := \min_{i\in\mathbb{N}} i$ with $x_i \neq y_i$, and $d(x,y) = 0$ if $x = y$.

This d is indeed a metric on $A^{\mathbb{N}}$, and in fact its an ultrametric (hw). Also d is a complete metric (hw) and for a fixed $a_o \in A$, the set of sequences which are eventually a_0 form a countable dense set. Thus $a^{\mathbb{N}}$ is Polish.

We may also talk about the topology (i.e. the collection of open sets of $A^{\mathbb{N}}$. For $2^{-n} < r \leq 2^{-(n-1)}$, the open ball

$$\begin{aligned} B(x,r) &:= \{y \in A^{\mathbb{N}} : d(y,x) < r\} \\ &= \{y \in A^{\mathbb{N}} : d(y,x) \leq 2^{-n}\} \\ &= \{y \in A^{\mathbb{N}} : y|_n = x|_n\} \\ &= [x|_n] \end{aligned}$$

where $n = 0, 1, \ldots, n-1$, and where the last term denotes the cylinder with base $x|_n \in A^{\mathbb{N}}$. More generally, for a finite word $w \in A^{<\mathbb{N}}$, let

$$[w] = \{y \in A^{\mathbb{N}} : y \supseteq w\} := \{y \in A^{\mathbb{N}} : y|_{ln(w)} = w\}$$

denote the **cylinder with base** w. Each cylinder is an open ball, as well as a closed ball (indeed a compliment of a cylinder is a countable union of disjoint cylinders), whose center is any element of it (the realtor's metric).

Thus, every open set is a union of cylinders, hence the cylinders form a countable basis for $A^{\mathbb{N}}$. When working with $A^{\mathbb{N}}$, we work with this basis. Cylinders are clopen, which makes $A^{\mathbb{N}}$ totally disconnected, in fact, 0-dimensional.

Proposition 2. $A^{\mathbb{N}}$ is compact $\iff$ A is finite,

Proof. Uses Konigs lemma, and is left as a (HW). □

The nice thing about the space $A^{\mathbb{N}}$ is that since it is so disconnected, it behaves like a discrete space, allowing us to do combinatorics on it, but also take limits!

2.2 Sigma Algebras

Definition 2.2.1. Let X be a set. A collection $\mathcal{A} \subseteq \mathcal{P}(X)$ is called an **algebra** (resp. σ**-algebra**) if $\emptyset \in \mathcal{A}$, and A is closed under compliments and finite unions (resp. countable unions).

Definition 2.2.2. A set X equipped with a σ-algebra $\mathcal{S}$ is called a **measurable space**.

Example 2. (a) For a nonempty set X, $\mathcal{P}(X)$ is a σ-algebra.

(b) Let X be a set. The collection $\mathcal{A}$ of finite and co-finite sets is an algebra (because finite sets are finite). The collection $\mathcal{S}$ countable and co-countable sets is a σ-algebra.

(c) In a metric/topological space, the collection of clopen sets is an algebra, ad we call it the algebra of clopen sets.

(d) For a finite non-empty set A, the clopen sets of $A^{\mathbb{N}}$ are exactly the finite disjoint union of cylinders, where the finiteness comes from the compactness of $A^{\mathbb{N}}$ (HW).

(e) A compliment of a box B in $\mathbb{R}^d$ os a finite disjoint union of boxes, so the collection of finite disjoint unions of boxes is an algebra (also we have a finite union of boxes is a finite disjoint union of boxes).

Remark 1. Notice that an arbitrary intersection of (σ)-algebras is itself a (σ)-algebra, i.e. if $\mathcal{A}_i$ is a (σ)-algebra for i in some index set I, then $\bigcap_{i\in I} \mathcal{A}_i$ is a (σ)-algebra.

This allows us to define the following notions:

Definition 2.2.3. Let X be a set and $\mathcal{C} \subseteq \mathcal{P}(X)$. The (σ)**-algebra generated by** $\mathcal{C}$ is the smallest (σ)0algebra containing $\mathcal{C}$. Namely,

$$\langle \mathcal{C} \rangle = \bigcap \{\mathcal{A} : \mathcal{A} \subseteq \mathcal{P}(X),\ \mathcal{A} \text{ is an algebra},\ \mathcal{A} \supseteq \mathcal{C}\}$$

and

$$\langle \mathcal{C} \rangle_\sigma = \bigcap \{\mathcal{A} : \mathcal{A} \subseteq \mathcal{P}(X),\ \mathcal{A} \text{ is an } \sigma\text{-algebra},\ \mathcal{A} \supseteq \mathcal{C}\}$$

are the generated algebra of $\mathcal{C}$ and σ-algebra of $\mathcal{C}$ respectively.

Definition 2.2.4. For a metric/topological space X, the σ-algebra $\mathcal{B}(X)$ generated by the open sets is called the **Borel σ-algebra** and its elements are called **Borel sets**.

The definitions of $\langle \mathcal{C} \rangle$ and $\langle \mathcal{C} \rangle_\sigma$ are top-down, and we give their bottom-up equivalent:

Proposition 3. Let X be a set and $\mathcal{C} \subseteq \mathcal{P}(X)$. Then:

(a) $\langle \mathcal{C} \rangle = \bigcup_{i \in \mathbb{N}} \mathcal{C}_n$ where $\mathcal{C}_0 := \mathcal{C}$ and

$$\mathcal{C}_{n+1} := \{B^c : B \in \mathcal{C}_n\} \cup \left\{ \bigcup_{i<k} B_i : B_i \in \mathcal{C}_n,\ k \in \mathbb{N} \right\}.$$

(b)

(c) $\langle \mathcal{C} \rangle_\sigma = \bigcup_{\alpha \in \omega_1} \mathcal{C}_\alpha$ where $\mathcal{C}_0 := \mathcal{C}$ and

$$\mathcal{C}_{n+1} := \{B^c : B \in \mathcal{C}_\beta, \beta < \alpha\} \cup \left\{ \bigcup_{i \in \mathbb{N}} B_i : B_i \in \bigcup_{\beta < \alpha} \mathcal{C}_\beta \right\},$$

where ω_1 is the smallest uncountable cardinal.

Proof. (a) is left as homework and (b) is optional. □

Remark 2. In a metric/topological space, for any countable basis $\mathcal{U}$, the σ-algebra generated by $\mathcal{U}$ is $\mathcal{B}(X)$.

Proof. Indeed, ever open set $\mathcal{O}$ is a union of sets in $\mathcal{U}$, hence a countable union of sets in $\mathcal{U}$, and as such $\mathcal{O} \in \langle \mathcal{U} \rangle_\sigma$ is a σ-algebra containing all open sets, hence $\mathcal{B}(X) \subseteq \langle \mathcal{U} \rangle_\sigma$. But also $\langle \mathcal{U} \rangle_\sigma \subseteq \mathcal{B}(X)$ because $\mathcal{U}$ is a collection of open sets. □

2.3 Measures and Premeasures

Definition 2.3.1. Let X be a set and $\mathcal{C} \subseteq \mathcal{P}(X)$. A function

$$\mu : \mathcal{C} \to [0, \infty]$$

is said to be **finitely additive** if

$$\mu\left(\bigsqcup_{i<k} A_i\right) = \sum_{i<k} \mu(A_i)$$

whenever $k \in \mathbb{N}$ and $\bigsqcup_{i<k} A_i \in \mathcal{C}$. Moreover, we say μ is **countably additive** if

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

whenever $A_i \in \mathcal{C}$ and $\bigsqcup_{i\in\mathbb{N}} A_i \in \mathcal{C}$.

Definition 2.3.2. For a measurable space $(X, \mathcal{S})$, a **measure** on $(X, \mathcal{S})$ is a countably additive function

$$\mu : \mathcal{S} \to [0, \infty]$$

such that $\mu(\emptyset) = 0$. The triple $(X, \mathcal{S}, \mu)$ is called a **measure space**.

Remark 3. People also deal with **finitely additive measures** on algebras, but a finitely additive measure (even on a σ-algebra) is not generally a measure (since it may not be countably additive).

Definition 2.3.3. A measure μ on a measurable space $(X, \mathcal{S})$ is called:

- A **probability measure** if $\mu(X) = 1$.
- **finite** if $\mu(X) < \infty$.
- **σ-finite** if $X = \bigcup_{n\in\mathbb{N}} B_n$, where $B_n \in \mathcal{S}$ and $\mu(B_n) < \infty$ for all $n \in \mathbb{N}$.

Proposition 4. Let $(X, \mathcal{S})$ be a measurable space.

(a) Any countable non-negative linear combination of measures on $(X, \mathcal{S})$ is a measure,i.e. if $\mu_1, \ldots, \mu_n$ are measures and $a_n \geq 0$, then

$$\sum_{i=1}^{n} a_i \mu_i$$

is a measure.

(b) Any convex combination of probability measures on $(X, \mathcal{S})$ is a probability measure,i.e. if $\mu_1, \ldots, \mu_n$ are measures and $a_n \geq 0$, with $\sum_{i=1}^n a_i = 1$, then

$$\sum_{i=1}^{n} a_i \mu_i$$

is a probability measure.

Example 3. (a) The **Dirac (delta) measure** (or point-measure): Let X be a set and fix some point $x_0 \in X$. Then define a meausre

$$\delta_{x_0} : \mathcal{P}(X) \to [0,1]$$

by:

$$\delta_{x_0}(B) := \begin{cases} 1, & \text{if } x_0 \in B \\ 0, & \text{otherwise.} \end{cases} \quad .$$

This is called the Dirac measure at x_0.

(b) The **zero measure**: On any set X, the zero measure is

$$\zeta : \mathcal{P}(X) \to \{0\}.$$

(c) **The counting measure:** On any set X, the counting measure is the map

$$\chi : \mathcal{P}(X) \to [0, \infty]$$

via

$$\chi(B) := \begin{cases} |B|, & \text{if } B \text{ is finite} \\ \infty, & \text{otherwise.} \end{cases} \quad .$$

Note that when X is countable, then

$$\chi = \sum_{x \in X} \delta_x.$$

Moreover, χ is finite when X is finite and σ-finite when X is countable. If X is uncountable then X is not σ finite.

(d) Given a set X, define a measure μ on the σ-algebra of countable and co-countable subsets of X as follows:

$$\mu(B) := \begin{cases} 0, \text{if } B \text{ is countable} \\ 1, \text{otherwise} \end{cases} .$$

If X is countable, then μ is the zero-measure.

Definition 2.3.4. Let $(X, \mathcal{S}, \mu)$ be a measure space. A set $B \in \mathcal{S}$ is called an **atom**(or μ-atom) if $\mu(B) > 0$ and for all $A \subseteq B$, with $A \in \mathcal{S}$, we have that either $\mu(A) = 0$ or $\mu(A) = \mu(B)$.

Definition 2.3.5. A measure space $(X, \mathcal{S}, \mu)$ is called:

- **Atomic**(or purely-atomic) if every positive measure set in $\mathcal{S}$ contains an atom.
- **Atomless** if there are no atoms.

Remark 4. The zero measure is both atomic and atomless. Also notice that all measures in the previous examples are atomic. To define interesting atomless measures, we'll need to define them on a algebra $\mathcal{A}$ and extend them to the σ-algebra generated by $\mathcal{A}$.

2.4 Finitely Additive Measures and Premeasures

Definition 2.4.1. Let $\mathcal{A}$ be an algebra on a set X. A function

$$\mu : \mathcal{A} \to [0, \infty]$$

is called a **finitely additive measure** (resp. a countably additive measure or **premeasure**) if $\mu(\emptyset) = 0$ and μ is finitely (resp. countably) additive.

Proposition 5 (Disjointification Trick)**.** For an algebra $\mathcal{A}$, any countable union $\bigcup_{n\in\mathbb{N}} A_n$ of sets $A_n \in \mathcal{A}$ is equal to a countable disjoint union $\bigsqcup_{n\in\mathbb{N}} A_n^{'}$ of sets $A_n^{'} \in \mathcal{A}$.

Proof. Take $A_0' = A_0$ and $A_k' = A_k \setminus \left[\bigcup_{i<n} A_i\right]$. $\square$

Proposition 6 (Properties of Finitely Additive Measures)**.** Let μ be a finitely additive measure on an algebra $\mathcal{A}$ on a set X. Then,

(a) μ is monotone: If $A \subseteq B$ then $\mu(A) \leq \mu(B)$ for all $A, B \in \mathcal{A}$.

(b) μ is countably super additive:

$$\mu(\bigsqcup_{n\in\mathbb{N}} A_n) \geq \sum_{n\in\mathbb{N}} \mu(A_n)$$

for all $A_n \in \mathcal{A}$ with $\bigsqcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

(c) μ is finitely-subadditive:

$$\mu(\bigcup_{n<N} A_n) \leq \sum_{n<N} \mu(A_n)$$

for all $A_n \in \mathcal{A}$. Moreover, if μ is a premeasure (i.e. countably additive) then it is countably subbaditive, i.e.

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} \mu(A_n)$$

for all $A_n \in \mathcal{A}$ with $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Proof. (a): $\mu(B) = \mu((B \setminus A) \cup A) = \mu(B \setminus A) + \mu(A) \geq \mu(A)$.

(b):

$$\begin{aligned}
\mu\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) &= \mu\left(\bigsqcup_{n\leq N} A_n \sqcup \bigsqcup_{n>N} A_n\right) \\
&= \sum_{n\leq N} \mu(A_n) + \mu(\bigsqcup_{n>N} A_n) \\
&\geq \sum_{n\leq N} \mu(A_n).
\end{aligned}$$

Taking $N \to \infty$ yields our desired inequality.

(c): We use the disjointification trick:

$$\begin{aligned}\mu\left(\bigcup_{n<N} A_n\right) &= \mu\left(\bigsqcup_{n<N} A_n^{'}\right)\\ &= \sum_{n<N} \mu(A_n^{'})\\ &\leq \sum_{n<N} \mu(A_n).\end{aligned}$$

Using the same logic we can prove countable subadditivity for premeasures.

□

2.5 Construction of Premeasures

2.5.1 Bernoulli Premeasures

Let $X = 2^{\mathbb{N}}$, where $2 := \{0,1\}$. Any probability measure on 2 is of the form $\nu_p(1) = p$ and $\nu_p(0) = 1 - p$ for $p \in (0,1)$.We will define a premeasure on the algebra $\mathcal{A}$ of clopen sets in $2^{\mathbb{N}}$ which from the homework is the algebra of finite disjoint unions of cylinders. This premeasure μ_p will satisfy

$$\mu_p([w]) = \nu_p(1)^{\text{number of 1's in } w} \cdot \nu_p(0)^{\text{number of 0's in } w}$$

and will be called the **bernoulli(p) measure**.

We first define μ_p on a cylinder $[w]$, where $w \in 2^{<\mathbb{N}}$ by:

$$\tilde{\mu_p}([w]) = p^{\text{number of 1's in } w} \cdot (1-p)^{\text{number of 0's in } w}.$$

For instance, $\tilde{\mu_p}([01100]) = p^2 \cdot (1-p)^3$. Then for each $B \in \mathcal{A}$, we "define "

$$\mu_p(B) := \sum_{n<N} \tilde{\mu_p}[w_n],$$

where $B = \bigsqcup_{n<N}[w_n]$. We first need to show this is well defined, i.e. that it does not depend on how B is partitioned as a disjoint union of cylinders.

Proposition 7. $\tilde{\mu_p}$ is finitely additive on equal-length cylinders, i.e. for any cylinder $[w]$ and $n \in \mathbb{N}$:

$$\tilde{\mu_p}([w]) = \sum_{u \in 2^n} \tilde{\mu_p}([wu]).$$

Proof. One can show this by induction on n, but it is enough to verify just the base case $n = 1$:

$$\tilde{\mu_p}([w0]) + \tilde{\mu_p}([w1]) = \tilde{\mu_p}([w]) \cdot (1-p) + \tilde{\mu_p}([w]) \cdot p = \tilde{\mu_p}([w]).$$

□

Proposition 8. Let $A \in \mathcal{A}$, and $\mathcal{D}_1$, $\mathcal{D}_2$ be two finite partitions of A into finite cylinders. Then,

$$\sum_{d_1 \in \mathcal{D}_1} \tilde{\mu_p}(d_1) = \sum_{d_2 \in \mathcal{D}_2} \tilde{\mu_p}(d_2).$$

Proof. Let $\mathcal{R}$ be a **common refinement** of $\mathcal{D}_1$ and $\mathcal{D}_2$ (i.e. take all the non-empty intersections of their cylinders). Clearly this is still a finite partition. Now, we may take very cylinder in $\mathcal{R}$ to have the same length (by splitting each cylinder into a finite partitions of cylinders bigger base length). Then,

$$\begin{aligned}\sum_{D_1 \in \mathcal{D}_1} \tilde{\mu_p}(D_1) &= \sum_{D_1 \in \mathcal{D}_1} \sum_{\substack{R \in \mathcal{R} \\ R \subseteq D_1}} \tilde{\mu_p}(R) \\ &= \sum_{R \in \mathcal{R}} \tilde{\mu_p}(R) \\ &= \sum_{D_2 \in \mathcal{D}_2} \sum_{\substack{R \in \mathcal{R} \\ R \subseteq D_2}} \tilde{\mu_p}(R) \\ &= \sum_{D_2 \in \mathcal{D}_2} \tilde{\mu_p}(D_2).\end{aligned}$$

□

Our second claim shows that μ_p is well defined on $\mathcal{A}$ and also implies the following corollary:

Corollary 2.5.1. μ_p is finitely additive.

Proof. Homework. □

Proposition 9. μ_p is a premeasure (i.e. countably additive)

Proof. This follows automatically from compactness. If a clopen set (hence closed, hence compact) is a disjoint union of other clopen (hence open) sets $\bigsqcup_{n \in \mathbb{N}} \mathcal{U}_n$, then all but finitely many of these cylinders have to be empty (because there is a finite subcover of $\{\mathcal{U}_n\}_{n \in \mathbb{N}}$). □

This construction equally works for $A^{\mathbb{N}}$, for any finite non-empty A, and any probability measure on A. We obtain a premeasure $\mu = \nu^{\mathbb{N}}$ on the algebra $\mathcal{A}$ of clopen subsets of $A^{\mathbb{N}}$ satisfying:

$$\mu([w]) = \nu(w_0) \cdot \nu(w_1) \cdot \dots \cdot \nu(w_{n-1}),$$

for every $w \in A^{\mathbb{N}}$ and $n \in \mathbb{N}$.

2.5.2 Lebesgue Premeasure on $\mathbb{R}^d$.

Analogously to Bernoulli measures on $A^{\mathbb{N}}$, we define a premeasure on the algebra $\mathcal{A}$ generated by the boxes in $\mathbb{R}^d$. Note that elements of $\mathcal{A}$ are finite disjoint unions of boxes, just like the clopen sets in $A^{\mathbb{N}}$, for a finite A, are finite disjoint unions of cylinders.

We firs define a preameasure on boxes:

$$\tilde{\lambda}(I_1 \times \dots I_d) := \prod_{n=1}^{d} ln(I_n),$$

where $ln(I_n)$ is equal to the difference between its right and left endpoint. As convention, we set $0 \cdot \infty = 0$.

We then "define" the potential measure on $\mathcal{A}$ by

$$\lambda(A) = \sum_{B \in \mathcal{P}} \tilde{\lambda}(B)$$

where $\mathcal{P}$ is any finite partition of A into boxes. As with Bernoulli, we need to show that this is well defined,i.e. that it does not depend on the choice of the partition $\mathcal{P}$. As with Bernoulli, we show that if a box is partitioned into a "homogeneous" collection of boxes, then $\tilde{\lambda}$ is finitely additive. The notion of "homogeneous" for boxes are **grid-partitions**, namely, a partition $\mathcal{P}$ of a box $B = I_1 \times \dots I_d$ into ones of the following form: Each I_k is partitioned into finitely many intervals

$$I_k = \bigsqcup_{n < N_K} I_k^{(n)}$$

and

$$\mathcal{P} = \left\{ I_1^{n_1} \times \dots \times I_d^{n_d} : (n_1, \dots, n_d) \in N_1 \times \dots N_d \right\}$$

where we view $N := \{0, 1, \dots, N-1\}$.

Proposition 10. If $\mathcal{P}$ is a grid-partition of a box B, then

$$\tilde{\lambda}(B) = \sum_{P \in \mathcal{P}} \tilde{\lambda}(P).$$

Proof. This is trivial in $d = 1$ and for $d > 1$ apply induction using the distributive law, i.e.

$$(a_1 + \cdots + a_k) \cdot (b_1 + \cdots + b_l) = \sum_{\substack{i \le k \\ j \le l}} a_i b_j.$$

$\square$

Proposition 11. If $\mathcal{P}_1$ and $\mathcal{P}_2$ are two finite partitions of a set $A \in \mathcal{A}$ into boxes, then

$$\sum_{P_1 \in \mathcal{P}_1} \tilde{\lambda}(P_1) = \sum_{P_2 \in \mathcal{P}_2} \tilde{\lambda}(P_2).$$

Proof. Take a grid-partition of A that is a common refinement of $\mathcal{P}_1$ and $\mathcal{P}_2$. The rest is left as homework. $\square$

Corollary 2.5.2. λ on $\mathcal{A}$ is well-defined and finitely additive.

Proposition 12. λ is countably additive on $\mathcal{A}$.

Remark 5. For $a, b \in \mathbb{R}$, we write $a \approx_\varepsilon b$ if $|a - b| \le \epsilon$.

Proof. Because a finitely additive measure is always countably supperadditive, it is enough to prove countable subadditivity. We prove this in the case where a bounded box B is written as a countable disjoint union of boxes

$$B = \bigsqcup_{n \in \mathbb{N}} B_n.$$

The general case follows easily from this and is left as a small homework exercise.

In the case of cylinders of $2^{\mathbb{N}}$, we used that B is compact and the B_n are open, but for boxes neither is true in general. However, we can approximate B by closed boxes and B_n by open boxes. Fix $\varepsilon > 0$. Let $B' \subseteq B$ be a closed box such that

$$\lambda(B') \approx_{\varepsilon/2} \lambda(B).$$

Also for each $n \in \mathbb{N}$, let $\tilde{B}_n \supseteq B_n$ be an open box such that

$$\lambda(\tilde{B}_n) \approx_{\varepsilon \cdot 2^{-(n+1)}} \lambda(B_n).$$

Then, $\{\tilde{B}_n\}_{n\in\mathbb{N}}$ is an open cover of the compact set B', so there is a finite subcover $\{\tilde{B}_n\}_{n<N}$. Then:

$$\begin{aligned}
\lambda(B) &\approx_{\varepsilon/2} \lambda(B') \\
&\leq \lambda\left(\bigcup_{n<N} \tilde{B}_n\right) \\
&\leq \sum_{n<N} \lambda(\tilde{B}_n) \\
&\leq \sum_{n\in\mathbb{N}} \lambda(\tilde{B}_n) \\
&\approx_{\varepsilon/2} \sum_{n\in\mathbb{N}} \lambda(B_n),
\end{aligned}$$

so:

$$\lambda(B) \leq \varepsilon + \sum_{n\in\mathbb{N}} \lambda(B_n)$$

which implies our desired inequality since ε was arbitrary. □

We call this premeasure λ on $\mathcal{A}$ the **Lebesgue premeasure**.

2.6 Caratheodory Extension

To define measures, we always define a premeasure on some algebra and apply the following theorem, where we call a premeasure μ on an algebra $\mathcal{A}$ on X **σ-finite** if X can be partitioned into countably many sets, $\{A_i\}_{i\in\mathbb{N}}$ such that $A_i \in \mathcal{A}$ and $\mu(A_i) < \infty$.

Theorem 2.6.1 (Caratheodory)**.** Every premeasure μ on an algebra on a set X admits an extension to a measure on the σ-algebra $\langle \mathcal{A} \rangle_\sigma$. Moreover:

- The outer measure μ^* is such and extension and any extension ν satisfies that $\nu \leq \mu^*$.
- If μ is σ-finite, then the extension is unique and equal to σ^*.

To prove this, we need the following notion:

Definition 2.6.1. Let $\mathcal{A}$ be an algebra on a set X and μ a premeasure on $\mathcal{A}$. The **outer measure of** μ is the function

$$\mu^* : \mathcal{P}(X) \to [0, \infty]$$

defined by: For every $S \subseteq X$.

$$\mu^*(S) := \inf \left\{ \sum_{n \in \mathbb{N}} \mu(A_n) : \bigcup_{n \in \mathbb{N}} A_n \supseteq S \text{ and } \{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A} \right\}.$$

Proposition 13. The outer measure has the following properties:

(a) Monotone: $\mu^*(A) \leq \mu^*(B)$ for $A \subseteq B \subseteq X$.

(b) Countably subbaditive: $\mu^*(\bigcup_{n \in \mathbb{N}} S_n) \leq \sum_{n \in \mathbb{N}} \mu^*(S_n)$ for all $S, S_n \subseteq X$.

Proof. (a) follows from the definition of μ^* because a cover of B is also a cover of A. The same is true for (b), since the union of covers of the S_n is a cover of $\bigcup_{i \in \mathbb{N}} S_i$. □

Lemma 2.6.1. For any premeasure μ on an algebra $\mathcal{A}$, the outer measure μ^* is on A is equal to μ, i.e.

$$\mu^*|_{\mathcal{A}} = \mu.$$

Proof. Let $A \in \mathcal{A}$. and let $\{A_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$ be a cover of A. By disjointification, we may assume that all the A_n are disjoint; also by replacing A_n with $A_n \cap A$, we may assume that $A = \bigsqcup_{n \in \mathbb{N}} A_n$. But then, by countable additivity of μ, we have that

$$\mu(A) = \sum_{n \in \mathbb{N}} \mu(A_n),$$

so even with the original A_n, we had that

$$\mu(A) \leq \sum_{n \in \mathbb{N}} \mu(A_n)$$

by monotonicity. □

2.6.1 Caratheodory's Theorem: Existence

Every premeasure μ on an algebra $\mathcal{A}$ on a set X admits an extension to a measure on the σ-algebra $\langle \mathcal{A} \rangle_\sigma$. In fact, μ^* is such an extension.

In orde to show μ^* is countably additive on $\langle \mathcal{A} \rangle_\sigma$, it is enough to show that it is finitely additive since outer measures are countable subadditive, and finite additivity implies countable superadditivity.

Caratheodory's Proof:

Definition 2.6.2. A set B **conserves** a set S if

$$\mu^*(S) = \mu^*(S \cap B) + \mu^*(B^c \cap S).$$

Notice by that subadditivity of μ^*, failure of this equality means the left hand side is stricly less than the right hand side.

Moreover, call B **conservative** if it conserves every set.

Let $\mathcal{M}$ denote the collection of all conservative sets. Then, the proof goes as follows:

(i) $\mathcal{A} \subseteq \mathcal{M}$.

(ii) $\mathcal{M}$ is a sigma algebra (hence contains $\langle \mathcal{A} \rangle_\sigma$).

(iii) μ^* is (almost by definition of $\mathcal{M}$) finitely additive on $\mathcal{M}$.

The verification of the proof is left as homework.

Tao's Proof:

This proof works only for σ-finite premeasures, so assume μ is σ-finite on $\mathcal{A}$. We will first prove the result assuming that μ is finite, and deduce the σ-finite case from this as a homework exercise.

We define a **pseudo-metric** (i.e. a metric where the axiom $d(x, y) = 0$ implies that $x = y$ does not hold):

$$d_{\mu^*} : \mathcal{P}(X) \to [0, \mu(X)]$$

by:

$$d_{\mu^*}(A, B) = \mu^*(A \triangle B).$$

Remark 6 (The Secret of Symmetric Differences)**.** $\mathcal{P}(X)$ with $\triangle$ is an abelian group, with $\emptyset$ as the identity, and each element as its own inverse.

Proof. Just think of $\mathcal{P}(X)$ as 2^X , then $\triangle$ is just coordinate wise addition mod 2. □

Proposition 14. d_{σ^*} is a pseudo-metric.

Proof. Symmetry hold by definition and as well as the fact that

$$d_{\mu^*}(A, A) = \mu^*(A\triangle A) = \mu^*(\emptyset) = 0.$$

As for the triangle inequality, let $A, B, C \in\subseteq X$ and observe:

$$A\triangle C = (A\triangle B)\triangle(B\triangle C) \subseteq (A\triangle B) \cup (B\triangle C).$$

Therefore,

$$\begin{aligned} d_{\mu^*}(A, C) &= \mu^*(A\triangle C) \\ &\leq \mu^*\left((A\triangle B) \cup (B\triangle C)\right) \\ &= \mu^*(A\triangle B) + \mu^*(B\triangle C) \\ &= d_{\mu^*}(A, B) + d_{\mu^*}(B, C). \end{aligned}$$

□

Let now $\mathcal{M} = \overline{\mathcal{A}}^{d_{\mu^*}}$, i.e. the closure of $\mathcal{A}$ inside $\mathcal{P}(X)$ with respect to our pseudo-metric d_{μ^*}. We will show that $\mathcal{M}$ is a σ-algebra (hence $\mathcal{M} \supseteq \mathcal{A}$) and μ^* is finitely additive on $\mathcal{M}$.

Let us first recall a familiar definition:

Definition 2.6.3. A function f from a metric space (X, d_X) to a metric space (Y, d_Y) is said to be K**-Lipschitz** if

$$d_Y(f(x), f(y)) \leq K \cdot d_X(x, y).$$

Proposition 15. The function

$$\mu^* : \mathcal{P}(X) \to [0, \infty]$$

via

$$A \mapsto \mu^*(A)$$

is continuous (with respect to d_{μ^*}) and in fact is 1-Lipschitz.

Proof. Just note that $\mu^*(A) = \mu^*(A \triangle \emptyset) = d_{\mu^*}(A, \emptyset)$, so:

$$|\mu^*(A) - \mu^*(B)| = |d_{\mu^*}(A, \emptyset) - d_{\mu^*}(B, \emptyset)| \leq d_{\mu^*}(A, B).$$

□

Proposition 16. The function

$$\mathcal{P}(X) \mapsto \mathcal{P}(X)$$

via

$$A \mapsto A^c$$

is continuous, and in fact an isometry (i.e. a distance preserving function).

Proof. Just note that $A \triangle B = A^c \triangle B^c$, so

$$d_{\mu^*}(A, B) = \mu^*(A^c \triangle B^c) = d_{\mu^*}(A^c, B^c).$$

□

This implies that $\mathcal{M}$ is closed under compliments: if $M \in \mathcal{M}$, then there exists a sequence $(A_n) \subseteq \mathcal{A}$ with $A_n \to M$, but by continuity, $\lim_{n\to\infty} A_n^c = M^c$, so $M^c \in \mathcal{M}$.

Proposition 17. The function

$$\mathcal{P}(X)^2 \to \mathcal{P}(X)$$

via

$$(A, B) \mapsto A \cup B$$

is continuous, in fact, 1-Lipschitz with respefct to the $d_{\mu^*} + d_{\mu^*}$ metric on $\mathcal{P}(X)^2$. The same holds for $(A, B) \mapsto A \cap B$ because its a composition of 1-Lipschitz functions $()^c$ of $\cup$.

Proof.

$$\begin{aligned} d_{\mu^*}(A_1 \cup B_1, A_2 \cup B_2) &= \mu^*\left((A_1 \cup B_1) \triangle (A_2 \cup B_2)\right) \\ &\leq \mu^*\left((A_1 \triangle A_2) \cup (B_1 \triangle B_2)\right) && \text{[Monotonicity]} \\ &\leq \mu^*(A_1 \triangle A_2) + \mu^*(B_1 \triangle B_2) && \text{[Subadditivity]} \\ &= d_{\mu^*}(A_1, A_2) + d_{\mu^*}(B_1, B_2). \end{aligned}$$

We leave it as an exercise to check that

$$(A_1 \cup B_1) \triangle (A_2 \cup B_2) \subseteq (A_1 \triangle A_2) \cup (B_1 \triangle B_2).$$

□

This implies that $\mathcal{M}$ is closed under finite unions, hence making it an algebra: Let $A, B \in \mathcal{M}$, then there exists two sequences $(A_n), (B_n) \subseteq \mathcal{A}$ such that $A_n \to A$ and $B_n \to B$, so by continuity:

$$\lim_n (A_n \cup B_n) = A \cup B,$$

hence $A \cup B \in \mathcal{M}$.

Proposition 18. μ^* is finitely additive on $\mathcal{M}$.

Proof. Let $A, B \in \mathcal{M}$ which are disjoint, in order to show that $\mu^*(A \sqcup B) = \mu^*(A) + \mu^*(B)$. There is some $(A_n), (B_n) \subseteq \mathcal{A}$ with $A_n \to A$ and $B_n \to B$. By the continuity of union,

$$\lim(A_n \cup B_n) = A \cup B = A \sqcup B.$$

By the continuity of μ^*, we have that $\lim \mu^*(A_n) = \mu^*(A)$, $\lim \mu^*(B_n) = \mu^*(B)$ and $\lim \mu^*(A_n \cup B_n) = \mu^*(A \sqcup B)$. Recall that we have that $\mu^*|_{\mathcal{A}} = \mu$ which is finitely additive, so we have that

$$\mu^*(A_n \cup B_n) = \mu^*(A_n) + \mu^*(B_n) - \mu^*(A_n \cap B_n).$$

But since intersection is continuous, we have that $\lim A_n \cap B_n = A \cap B = \emptyset$. Once again by the continuity of μ^*, we have that

$$\lim \mu^*(A_n \cap B_n) = \mu^*(A \cap B) = \mu^*(A \cap B) = \mu^*(\emptyset) = 0.$$

Therefore,

$$\mu^*(A \sqcup B) = \lim \mu^*(A_n \cup B_n) = \lim \mu^*(A_n) + \lim \mu^*(B_n) = \mu^*(A) + \mu^*(B).$$

□

Proposition 19. $\mathcal{M}$ contains all countable unions of sets in $\mathcal{A}$.

Proof. Let $(A_n) \subseteq \mathcal{A}$, to show that $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$. By disjointification, we may assume that the A_n are pairwise disjoint. It is enough to show that

$$\lim_{n\to\infty} \left(\bigsqcup_{k\leq n} A_k \right) = \bigsqcup_{k\in\mathbb{N}} A_k =: A.$$

Observe that

$$d_{\mu^*}\left(\bigsqcup_{k\leq n} A_k, A\right) = \mu^*\left(\bigsqcup_{k>n} A_k\right) \leq \sum_{k>n} \mu^*(A_k) \to 0$$

as $n \to \infty$ because the series

$$\sum_{k\in\mathbb{N}} \mu^*(A_k)$$

converges; indeed: for all $n \in \mathbb{N}$,

$$\sum_{k\leq n} \mu^*(A_k) = \sum_{k\leq n} \mu(A_k) = \mu\left(\bigsqcup_{k\leq n} A_k \leq \mu(X) < \infty\right),$$

which is the only place we use the finiteness of μ. □

This implies that $\mathcal{M}$ is clsoed under countable unions, thus making it a sigma algebra. indeed: if $(M_n) \subseteq \mathcal{M}$, let $\varepsilon > 0$ and take $A_n \in \mathcal{A}$ so $A_n \approx_{\varepsilon\cdot 2^{-(n+1)}} M_n$, i.e.

$$d_{\mu^*}(A_n, M_n) \leq 2^{-(n+1)} \cdot \varepsilon.$$

But then, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$ and

$$d_{\mu^*}\left(\bigcup_n A_n, \bigcup_n M_n\right) \leq \sum_{n\in\mathbb{N}} d_{\mu^*}(A_n, M_n) \leq \varepsilon$$

So $\bigcup_{n\in\mathbb{N}}$ is ε-close to an element of $\mathcal{M}$. But ε is arbitrary and $\mathcal{M}$ is closed (topologically speaking), hence $\bigcup_{n\in\mathbb{N}} M_N \in \mathcal{M}$.

2.6.2 Caratheodory's Extension: Uniqueness

Let $\mathcal{A}$ be an algebra on a set X and μ a premeasure on $\mathcal{A}$. Then for each extension ν of μ to a measure of $\langle\mathcal{A}\rangle_\sigma$, we have that $\nu \leq \mu^*$. If μ is σ-finite, then $\nu = \mu^*$.

Proof. Since μ^* is defined as the infimum over covers by sets in $\mathcal{A}$, we fix a set $S \in \langle\mathcal{A}\rangle_\sigma$ and a cover $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ of S, and we show that

$$\nu(S) \leq \sum_{n\in\mathbb{N}} \mu(A_n).$$

But this follows from the countable subadditivity of ν:

$$\nu(S) \leq \nu\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n\in\mathbb{N}} \nu(A_n) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

This gives us that $\nu \leq \mu^*$.

Now assume ν is σ-finite. It is actually enough to show that $\nu = \mu^*$ assuming μ is finite because given a witness to σ-finiteness, i.e. a partition $X = \bigsqcup_{n\in\mathbb{N}} X_n$ with each $X_n \in \mathcal{A}$ and $\mu(X_n) < \infty$, the fact that $\nu|_{X_n} = \nu^*|_{X_n}$ for all n implies that

$$\begin{aligned}
\nu(S) = \nu\left(\bigsqcup_n (S \cap X_n)\right) &= \sum_n \nu(S \cap X_n) \\
&= \sum_n \mu^*(S \cap X_n) \\
&= \mu^*\left(\bigsqcup_n (S \cap X_n)\right) \\
&= \mu^*(S)
\end{aligned}$$

For each $S \in \langle \mathcal{A} \rangle_\sigma$.

Thus, suppose that μ is finite. We show that the function

$$S \mapsto \nu(S) : \langle \mathcal{A} \rangle_\sigma \to [0, \mu(X)]$$

is continuous with respect to the pseudo-metric d_{μ^*}. indeed, it is 1-lipschitz:

$$\begin{aligned}
|\nu(S_1) - \nu(S_2)| &\leq \nu(S_1 \setminus S_2) + \nu(S_2 \setminus S_1) \\
&= \nu(S_1 \triangle S_2) \\
&\leq \mu^*(S_1 \triangle S_2) \\
&= d_{\mu^*}(S_1, S_2).
\end{aligned}$$

So ν and μ^* are continuous functions on $\langle \mathcal{A} \rangle_\sigma$ which coincide on a set $\mathcal{A}$ which is dense in $\langle \mathcal{A} \rangle_\sigma$ with respect to d_{μ^*} since $\langle \mathcal{A} \rangle_\sigma \subseteq \bar{\mathcal{A}}^{d_{\mu^*}}$. Thus, $\nu = \mu^*$ everywhere on $\langle \mathcal{A} \rangle_\sigma$. □

As such, there are unique measures extending the Bernoulli and Lebesgue premeasures, and we call them Bernoulli and Lebesgue measures. By Bernoulli, we mean the premeasure obtained on clopen sets on $A^{\mathbb{N}}$ for finite A and any probability measure m on A. The corresponding Bernoulli measure is denoted by $m^{\mathbb{N}}$.

Definition 2.6.4. For a metric/topological space X, a **Borel measure** is any measure defined on the borel σ-algebra $\mathcal{B}(X)$.

Example 4. ADD EXAMPLES

2.7 Null and Measurable Sets

Definition 2.7.1. Let $(X, \mathcal{B}, \mu)$ be a measure space. A set $A \subseteq X$ is called **μ-null** if there is some $B \in \mathcal{B}$ such that $A \subseteq B$ and $\mu(B) = 0$. Denote the family of all μ-null sets by $Null_\mu$.

Remark 7. μ-null sets form a **σ-ideal**, i.e. they are closed under subsets (downwards) and under countable unions. In particular, if Z is μ-null then $\mathcal{P}(Z) \subseteq Null_\mu$.

Proof. If the Z_n are μ-null, then $Z_n \subseteq \tilde{Z}_n \in \mathcal{B}$ and $\mu(\tilde{Z}_n) = 0$, so

$$\bigcup_{n \in \mathbb{N}} Z_n \subseteq \bigcup_{n \in \mathbb{N}} \tilde{Z}_n$$

and

$$\mu\left(\bigcup_{n \in \mathbb{N}} \tilde{Z}_n\right) = \sum_{n \in \mathbb{N}} \mu(\tilde{Z}_n) = 0.$$

□

Definition 2.7.2. For any sets $A, B \subseteq X$, write $A =_\mu B$ if $A \triangle B$ is μ-null. Call a set $A \subseteq X$ **μ-measurable** if $A =_\mu B$ for some $B \in \mathcal{B}$. Denote by $Meas_\mu$ the collection of all μ-measurable sets.

Proposition 20. Let $(X, \mathcal{B}, \mu)$ be a measure space. Then, $Meas_\mu$ is a σ-algebra. In fact,

$$Meas_\mu = \langle \mathcal{B} \cup Null_\mu \rangle_\sigma.$$

Proof. For complements, we have that $A \triangle B$ is null if and only if $A^c \triangle B^c$ is null because $A \triangle B = A^c \triangle B^c$. As for countable unions, if $A_n \triangle B_n$ is null and $B_n \in \mathcal{B}$, then:

$$\left(\bigcup_{n\in\mathbb{N}} A_n\right) \triangle \left(\bigcup_{n\in\mathbb{N}} B_n\right) \subseteq \bigcup_{n\in\mathbb{N}} (A_n \triangle B_n),$$

and the right hand side is null. Thus, $Meas_\mu \supseteq \langle \mathcal{B} \cup Null_\mu \rangle_\sigma$ and the other inclusion follows by the definition of μ-measurable sets (notice $M = B \triangle (M \triangle B)$. □

Remark 8. One can show that $Meas_\mu$ is what we obtain in both Caratheodory's and Tao's proof of Caratheodory extension (left as homework).

Proposition 21. Let $(X, \mathcal{B}, \mu)$ be a measure space. Then,

$$\{B \sqcup Z : B \in \mathcal{B}, \text{ and } Z \text{ is } \mu\text{-null}\} = Meas_\mu = \{B \backslash Z : B \in \mathcal{B}, \text{ and } Z \text{ is } \mu\text{-null}\}$$

Proof. Since $B \cup Z$ and $B \setminus Z$ are μ-measurable, it is enough to show that every μ-measureable set is of those two forms. Let M be a μ-measurable set, so $M \triangle B =: Z$ is μ-null for some $B \in \mathcal{B}$. Thus, $M = B \triangle Z$. Let $\tilde{Z} \supseteq Z$ be in $\mathcal{B}$ and such that $\mu(\tilde{Z}) = 0$. Let

$$B' = B \setminus \tilde{Z}$$

and

$$\tilde{B} = B \cup \tilde{Z}.$$

Then,

$$B' \sqcup (B \cap (\tilde{Z} \setminus Z)) \sqcup (B^c \cap Z) = M = \tilde{B} \setminus (B \cap Z) \setminus (B^c \cap (\tilde{Z} \setminus Z))$$

□

Corollary 2.7.1. For any μ-measurable set M, there are some $B_0, B_1 \in \mathcal{B}$ such that $B_1 \supseteq M \supseteq B_0$ and $\mu(B_0) = \mu(B_1)$, i.e. $B_0 \triangle M$ and $M \triangle B_1$ are μ-null.

Definition 2.7.3. A measure space $(X, \mathcal{B}, \mu)$ is called **complete** if $\mathcal{B} = Meas_\mu$.

Proposition 22. Every measure μ on a measurable space $(X, \mathcal{B})$ admits a unique **completion**, i.e. a unique extension to a measure on $Meas_\mu$.

Proof. **Existence:** Let M be μ-measurable so $M = B \triangle Z$ for some $B \in \mathcal{B}$ and Z is μ-null. Then define the extension as

$$\overline{\mu}(M) := \mu(B).$$

This is well defined since if $M = B_0 \triangle Z_0 = B_1 \triangle Z_1$, where $B_i \in \mathcal{B}$ and Z_i are μ-null, then

$$B_0 \triangle B_1 = (M \triangle Z_0) \triangle (M \triangle Z_1) = Z_0 \triangle Z_1 \subseteq Z_0 \cup Z_1,$$

so, $B_0 =_\mu B_1$ and hence $\mu(B_0) = \mu(B_1)$.

Uniqueness: Any extension ν satisfies that $\nu(Z) = 0$ for all $Z \in Null_\mu$ by monotonicity,so whenever $M = B \triangle Z$ with $B \in \mathcal{B}$ and $Z \in Null_\mu$, we must have that

$$\nu(M) = \nu(B) + \nu(Z) - 2\nu(Z \cap B) = \nu(B) = \mu(B).$$

□

2.8 Non-measurable Sets

We will give an example of a non-Lebesgue-measurable subset of $\mathbb{R}$.

Definition 2.8.1. Let E be an equivalence relation on a set X. A **transversal** for E is a set $Y \subseteq X$ which meets each E-class in exactly one point. A **selector** for E is a map

$$s : X \to X$$

such that $s(x) \in [x]_E$ and

$$xEy \iff s(x) = s(y),$$

for all $x, y \in X$. For a selector s, we can get a transversal $Y := s(X)$, and vice versa, from a transversal Y, we can get a selector by setting $s(x)$ to be equal to the unique $y \in Y \cap [x]_E$.

Selectors and transversals exist by the Axiom of Choice, but this typically results in ill-behaved functions and sets, for example, non-measureability.

Example 5. Let $E_{\mathbb{Q}}$ denote the **Vitali equivalence relation** on $\mathbb{R}$ defined by

$$xE_{\mathbb{Q}}y \iff y - x \in \mathbb{Q}.$$

This is simply the coset equivalence relation of $\mathbb{Q}$ as a subgroup of $\mathbb{R}$ under addition. Also, this is the orbit equivalence relation of the action of $\mathbb{Q}$ on $\mathbb{R}$ by translation. For each $x \in \mathbb{R}$, the class $[x]_{E_{\mathbb{Q}}} = x + \mathbb{Q}$, in particular, it intersects $[0, 1]$.

We claim that any transversal $Y \subseteq [0, 1]$ of $E_{\mathbb{Q}}$ is non-measureable with respect to the Lebesgue measure λ. Indeed, observe that

$$[0,1] \subseteq \bigsqcup_{q\in\mathbb{Q}\cap[-1,1]} q + Y \subseteq [-1, 2].$$

If Y was measurable, so would its translates $q + Y$ (Lebesgue measure is translation-invariant) and so $\lambda(y + Y) = \lambda(y)$. Therefore,

$$\begin{aligned} 1 = \lambda\left([0,1]\right) &\leq \lambda\left(\bigsqcup_{q\in\mathbb{Q}\cap[-1,1]} q + Y\right) \\ &= \sum_{q\in\mathbb{Q}\cap[-1,1]} \lambda(q + Y) \\ &= \infty \cdot \lambda(Y) \\ &\leq \lambda\left([-1, 2]\right) \\ &= 3, \end{aligned}$$

a contradiction (think about it).

Remark 9. ADD REMARK

2.9 Pocket Tools for Working with Measures

Proposition 23 (Monotone Convergence)**.** Let $(X, \mathcal{B}, \mu)$ be a measure space.

(a) $\mu(\cup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A_n)$ for all μ-measurable A_n with $A_n \subseteq A_{n+1}$.

(b) $\mu(\cap_{n\in\mathbb{N}} B_n) = \lim_{n\to\infty} \mu(B_n)$ for all μ-measurable A_B with $B_n \supseteq B_{n+1}$ with $\mu(B_0) < \infty$.

Remark 10. Caution about (b)

Proof. (a): We disjointify: $A_0^{'} := A_0$ and $A_n^{'} = A_n \setminus A_{n-1}$, so

$$\bigsqcup_{n\in\mathbb{N}} A_n^{'} = \bigcup_{n\in\mathbb{N}} A_n,$$

hence

$$\begin{aligned}
\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) &= \mu\left(\bigsqcup_{n} A_n^{'}\right) \\
&= \sum_{n\in\mathbb{N}} \mu(A_n^{'}) \\
&= \lim_{N\to\infty} \sum_{n\leq N} \mu(A_n^{'}) \\
&= \lim_{N\to\infty} \mu\left(\bigsqcup_{n\leq N} A_n^{'}\right) \\
&= \lim_{N\to\infty} \mu(A_N).
\end{aligned}$$

(b): The sets $A_n := B_0 \setminus B_n$ are increasing, so by (a), we have that

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} (\mu(B_0) - \mu(B_n)) = \mu(B_0) - \lim_{n\to\infty} \mu(B_n),$$

where we used that $\mu(B_n) \leq \mu(B_0) < \infty$ in the second equality. On the other hand:

$$\mu\left(\bigcup_{n} A_n\right) = \mu\left(B_0 \setminus \bigcap_{n} B_n\right) = \mu(B_0) - \mu\left(\bigcap_{n} B_n\right).$$

Rearranging the previous two equations yields our equality. □

Lemma 2.9.1 (Borel-Cantelli Lemmas)**.** Let $(X, \mathcal{B}, \mu)$ be a measure space. Let (A_n) be a sequence of μ-measurable sets.

(a) If $\sum_{n\in\mathbb{N}} \mu(A_n) < \infty$, then almost every $x \in X$ is eventually not in A_n, i.e. the set

$$\limsup A_n = \{x \in X : \exists_n^\infty, x \in A_n\} = \bigcap_{m\in\mathbb{N}} \bigcup_{n\geq m} A_m$$

is μ-null (where $\exists_n^\infty \equiv \forall m \exists n \geq m$).

(b) Measure Compactness: Suppose $\mu(X) < \infty$. Then if there is some $\delta > 0$ such that $\mu(A_n) \geq \delta$ for all $n \in \mathbb{N}$ then

$$\mu\left(\limsup A_n\right) \geq \delta.$$

Proof. (a): note taht $\limsup A_n \subseteq \bigcup_{n \geq m} A_n$ for all $m \in \mathbb{N}$ (since lim sup is decreasing) so:

$$\mu(\limsup A_n) \leq \mu\left(\bigcup_{n \geq m} A_n\right) \leq \sum_{n \geq m} \mu(A_n) \to 0$$

as $m \to \infty$.

(b): Since $\mu(X) < \infty$, then

$$\mu(\limsup A_n) = \mu\left(\bigcap_{m \in \mathbb{N}} \bigcup_{n \geq m} A_n\right) = \lim_{m \to \infty} \mu\left(\bigcup_{n \geq m} A_n\right) \geq \delta$$

□

Definition 2.9.1. Let $(X, \mathcal{B}, \mu)$ be a measure space. A sequence (V_n) of μ-measurable sets is called **vanishing** (resp. **almost vanishing**) if (V_n) is decreasing and $\bigcap_{n \in \mathbb{N}} V_n$ is empty (resp. null).

Proposition 24. Let $\mathcal{F}$ be a collection of μ-measurable sets that is closed under countable unions. If $\mathcal{F}$ contains positive measure sets of arbitrarily small measure, then $\mathcal{F}$ contains an almost vanishing sequence of positive measure sets.

Proof. For each $n \in \mathbb{N}$, let $A_n \in \mathcal{F}$ be a positive measure set with $\mu(A_n) \leq 2^{-n}$. The sets A_n may not be decreasing,but the sets $V_n := \bigcup_{m \geq n} A_m$ are decreasing and

$$\bigcap_{n \in \mathbb{N}} V_n = \limsup A_n$$

which is null by Borel-cantelli because

$$\sum_{n \in \mathbb{N}} \mu(A_n) < \infty.$$

□

Definition 2.9.2. Let $(X, \mathcal{B}, \mu)$ be a measure space and P a property of points in X. Then, we say that P holds **almost everywhere**(abbreviated as **a.e**) in X if the set $\{x \in X : \ x \text{ satisfies } P\}$ is co-null.

2.10 Measure Exhaustion

Definition 2.10.1. In a measure space, call a collection $\mathcal{C}$ of sets **almost disjoint** if the pairwise intersections of sets in $\mathcal{C}$ are null.

Proposition 25 (Countable Pigeonhole principle)**.** Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space. Then any almost disjoint collection $\mathcal{C}$ of μ-measurable positive measure sets is countable.

Proof. We first prove it for the $\mu(X) < \infty$ case. Then, for each $n \in \mathbb{N}^+$, the set

$$\mathcal{C}_n = \{C \in \mathcal{C} : \mu(C) \geq 1/n\}$$

is finite (in fact it has at most $n \cdot \mu(X)$ elements) and $\mathcal{C} = \bigcup_{n \in \mathbb{N}^+} \mathcal{C}_n$, so $\mathcal{C}$ is countable.

For the general σ-finite case, let $X = \bigsqcup_{n \in \mathbb{N}} X_i$ where each $X_n \in \mathcal{B}$ and is of finite measure. Define

$$\mathcal{D}_n = \{C \in \mathcal{C} : \mu(C \cap X_n) > 0\}.$$

Then by the finite case, each $\mathcal{D}_n$ is countable and since $\mathcal{C} = \bigcup_{n \in \mathbb{N}} \mathcal{D}_n$, it is countable. $\square$

Proposition 26 (Transfinite Measure Exhaustion)**.** Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space and let $(A_\alpha)_{\alpha < \omega_1}$ be an increasing sequence of μ-measurable sets where ω_1 is the first uncountable ordinal. Then the sequence almost stabilizes at some countable ordinal γ, i.e. for all $\alpha \geq \gamma$, $A_\alpha =_\mu A_\gamma$.

Proof. We fist disjointify: $A'_\alpha := A_\alpha \setminus [\bigcup_{\beta < \alpha} A_\beta]$. So $\{A'_\alpha\}_{\alpha < \omega_1}$ is an almost disjoint collection, hence all but countably many of A_α are null by countable pigeonhole, i.e. there is some countable ordinal γ such that A'_α is null for all $\alpha > \gamma$, hence $A_\alpha =_\mu A_\gamma$ because $A_\alpha \setminus A_\gamma = \bigcup_{\gamma < \beta < \alpha} A'_\beta$ is null being a countable union of null sets. $\square$

Remark 11. This allows us to run transfinite algorithms which at each step handle a positive measure set. Then we know the algorithm will stop at a countable stage, having handled a co-null set.

We now discuss an important application. In a measure space with atoms, we can't achieve every value of measure between 0 and $\mu(X)$, but this is the only abstraction.

Theorem 2.10.1 (Sierpinski Theorem)**.** In an atomless measure space $(X, \mathcal{B}, \mu)$, every value $0 < r \leq \mu(X)$ is achieved, i.e. there is a $B \in \mathcal{B}$ with $\mu(B) = r$.

First we prove a more humble statement.

Proposition 27. Every positive measure set Y contains positive measure sets of arbitrarily small measure.

Proof. Since Y is not an atom there is some $X_\emptyset \subseteq Y$ with $\mu(X_\emptyset) < \mu(Y)$. We build a sequence $(X_s)_{s \in 2^{<\mathbb{N}}}$ of positive measure sets such that $X_s = X_{s0} \sqcup X_{s1}$ as follows: if X_s is already defined, it is not an atom, so there is some $X_{s0} \subseteq X_s$ in $\mathcal{B}$ with $0 < \mu(X_{s0}) < \mu(X_s)$. Let $X_{s1} = X_s \setminus X_{s0}$. For each $s \in 2^{<\mathbb{N}}$, one of X_{s0} and X_{s1} must have measure at most half of $\mu(X_s)$. Looking at the tree we get from the sequence, this gives an infinite branch $(X_{s_n})_{n \in \mathbb{N}}$ in the tree of positive measure with $\mu(X_{s_n}) \leq 2^{-n}\mu(X_\emptyset)$. $\square$

Iteratively using the previous proposition, we now explicitly build a set $B \in \mathcal{B}$ with $\mu(B) = r$. We present two proofs, one via transfinite exhaustion and the other via a 1/2-greedy algorithm, with the latter one being the more preferable one of the two.

Proof Via transfinite exhaustion. Define a sequence $(A_\alpha)_{\alpha<\omega} \subseteq \mathcal{B}$ of pairwise disjoint sets such that $\mu\left(\bigsqcup_{\alpha<\beta} A_\alpha\right) \leq r$ for each $\beta < \omega_1$ by induction as follows: If $(A_\alpha)_{\alpha<\beta}$ is already defined, let A_β be a positive measure subset of $X \setminus \bigcup_{\alpha<\beta} A_\alpha$ of measure at most $r - \mu\left(\bigsqcup_{\alpha<\beta} A_\alpha\right)$ if $r - \mu\left(\bigsqcup_{\alpha<\beta} A_\alpha\right) > 0$. Otherwise, put $A_\beta := \emptyset$. Now the proof of countable pigeonhole for measures (using the condition that $\mu\left(\bigsqcup_{\alpha<\beta} A_\alpha\right) < r$, for all $\beta < \omega_1$, instead of the finiteness of μ) gives that all but countably many of the A_α are null, i.e. there is some $\beta < \omega_1$ with A_α being null for all $\alpha \geq \beta$. Thus, $\mu\left(\bigsqcup_{\alpha<\beta} A_\alpha\right) = r$. $\square$

proof via 1/2*-greedy algorithm.* We inductively build a sequence $(B_n)_{n\in\mathbb{N}} \subseteq \mathcal{B}$ of pairwise disjoint sets such that $\mu(\bigsqcup_{i\leq n} B_i) \leq r$. Suppose that $(B_i)_{i<n}$ is already defined, then take $B_n \in \mathcal{B}$ to be any set such that

$$\mu(B_n) \geq \frac{1}{2} \sup \left\{ B \in \mathcal{B} : B \subseteq X \setminus \bigsqcup_{i<n} B_i, \text{ and } \mu(B) \leq r - \mu(\bigsqcup_{i<n} B_i) \right\}.$$

Now that $(B_n)_{n\in\mathbb{N}}$ is defined, by monotone convergence,

$$\sum_{n\in\mathbb{N}} \mu(B_n) = \mu\left(\bigsqcup_{n\in\mathbb{N}} B_n\right) \leq r,$$

namely, the series is summable and so $\lim_{n\to\infty} \mu(B_n) = 0$. We check that the set $B_\infty := \bigsqcup_{n\in\mathbb{N}} B_n$ has measure r. Indeed, otherwise, $\mu(B_\infty) < r$, so by the previous proposition, there is some $B' \subseteq X \setminus B_\infty$ in $\mathcal{B}$ such that $0 < \mu(B') < r - \mu(B_\infty)$. But taking n to be large enough so that $\mu(B_n) < 1/2\mu(B')$, we get a contradiction with the choice of B_n. □

2.11 Approximating Measurable Sets

2.11.1 99% Lemma

We begin with the following observation.

Proposition 28 (Percentage of Carrots in Soup)**.** Let $(X, \mathcal{B}, \mu)$ be a measure space and let A, B be μ-measurable sets with $0 < \mu(B) < \infty$. Then for any (percentage) $p \in (0, 1)$ and any (finite or countable) partition $B = \bigsqcup_{n<N} B_i$, where $N \in \mathbb{N} \cup \infty$, we have that if

$$\frac{\mu(A \cap B)}{\mu(B)} \geq p$$

then

$$\frac{\mu(A \cap B_n)}{\mu(B_n)} \geq p$$

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

Proof. Notice that

$$\frac{\mu(A \cap B)}{\mu(B)} = \sum \cdot \frac{\mu(B_n)}{\mu(B)} \cdot \frac{\mu(A \cap B_n)}{\mu(B_n)}$$

and we know that

$$\sum \frac{\mu(B_n)}{\mu(B)} = 1$$

so it is a convex combination. □

Lemma 2.11.1 (99% Lemma)**.** Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space and let $\mathcal{C} \subseteq \mathcal{B}$ be a collection of sets whose finite disjoint unions form an algebra generating $\mathcal{B}$. Then each positive measure set $M \subseteq X$ admits a set $C \in \mathcal{C}$ whose 99% is M, i.e. For all $\varepsilon > 0$, there is some $C \in \mathcal{C}$ such that

$$\frac{\mu(M \cap C)}{\mu(C)} \geq 1 - \varepsilon.$$

Proof. By the uniqueness of Caratheodory theorem, we have that $\mu = (\mu|_{\langle \mathcal{C} \rangle_\sigma})^*$ and thus

$$\mu(M) = \inf \left\{ \sum_{k \in \mathbb{N}} \mu C_k : \bigsqcup_k C_k \supseteq M, \text{ and } \{C_k\} \subseteq \mathcal{C} \right\}.$$

Using σ-finiteness, M has a μ-measurable subset of positive finite measure, so by shrinking M, we may assume that $\mu(M) < \infty$. Then, there exists some $\{C_k\} \subseteq \mathcal{C}$ such that $\bigsqcup_{k \in \mathbb{N}} C_k \supseteq M$, and that

$$\frac{\mu(M)}{\mu(\bigsqcup_{k \in N} C_k)} \geq 1 - \varepsilon,$$

so by Carrot-soup observation there is some $k \in \mathbb{N}$ such that

$$\frac{\mu(M \cap C_k)}{\mu(C_k)} \geq 1 - \varepsilon.$$

$\square$

Example 6. Two familiar examples are:

(a) For $(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right), \lambda)$ we take $\mathcal{C}$ to be the collection of boxes, therefore we get that every positive measure set contains 99% of a box C.

(b) For $(A^{\mathbb{N}}, \mathcal{B}\left(A^{\mathbb{N}}\right), \mu)$, where A is finite and μ is Bernoulli, we take $\mathcal{C}$ to be the collection of cylinders, and thus every positive measure set contains 99% of a cylinder C.

Remark 12. In both of these examples, we can take the box/cylinder C to be arbitrarily small (both small in diameter and small in measure because each box/cylinder partitions into arbitrarily small finitely many) boxes/cylinders, hence the carrot-soup observation applies.

2.11.2 Application: Ergodicity

Definition 2.11.1. Let (X, μ) be a measure space and E be an equivalence relation on X. The relation E is called **ergodic** (with respect to μ) if every E-invariant (i.e. union of E-classes) μ-measurable set is null or conull. In other words, X is not decomposable into two E-invariant positive measure sets.

Example 7. We start with a couple relevant examples of equivalence relations.

(a) Let Γ be a countable group acting on a measure space $(X, \mathcal{B}, \mu)$ so that $\gamma \cdot B \in \mathcal{B}$ for all $\gamma \in \Gamma$, and $B \in \mathcal{B}$. For instance, translations actions $\mathbb{Z} \curvearrowright \mathbb{R}$ or $\mathbb{Q} \curvearrowright \mathbb{R}$, or dialations $(\mathbb{Q}^+, \cdot) \curvearrowright \mathbb{R}^d$. Then the orbit equivalence relation on X of this action, denoted E_Γ and is defined by:

$$xE_\gamma y \iff x \text{ and } y \text{ are in the same } \Gamma\text{-orbit.}$$

(b) Let $(X, \mathcal{B}, \mu)$ be a measure space and $T : X \to X$ not necessarily a bijection. We will assume that T is "μ-measurable". Its orbit equivalence relation, denoted E_T, is defined by

$$xE_T y \iff T^n x = T^m y \text{ for some } n, m \in \mathbb{N}.$$

We draw an edge from $X \to T(x)$. Then the T-orbits are exactly the connected components of this graph, which is the graph of T as a subset of $X \times X$.

Example 8. Now we present some Ergodic and non Ergodic equivalence relations.

(a) Non-ergodic: Let $Z \curvearrowright \mathbb{R}$ by translation: $z \times r = z + r$, for $z \in Z$ and $r \in \mathbb{R}$. Then the orbit equivalence relation is just the coset equivalence relation $Z \subseteq \mathbb{R}$. The orbit of $y \in \mathbb{R}$ is $x + \mathbb{Z}$. Then, $A := (0, 1/2) + \mathbb{Z} = \bigsqcup_{n \in \mathbb{Z}} (n, n+1/2)$ is $E_\mathbb{Z}$-invariant, but it and its compliment have positive measure, so $E_\mathbb{Z}$ is not λ-ergodic, where λ is the Lebesgue measure. Note that $E_\mathbb{Z}$ admits a measureable transversal, for instance $[0, 1)$.

(b) Ergodic: Let $\mathbb{Q} \curvearrowright \mathbb{R}$ by translation, so its orbit equivalence relation $E_\mathbb{Q}$ is the coset equivalenence of $Q \subseteq \mathbb{R}$. Recall that $E_\mathbb{Q}$ does not admit a measurable transversal(via the Vitali construciton) and the reason fir this is that $E_\mathbb{Q}$ is ergodic, which we will prove using the 99% lemma.

Proposition 29. $E_{\mathbb{Q}}$ is ergodic.

Proof. Suppose otherwise. So there is a positive measure set $A \subseteq \mathbb{R}$ with $B := \mathbb{R} \setminus A$ being of positive measure. By the 99% lemma, there is a positive interval J whose 99% is B. Once again by the 99% lemma, there is a positive measure interval I whose 99% is A and moreover, $lh(I) < lh(J)$. Using that rationals are dense, we can cover at least half of J by finitely many pairwise disjoint rational translates of I, i.e. $\bigsqcup_{i<k}(q_i + I) \subseteq J$ and $\mu(\bigsqcup_{i<K} q_i + I) \geq 1/2\lambda(J)$. Since $q_i + A = A$ for all i, we have that 99% of each $q_i + I$ is still A. So at least half of 99% of J is A, contradicting that at most 1% of J is A. $\square$

2.12 Regularity of Measures

Definition 2.12.1. Let $(X, \mathcal{B}, \mu)$ be a measure space and X a metric space.Then μ is called **regular** if each μ-measurable set M satisfies the following two conditions:

- **Outer regularity:** $\mu(M) = \inf\{\mu(\mathcal{U}) : \mathcal{U} \supseteq M \text{ is open}\}$,
- **Inner regularity:** $\mu(M) = \sup\{\mu(C) : C \subseteq M \text{ is closed}\}$.

Moreover, μ is called **strongly regular** if

$$0 = \inf\{\mu(\mathcal{U} \setminus M) : \mathcal{U} \supseteq M \text{ is open}\} = \inf\{\mu(M \setminus C) : C \subseteq M \text{ is closed}\}.$$

Remark 13. All finite regular measures are strongly regular.

Proposition 30. If μ is strongly regular, then every measurable set M is $=_\mu$ to a G_δ set and $=_\mu$ to a F_σ set; More precisely, there is a G_δ set G and a F_σ set F such that $F \subseteq M \subseteq G$ and $F =_\mu M =_\mu G$.

Proof. By strong regularity, for each $n \in \mathbb{N}^+$, there is an open set U_n and a closed set C_n such that $C_n \subseteq M \subseteq U_n$ and $\mu(M \setminus C_n)$, $\mu(U_n \setminus M)$ are both at most $1/n$. Let

$$G := \bigcap_{n \in \mathbb{N}} U_n$$

and

$$F := \bigcup_{n \in \mathbb{N}} C_n.$$

Thus, $F \subseteq M \subseteq G$ and

$$\mu(M \setminus F) \leq \mu(M \setminus C_n) \leq 1/n \to 0$$

and

$$\mu(G \setminus M) \leq \mu(U_n \setminus M) \leq 1/n \to 0,$$

as $n \to \infty$, and so $\mu(F) = \mu(M) = \mu(G)$. □

Theorem 2.12.1. Every finite Borel measure μ on a metric space X is strongly regular.

Proof. Let $\mathcal{S}$ be the collection of all μ-measurable sets $M \subseteq X$ which satisfy:

$$\begin{aligned} 0 &= \inf\{\mu(U \setminus M) : U \supseteq M \text{ is open}\} \\ &= \inf\{\mu(M \setminus C) : C \subseteq M \text{ is closed}\}. \end{aligned}$$

First, we claim $\mathcal{S}$ contains all open sets.

Recall that open sets are F_σ in metric spaces, so for an open set $U \subseteq X$, we have that

$$U = \bigsqcup_{n \in \mathbb{N}} C_n$$

where the C_n are closed. Replacing each C_n with

$$\bigcup_{i \leq n} C_i,$$

we may assume that U is an increasing union of C_n's. But then by monotone convergence, $\mu(U) = \lim_{n \to \infty} \mu(C_n)$.

Secondly, we show $\mathcal{S}$ is an algebra: indeed, complement of open/closed is closed/open. Also finite unions of open/closed is open/closed.

Furthermore, $\mathcal{S}$ is closed under countable unions, and is hence a σ-algebra. Let $M := \bigcup_{n \in \mathbb{N}} M_n$ where $M_n \in \mathcal{S}$. Since $\mathcal{S}$ is closed under finite unions, we may replace M_n with $\bigcup_{i \leq n} M_i$ and assume that M is an increasing union of M_n's.

For outer regularity, let $U_n \subseteq M_n$ be open an such that

$$\mu(U_n \setminus M_n) \leq \varepsilon \cdot 2^{-(n+1)}$$

. Then, $U := \bigcup_{n \in \mathbb{N}} U_n$ is open, and

$$\mu(U \setminus M) \leq \mu\left(\bigcup_{n \in \mathbb{N}} [U_n \setminus M_n]\right) \leq \sum_{n \in \mathbb{N}} \mu(U_n \setminus M_n) \leq \varepsilon.$$

For inner regularity, let $C_n \subseteq M_n$ be closed and such that $\mu(C_n) \approx_{1/n} \mu(M_n)$. Because M is an increasing union, by monotone convergence we have that $\mu(M) = \lim \mu(M_n)$ and so for large enough $n \in \mathbb{N}$, we have that

$$\begin{aligned} |\mu(M) - \mu(C_n)| &\leq |\mu(M) - \mu(M_n)| + |\mu(M_n) - \mu(C_n)| \\ &\leq \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon. \end{aligned}$$

Thus, $\mathcal{S}$ contains all Borel sets. For a μ-measurable $M \subseteq X$, let $B_0 \subseteq M \subseteq B_1$ be borel sets with $B_0 =_\mu M =_\mu B_1$ and let $U \supseteq B_1$ be an open set such that $\mu(U) \approx_\epsilon \mu(B_1) = \mu(M)$, and $C \subseteq B_0$ a closed set such that $\mu(C) \approx_\varepsilon \mu(B_0) = \mu(M)$. So $\mathcal{S}$ contains all μ-measurable sets. $\square$

Remark 14. It is not true that σ-finite Borel measures on metric spaces are regular.

Example 9. σ-finite counter example:

$$ADD\ MEEE$$

Definition 2.12.2. Let X be a Hausdorff topological space (e.g. a metric space) and let μ be a Borel meausre on X. We say that μ is

- **σ-finite by open sets** if $X = \bigcup_{n\in\mathbb{N}} U_n$ where U_n is open and has finite μ-measure.
- **finite on compact sets** if each compact set has finite μ-measure.
- **locally finite** if every point $x \in X$ admits a neighbourhood V (i.e. $x \in Int(V)$) of finite μ-measure (in particular, an open neighbourhood of finite measure).

Corollary 2.12.1. For a metric space X, every Borel measure that's σ-finite by open sets is strongly regular.

Proof. Let $X = \bigcup_{n\in\mathbb{N}} U_n$ where U_n is open and of finite measure. Let $M \subseteq X$ be μ-measurable. For each $n \in \mathbb{N}$, viewing U_n as a metric space with $\mu|_{U_n}$ a finite Borel measure, we get a set $V_n \subseteq U_n$ open relative to U_n (hence open in X since U_n is open) such that $V_n \supseteq M \cap U_n$, and $\mu(V_n) \approx_{\varepsilon\cdot 2^{-(n+1)}} \mu(M \cap U_n)$. Thus, $V = \bigcup_{n\in\mathbb{N}} V_n$ is open in X and

$$\mu(V \setminus M) \leq \mu\left(\bigcup_n [V \setminus M]\right) \leq \varepsilon.$$

This handles strong regularity outer regularity.

For strong inner regularity, let $U \supseteq M^C$ be an open set with $\mu(U \backslash M^c) \leq \varepsilon$. But U^c is closed and $U \setminus M^c = M \setminus U^c$, hence $\mu(M \setminus U^c) \leq \varepsilon$. □

Since we will use the other two conditions as well, lets sort out the relationship between them.

Proposition 31. Let X be a Hausdorff topological space. For a Borel measure μ, then consider the following conditions:

(1) μ is finite on compact sets.

(2) μ is locally finite.

(3) μ is σ-finite by open sets.

Then, (3) $\Rightarrow$ (2) $\Rightarrow$ (1) always. Moreover, if X is locally compact(i.e. for all $x \in X$ there is an open U and compact K such that $x \in U \subseteq K$.), (1) $\Rightarrow$ (2). If X is second countable, then (2) $\Rightarrow$ (3).

Proof. (3) $\Rightarrow$ (2): If $X = \bigcup_{n \in \mathbb{N}} U_n$, where U_n is open and of finite measure, then for every $x \in X$, x is contained in some U_n.

(2) $\Rightarrow$ (1) : Let K be a compact set and for each $x \in K$, let U_x be an open neighbourhood of finite measure containing x. Then, the cover $\{U_x\}_{x \in K}$ admits a finite subcover $U_{x_1}, \ldots U_{x_n}$, so $K \subseteq \bigcup_{i \leq n} U_{x_i}$ and this union has finite measure.

(1) $\Rightarrow$ (2) : Suppose X is locally compact and (1) holds. Then every point $x \in X$ has a compact neighbourhood(equivalent to previous definition when X is Hausdorff) and compact sets have finite measure.

(2) $\Rightarrow$ (3): Suppose X is second countable and μ is locally finite. Let $\{U_n\}_{n \in \mathbb{N}}$ be a countable basis for X. Then for each $x \in X$, there is an open neighbourhood U of finite measure containing x, hence there is some $n_x \in \mathbb{N}$ such that $x \in U_{n_x} \subseteq U$, so U_{n_x} has finite measure. But then $X = \bigcup_{x \in X} U_{n_x}$ and this is a countable union. □

Thus, since $\mathbb{R}^d$ and $A^{\mathbb{N}}$(for a finite A) are locally compact second countable metric spaces, all these notions coincide for Borel measures on them. In particular:

Corollary 2.12.2. The Lebesgue measure on $\mathbb{R}^d$ and the Bernoulli measure on $A^{\mathbb{N}}$. with $|A| < \infty$, are strongly regular.

2.13 Tightness

Definition 2.13.1. A Borel measure μ on a Hausdorff topological space is called **tight** if for every μ-measurable set $M \subseteq X$,

$$\mu(M) = \sup\{\mu(K) : K \subseteq M \text{is compact}\}.$$

Before the next theorem, let us recall some equivalent definitions of compactness for metric spaces

Theorem 2.13.1 (Compactness in metric spaces)**.** For a metric space X, the following are equivalent:

(1) X is compact (every open cover has a finite subcover).

(2) X is sequentially compact (every sequence has a convergent subsequence).

(3) X is complete and totally bounded(for every $\varepsilon > 0$ we can cover the whole space with finitely many ε-balls).

Corollary 2.13.1. In a complete metric space, compact is equivalent to closed and totally bounded.

Theorem 2.13.2. Finite Borel measures on Polish spaces are tight.

Proof. Since we know that a finite Borel measure μ on a polish space X is strongly regular, every μ-measurable set can be approximated from below by closed sets, so it is enough to show that closed sets can be approximated from below by compact sets. Let $C \subseteq X$ be a closed set. Since C is Polish with the same metric, we may assume $X = C$. For a polish space X and a finite Borel measure μ on X, it is enough to show that for each $\varepsilon > 0$, there is a compact $K \subseteq X$ with $\mu(K) \approx_\varepsilon \mu(X)$. Fix $\varepsilon > 0$. Let $\varepsilon_n = 1/n$ and for each $n \in \mathbb{N}^+$, let $\{B_l^{\varepsilon_n}\}_{l \in \mathbb{N}}$ be a countable cover of X with closed balls of radius at most ε_n(such a cover exists by separability). Because

$$X = \bigcup_{L \in \mathbb{N}} \left(\bigcup_{l \leq L} B_l^{\varepsilon_n} \right),$$

(the outermost union is increasing), we have that

$$\mu(X) \approx_{\varepsilon \cdot 2^{-(n+1)}} \mu\left(\bigcup_{l < L_n} B_l^{\varepsilon_n} \right)$$

for L_n large enough, by monotone convergence. Put $C_n := \bigcup_{l \le L_n} B_l^{\varepsilon_n}$, so C_n is closed and $K := \bigcap_{n\in\mathbb{N}} C_n$ is still closed but also totally bounded by definition, hence compact. Finally,

$$\mu(X \setminus K) \le \mu\left(\bigcup_{n\in\mathbb{N}} [X \setminus C_n]\right) \le \varepsilon.$$

□

Corollary 2.13.2 (Strong regularity and tightness for locally finite measures.)**.** Let X be a Polish Space. Then, every locally finite Borel measure on X is strongly regular and tight.

Proof. Polish spaces are second countable, and so local finteness is equivalent to σ-finiteness by open sets,i.e. $X = \bigcup_{n\in\mathbb{N}} U_n$, where U_n is open and finite measure. Thus, μ is strongly regular by a previous result for metric spaces and we only need to show tightness.

From DST, we know that open subsets of Polish spaces are Polish (with a different equivalent metric), so on each U_n, we know that μ is tight. We leave the rest of the proof as homework. □

CHAPTER 3

Measurable Functions and Integration

3.1 Measurable Functions

Definition 3.1.1. Let $(X, \mathcal{I})$ and $(Y, \mathcal{J})$ be measurable spaces. A function $f : X \to Y$ is said to be:

(a) $(\mathcal{I}, \mathcal{J})$**-measurable** if $f^{-1}(J) \in \mathcal{I}$ for every $J \in \mathcal{J}$.

(b) $\mathcal{I}$**-measurable** if Y is a metric space and f is $(\mathcal{I}, \mathcal{B}(Y))$-measurable.

(c) **Borel** if X and Y are both metric spaces and f is $(\mathcal{B}(X), \mathcal{B}(Y))$-measurable.

(d) μ**-measurable** if μ is a measure on $(X, \mathcal{I})$, Y is a metric space and f is $Meas_\mu$-measurable, i.e. $f^{-1}(B)$ is μ-measurable for each Borel $B \subseteq Y$.

Remark 15. For functions $f : \mathbb{R} to \mathbb{R}$, we will view the left $\mathbb{R}$ as the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ and the right $\mathbb{R}$ as a metric space, so the definition of λ-measurable is asymmetric: The preimage of Borel sets are λ-measurable. This is done because we get to call more functions measurable since the theory works for them.

Proposition 32. Let $(X, \mathcal{I})$ and $(Y, \mathcal{J})$ be measurable spaces and $f : X \to Y$. If for some $\mathcal{J}_0 \subseteq \mathcal{J}$ which generates $\mathcal{J}$ as a σ-algebra, we have that $f^{-1}(J_0) \in \mathcal{I}$ for all $J_0 \in \mathcal{J}_0$, then f is $(\mathcal{I}, \mathcal{J})$-measurable.

Proof. Let $\mathcal{S} = \{J \in \mathcal{J} : f^{-1}(J) \in \mathcal{I}\}$ and observe that $\mathcal{S} \supseteq \mathcal{J}_0$ and $\mathcal{S}$ is a σ-algebra since preimages respect union and complements. Hence, $\mathcal{S} = \mathcal{J}$. $\square$

Corollary 3.1.1. Let $(X, \mathcal{I})$ be a measurable space, and Y a metric space. Let $f : X \to Y$. If $f^{-1}(V) \in \mathcal{I}$ for each open $V \subseteq Y$, then f is $\mathcal{I}$-measurable. In particular, continuous functions are Borel because the preimage of open sets are open.

The following is one of the reasons for building measure theory.

Theorem 3.1.1. Pointwise limits of measurable functions are measurable. More precisely, if $(X, \mathcal{I})$ is a measurable space and Y is a separable metric space, then $\lim_{n\to\infty} f_n$ is $\mathcal{I}$-measurable for any sequence of $\mathcal{I}$-measurable functions $f_n : X \to Y$ for which $\lim_{n\to\infty} f_n$ exists for each $x \in X$.

Proof. By the last corollary, it is enough to show that $f^{-1}(U) \in \mathcal{I}$ for each open $U \subseteq Y$. Note that openness of U gives the following: for $x \in X$, if $f(x) \in U$, then $\forall_n^\infty f_n(x) \in U$ ($\forall_n^\infty := \exists m \forall n \geq m$). If the converse were also true we would be done because then

$$f^{-1}(U) = \bigcup_{m\in\mathbb{N}} \bigcap_{n\geq m} f_n^{-1}(U).$$

Unfortunately, the converse isn't true, for instance, take $U = (0,1) \subseteq \mathbb{R}$ and $f(x) := 1/n$. So, $f_n(x) \in U$ for all $n \in \mathbb{N}$, but the limit is $0 \notin U$. The converse holds for closed sets but U is open. However, using separability, we can find a presentation of U, which behaves as both closed and open.

We claim that

$$\bigcup_{k\in\mathbb{N}} V_k = U = \bigcup_{k\in\mathbb{N}} \overline{V}_k,$$

for some open $V_k \subseteq Y$. Let $\mathcal{D} \subseteq Y$ be a countable dense set and

$$\mathcal{V} := \left\{ B_{1/n}(y) : y \in \mathcal{D},\ n \in \mathbb{N}^+,\ \overline{B(y)_{1/n}} \subseteq U \right\}.$$

Clearly, $\mathcal{V}$ is countable. Note that if $V \in \mathcal{V}$, then $\overline{V} \subseteq U$, and so it is enough to show htat

$$U = \bigcup_{V\in\mathcal{V}} V.$$

Fix $y \in U$. Then for $n \in \mathbb{N}^+$ large enough, $\overline{B_{1/n}(y)} \subseteq U$. Let $y' \in \mathcal{D}$ such that $y' \in B_{1/2n}(y)$. Equivalently, $y \in B_{1/2n}(y')$. Moreover,

$$\overline{B_{1/2n}(y')} \subseteq \overline{B_{1/n}(y)} \subseteq U.$$

Thus, $B_{1/2n}(y') \in \mathcal{V}$ and so

$$y \in \bigcup_{V \in \mathcal{V}} V.$$

Finally, we can have that for all $x \in X$,

$$f(x) \in U \iff \exists k \forall_n^\infty f_n(x) \in \overline{V_k}.$$

$(\Rightarrow)$: $f(x) \in U = \bigcup_{k \in \mathbb{N}} V_k$ and so there is some k such that $f(x) \in V_k$. Therefore, $\exists k \forall_n^\infty f_n(x) \in V_k$.

$(\Leftarrow$:) suppose that $\exists k \forall_n^\infty f_n(x) \in V_k$. Then, there is some k such that $f(x) \in \overline{V_k}$ (by closedness). and so $f(x) \in U$. Therefore,

$$f^{-1}(U) = \bigcup_{k \in \mathbb{N}} \bigcup_{m \in \mathbb{N}} \bigcap_{n \geq m} f_n^{-1}(\overline{V_k}) \in \mathcal{I}.$$

□

Proposition 33. Let X, Y be metric spaces, where Y is second countable. Let μ be a strongly regular Borel measure on X. Let $f : X \to Y$ be a μ-measurable function. Then,

(a) f is Borel on a co-null Borel set, i.e. $f : |_{X'} : X' \to Y$ is a Borel function for some co-null Borel X' (Note: $\mathcal{B}(X') = \{B \in \mathcal{B}(X) : B \subseteq X'\}$).

(b) **Luzin's Theorem:** For all $\varepsilon > 0$, f is continuous on a closed set C with $\mu(X \setminus C) < \varepsilon$, i.e. $f|_C : C \to Y$ is continuous.

Proof. Let $\{V_n\}$ be a countable basis for Y, so it generates $\mathcal{B}(Y)$ as a σ-algebra.

(a): $f^{-1}(V_n)$ is μ-measurable, hence $f^{-1}(V_n) =_\mu B_n$ for some Borel $B_n \subseteq X$. Let

$$Z = \bigcup_{n \in \mathbb{N}} (f^{-1}(V_n) \triangle B_n),$$

so Z is null, hence $Z \subseteq \tilde{Z}$ where $\tilde{Z}$ is Borel and still null. Put $X' := X \setminus \tilde{Z}$. So X is Borel and co-null. then,

$$(f|_{X'})^{-1}(V_n) = f^{-1}(V_n) \cap X' = B_n \cap X'$$

which is Borel. So $F|_{X'}$ is Borel.

(b): $f^{-1}(V_n)$ is μ-measurable, hence by strong outer regularity there is some open $U_n \subseteq X$ such that

$$\mu(U_n \triangle f^{-1}(V_n)) \leq \varepsilon \cdot 2^{-(n+2)}.$$

Let

$$Z = \bigcup_{n \in \mathbb{N}} (f^{-1}(V_n) \triangle U_n).$$

So,

$$\mu(Z) \leq \varepsilon/2.$$

Again by outer regularity, there is an open set $\tilde{Z} \supseteq Z$ with $\mu(\tilde{Z} \setminus Z) \leq \varepsilon/2$, so $\mu(\tilde{Z}) \leq \varepsilon$. Take $C := X \setminus \tilde{Z}$, so it is closed and $\mu(X \setminus C) \leq \varepsilon$. Moreover,

$$(f|_C)^{-1}(V_n) = f^{-1}(V_n) \cap C = U_n \cap C,$$

so $f|_C)^{-1}(V_n)$ is open relative to C and hence $f|_C$ is continuous. □

3.2 Pushfoward Measures

Definition 3.2.1. Let $(X, \mathcal{I})$ and $(Y, \mathcal{J})$ be measurable spaces and let μ be a measure on $\mathcal{I}$. Let $f : X \to Y$ be an $(\mathcal{I}, \mathcal{J})$-measurable function. Then, the ***f*-pushfoward of** μ is the measure $f_*\mu$ defined by: For $J \in \mathcal{J}$,

$$f_*\mu(J) := \mu\left(f^{-1}(J)\right).$$

Example 10. Here are a few examples of pushfoward measures.

(a) Let $S^1 \subseteq \mathbb{C}$ denote the unit circle, which is usually considered a group under complex multiplication. This is identified with the group $(\mathbb{R}/\mathbb{Z}, +)$ as follows: $\mathbb{R}/\mathbb{Z} \cong [0, 1)$, since $[0, 1)$ is a transversal for the coset equivalence relation of $\mathbb{Z} \leq \mathbb{R}$. Define $f : [0, 1) \to S^1$ by the mapping $x \mapsto e^{2\pi i x}$, which is a group isomoprhism between $(\mathbb{R}/\mathbb{Z}, +)$ and $(S^1, \cdot)$. Then, $f_*\lambda$ is a Borel measure on S^1.

(b) Let $\mathbb{R}_{>0}$ be the group of positive reals under multiplication. Consider $f : \mathbb{R} \to \mathbb{R}_{>0}$ by $x \mapsto e^x$, and take the pushfoward measure $f_*\lambda$. In particular, for $(a, b) \subseteq \mathbb{R}_{>0}$,

$$f_*\lambda((a, b)) = \lambda((\log a, \log b)) = \log b - \log a.$$

Definition 3.2.2. Let $(X, \mathcal{I}, \mu)$ be a measure space.

- For an $(\mathcal{I}, \mathcal{I})$ measurable function $f : X \to X$, we say that μ is ***f*-invariant** or that f **preserves** if $f_*\mu = \mu$,i.e. $\mu(B) = \mu(f^{-1}(B))$ for all $B \in \mathcal{I}$.
- For a group action $\Gamma \curvearrowright X$ such that each group element acts as an $(\mathcal{I}, \mathcal{I})$-measurable function, we say that μ is **Γ-invariant** or Γ **preserves** μ if for each $\gamma \in \Gamma$, $\gamma_*\mu = \mu$.

Example 11. Let $S : A^{\mathbb{N}} \to A^{\mathbb{N}}$ be the **left-shift map**, i.e. $(x_n) \mapsto (x_{n+1})$. Any Bernoulli measure $\nu^{\mathbb{N}}$ is shift-invariant: indeed, it suffices to check on cylinders $[w] : s^{-1}([w]) = \bigsqcup_{a \in A}[aw]$ so

$$\begin{aligned} \nu^{\mathbb{N}}(s^{-1}([w])) &= \sum_{a \in A} \nu^{\mathbb{N}}([aw]) \\ &= \sum_{a \in A} \nu(a) \cdot \nu^{\mathbb{N}}([w]) \\ &= \nu^{\mathbb{N}}([w]) \end{aligned}$$

Definition 3.2.3. A **topological group** is a group G equipped with a topology making multiplication and inverse continuous. A Borel measure μ on G is called **left-invariant** (resp. **right-invariant**) if it is invariant under the left-translation (resp. right-translation) action $G \curvearrowright G$, i.e. $\mu(g \cdot B) = \mu(B)$ (resp. $\mu(B \cdot g) = \mu(B)$).

Theorem 3.2.1 (Haar)**.** Every locally compact (Hausdorff) group admits a unique (up to scaling) locally finite (or equivalently finite on compact sets) left-invariant Borel measure (also a right-invariant Borel measure). This measure is called a **left Haar measure** (resp. **right Haar measure**).

Example 12. Here are a few examples of Haar measures:

(a) For $(\mathbb{R}^d, +)$, the Lebesgue measure is a Haar measure.

(b) For $(\mathbb{R}_{>0}, \cdot)$, the pushfoward of Lebesegue by $x \mapsto e^x : \mathbb{R} \to \mathbb{R}_{>0}$ is a Haar measure because this function is a topological group isomorphism and Lebesgue is Haar for $(\mathbb{R}, +)$.

(c) For $(S^1, \cdot)$, the pushfoward of Lebesgue bt $x \mapsto e^{2\pi i x} : [0, 1) \to S^1$ is Haar because this function is a topological group isomorphism and the Lebesgue measure is Haar on $\mathbb{R}/\mathbb{Z} \cong [0, 1)$. Note that this measure is a probability meausre, hence this is the unique probability Haar measure.

(d) Consider the group $(\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}} \cong 2^{\mathbb{N}}$ as a compact group with the same topology as $2^{\mathbb{N}}$, under coordinate wise addition modulo 2. The Bernoulli half measure is the Haar probability measure on $(\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$. Incidentally, this is also the pushfoward of the Lebesgue measure on $[0,1) \cong \mathbb{R}/\mathbb{Z}$ by:

$$f : [0,1) \to 2^{\mathbb{N}} \cong (\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$$

by mapping points to their binary representation.

Example 13. One example of a non-Haar measure is the following: Consider the group $\mathrm{GL}_n(\mathbb{R})$ of all invertible $n \times n$ real matrices under multiplication, this group is locally compact when viewed as a subset of $\mathbb{R}^{n^2}$: indeed, $M \in \mathrm{Gl}_n(\mathbb{R}) \iff \det(M) \neq 0$, and the latter is an open condition, so $\mathrm{GL}_n(\mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$. Furthermore, its compliment, the $\det = 0$ set, is "lower dimensional " closed set, and one can show that it is null. so $\mathrm{GL}_n(\mathbb{R})$ is a Lebesgue conull open subset of $\mathbb{R}^{n^2}$. However, the Lebesgue measure on $\mathrm{Gl}_n(\mathbb{R})$ is not Haar because multiplication by a matrix such as $\mathrm{diag}(2,\dots,2)$ scales the Lebesgue measure by 2. The Haar measure on $\mathrm{Gl}_n(\mathbb{R})$ is defined using the Jacobian, i.e. the integral with $1/\det$ in it.

We showed that the translation actions $\mathbb{Q} \curvearrowright \mathbb{R}$ and $\bigoplus_{n\in\mathbb{N}}(\mathbb{Z}/2\mathbb{Z}) \curvearrowright \prod_{n\in\mathbb{N}} \mathbb{Z}/2\mathbb{Z}$ are ergodic, and one can also similarly show that for an irrational $\lambda \in \mathbb{R} \setminus \mathbb{Q}$, the rotation by $2\pi\lambda$ on S^1 is ergodic, which is the same as the translation aciton of the (dense) subgroup $\langle e^{2\pi i\lambda} \rangle \leq S^1$ on S^1. The following shows that this is a general phenomenon:

Theorem 3.2.2. Let G be a locally compact (Hausdorff) group and $\Gamma \leq G$ a dense subgroup. Then the (left) translation action $\Gamma \curvearrowright G$ is ergodic with respect to any Haar-measure.

3.3 Borel/Measure Isomoprhism Theorems

The following is one of the basic theorems in DST, which is used by many mathematicians (e.g. ergodic and probability theorists) all the time without mention.

Theorem 3.3.1 (Borel Isomoprhism Theorem)**.** Any two uncountable Polish spaces are Borel isomoprhic, i.e. there is a bijective function $f : X \to Y$ such that f and f^{-1} are Borel.

We provide a proof sketch:

For an uncountable polish space X, it is enough to show that X is Borel isomorphic to $2^{\mathbb{N}}$. By the Borel version of the Cantor-Schroder-Berstein theorem, it is enough to show that there is Borels injections $2^{\mathbb{N}} \to X$ and $X \to 2^{\mathbb{N}}$. The first injection is called the **Cantor-Bendixson theorem:** For each uncountable Polish X, there is a continuous embedding $2^{\mathbb{N}} \to X$.

Lemma 3.3.1 (Binary Representation)**.** Any second countable metric space X admits a Borel injection

$$b : X \to 2^{\mathbb{N}}$$

which we call a **binary representation** map.

Proof. Let (U_n) be a countable basis for X, so it separates points. then, define

$$b : X \to 2^{\mathbb{N}}$$

via

$$x \mapsto (\mathbb{1}_{U_n}(x))_{n\in\mathbb{N}}.$$

To check that B is Borel, it is enough to observe that $b^{-1}(V_n) = U_n$ is Borel, where $V_n := \{x \in 2^{\mathbb{N}} : x(n) = 1\}$ because $\{V_n\}$ generates $\mathcal{B}\left(2^{\mathbb{N}}\right)$ (indeed, every cylinder in $2^{\mathbb{N}}$ is a finite intersection of these V_n and their complements) $\square$

This finishes the sketch of the Borel isomorphism theorem. $\diamond$

Definition 3.3.1. A measurable space $(X, \mathcal{I})$ is called a **standard Borel space** if there is a polish metric on X such that $\mathcal{I} = \mathcal{B}(X)$. In other words, X was a Polish space, but we forgot its topology and only kept the Borel σ-algebra.

Thus, the Borel isomorphism theorem says that there is only one (up to isomorphism) standard Borel space.

Definition 3.3.2. Let $(X, \mathcal{I}, \mu)$ and $(Y, \mathcal{J}, \nu)$ be measure spaces. A function $f : X \to Y$ is called a **measure isomorphism** if there are co-null sets $X' \subseteq X$ and $Y' \subseteq Y$ such that $(f|_{X'}) : X' \to Y'$ is a bjiection, and such that $(f|_{X'})$ is $(\mathcal{I}, \mathcal{J})$-measurable, and $(f|_{X'})^{-1}$ is $(\mathcal{J}, \mathcal{I})$-measurable, and $f_*\mu = \nu$.

Theorem 3.3.2 (Measure Isomorphism)**.** Every atomless Borel probability measure μ on a Polish space X is isomorphic to $([0, 1], \lambda)$. In fact, there is a Borel isomorphism $f : X \to [0, 1]$ with $f_*\mu = \lambda$.

Proof. Because X is atomless, then each singleton is μ-null, and so X must be uncountable since otherwise $\mu(X) = 0$. By the Borel isomorphism theorem, there is a Borel isomorphism $g : X \to [0, 1]$. so by replacing X with $[0, 1]$ and μ with $g_*\mu$, we may assume that μ is an atomless Borel probability measure on $[0, 1]$.

Let $f : [0, 1] \to [0, 1]$ be defined by $x \mapsto \mu([0, x])$. This is an increasing (maybe not strictly) and continuous function; indeed it is right continuous because of downward monotone convergence, i.e.

$$\mu([0, x]) = \lim_{x_n \to x^+} \mu([0, x_n]),$$

and left continuous because of upward monotone converge and atomlessness, i.e.

$$\mu([0, x]) = \mu([0, x)) = \lim_{x_n \to x^-} \mu([0, x_n]).$$

Furthermore, $f^{-1}([0, y]) = [0, x]$ where $\mu([0, x]) = y$, and this x is a maximum such, so,

$$\lambda([0, y]) = y = \mu(f^{-1}[0, y]) = f_*\mu([0, y]).$$

Since the sets $[0, y]$ generate the Borel σ-algebra, the measures λ and $f_*\mu$ coincide. It remains to show that f is a bijection on a conull set. f is not bijective as a whole because there might be intervals $(a, b]$ on which f is constant, but then $\mu((a, b]) = 0$ and there are only countably many such maximal intervals because these are pairwise disjoint and $[0, 1]$ is separable. So the union Z of all these maximal intervals $(a, b]$ is still μ-null and $f|_{X \setminus Z} : X \setminus Z \to [0, 1]$ is a bijection. □

Corollary 3.3.1. Every σ-finite Borel measure μ on a Polish space X is isomorphic to $(\mathbb{R}, \lambda)$.

Proof. Write X as a countable disjoint union of finite positive measure pieces and use that each piece is isomorphic to an interval in $\mathbb{R}$. Details left as an exercise. □

Definition 3.3.3. A measure space $(X, \mathcal{B}, \mu)$ is called **standard** if μ is σ-finite and $(X, \mathcal{B})$ is standard Borel.

Remark 16. In dynamics and probability theory, one mainly works with standard probability spaces (since we know that the atomless ones are all isomorphic).

We can restate as follows:

Theorem 3.3.3. A standard atomless infinite measure space is isomorphic to $(\mathbb{R}, \lambda)$ and a standard atomless probability measure is isomorphic to $([0,1], \lambda)$.

3.4 Integration

Given a measurable space $(X, \mathcal{B})$, we denote $L := L(X, B)$ as the set of all $\mathcal{B}$-measurable functions $X \to \overline{\mathbb{R}} = [-\infty, \infty]$, and

$$L^+ := L^+(X, B) := \{f \in L(X, \mathcal{B}) : f \geq 0\}.$$

Note that $L(X, B)$ is a vector space and $L^+(X, \mathcal{B})$ is closed under non-negative linear combinations. Both are closed under products and limits.

Definition 3.4.1. An **integral** on $L^+(X, \mathcal{B})$ is a countably additive linear functional

$$I : L^+ \to [0, \infty].$$

In other words:

(i) $I(\alpha \cdot f + \beta \cdot g) = \alpha I(f) + \beta I(g)$ for all $\alpha, \beta \geq 0$ and $f, g \in L^+$.

(ii) $I\left(\sum_{n\in\mathbb{N}} f_n\right) = \sum_{n\in\mathbb{N}} I(f_n)$.

Proposition 34. Every integral I on $L^+(X, \mathcal{B})$ defines a measure μ_I on $(X, \mathcal{B})$ by:

$$\mu_I(B) := I(\mathbb{1}_B).$$

Proof. Indeed, $\mu_I(\emptyset) = I(\mathbb{1}_\emptyset) = I(0) = 0$ and if $B = \bigsqcup_{n\in\mathbb{N}} B_n$, then $\mathbb{1}_B := \sum_{n\in\mathbb{N}} \mathbb{1}_{B_n}$, and so $\mu_I(B) = I(\sum_{n\in\mathbb{N}} \mathbb{1}_{B_n}) = \sum_n I(\mathbb{1}_{B_n}) = \sum_n \mu_I(B_n)$. □

We would like to solve the inverse problem, indeed we will prove:

Theorem 3.4.1. Every measure μ on $(X, \mathcal{B})$ admits a unique integral

$$I : L^+(X, \mathcal{B}) \to [0, \infty]$$

such that $\mu_I = \mu$. This integral I is called the **integral over** μ, denoted

$$\int f \mathrm{d}\mu := \int f(x) \mathrm{d}\mu(x) := I(f).$$

We shall eventually prove this theorem by gradually defining the integral.

Remark 17. Once I is defined on L^+, we will be able to also define it on a subspace of $L(X, \mathcal{B})$ using the fact that for each $f \in L(X, \mathcal{B})$, we have that

$$f = f_+ - f_-,$$

where

$$f_-(x) := \begin{cases} -f(x), \text{ if } f(x) \leq 0 \\ 0, \text{otherwise} \end{cases},$$

and

$$f_+(x) := \begin{cases} f(x), \text{ if } f(x) \geq 0 \\ 0, \text{otherwise} \end{cases}.$$

Definition 3.4.2. A function $f \in L(X, \mathcal{B})$ is called **simple** if it is a finite linear combination of indicator function of sets from $\mathcal{B}$, so **non-negative simple functions** are the non-negative linear combinations of indicator functions. Denote by $S(X, \mathcal{B})$ (resp. $S^+(X, \mathcal{B})$) the subspace of simple (resp. non-negative simple) functions. Note that a non-negative simple function f is of the form

$$f = \sum_{i<n} \alpha_I \cdot \mathbb{1}_{A_i}$$

for some $n \in \mathbb{N}$, $\alpha_i \geq 0$, $A_i \in \mathcal{B}$.

Remark 18. Notice that a function $f \in L(X, \mathcal{B})$ is simple if and only if $f(x)$ is finite.

Definition 3.4.3. For a simple function f with $f(x) = \{\alpha_0, \dots \alpha_{n-1}\}$, the representation

$$f = \sum_{i<n} \alpha_i \mathbb{1}_{f^{-1}(\alpha_i)}$$

is called **standard** . In particular, $X = \bigsqcup_{i<n} f^{-1}(\alpha_i)$.

Definition 3.4.4. Given a measure μ on $(X, \mathcal{B})$, we define its integral on $S^+(X, \mathcal{B})$ by setting for each $f \in S^+(X, \mathcal{B})$,

$$\int f \mathrm{d}\mu := \sum_{i<n} \alpha_i \cdot \mu(A_i). \tag{$\star$}$$

where $f = \sum_{i<n} \alpha_i \mathbb{1}_{A_i}$ is some representation of f.

It is left as a homework exercise to show that this is well-defined (i.e. does not depend on the representation).

Remark 19. We could alternatively define the integral for the standard representation and then prove that ($\star$) holds for any representation.

Proposition 35. The integral on S^+ satisfies the following:

(i) Linearity: $\int(\alpha f + \beta g)\mathrm{d}\mu = \alpha \int f \mathrm{d}\mu + \beta \int g \mathrm{d}\mu$.

(ii) Non-negativity: If $f \geq 0$, then $\int f \mathrm{d}\mu \geq 0$. Equivalently, if $f \leq g$, then $\int f \mathrm{d}\mu \leq \int g \mathrm{d}\mu$.

(iii) If $\int f \mathrm{d}\mu = 0$, then $f = 0$ almost everywhere.

(iv) Each $f \in S^+$ defines a measure μ_f on $(X, \mathcal{B})$ by:

$$\mu_f(B) := \int (\mathbb{1}_B \cdot f)\mathrm{d}\mu =: \int_B f \mathrm{d}\mu$$

Proof. (i)-(iii) follow from the definitions, and so we prove (iv) only. (i) implies that $\mu_f(\emptyset) = 0$ and so we verify countable additivity. Let $B = \bigsqcup_{n\in\mathbb{N}} B_n$, so $\mathbb{1}_B = \sum_{n\in\mathbb{N}} \mathbb{1}_{B_n}$. Let $f = \sum_{i<m} \alpha_i \mathbb{1}_{A_i}$. Then:

$$\begin{aligned}
\mu_f(B) = \int \mathbb{1}_B \cdot f \ \mathrm{d}\mu &= \int \sum_{i<m} \alpha_i \mathbb{1}_{A_i \cap B} \ \mathrm{d}\mu \\
&= \sum_{i<m} \alpha_i \cdot \mu(A_i \cap B) \\
&= \sum_{i<m} \sum_{n\in\mathbb{N}} \alpha_i \cdot \mu(A_i \cap B_n) \\
&= \sum_{n\in\mathbb{N}} \int (\mathbb{1}_{B_n} \cdot f) \ \mathrm{d}\mu \\
&= \sum_{n\in\mathbb{N}} \mu_f(B_n).
\end{aligned}$$

□

Proposition 36. This proposition deals with approximating functions by simple functions.

(a) For each $f \in L^+(X, \mathcal{B})$, there is an increasing function $f_n \nearrow f$ of non-negative simple functions such that the convergence is uniform on every $B \subseteq X$ on which f is bounded, i.e.

$$\|f_n|_B - f|_B\|_u \longrightarrow 0.$$

where $\|\cdot\|_u$ is the uniform norm.

(b) For each $f \in L(X, \mathcal{B})$, there is a sequence $f_n \to f$ of simple functions such that $|f_n| \nearrow |f|$ and the convergence is uniform on every set $B \subseteq X$ on which f is bounded.

Proof. (b): Follows by applying (a) to f^+ and f^-, thus obtaining a non-negative simple function $f_n^+ \nearrow f^+$ and $f_n^- \nearrow f^-$ with the uniform convergence statement satisfied, so the functions $f_n := f_n^+ - f_n^-$ are desired.

(a): To define f_n, we cut off f at $\leq 2^n$ and split $[0, 2^n]$ into pieces of size 2^{-n}, so in total 2^{2n} pieces. Let $A_k := f^{-1}\left((2^{-n}k, 2^{-n}(k+1)]\right)$ and let $B_k := f^{-1}\left((2^{-n}k, \infty]\right)$. So $B_0 \supset B_1 \supseteq \ldots$ and $A_k = B_{k+1} \setminus B_k$. We set

$$f_n := \sum_{k=0}^{2^{2n}-1} (2^{-n}k)\mathbb{1}_{A_k} = \sum_{k=1}^{2^{2n}} 2^{-n}\mathbb{1}_{B_k}$$

Note that $f_n \nearrow f$ since the sets $X_n := f^{-1}\left([0, 2^n]\right) \nearrow (X \setminus \{x \in X : f(x) = \infty\})$ so (f_n) converges to f on the increasing union $\bigcup_{n \in \mathbb{N}} X_n$ and on $X_\infty := \{x \in X : f(x) = \infty\}$, $f_n|_{X_\infty} \equiv 2^n$, so then $f_n|_{X_\infty} \nearrow \infty = f|_{X_\infty}$. As for uniform convergence, we have that if f is bounded on $B \subseteq X$, then $B \subseteq X_n$ for large enough $n \in \mathbb{N}$ and so

$$\|f_n|_{X_n} - f|_{X_n}\|_u \leq 2^{-n} \to 0.$$

$\square$

Definition 3.4.5. For $f \in L^+(X, \mathcal{B})$, we define

$$\int f \, d\mu := \sup\left\{\int s \, d\mu : \; 0 \leq s \leq f \text{ is simple}\right\}.$$

Proposition 37. Let $f, g \in L^+(X, \mathcal{B})$.

(a) If $f \leq g$, then $\int f \, d\mu \leq \int g \, d\mu$.

(b) $\int a \cdot f \, d\mu = a \cdot \int f \, d\mu$ for all $a \geq 0$.

(c) $\int f \, d\mu = 0 \iff f = 0$ almost everywhere.

Proof. (c): $(\Rightarrow)$: We prove the contrapositive. Suppose $X_0 := \{x \in X : f(x) > 0\}$ has positive measure, we show that $\int f \, d\mu > 0$. Indeed, $X_0 = \bigcup_{n\in\mathbb{N}} X_n$ where $X_n = \{x \in X : f(x) > 1/n\}$, so for some $n \in \mathbb{N}^+$, X_n has positive measure. But $s := 1/n \cdot \mathbb{1}_{X_n} \leq f$, so

$$\int f \, d\mu \geq \int s \, d\mu = 1/n \cdot \mu(X_n) > 0.$$

□

Remark 20. $\int (f+g) \, d\mu = \int f \, d\mu + \int g \, d\mu$ is not immediate because we don't know wether every simple function $s \leq f + g$ splits $f = \hat{f} + \hat{g}$ into a sum of simple functions $\hat{f} \leq f$ and $\hat{g} \leq g$. To prove this, we need to replace sup with limits of sequences.

Theorem 3.4.2 (Monotone Convergence Theorem)**.** *Let $f_n, f \in L^+(X, \mathcal{B})$. If $f_n \nearrow f$, then $\int f_n \, d\mu \nearrow \int f \, d\mu$.*

Proof. We may assume without loss of generality that $f > 0$ on all X by restricting to $\{x \in X : f(x) > 0\}$. Because $f_n \leq f$, we have that $\int f_n \, d\mu \leq \int f \, d\mu$ and $\int f_n \, d\mu$ is increasing. We need to show that

$$\lim_n \int f_n \, d\mu \geq \int f \, d\mu.$$

We first fix a simple function $0 \leq s \leq f$. and aim to show that $\lim_n \int f_n \, d\mu \geq \int s \, d\mu$. For this, it suffices to fix $\varepsilon > 0$ and show that

$$\lim_n \int f_n \, d\mu \geq \int (1-\varepsilon) s \, d\mu,$$

because $\int (1-\varepsilon) s \, d\mu = (1-\varepsilon) \int s \, d\mu \nearrow 1$ as $\varepsilon \to 0$. Because $(1-\varepsilon)s < f$ for all $x \in X$, we have that $X = \bigcup_{n\in\mathbb{N}} X_n$ where $X_n = \{x \in X : f_n(x) > (1-\varepsilon)s\}$ and the X_n's are increasing. By monotonicity,

$$\int f_n \, d\mu \geq \int_{X_n} f_n \, d\mu \geq \int_{X_n} (1-\varepsilon) s \, d\mu = \mu_{(1-\varepsilon)s}(X_n)$$

and $\mu_{(1-\varepsilon)s}(X_n) \nearrow_{n\to\infty} \mu_{(1-\varepsilon)s}(X) = \int (1-\varepsilon) s \, d\mu$ by monotone convergence for measures, and thus the result follows by taking the supremum over all simple functions $s \leq f$. □

Corollary 3.4.1 (Countable Additivity of Integrals)**.** The integral on $L^+(X, \mathcal{B})$ is countably additive, i.e.

$$\int \sum_{n\in\mathbb{N}} f_n \; d\mu = \sum_{n\in\mathbb{N}} \int f_n d\mu.$$

Proof. We first show that $\int f + g \; d\mu = \int f \; d\mu + \int g \; d\mu$. Let $f_n \nearrow f$ and $g_n \nearrow g$ be simple functions, so $f_n + g_n \nearrow f + g$. By monotone convergence theorem, we have :

$$\begin{aligned} \int f + g \; d\mu &= \lim_n \int (f_n + g_n) \; d\mu \\ &= \lim_n \int f_n \; d\mu + \lim_n \int g_n \; d\mu \\ &= \int f \; d\mu + \int g \; d\mu, \end{aligned}$$

where the last and first equalities follow from MCT, and the third equality from the fact that the integral on simple functions is linear.

Now for an infinite sum $\sum_{n\in\mathbb{N}} f_n$, observe that

$$\sum_{n\leq N} f_n \nearrow_{N\to\infty} \sum_{n\in\mathbb{N}} f_n,$$

so once again by MCT, we have that:

$$\sum_{n\in\mathbb{N}} \int f_n \; d\mu = \lim_{N\to\infty} \sum_{n<N} \int f_n \; d\mu = \lim_{N\to\infty} \int \sum_{n<N} f_n \; d\mu = \int \sum_{n\in\mathbb{N}} f_n \; d\mu.$$

□

Corollary 3.4.2 (Functions Defining Measures)**.** each $f \in L^+(X, \mathcal{B})$ defines a measure μ_F on $(X, \mathcal{B})$ by

$$\mu_F(B) := \int \mathbb{1}_B \cdot f \; d\mu =: \int_B f \; d\mu.$$

Proof. Clearly $\mu_f(\emptyset) = 0$ so let $B = \bigsqcup_{n\in\mathbb{N}} B_n$ with all sets in B. Then $\mathbb{1}_B = \sum_{n\in\mathbb{N}} \mathbb{1}_{B_n}$, so $\mathbb{1}_B \cdot f = \sum_{n\in\mathbb{N}} (\mathbb{1}_{B_n} \cdot f)$, so by countable additivity:

$$\mu_f(B) = \int \mathbb{1}_B \cdot f \; d\mu = \int \sum_{n\in\mathbb{N}} \mathbb{1}_{B_n} \cdot f \; d\mu = \sum_{n\in\mathbb{N}} \int \mathbb{1}_{B_n} \cdot f \; d\mu = \sum_{n\in\mathbb{N}} \mu_f(B_n).$$

□

Lemma 3.4.1 (Fatou's Lemma)**.** Let $\{f_n\} \subseteq L^+(X, \mathcal{B})$. Then,

$$\int \liminf f_n \ d\mu \leq \liminf \int f_n \ d\mu.$$

Proof. $\liminf f_n := \lim_{n\to\infty}(\inf_{i\geq n} f_i)$, so

$$\inf_{i\geq n} f_i \nearrow_{n\to\infty} \liminf_{n\to\infty} f_n,$$

so by MCT:

$$\lim_{n\to\infty} \int \inf_{i\geq n} f_i \ d\mu = \int \liminf_{n\to\infty} f_n \ d\mu.$$

Now, by monotonicity, we have that

$$\int \inf_{i\geq n} f_i \ d\mu \leq \int f_m \ d\mu$$

for all $m \geq n$, so

$$\int \inf_{i\geq n} f_i \ d\mu \leq \inf_{m\geq n} \int f_m \ d\mu,$$

hence

$$\int \liminf_{n\to\infty} f_n \ d\mu \leq \lim_{n\to\infty} \inf_{m\geq n} \int f_m \ d\mu = \liminf_{n\to\infty} \int f_n \ d\mu.$$

□

Example 14 (strict inequality in Fatou)**.** For $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, let $f_n := \mathbb{1}_{[n,n+1)}$, so $f_n \to 0$, but $\int f_n \ d\lambda = 1$ for all $n \in \mathbb{N}$.

Definition 3.4.6. Let $(X, \mathcal{B}, \mu)$ be a measure space, where $\mathcal{B} = Meas_\mu$, i.e. μ is complete. A μ-measurable function is said to be **μ-integrable** if $\int |f| \ d\mu < \infty$. In this case,

$$\int f \ d\mu := \int f_+ \ d\mu - \int f_- \ d\mu$$

Remark 21. Note that $|f| = f_+ + f_-$ so $\int |f| \ d\mu - \int f_+ \ d\mu + \int f_- \ d\mu$. In particular, $\left|\int f \ d\mu\right| \leq \int |f| \ d\mu$.

Remark 22. If f is a complex-valued function, then

$$\int f\mu = \int Re(f) \ d\mu + i \cdot \int Im(f) \ d\mu.$$

We denote the space of μ-integrable μ-measurable functions by $L^1(X,\mu) := L^1(X,\mathcal{B},\mu)$. Commonly, $L^+(X,\mu)$ denotes the quotient of the set of all integrable functions by the equivalence relation $f = g$ almost everywhere. We will typically use $L^1(X,\mu)$ as the literal space of all μ-integrable functions, with the understanding that we could have used the quotient.

We define a (pseudo)norm on $L^1(X,\mu)$ by

$$\|f\|_1 := \int |f| \, d\mu.$$

We call this the L^1**-norm**.

Proposition 38. $L^1(X,\mu)$ equipped with $\|\cdot\|_1$ is a (pseudo)normed vector space, i.e. for all $f, g \in L^1(X,\mu)$, we have:

(i) $\|f\|_1 \geq 0$ and $\|f\|_1 = 0$ if and only if $f = 0$ almost everywhere.

(ii) $\|\alpha \cdot f\|_1 = |\alpha| \cdot \|f\|_1$, for all $\alpha \in \mathbb{R}$.

(iii) Triangle inequality: $\|f+g\|_1 \leq \|f\|_1 + \|g\|_1$.

Proof. For (iii), note that

$$\begin{aligned} \|f+g\|_1 = \int |f+g| \, d\mu &\leq \int |f| + |g| \, d\mu \\ &= \int |f| \, d\mu + \int |g| \, d\mu \\ &= \|f\|_1 + \|g\|_1 . \end{aligned}$$

□

As always, a (pseudo) normed vector space is also a (pseudo) metric space by

$$d_1(f,g) := \|f-g\|_1 .$$

Thus, we say that a sequence $(f_n) \subseteq L^1(X,\mu)$ **converges in L^1 norm** to $f \in L^1(X,\mu)$ if it converges in the (pseudo) metric d_1, i.e. $f_n \to_{L^1} f$ if

$$d_1(f, f-n) := \|f - f_n\|_1 \to 0$$

as $n \to \infty$.

Proposition 39. $f_n \to_{l^1} f$ implies (in particular) $\int f_n \, d\mu \to \int f \, d\mu$.

Proof.

$$\left|\int f\ d\mu - \int f_n\ d\mu\right| = \left|\int (f - f_n)\ d\mu\right| \leq \int |f - f_n|\ d\mu = \|f - f_n\|_1 \to 0.$$

□

Remark 23. When μ is the counting measure on X, then , we define $l^1(x) := L^1(X, \mu)$.

Example 15. Here are a couple examples of l^1 with different spaces.

(a) $L^1(\mathbb{N})$ is exactly the space of absolutely summable sequences. Indeed, for $f \in L^1(\mathbb{N})$ and μ a counting measure on $\mathbb{N}$,

$$\int f\ d\mu = \sum_{n\in\mathbb{N}} f(n),$$

so $\|f\|_1 := \sum_{n\in\mathbb{N}} |f(n)|$.

(b) Let $d \in \mathbb{N}$, so $d = \{0, \ldots, d-1\}$. What is $l^1(d)$? It is just $\mathbb{R}^d$ with the 1-norm:

$$\|f\|_1 := \sum_{i<d} |f(i)|.$$

3.4.1 Pointwise VS L^1 Convergence

Example 16 (Disagreement)**.** All examples below are in $L^1(\mathbb{R}, \lambda)$.

(a) Let $f_n := \mathbb{1}_{[n,n+1)}$ and $f \equiv 0$. We have that $f_n \to 0$ pointwise but $\int f_n\ d\mu = 1$ for all $n \in \mathbb{N}$ while $\int f\ d\mu = 0$. Also, $\|f_n - f_m\|_1 = 1$ for $n \neq m$.

(b) Let $f_n := n \cdot \mathbb{1}[0, 1/n]$ for $n \in \mathbb{N}^+$. Then, $f_n \to 0$ pointwise but $\int f\ d\mu = 1/n \cdot n = 1$, hence f_n does not converge to 0 in the L^1 norm.

(c) Here is an example of $f_n \to_{L^1} 0$ but not at any point. Take $f_1 = \mathbb{1}_{[0,1]}$ and $f_n = \mathbb{1}_{[j/2^k,(j+1)/2^k]}$, where $n = 2^k + j$ with $0 \leq j < 2^k$. Now, $\|f_n - 0\|_1 = \int f_n\ d\lambda = 1/2^k$ if f_n belongs to the kth group, so $f_n \to_{L^1} 0$. But $(f_n(x))$ diverges for each $x \in [0, 1]$ because $(f_n(x))$ has infinitely many 0's and 1's.

We will discuss example (c) and how to fix it later once we have introduced convergence in measure. However, we can still fix examples (a), (b) right now. Note that in these two examples, there is no integrable $g \geq 0$ dominating all of the $|f_n|$. Indeed, in (a) sucha g has to be greater than or equal to $\mathbb{1}_{[0,\infty)}$, and in (b), it has to be $\geq \max\{f_n\} \sim 1/x$.

Theorem 3.4.3 (Dominated Convergence Theorem)**.** Let f_n and f be μ-measurable functions with $f_n \to f$ almost everywhere. If there is a dominating $g \in L^1(X, \mu)$ (i.e. $g \geq 0$ and $|f_n| \leq g$ for all $n \in \mathbb{N}$), then $f_n, f \in L^1(X, \mu)$ and

$$\int f_n \, d\mu \to \int f \, d\mu \tag{$\star$}$$

as $n \to \infty$. In fact, $f_n \to_{l^1} f$ as $n \to \infty$.

Proof. The condition $|f_n| \leq g$ implies that $|f| \leq g$ almost everywhere, hence $f_n, f \in L^1$. Fatou's lemma applied to $|f_n|$ gives

$$\int |f| \, d\mu \leq \liminf \int |f_n| \, d\mu.$$

Other non-negative sequences are $g + f_n$ and $g - f_n$, and Fatou applied to each yields:

$$\begin{aligned}
\int g d\mu + \int f d\mu &= \int g + f d\mu \\
&\leq \liminf \int (g + f_n) d\mu \\
&= \int g d\mu + \liminf \int f_n d\mu
\end{aligned}$$

and

$$\begin{aligned}
\int g d\mu - \int f d\mu &= \int g - f d\mu \\
&\leq \liminf \int (g - f_n) d\mu \\
&= \int g d\mu + \liminf \left(- \int f_n d\mu \right) \\
&= \int g d\mu - \limsup \int f_n d\mu.
\end{aligned}$$

Some algebraic of the previous two chains of inequalities yields:

$$\limsup \int f_n d\mu \leq \int f d\mu \leq \int \liminf \int f_n d\mu,$$

and as such, $\lim \int f_n d\mu = \int f d\mu$.

To get the L^1 convergence, apply $(\star)$ to $|f - f_n|$. Indeed,

$$|f - f_n \leq |f| + |f_n| \leq 2g,$$

hence by $(\star)$:

$$\|f - f_n\|_1 = \int |f - f_n| d\mu \to \int \lim |f - f_n| d\mu = 0,$$

so $f_n \to_{L^1} f$. □

3.4.2 L^1 as a Pseudo-Metric Space

We analyze dense families in L^1 as well as whether the L^1 is complete.

Proposition 40. Simple functions are dense in L^1 (in the L^1 metric).

Proof. If $f \in L^1(X, \mu)$ then there exists a sequence (s_n) of simple functions such that $s_n \to f$ point-wise and $|s_n| \leq |f|$, hence by dominating convergence theorem, $s_n \to_{L^1} f$ as $n \to \infty$. □

Definition 3.4.7. A measure space $(X, \mathcal{B}, \mu)$ is said to be **countably generated** if $Meas_\mu$ is separable as a pseudo metric space with respect to the metric $d_\mu(A, B) := \mu(A \triangle B)$, i.e. there exists a countable $\mathcal{Q} \subseteq Meas_\mu$ such that for each $M \in Meas_\mu$ and $\varepsilon > 0$, there exists a $A \in \mathcal{Q}$ with $d_\mu(M, A) < \varepsilon$.

Proposition 41. For a σ-finite measure space $(X, \mathcal{B}, \mu)$, if $\mathcal{B}$ is countably generated as a σ-algebra, then $(X, \mathcal{B}, \mu)$ is countably generated.

Proof. Follows from the uniqueness part of Caratheodory theorem. Details done on midterm. □

Remark 24. The converse of the last proposition is not true: There are countably generated $(X, \mathcal{B}, \mu)$ such that $\mathcal{B}$ is not countably generated as a σ-algebra.

Proposition 42. If a measure space $(X, \mathcal{B}, \mu)$ is countably generated (i.e. $Meas_\mu$ is separable) then there is a countable collection of simple functions which is dense in $L^1(X, \mu)$ in the L^1-metric. In particular, $L^1(X, \mu)$ is separable.

Example 17. Consider the following examples:

(a) $L^1(\mathbb{R}^d, \lambda)$ is separable because $\mathbb{R}^d$ is second countable and $\mathcal{B} = \mathcal{B}\left(\mathbb{R}^d\right)$.

(b) $L^1(A^{\mathbb{N}}, \mu)$, where A is a non-empty finite dictionary and μ is a Bernoulli measure, is separable because $A^{\mathbb{N}}$ is second countable and $\mathcal{B} = \mathcal{B}\left(A^{\mathbb{N}}\right)$.

(c) $l^1(X)$ for some set X is separable if and only if X is countable. Indeed, if X is countable, then $\mathcal{B} = \mathcal{P}(X)$ is generated by singletons. If X is uncountable, then the family $\{\mathbb{1}_x\}_{x \in X}$ is discrete and uncountable, hence $L^1(x)$ is not separable.

(d) For any σ-finite Borel measure μ on a second countable metric space X, we have that it has a separable $L^1(X, \mu)$ space.

We now discuss the completeness of $L^1(X, \mu)$, for which we first give a criterion of completeness for normed vector spaces.

Definition 3.4.8. Let $(V, \|\cdot\|)$ be a (pseudo)-normed vector space, viewed as a metric space with metric $d(f, g) := \|f - g\|$ for all $f, g \in \mathbb{F}$. For a series $\sum_{n \in \mathbb{N}} f_n$ of elements in V, we say that it:

- **converges in norm** if there is some $f \in V$ such that $\sum_{n \leq N} f_n \to f$ as $N \to \infty$ in norm, i.e. $\left\|f - \sum_{n \leq N} f_n\right\| \to 0$ as $N \to \infty$. In this case, we simply write $\sum_{n \in \mathbb{N}} f_n = f$.
- **absolutely converges** if $\sum_{n \in \mathbb{N}} \|f\| \leq \infty$.

Lemma 3.4.2 (Criterion for Completeness)**.** Let $(V, \|\cdot\|)$ be a (pseudo) normed vector space. Then V is complete if and only if every absolutely convergent series converges in norm.

Proof. $(\Rightarrow)$: Suppose V is complete and let $\sum_{n \in \mathbb{N}} f_n$ be an absolutely convergent series. But then, the sequence

$$g_N := \sum_{n \leq N} f_n$$

is Cauchy: For $M \geq N$,

$$\|g_N - g_M\| = \left\| \sum_{n=N+1}^{M} f_n \right\| \leq \sum_{n=N+1}^{M} \|f_n\| \leq \sum_{n>N} \|f_n\| \to 0$$

as $N \to 0$ because the tail of a convergent series converges to 0. Hence, (g_N) has a limit f so $\sum_{n\in\mathbb{N}} f_n = f$.

$(\Leftarrow)$: Let (f_n) be a Cauchy sequence. Then we consider the series

$$\sum_{n\in\mathbb{N}} (f_{n+1} - f_n)$$

because its partial sums

$$\sum_{n<N} (f_{n+1} - f_n) = f_N - f_0$$

converge if and only if (f_n) converges. Thus it is enough to show that this series converges absolutely, i.e.

$$\sum_{n\in\mathbb{N}} \|f_{n+1} - f_n\| < \infty.$$

But this may not be true (e.g. $f_n := n^{-1/2}$ in $V := \mathbb{R}$), so we use the **acceleration trick for Cauchy sequences**, i.e. we recall that a Cauchy sequence converges if and only if it has a convergent subsequence, and we pass to a subsequence (f_{n_k}) such that

$$\left\| f_{n_{k+1}} - f_{n_k} \right\| \leq 2^{-K}$$

(which we can do using the Cauchy property)/ So without loss of generality, we assume that (f_n) had this property to begin with, i.e. $\|f_{n+1} - f_n\| \leq 2^{-n}$. Then the series $\sum_{n\in\mathbb{N}}(f_{n+1} - f_n)$ converges absolutely since the terms are summable. Hence, it converges in norm by assumption, and therefore so does the sequence (f_n). $\square$

Theorem 3.4.4. For any measure sapce $(X, \mathcal{B}, \mu)$, the space $L^1(X, \mathcal{B}, \mu)$ is complete.

Proof. Let $\sum_{n\in\mathbb{N}} f_n$ be an absolutely convergent series, i.e. $\sum_{n\in\mathbb{N}} \|f_n\| < \infty$. Then $g := \sum_{n\in\mathbb{N}} |f_n|$ dominates the sequence of partial sums $\sum_{n<N} f_n$ and g is in $L^1(X,\mu)$ since

$$\|g\|_1 = \int \sum_{n\in\mathbb{N}} |f_n| \, d\mu = \sum \int |f_n| \, d\mu = \sum_{n\in\mathbb{N}} \|f_n\| < \infty.$$

In particular, the function $\sum_{n\in\mathbb{N}} |f_n|$ is finite almost everywhere (since it is integrable) hence the series $\sum_{n\in\mathbb{N}} f_n(x)$ converges absolutely (i.e. $\sum_{n\in\mathbb{N}} |f_n(x)| < \infty$) for almost every $x \in X$, hence converges almost everywhere, so let $f(x) := \sum_{n\in\mathbb{N}} f_n(x)$ almost everywhere. Then f is measurable being the limit of partial sums (hence measurable functions) and

$$|f| = \left|\sum_{n\in\mathbb{N}} f_n\right| \leq \sum_{n\in\mathbb{N}} |f_n| = g.$$

As such, we may finally apply dominated convergence theorem to the sequence $(\sum_{n<N} f_n)_{N\in\mathbb{N}}$ of partial sums and get that these partial sums converge to f in the L^1-metric. $\square$

3.4.3 Properties of Integrable Functions

Proposition 43 (Chebyshev's inequality)**.** For each $f \in L^1(X,\mu)$ and $\alpha \in (0,\infty]$, we have that, for $A_\alpha := \{x \in X : |f(x)| > \alpha\}$,

$$\mu(A_\alpha) \leq \frac{1}{\alpha} \cdot \|f\|_1 .$$

Proof. Consider:

$$\|f\|_1 = \int |f| \; d\mu \geq \int_{A_\alpha} |f| \, d\mu \geq \int_{A_\alpha} \alpha \, d\mu = \alpha \cdot \int_{A_\alpha} d\mu = \alpha \cdot \mu(A_\alpha).$$

$\square$

Corollary 3.4.3. For each $f \in L^1$, its support $supp(f) := \{x \in X : f(x) \neq 0\}$ is σ-finite for μ, i.e. $supp(f) = \bigcup_{n\in\mathbb{N}} B_n$, where B_n is measurable and of finite measure.

Proof. Let $B_n := \{x \in X : |f(x)| > 1/n\}$, for $n \in \mathbb{N}^+$. Then, $supp(f) = \bigcup_{n\in\mathbb{N}} B_n$ and $\mu(B_n) \leq n \cdot \|f\|_1 < \infty$. $\square$

Recall that for any $f \in L^1$, we have a measure $\mu_{|f|}(A) := \int_A |f| d\mu$.

Proposition 44 (99 % Boundedness)**.** For any $f \in L^1(X, \mu)$ and $\varepsilon > 0$, there is a measurable $X' \subseteq X$ such that $f|_{X'}$ is bounded and

$$\mu_{|f|}(X \setminus X') = \int_{X \setminus X'} |f| d\mu < \varepsilon$$

Proof. Let $X_n = \{x \in X : |f(x)| \leq n\}$. Without loss of generality, change f so that $|f| < \infty$(because it is finite almost everywhere this is fine). Then, $X = \bigcup_{n \in \mathbb{N}}^{\uparrow} X_n$ so by increasing monotone convergence for $\mu_{|f|}$, we have that

$$\lim_{n \to \infty} \mu_{|f|}(X_n) = \mu_{|f|}(X) < \infty,$$

so $\mu_{|f|}(X \setminus X_n) \to 0$ as $n \to \infty$, so for large enough $n \in \mathbb{N}$, we have that $\mu(X \setminus X_n) < \varepsilon$, so take $X' := X_n$. $\square$

Definition 3.4.9. Let $(X, \mathcal{B}, \mu)$ be a measurable space and μ, ν be measures on $(X, \mathcal{B})$. We say that ν is **absolutely continuous** with respect to μ, and write $\nu << \mu$, if every μ-null set is ν-null.

Example 18. For any $f \in L^1(X, \mu)$, the measure $\mu_{|f|}$ is finite and $\mu_{|f|} << \mu$ since if $B \subseteq X$ is μ-null, then $\mu_{|f|}(B) = \int_B |f| d\mu = 0$.

The following proposition justifies the terminology absolutely continuous:

Proposition 45. Let μ, ν be measures on $(x, \mathcal{B})$. If ν is finite, then $\nu << \mu$ if and only if for every $\varepsilon > 0$ there exists some $\delta > 0$ such that if $\mu(B) < \delta$, then $\nu(B) < \varepsilon$ for all $B \in \mathcal{B}$.

Proof. ($\Leftarrow$): Trivial because $0 < \varepsilon$.

($\Rightarrow$) : We prove the contrapositive. Assume there is some $\varepsilon > 0$ such that for all $\delta > 0$ there is some $B_\delta \in \mathcal{B}$ such that $\mu(B_\delta) < \delta$ but $\nu(B_\delta) \geq \varepsilon$. By the first application of Borel-Cantelli (Proposition 24??) applied to the collection

$$\mathcal{C} := \{B \in \mathcal{B} : \nu(B) \geq \varepsilon\},$$

and μ, this collection admits a μ-almost vanishing sequence, i.e. a decreasing sequence $(B_n) \subseteq \mathcal{C}$ with $\mu\left(\bigcap_{n \in \mathbb{N}}^{\downarrow} B_n\right) = 0$. But because ν is a finite measure, decreasing monotone convergence applied to (B_n) and ν yields

$$\nu\left(\bigcap_{n \in \mathbb{N}}^{\downarrow} B_n\right) = \lim \nu(B_N) \geq \varepsilon > 0,$$

so it can't be that $\nu << \mu$. $\square$

Corollary 3.4.4 (Absolute continuity of integrable functions)**.** For any $f \in L^1(X, \mu)$, we have that for each $\varepsilon > 0$ there exists a $\delta > 0$ such that whenever $\mu(B) < \delta$, we have that $\int_B |f| d\mu < \varepsilon$.

Remark 25. This proposition can also be shown directly from just 99% boundedness, and this is left as homework.

3.5 Convergence in Measure

Recall the example given of L^1-convergence but not pointwise, defined by: $f_n = \mathbb{1}_{[j/2^k,(j+1)/2^k]}$ where $n = 2^k + j$ with $0 \le j < 2^k$. We had that $\|f_n - 0\|_1 = \|f_n\|_1 \to 0$ so $f_n \to_{L^1} 0$ but (f_n) does not converge pointwise. However, there are subsequences of (f_n) which do converge to 0 almost everywhere, for instance (f_{2^n}). Turns out that this is a general phenomenon: Every L^1-convergent sequence admits a subsequence converging almost everywhere, To prove this, we will need an intermediate notion of convergence, called convergence in measure.

Definition 3.5.1. For a measure space (X, μ), μ-measurable functions $f, g : x \to \overline{\mathbb{R}}$, and $\alpha > 0$, we define the following:

$$\Delta_\alpha := \{x \in X : |f(x) - g(x)| \ge \alpha\},$$

$$\delta_\alpha(f, g) := \mu\left(\Delta_\alpha(f, g)\right).$$

Remark 26. For μ-measurable sets $A, B \subseteq X$, $\Delta_\alpha(\mathbb{1}_A, \mathbb{1}_B) = A \triangle B$, and $d_\mu(A, b) := \mu(A \triangle B) = \delta_\alpha(\mathbb{1}_A, \mathbb{1}_B)$, for all $0 < \alpha \le 1$.

The function δ_α does not satisfy the triangle-inequality: Let $f \equiv 0$, $g \equiv 1$, $h \equiv 2\text{m}$, then $\delta_2(f, g) = 0 = \delta_2(g, h)$ but $\delta_2(f, g) = \mu(X)$ so δ_α is not a pseudo-metric.

However, the family $\{\delta_\alpha\}_{\alpha>0}$ is "kind of pseudo-metric".

Proposition 46 (Additive triangle inequality)**.** For all $\alpha, \beta > 0$ and $f, g, h : X \to \overline{\mathbb{R}}$ all μ-measurable, we have that:

$$\Delta_{\alpha+\beta}(f, h) \subseteq \Delta_\alpha(f, g) \cup \Delta_\beta(g, h),$$

$$\delta_{\alpha+\beta}(f, h) \le \delta_\alpha(f, g) + \delta_\beta(g, h).$$

Proof. For each $x \in X$, we have that $x \in \Delta_{\alpha+\beta}(f,h)$ if and only if $|f(x) - g(x)| \geq \alpha + \beta$. And so by the real triangle inequality we have that $|f(x) - g(x)| + |g(x) - h(x)| \geq \alpha + \beta$. Thus $|f(x) - g(x)| \geq \alpha$ or $|g(x) - h(x)| \geq \beta$, so $x \in \Delta_\alpha(f,g) \cup \Delta_\beta(f,g)$. □

Definition 3.5.2. For a measure space (X, μ) and μ-measurable functions (f_n) and f, we say that (f_n) **converges in measure** to f, denoted as $f_n \to_\mu f$ if

$$\lim_{n\to\infty} \delta_\alpha(f_n, f) = 0$$

for all $\alpha > 0$.

Example 19. Let us recall some previous examples.

(a) Let $f_n := \mathbb{1}_{[n,n+1)}$, then $f_n \to 0$ pointwise but not in measure and not in L^1!

(b) Let $f_n := n^2 \cdot \mathbb{1}_{(0,1/n)}$. Then $f_n \to 0$ pointwise but not in L^1 because $\int f_n d\lambda = n$. However, $f_n \to_\lambda 0$ because for each α, $\delta_\alpha(f_n, 0) = 1/n$ for all large enough n.

As such a moral one can get from these examples is that convergence in measure does not detect how badly f_n differs from the limit, but how large is the place where they differ.

(c) Let $f_n = \mathbb{1}_{[j/2^k,(j+1)/2^k]}$ where $n = 2^k + j$ with $0 \leq j < 2^k$, Then as we saw, f_n converges to 0 in L^1, but it does not converge pointwise. Also, $f_n \to_\lambda 0$ because $\delta_\alpha(f_n, 0) = 2^{-k} \to 0$, if n is in the kth group and $\alpha \leq 1$.

The following two facts are the only general implications between these three modes of convergence:

Proposition 47. For any measure space (X, μ), if $f_n \to_{L^1} f$, then $f_n \to_\mu f$.

Proof. Suppose $f_n \to_{L^1} f$ and fix $\alpha > 0$. Then by Chebyshev:

$$\delta_\alpha(f_n, f) = \mu\left(\{x \in X : |f_n(x) - f(x)| \geq \alpha\}\right) \leq \frac{1}{\alpha} \cdot \|f - f_n\|_1 \to 0$$

as $n \to \infty$. □

Proposition 48 (Switch of quantifiers trick)**.** Let (X, μ) be a finite measure space. Let $P_n \subseteq x$ be an increasing sequence of μ-measurable sets. For every $\varepsilon > 0$, we have that: If for all $x \in X$ there exists some $n \in \mathbb{N}$ such that $x \in P_n$, then, there exists an $n \in \mathbb{N}$ for all $x \in X \setminus Z$ for some $\mu(Z) \leq \varepsilon$ such that $x \in P_n$. (alternatively we could have used the symbol $\forall^{\mu}_{-\varepsilon}$ to denote that for all $x \in X \setminus Z$ for some $\mu(Z) \leq \varepsilon$.)

Proof. We know that since X is finite, $\bigcup^{\uparrow}_{n\in\mathbb{N}} P_n = X$, so $\lim \mu(P_n) = \mu(X)$, hence for a large enough $n \in \mathbb{N}$, we have that $\mu(P_n) \geq \mu(X) - \varepsilon$. □

By a switch of quantifiers trick, one can also prove Egorov's theorem about almost uniform convergence.

Now, let us study convergence in measure.

Proposition 49 (Almost uniqueness of limit)**.** In any measure space if (X, μ), $f_n \to_\mu f$ and $f_n \to_\mu g$, then $f = g$ almost everywhere.

Proof. For each $\alpha > 0$, we have that $\delta_\alpha(f, g) \leq \delta_{\alpha/2}(f, f_n) + \delta_{\alpha/2}(f_n, g) \to 0$ as $n \to \infty$ so $\delta_\alpha(f, g) = 0$ for all $\alpha > 0$. Hence, $f = g$ almost everywhere since $\{x \in X : |f(x) - g(x)| > 0\} = \bigcup_{n\in\mathbb{N}} \Delta_{1/n}(f, g)$ and the latter is null. □

Definition 3.5.3. Call a sequence (f_n) **Cauchy in measure** if for each $\alpha > 0$, we have that $\delta_\alpha(f_n, f_m) \to 0$ as $\min(n, m) \to \infty$.

Proposition 50. Similarly to the case of regular Cauchy sequences, we have that:

(a) If $f_n \to_\mu f$ then (f_n) is Cauchy in measure.

(b) If (f_n) is Cauchy in measure and admits a subsequence $f_{n_k} \to_\mu f$ as $k \to \infty$, then $f_n \to_\mu f$ as $n \to \infty$.

Proof. Left as homework. □

Theorem 3.5.1 (Completeness of convergence in measure)**.** Every sequence (f_n) which is Cauchy in sequence converges in measure, i.e. there exists a μ-measurable function f such that $f_n \to_\mu f$. Moreover, $f_{n_k} \to_\mu f$ almost everywhere for some subsequence (f_{n_k}).

Proof. Note that by part (b) of the previous proposition, we may restrict to any subsequence (thanks to acceleration).

First, we claim that without loss of generality, $\delta_{2^{-n}}(f_n, f_{n+1}) \leq 2^{-n}$ for all n, by restricting to a subsequence.

We define (n_k) recursively: let $n_{-1} = 0$, and choose $n_k > n_{k-1}$ such that $\delta_{2^{-k}}(f_{n_k}, f_n) \leq 2^{-k}$ for all $n \geq n_k$. Such a n_k exists by the Cauchy condition with $\alpha := 2^{-k}$.

We now show that for almost any $x \in X$, $(f_n(x))$ is Cauchy.

Our second claim is that if $x \notin B_N := \bigcup_{n \geq N} \Delta_{2^{-n}}(f_n, f_{n+1})$ then for all $m \geq n \geq N$, we have

$$|f_n(x) - f_m(x)| \leq 2^{-(n-1)} \to 0$$

as $n \to \infty$, so $(f_n(x))$ is Cauchy. Indeed,

$$|f_n(x) - f_m(x)| \leq \sum_{i=n}^{m-1} |f_i(x) - f_{i+1}(x)| \leq \sum_{i=n}^{m-1} 2^{-i} \leq \sum_{i=n}^{\infty} 2^{-i} = 2^{-(n-1)}.$$

But,

$$\mu(B) \leq \sum_{n \geq N} \mu\left(\Delta_{2^{-n}}(f_n, f_{n+1})\right) = \sum_{n \geq N} \delta_{2^{-n}}(f_n, f_{n+1}) \leq \sum_{n \geq N} 2^{-n} = 2^{-(N-1)},$$

which is summable, so by Borel-Cantelli we have that almost everywhere $x \in X$ is eventually not in B_N, i.e. there is some N such that $x \notin \bigcup_{n \geq N} B_n = B_N$ (the last equality is due to (B_m) being increasing) thus by our second claim, $(f_n(x))$ is Cauchy and let $f(x)$ denote the limit. The function $f : X \to \overline{\mathbb{R}}$ (defined almost everywhere) is μ-measurable being an almost everywhere pointwise limit of μ-measurable functions f_n. It remains to show that $f_n \to_\mu f$. To this end, fix $\alpha > 0$, and chose $N \in \mathbb{N}$ so that $2^{-(N-2)} \leq \alpha$. Then $\Delta_\alpha(f_N, f) \subseteq \Delta_{2^{-(N-2)}}(f_N, f)$, and by our second claim, if $x \notin B_N$, then

$$|f_N(x) - f(x)| = \lim_{m \to \infty} |f_N(x) - f_m(x)| \leq 2^{-(N-1)} < 2^{-(N-2)},$$

so $x \notin \Delta_{2^{-(N-2)}}(f_N, f)$. In other words, $\Delta_{2^{-(N-2)}}(f_N, f) \subseteq B_N$, so

$$\delta_\alpha(f_N, f) \leq \delta_{2^{-(N-2)}}(f_N, f) \leq \mu(B_N) \leq 2^{-(N-1)} \to 0$$

as $N \to \infty$, so $f_N \to_\mu f$ as $N \to \infty$. $\square$

Corollary 3.5.1. In any measure space (X, μ), if $f_n \to_\mu f$ then $f_{n_k} \to f$ almost everywhere for some subsequence (f_{n_k}). In particular, if $f_n \to_{L^1} f$ then $f_{n_k} \to f$ almost everywhere for some subsequence (f_{n_k}).

Proof. Since $f_n \to_\mu f$, then (f_n) is Cauchy in measure, so for some μ-measurable function g, $f_n \to_\mu g$ and $f_{n_k} \to g$ almost everywhere for some subsequence (f_{n_k}). But limits are almost unique in convergence in measure so $g = f$ almost everywhere. □

3.6 Product Measures

Let $(X, \mathcal{I})$ and $(Y, \mathcal{J})$ be measurable spaces. Recall that $I \otimes J$ denotes the σ-algebra generated by rectangles, i.e. sets of the form $I \times J$ where $I \in \mathcal{I}$ and $J \in \mathcal{J}$.

Theorem 3.6.1. For any measure spaces $(X, \mathcal{I}, \mu)$ and $(Y, \mathcal{J}, \nu)$, there is a measure ρ on $(X \times y, \mathcal{I} \otimes \mathcal{J})$ such that $\rho(I \times J) = \mu(I) \cdot \nu(J)$ for all rectangles $I \times J$. If μ and ν are σ-finite, then such a measure ρ is unique and is called the **product measure** if μ and ν, denoted by $\mu \times \nu$.

Proof. Let $\mathcal{A}$ denote the algebra generated by all rectangles, hence $\mathcal{A}$ consists of finite disjoint unions of rectangles (because the intersection of two rectangles is a rectangle and the compliment of a rectangle is a disjoint union of three rectangles). To prove the theorem, it is enough to show that the formula

$$\rho(I \times J) = \mu(I) \cdot \nu(J)$$

defines a (countably additive) premeasure on $\mathcal{A}$ and then apply Caratheodory's theorem (noting that if μ and ν are σ-finite, then so would the premeasure ρ: Indeed, if $X = \bigcup_{n\in\mathbb{N}} I_n$ where $\mu(I_n) < \infty$ and $Y = \bigcup_{m\in\mathbb{N}} J_m$ with $\nu(I_m) < \infty$, then $X \times Y = \bigcup_{n,m} I_n \times J_m$ and $\rho(I_n \times J_m) = \mu(I_n) \cdot \nu(J_m) < \infty$). To this end, it suffices to show that

$$\mu(I) \cdot \nu(J) = \rho(I \times J) = \sum_{n\in\mathbb{N}} \rho(I_n \times J_n) = \sum_{n\in\mathbb{N}} \mu(I_n) \cdot \nu(J_n) \tag{1}$$

whenever

$$I \times J = \bigsqcup_{n\in\mathbb{N}} I_n \times J_n. \tag{2}$$

Assuming (2), note that $\mathbb{1}_{I\times J} = \sum \mathbb{1}_{I_n \times J_n}$ and $\mathbb{1}_{U\times V}(x, y) = \mathbb{1}_U(x) \cdot \mathbb{1}_V(y)$, thus we have that $\mathbb{1}_I(x) \cdot \mathbb{1}_J(y) = \sum_{n\in\mathbb{N}} \mathbb{1}_{I_n}(x) \cdot \mathbb{1}_{J_n}(x)$. For each $x \in X$, integrating over Y, we get:

$$\begin{aligned}\mathbb{1}_I \cdot \nu(J) &= \int_Y \mathbb{1}_I(x) \cdot \mathbb{1}_J(y) d\nu(y) \\ &= \int_Y \sum_n \mathbb{1}_{I_n}(x) \cdot \mathbb{1}_{J_n}(x) d\nu(y) \\ &= \sum_n \int_Y \mathbb{1}_{I_n}(x) \cdot \mathbb{1}_{J_n}(x) d\nu(y) \\ &= \sum_n \mathbb{1}_{I_n}(x) \cdot \nu(J).\end{aligned}$$

Now we integarate over X:

$$\begin{aligned}\mu(I) \cdot \nu(J) &= \int_X \mathbb{1}_I(x) \cdot \nu(J) d\mu(x) \\ &= \int_X \sum_{n\in\mathbb{N}} \mathbb{1}_{I_n}(x) \nu(J_n) d\mu(x) \\ &= \sum_{n\in\mathbb{N}} \int_X \mathbb{1}_{I_n}(x) \nu(J_n) d\mu(x) \\ &= \sum_{n\in\mathbb{N}} \mu(I_n) \cdot \nu(J_n).\end{aligned}$$

$\square$

Remark 27. By the uniqueness part in Caratheodory's theorem, the Lebesgue measure λ_d on $\mathcal{B}\left(\mathbb{R}^d\right)$ that we defined is equal to the product of length d. $\lambda_1 \times \ldots \lambda_1$,where λ_1 is the Lebesgue measure on $\mathcal{B}(\mathbb{R})$. Similarly, $\lambda_3 = \lambda_1 \times \lambda_2$.

3.6.1 Fubini-Tonelli Theorem

Definition 3.6.1. Let X, Y, Z be sets and $x_0 \in X$, $y_0 \in Y$.

(a) For a set $R \subseteq X \times Y$, call the sets

$$R_{x_0} := \{y \in Y : (x_0, y) \in R\}$$

$$R^{y_0} := \{x \in X : (x, y_0) \in R\}$$

the **vertical** and **horizontal fibres** of R at x_0 and y_0, respectively.

(b) For a function $f : X \times Y \to Z$, call the functions

$$\begin{aligned} f_{x_0} : Y &\longrightarrow Z \\ y &\mapsto f(x_0, y) \end{aligned}$$

and

$$\begin{aligned} f^{y_0} : X &\longrightarrow Z \\ x &\mapsto f(x, y_0) \end{aligned}$$

the **vertical and horizontal fibres** of f at x_0 and y_0 respectively.

The following is one of the most useful theorems in math:

Theorem 3.6.2 (Fubini-Tonelli)**.** Let $(X, \mathcal{I}, \mu)$ and $(Y, \mathcal{J}, \nu)$ be σ-finite measure spaces. Let $f : X \times Y \to \overline{\mathbb{R}}$ be a $\mu \times \nu$ measurable function. Then:

(a) $f_x : Y \to \overline{\mathbb{R}}$ and $f^y : X \to \overline{\mathbb{R}}$ are ν-measurable and μ-measurable for μ almost every $x \in X$ and ν-almost every $y \in Y$, respectively.

(b) **Tonelli:** If $f \geq 0$, then:

(i) the functions $x \mapsto \int_Y f_x d\nu$ and $y \mapsto \int_X f^y d\mu$ are μ-measurable and ν-measurable respectively.

(ii) Moreover,

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

(c) **Fubini:** If f is $\mu \times \nu$-integrable, then

(i) the functions $x \mapsto \int_Y f_x d\nu$ and $y \mapsto \int_X f^y d\mu$ are μ-measurable and μ-integrable and ν-measurable and ν-integrable respectively.

(ii) Moreover,

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

Remark 28. Usually, we first apply Tonelli's theorem to $|f|$ to verify that $\int_{X \times Y} |f|\, d\mu \times \nu < \infty$, and then we apply Fubini's theorem to f.

Example 20. For sets N, M (e.g. $N = \mathbb{N} = M$) and counting measures μ and ν on N and M respectively, the Fubini-Tonelli theorem says something we already know: Let $(a_{nm})_{n\in N, m\in M}$ be a matrix of reals.

Tonelli tells us that if $a_{nm} \geq 0$ for all $n \in N$ and $m \in M$, then

$$\sum_{n\in N}\sum_{m\in M} a_{nm} = \sum_{m\in M}\sum_{n\in N} a_{nm}.$$

Moreover, Fubini tells us that if (a_{nm}) is absolutely summable, i.e.

$$\sum_{(n,m)\in N\times M} |a_{nm}| < \infty,$$

then

$$\sum_{n\in N}\sum_{m\in M} a_{nm} = \sum_{(n,m)\in N\times M} a_{nm} = \sum_{m\in M}\sum_{n\in N} a_{nm}.$$

We will prove a version of Fubini-Tonelli theorem with $\mathcal{I}\otimes\mathcal{J}$-measurability, and leave it as a exercise to deduce the $\mu \times \nu$-measurable version.

Proposition 51. Let $(X, \mathcal{I})$ and $(Y, \mathcal{J})$ be measurable spaces.

(a) For each $\mathcal{I}\otimes\mathcal{J}$-measurable set $R \subseteq X \times Y$, i.e. $R \in \mathcal{I}\times\mathcal{J}$, we have that all fibres R_x and R^y are $\mathcal{J}$-measurable and $\mathcal{I}$-measurable respectively, i.e. $R_x \in \mathcal{J}$ and $R^y \in \mathcal{I}$.

(b) For each $\mathcal{I}\otimes\mathcal{J}$-measurable function $f : X \times Y \to Z$, where $(Z, \mathcal{K})$ is some measurable space, all fibres f_x and f^y are $\mathcal{J}$-measurable and $\mathcal{I}$-measurable respectively.

Proof. (a): Let $\mathcal{C} := \{R \in \mathcal{I}\otimes\mathcal{J} : \forall x \forall y,\ R_x \in \mathcal{J} \text{ and } R^y \in \mathcal{I}\}$. Then by definition $\mathcal{C}$, contains all rectangles $I \times J$, hence

$$(I \times J)_x = \begin{cases} J \text{ if } x \in I \\ \emptyset, \text{ otherwise.} \end{cases}$$

and similarly for $y \in Y$. Also $\mathcal{C}$ is closed under compliments and countable unions because these operations commute when taking fibres, i.e.

$$\left(\bigcup_n R_n\right)_x = \bigcup_n (R_n)_x$$

and $(R^c)_x = (R_x)^c$.

(b): Given $W \in \mathcal{K}$, we verify that $f_x^{-1}(W)$ and $(f^y)^{-1}(W)$ are $\mathcal{J}$-measurable and $\mathcal{I}$-measurable, respectively, because preimages commute with fibres: $f_x^{-1}(W) = (f^{-1}(W))_x$ and $(f^y)^{-1}(W) = (f^{-1}(W))^y$, which are in $\mathcal{J}$ and $\mathcal{I}$ by part (a). $\square$

Theorem 3.6.3 (Fubini-Tonelli for Sets)**.** Let $(X, \mathcal{I}, mu)$ and $(Y, \mathcal{J}, \nu)$ be σ-finite measure spaces. Let $\mathcal{R} \subseteq \mathcal{I} \otimes \mathcal{J}$ (i.e. $\mathbb{1}_{\mathcal{R}}$ is measurable). Then:

(i) The functions $x \mapsto \nu(\mathcal{R}_x)$ and $y \mapsto \mu(\mathcal{R}^y)$ are $\mathcal{I}$ and $\mathcal{J}$-measurable, respectively.

(ii) $\int_X \nu(\mathcal{R}_x) d\mu(x) = \mu \times \nu(\mathcal{R}) = \int_Y \mu(\mathcal{R}^y) d\nu(y)$.

Proof. Firstly, we may assume without loss of generality that μ and ν are finite by the usual argument of writing $X = \bigsqcup_{n\in\mathbb{N}} X_n$ and $Y = \bigsqcup_{m\in\mathbb{N}} Y_m$ for some $X_n \in \mathcal{I}$ and $Y_m \in \mathcal{J}$ of finite measure so $X \times Y = \bigsqcup_{m,n\in\mathbb{N}} X_n \times Y_m$, and $\mathcal{R} = \bigsqcup_{n,m} (\mathcal{R} \cap X_n \times Y_m)$ and using closure of limits for (i) and countable additivity for (ii).

Let $\mathcal{C}$ be the collection of $R \in \mathcal{I} \otimes \mathcal{J}$ that satisfy (i) and (ii). We aim to show that $\mathcal{C}$ is a σ-algebra containing the algebra $\mathcal{A}$ generated by rectangles $I \times J$ with $I \in \mathcal{I}$ and $J \in \mathcal{J}$.

$\mathcal{C}$ contains rectangles: Indeed, for $I \in \mathcal{I}$ and $J \in \mathcal{J}$, the function $x \mapsto \nu((I \times J)_x) = \nu(J) \times \mathbb{1}_I(x)$ which is clearly measurable and same for the y-fibres. For (ii), observe that

$$\int_X \nu((I \times J)_x) d\mu(x) = \int_X \nu(J) \cdot \mathbb{1}_I(x) d\mu(x) = \nu(J) \cdot \mu(I) = \mu \times \nu(I \times J),$$

and same for the y-fibres.

$\mathcal{C}$ is closed under disjoint unions: This follows from the fact that measures are finitely additive, and measurable functions are closed under addition and the integral is linear.

Therefore, $\mathcal{C}$ contains the algebra $\mathcal{A}$ because each element is a finite disjoint union of rectangles.

Using closedness under limits of measurable functions and monotone convergence theorem, we also get that $\mathcal{C}$ is closed under countable disjoint unions, and using the finiteness of μ,ν, and $\mu \times \nu$, we can also deduce that $\mathcal{C}$ is closed under complements, for instance $\mu \times \nu(R^c) = \mu \times \nu(X \times Y) - \mu \times \nu(R)$. But in order to conclude that $\mathcal{C}$ is a σ-algebra, we still need to show it is closed under finite intersections (to disjointify a countable union, this is needed), which is hard to show because the measure of an intersection is not expressible by the measures of sets. Instead, we appeal to the soon to be introduced Monotone class lemma, an see that it is enough to verify that $\mathcal{C}$ is closed under countable increasing unions and countable decreasing intersections ,i.e. it is a monotone class, because then, $\mathcal{C} \supseteq \left\langle \mathcal{A} \right\rangle_\sigma = \mathcal{I} \otimes \mathcal{J}$, hence $\mathcal{C} = \mathcal{I} \otimes \mathcal{J}$.

$\mathcal{C}$ is a monotone class: For (i) use the monotone convergence properties of measures (including the decreasing convergence due to finiteness of measures) and closedness of measurable functions under pointwise limits For (ii), for increasing unions apply monotone convergence theorem, and for decreasing intersections, apply dominated convergence theorem. $\square$

Definition 3.6.2. A collection $\mathcal{C}$ of subsets of a set X is called a **monotone class** if it is closed under countable increasing unions and countable decreasing intersections. The **monotone class generated by a collection $\mathcal{A}$** is the $\subseteq$-least monotone class containing $\mathcal{A}$, i.e. the intersection of all monotone classes containing $\mathcal{A}$

Lemma 3.6.1 (Monotone Class Lemma)**.** Let $\mathcal{C} \subseteq \mathcal{P}(X)$ be a monotone class. If $\mathcal{C}$ contains an algebra $\mathcal{A} \subseteq \mathcal{P}(X)$, then $\mathcal{C} \supseteq \left\langle \mathcal{A} \right\rangle_\sigma$.

Proof. By shrinking $\mathcal{C}$, we may assume without loss generality that $\mathcal{C}$ is the monotone class generated by $\mathcal{A}$. Then we show that $\mathcal{C} = \left\langle \mathcal{A} \right\rangle_\sigma$. For this we need to show that $\mathcal{C}$ is closed under complements and countable unions, but a countable union $\bigcup_{n\in\mathbb{N}} C_n = \bigcup^{\uparrow}_{n\in\mathbb{N}}(\bigcup_{m\leq n} C_m)$ and $\mathcal{C}$ is closed under countable increasing unions, so it is enough to show that $\mathcal{C}$ is closed under finite unions and complements, i.e. is an algebra.

Complements: Let $\mathcal{S} := \{S \in \mathcal{C} : S^c \in \mathcal{C}\}$ and show that $\mathcal{S} \supseteq \mathcal{A}$ and that it is a monotone class. But for $A \in \mathcal{A}$, A^c is also in $\mathcal{A}$ so $A \in \mathcal{S}$. As or countable increasing unions and countable decreasing unions, observe that the complements of a countable increasing union of sets is a countable decreasing intersection of their complement, and that the complements of a decreasing intersection of sets is a increasing union of their complements, so $\mathcal{C}$ being closed under these operations then so is $\mathcal{S}$, hence $\mathcal{S} = \mathcal{C}$.

Finite unions: For each $U \in \mathcal{C}$, let $\mathcal{S}(U) = \{V \in \mathcal{C} : U \cup V \in \mathcal{C}\}$. We need to show that for each $U \in \mathcal{C}$, the collection $\mathcal{S}(U) \supseteq \mathcal{A}$ is a monotone class, because then $\mathcal{S}(U) = \mathcal{C}$ hence $\mathcal{C}$ is closed finite unions. To check that $\mathcal{S}(U)$ is a monotone class, suppose $\{V_n\} \subseteq \mathcal{S}(U)$ and is increasing, and observe that $U \cup \bigcup^{\uparrow}_{n\in\mathbb{N}} V_n = \bigcup^{\uparrow}_n (U \cup V_n)$ hence it is in $\mathcal{S}(U)$ because $\mathcal{C}$ is closed under countable increasing unions. Suppose $\{V_n\}$ is now decreasing, then $U \cup \bigcap_n^{\downarrow} V_n = \bigcap_n^{\downarrow}(U \cap V_n)$ and the latter is in $\mathcal{C}$ since $\mathcal{C}$ is closed under countable decreasing intersections.

Finally, to show that $\mathcal{A} \subseteq \mathcal{S}(U)$, we fix $A \in \mathcal{A}$ and show that $A \in \mathcal{S}(U)$. But the latter is equivalent to $U \in \mathcal{S}(A)$. We in fact show that $\mathcal{S}(A) = \mathcal{C}$ hence $U \in \mathcal{S}(A)$. We already know that $\mathcal{S}(A)$ is a monotone class (we proved it for all $U \in \mathcal{C}$). Also $\mathcal{A} \subseteq \mathcal{S}(A)$ because $\mathcal{A}$ is closed under finite unions and $\mathcal{C} \supseteq \mathcal{A}$. Thus, $\mathcal{S}(A)$ contains the monotone class generated by $\mathcal{A}$, hence $\mathcal{S}(A) = \mathcal{C}$. □

Theorem 3.6.4 (Fubini-Tonelli for $\mathcal{I} \otimes \mathcal{J}$)**.** Let $(X, \mathcal{I}, \mu)$ and $(Y, \mathcal{J}, \nu)$ be σ-finite measure spaces. Let $f : X \times Y \to \overline{\mathbb{R}}$ be a $\mathcal{I} \otimes \mathcal{J}$-measurable function. Then:

(a) $f_x : Y \to \overline{\mathbb{R}}$ and $f^y : X \to \overline{\mathbb{R}}$ are $\mathcal{J}$ and $\mathcal{I}$-measurable for all $x \in X$ and $y \in Y$.

(b) **Tonelli:** If $f \geq 0$, then:

(i) The maps $x \mapsto \int_Y f_x d\mu$ and $y \mapsto \int_X f^y d\mu$ are $\mathcal{I}$ and $\mathcal{J}$-measurable.

(ii) We may switch the order of integration as follows:

$$\int_X \int_Y f_x(y) d\mu(y) d\nu(x) = \int_{X \times Y} f d(\mu \times \nu) = \int_Y \int_X f^y(x) d\mu(x) d\nu(y)$$

(c) **Fubini:** If f is $\mu \times \nu$-integrable, then:

(i) The maps $x \mapsto \int_Y f_x d\mu$ and $y \mapsto \int_X f^y d\mu$ are $\mathcal{I}$ and $\mathcal{J}$-integrable.

(ii) We may switch the order of integration as follows:

$$\int_X \int_Y f_x(y) d\mu(y) d\nu(x) = \int_{X \times Y} f d(\mu \times \nu) = \int_Y \int_X f^y(x) d\mu(x) d\nu(y)$$

Proof. We have already proven (a) and we know (b) for indicator functions, which implies (b) and (c) for simple functions by the linearity of the integral.

To conclude (b) for all $f \geq 0$, write f as an increasing limit of non-negative simple functions and in (i) use the closedness of measurable functions under pointwise limits and monotone convergence theorem. For (ii), just use the monotone convergence theorem three times. Finally, for (c) write $f = f_+ - f_-$, so f_+ and f_- are $\mu \times \nu$ integrable and apply (b) to f_+ and f_- individually, observing te finiteness of $\int f_+ d\mu \times \nu$ implies that the functions $x \mapsto \int f_x d\nu$ and $y \mapsto \int f^y d\mu$ are finite almost everywhere. Get (c) for f by the linearity of the integral. □

The $\mu \times \nu$-measurable version of Fubini-Tonelli follow from the $\mathcal{I} \otimes \mathcal{J}$ version, and is left as a homework exercise.

3.6.2 Infinite Products

We already learned how to construct the product of two measure spaces, and hence also finitely many spaces by induction. We would like to extend this to infinite products. First, let us recall/learn the product topology:

Definition 3.6.3. If $X_{i i \in I}$, where I is any index set, is a collection of topological spaces, then the product of these spaces is the set

$$\prod_{i \in I} X_i := \{(x_i)_{i \in I} : x_i \in X_i, \text{ for all } i \in I\}$$

and the **product topology** on this set is the one generated by **open cylinders:**

$$[i_1 \mapsto U_{i_1}, i_1 \mapsto U_{i_2}, \ldots, ik \mapsto U_{i_k}] := \left\{ (x_i)_{i \in I} \in \prod_{i \in I} X_i : \ \forall 1 \leq j \leq k, \ x_{i_j} \in U_{i_j} \right\}$$

for some finitely many indices $i_1, \ldots, i_k \in I$ and open $U_{i_j} \subseteq X_{i_j}$. In other words, in the product topology, the open sets are the arbitrary unions of open cylinders.

Index