

Emory University
MATH 512 Analysis II: Measure Theory
Learning Notes

Jiuru Lyu

April 29, 2026

Contents

1	Motivation and Prerequisites	3
1.1	Motivation of the Lebesgue Integration	3
1.2	Measuring Sets	6
1.3	Metric Spaces	8
1.4	Topology	10
1.5	Upper and Lower Bounds	12
2	Borel Sets and Functions	16
2.1	σ -Algebra	16
2.2	Borel Sets and Functions	18
2.3	Borel σ -Algebra on Euclidean Spaces	19
2.4	Extended Real Line and Limits	23
3	Measure and Abstract Integration	25
3.1	Measures	25
3.2	lim sup of Sets	28
3.3	Integration of Positive Functions	29
3.4	Integration of Real and Complex Function	40
4	Lebesgue-Stieltjes Measures	48
4.1	Increasing Functions and Distribution Functions	48
4.2	Outer Measure	51
4.3	Borel Measures on \mathbb{R}	55
4.4	The Monotone Class Theorem	59

5	Product Measures	64
5.1	Product Measures	64
5.2	Tonelli's and Fubini's Theorem	66
5.3	Lebesgue Measure on \mathbb{R}^n	69
6	L^p Spaces and Fourier Transforms	73
6.1	Norms, Banach, and L^p Spaces	73
6.2	Approximation of Sets and Functions	80
6.3	Continuity of Translation	83
6.4	Fourier Series and the Fejér Kernel	85
6.5	Convolutions	90
6.6	Fourier Integrals	93
7	Hilbert Space	97
7.1	Inner Product	97
7.2	Orthogonal Complement and Projection	102
7.3	Orthonormal Sets and Bases	107
8	Bonus Topics	110
8.1	Cantor Set and Cantor Function	110

1 Motivation and Prerequisites

1.1 Motivation of the Lebesgue Integration

Riemann integration has some flaws even for bounded functions on bounded domains.

Question: When can we exchange limits and integrals?

For Riemann integrals, we have

Theorem 1.1.1

If $f_n : [a, b] \rightarrow \mathbb{R}$ are Riemann integrable functions and if they *convergence uniformly* to f , i.e., $\forall \varepsilon > 0, \exists N \in \mathbb{N}$ s.t. $\forall n \geq N, \forall x \in [a, b], |f(x) - f_n(x)| < \varepsilon$. Then, f is Riemann integrable, and

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx = \int_a^b f(x) dx.$$

Remark. While this is a useful theorem, it is limited in space. Uniform convergence is a very strong condition.

Example 1.1.2

Let $f_n(x) = x^n$ on $[0, 1]$. They converge to the function

$$f(x) = \begin{cases} 0 & x \in [0, 1) \\ 1 & x = 1 \end{cases}$$

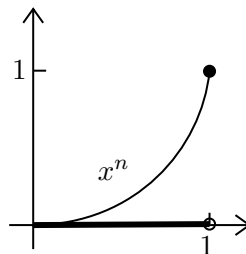
pointwise, but not uniformly. Nonetheless,

$$\lim_{n \rightarrow \infty} \int_0^1 x^n dx = \int_0^1 \lim_{n \rightarrow \infty} x^n dx = \int_0^1 f(x) dx.$$

Proof 1.

- Pointwise convergence. ✓
- Convergence is not uniform:

$$\begin{aligned} \sup_{x \in [0, 1]} |x^n - f(x)| &= \sup_{x \in [0, 1]} |x^n - 0| \\ &= \sup_{x \in [0, 1]} |x^n| \\ &= 1. \end{aligned}$$



Hence, we don't have uniform convergence.

- Integral and limit can exchange:

$$\begin{aligned}\lim_{n \rightarrow \infty} \int_0^1 x^n dx &= \lim_{n \rightarrow \infty} \left[\frac{1}{n+1} x^{n+1} \right]_0^1 \\ &= \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0 \\ \int_0^1 f(x) dx &= 0.\end{aligned}$$

Hence,

$$\lim_{n \rightarrow \infty} \int_0^1 x^n dx = 0 = \int_0^1 f(x) dx.$$

Q.E.D. ■

Remark. There are two conclusions in the above Theorem 1.1.1:

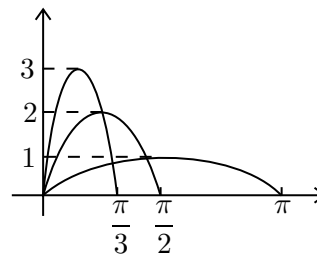
- limiting function is Riemann integrable, and
- its integral is equal to the specific value.

If we replace uniform convergence with pointwise convergence, both of these conclusions can fail.

Example 1.1.3 Pointwise Convergence Does Not Guarantee Convergence of the Integrals

Consider the following continuous functions:

$$f_n(x) = \begin{cases} n \sin(nx) & 0 < x < \frac{\pi}{n} \\ 0 & \text{o/w.} \end{cases}$$



f_n converge pointwise to 0. But, for each n ,

$$\int_0^\pi f_n(x) dx = 2.$$

In particular,

$$\lim_{n \rightarrow \infty} \int_0^\pi f_n(x) dx = 2 \neq 0 = \int_0^\pi f(x) dx = \int_0^\pi 0 dx.$$

Notation (Characteristic Function)

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A, \end{cases}$$

for a set A . χ_A is the characteristic function of the set A .

Example 1.1.4 Pointwise Convergence Does Not Guarantee Riemann Integrable Limiting Function

Let $\{q_n\}_{n=1}^{\infty}$ be some enumeration of the countable set $\mathbb{Q} \cap [0, 1]$. i.e., a sequence *s.t.*

$$\{q_n \mid n \in \mathbb{N}\} = \mathbb{Q} \cap [0, 1]$$

and $q_n \neq q_k$ if $n \neq k$. For each $n \in \mathbb{N}$, define

$$B_n = \{q_1, q_2, \dots, q_n\}.$$

Then,

- $\forall n \in \mathbb{N}$, χ_{B_n} is pointwise continuous and thus Riemann integrable, and

$$\int_0^1 \chi_{B_n}(x) dx = 0.$$

- $\chi_{\mathbb{Q} \cap [0,1]}$ is pointwise limit of χ_{B_n} , but it is not Riemann integrable because

$$\int_0^1 \chi_{\mathbb{Q} \cap [0,1]}(x) dx = 0 \neq 1 = \int_0^1 \chi_{\mathbb{Q} \cap [0,1]}(x) dx.$$

Proof 2.

- χ_{B_n} is continuous except at finitely many points: q_1, q_2, \dots, q_n . Therefore, it is Riemann integrable with integral equal to 0.
- For $\chi_{\mathbb{Q} \cap [0,1]}$, its minimum and maximum value on *any* interval is 0 and 1, respectively.

Q.E.D. ■

These two issues are fixed as follows: In this semester, we will see a more general notation of integrability (Lebesgue integration) and prove the following theorem:

Theorem 1.1.5 Dominated Convergence Theorem (DCT), Special Case

Let $f_n : [a, b] \rightarrow \mathbb{R}$ be integrable functions, which are uniformly bounded:

$$\sup_{n \in \mathbb{N}} \sup_{x \in [a, b]} |f_n(x)| < \infty.$$

If f_n converge pointwise to $f : [a, b] \rightarrow \mathbb{R}$, then f is integrable and

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx = \int_a^b f(x) dx.$$

For example, in Lebesgue integration, $\chi_{\mathbb{Q} \cap [0, 1]}$ is integrable, and

$$\int_0^1 \chi_{\mathbb{Q} \cap [0, 1]}(x) dx = \lim_{n \rightarrow \infty} \int_0^1 \chi_{B_n}(x) dx = 0.$$

1.2 Measuring Sets**Example 1.2.1 Integration is Related to Measuring the Size of Sets**

The length of interval:

$$\int \chi_{[c, d]}(x) dx = d - c.$$

More generally, we have the following definition:

Definition 1.2.2 (Lebesgue Measure and Measurable Sets). If χ_A is integrable, the quantity

$$m(A) := \int \chi_A(x) dx$$

will be called the *Lebesgue measure* of A , and A will be called *measurable*.

Example 1.2.3

$$m(\mathbb{Q} \cap [0, 1]) = \int \chi_{\mathbb{Q} \cap [0, 1]}(x) dx = 0.$$

Definition 1.2.4 (Translation Invariant). Since translation preserves the length of intervals, we should expect the same property for Lebesgue measure. Let $A \subseteq \mathbb{R}$, $\forall t \in \mathbb{R}$, denote its translation by

$$t + A = \{t + x \mid x \in A\}.$$

We expect $m(A) = m(t + A)$.

Definition 1.2.5 (σ -additivity for Measures).

- If $\{A_n\}_{n=1}^{\infty}$ is a sequence of pairwise disjoint sets and

$$f_n = \sum_{j=1}^n \chi_{A_j},$$

we have pointwise convergence:

$$\lim_{n \rightarrow \infty} f_n(x) = \sum_{j=1}^{\infty} \chi_{A_j} = \chi_{\bigcup_{j=1}^{\infty} A_j}$$

- If we can justify the exchange of limits and integrals, we would have

$$\begin{aligned} \int \chi_{\bigcup_{j \in \mathbb{N}} A_j}(x) \, dx &= \lim_{n \rightarrow \infty} \int f_n(x) \, dx = \lim_{n \rightarrow \infty} \int \sum_{j=1}^n \chi_{A_j}(x) \, dx \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^n \int \chi_{A_j}(x) \, dx \end{aligned}$$

That is,

$$m\left(\bigcup_{j \in \mathbb{N}} A_j\right) = \lim_{n \rightarrow \infty} \sum_{j=1}^n m(A_j) = \sum_{j=1}^{\infty} m(A_j).$$

This property is called σ -*additivity*, and it will be a part of the definition of Lebesgue measure.

The Lebesgue measure cannot be defined on all subsets of \mathbb{R} . The motivation above leads to an apparent paradox.

Theorem 1.2.6 Vitali

Let $\{q_n\}_{n=1}^{\infty}$ be an enumeration of $\mathbb{Q} \cap [-1, 1]$. There exists a set $V \subset [0, 1]$ s.t.

$$(q_j + V) \cap (q_k + V) = \emptyset \quad \text{for any } j \neq k$$

and

$$[0, 1] \subset \bigcup_{j \in \mathbb{N}} (q_j + V) \subset [-1, 2].$$

Corollary 1.2.7 There is no translation invariant function $m : \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty]$, which is σ -additive and obey

$$m([a, b]) = b - a \quad \text{for any } a < b.$$

Proof 1. Assume such m exists.

- σ -additivity \implies monotonicity.

Let $A \subset B$ [WTS: $m(A) < m(B)$]. Use σ -additivity, write $A_1 = A$, $A_2 = B \setminus A$, $A_n = \emptyset$ for $n \geq 3$.

$$m(B) = m(A) + m(B \setminus A) + 0 \geq m(A).$$

- For the Vitali set $[0, 1] \subset \bigcup_{j \in \mathbb{N}} (q_j + V) \subset [-1, 2]$, by monotonicity and σ -additivity,

$$\begin{aligned} m([0, 1]) &\leq m\left(\bigcup_{j \in \mathbb{N}} (q_j + V)\right) \leq m([-1, 2]) \\ 1 &\leq \sum_{j \in \mathbb{N}} m(q_j + V) \leq 3. \end{aligned}$$

By translation invariance:

$$1 \leq \sum_{j \in \mathbb{N}} m(V) \leq 3.$$

1. If $m(V) = 0$:

$$1 \leq 0 \leq 3 \quad *$$

2. If $m(V) > 0$:

$$\sum_{j \in \mathbb{N}} m(V) = \infty \quad \text{and} \quad 1 \leq \infty \leq 3 \quad *$$

Hence, such m cannot exist.

Q.E.D. ■

1.3 Metric Spaces

Definition 1.3.1 (Metric Space of Continuous Functions).

$$\begin{aligned} \mathcal{C}([a, b]) &= \mathcal{C}([a, b], \mathbb{C}) = \{f : [a, b] \rightarrow \mathbb{C} \mid f \text{ is continuous}\} \\ d_\infty(f, g) &= \max_{x \in [a, b]} |f(x) - g(x)| \end{aligned}$$

- $\mathcal{C}([a, b])$ is a metric space with metric d_∞ .
- $\mathcal{C}([a, b])$ is also complete.

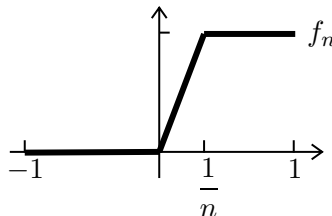
Example 1.3.2 Another Metric on Functions

$$d_1(f, g) = \int_a^b |f(x) - g(x)| dx \quad \text{on } \mathcal{C}([a, b]).$$

$\mathcal{C}([a, b])$ is a metric space with metric d_1 but not complete.

Proof 1. Consider functions

$$f_n(x) = \begin{cases} 0 & x < 0 \\ nx & 0 \leq x \leq \frac{1}{n} \\ 1 & x > \frac{1}{n} \end{cases}$$



on $[-1, 1]$.

Claim f_n is a Cauchy sequence w.r.t. d_1 , but it is not convergent in $\mathcal{C}([-1, 1])$ w.r.t. d_1 .

Define

$$g(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

Then,

$$d_1(f_n, g) = \frac{1}{2n} \quad (\text{area of the triangle})$$

[Note: we slightly abuse the notation here, as $g \notin \mathcal{C}([-1, 1])$, but we can confirm that $d_1(f_n, g)$ is well-defined.] So,

$$d_1(f_n, f_m) \leq d_1(f_n, g) + d_1(f_m, g) = \frac{1}{2n} + \frac{1}{2m}$$

Hence, f_n is Cauchy.

Now, for the sake of contradiction, if f_n converges to some $f \in \mathcal{C}([-1, 1])$, then $d_1(f_n, f) \rightarrow 0$ when $n \rightarrow \infty$. But we also know $d_1(f_n, g) \rightarrow 0$ when $n \rightarrow \infty$. So,

$$d_1(f, g) \leq d_1(f_n, f) + d_1(f_n, g) \rightarrow 0.$$

That is, $d_1(f, g) = 0$. i.e.,

$$\int_{-1}^1 |f(x) - g(x)| dx = 0.$$

So,

$$\int_{-1}^0 |f(x) - g(x)| dx = 0 \implies f(x) = g(x) \quad \text{on } [-1, 0].$$

[We are using: If h is continuous on $[a, b]$, $h \geq 0$, $\int_a^b h(x) dx = 0$, then $h = 0$.]

Similarly, fix $\varepsilon > 0$, we have

$$\int_{\varepsilon}^1 |f(x) - g(x)| dx = 0 \implies f(x) = g(x) \quad \text{on } [\varepsilon, 1].$$

Combining the two cases, we get $f = g$.

Since g is not continuous, f is also not continuous. *

So, f_n cannot be convergent in $\mathcal{C}([-1, 1])$, and the metric space is not complete. Q.E.D. ■

1.4 Topology

Definition 1.4.1 (Topology). A *topology* on set X is a set $\mathcal{T} \subset \mathcal{P}(X)$ with the following properties:

- $\emptyset, X \in \mathcal{T}$,
- Arbitrary union of elements of \mathcal{T} is in \mathcal{T} .
- A finite intersection of elements of \mathcal{T} is in \mathcal{T} .

Theorem 1.4.2 Topology Induced by Metric

If X is a metric space with metric d , a topology induced by metric d is

$$\mathcal{T} = \{A \subset X \mid A \text{ is open w.r.t. } d\}.$$

Proof 1. Recall: A is open in X means: (a) $\forall x \in A, \exists \varepsilon > 0$ s.t. $\forall y \in X, d(x, y) < \varepsilon \implies y \in A$, or (b) $\forall x \in A, \exists \varepsilon > 0$ s.t. $B_\varepsilon(x) \subset A$. Now, we can proceed with the trivial proof:

- \emptyset and X are open.
- Arbitrary union of open sets is open.
- A finite intersection of open sets is open.

Q.E.D. ■

Example 1.4.3 Different Metrics May Generate the Same Topology

- In \mathbb{R} :

$$d(x, y) = |x - y|$$

- In \mathbb{R}^n :

$$d_2(x, y) = \sqrt{\sum_{j=1}^n |x_j - y_j|^2}$$

$$d_1(x, y) = \sum_{j=1}^n |x_j - y_j|$$

$$d_\infty(x, y) = \max_j |x_j - y_j|$$

All d_1 , d_2 , and d_∞ generates the same topology.

Lemma 1.4.4 If d, \tilde{d} are metrics on X s.t. $\exists a, b > 0$ s.t. $\forall x, y \in X$

$$ad(x, y) \leq \tilde{d}(x, y) \leq bd(x, y),$$

then d and \tilde{d} generate the same topology.

Definition 1.4.5 (Topological Equivalence). If two metrics induce the same topology, they are said to be *topologically equivalent*.

Example 1.4.6

Consider a circle on $\mathbb{R}^2 : [0, 2\pi) \rightarrow \{x^2 + y^2 = 1\}$. Extending it to a unit circle in \mathbb{C} :

$$\{e^{it} \mid t \in [0, 2\pi)\}.$$

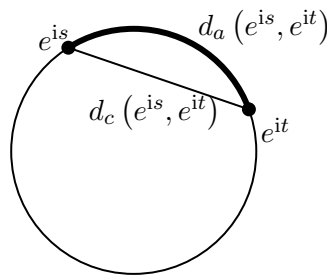
Then,

- chord metric:

$$d_c(e^{is}, e^{it}) = |e^{is} - e^{it}|$$

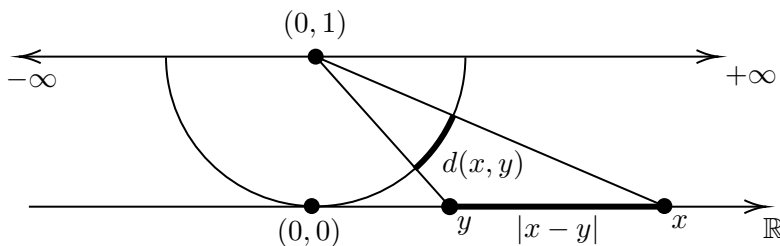
- arc length metric:

$$d_a(e^{is}, e^{it}) = \min \{|t - s|, 2\pi - |t - s|\}.$$



d_c and d_a generate the same topology on the circle.

Example 1.4.7 Extending the Real Numbers



$$d(x, y) = |\arctan x - \arctan y|.$$

d generates the same topology on \mathbb{R} as the absolute value metric does. With the metric d , we can work on the *extended real line*:

$$\widehat{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\},$$

with $\arctan(\pm\infty) = \pm\frac{\pi}{2}$. Then,

$$d(x, y) = |\arctan x - \arctan y|$$

is a metric on $\widehat{\mathbb{R}}$. Note that, also, $\arctan : \widehat{\mathbb{R}} \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Since $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ is compact, we know that $\widehat{\mathbb{R}}$ is a compact, complete metric space.

1.5 Upper and Lower Bounds

Consider a sequence $x_n \in \mathbb{R}$. A natural question to ask is: Is the sequence bounded? Denote

$$a_1 = \inf_n x_n, \quad a_2 = \inf_{n \geq 2} x_n, \quad b_2 = \sup_{n \geq 2} x_n, \quad b_1 = \sup_n x_n$$

Then,

$$a_1 \leq a_2 \leq b_2 \leq b_1.$$

Similarly, we can define

$$a_k = \inf_{n \geq k} x_n, \quad b_k = \sup_{n \geq k} x_n.$$

We have

$$a_k \leq a_{k+1} \leq b_{k+1} \leq b_k.$$

They form *nested compact intervals*. From previous analysis courses, we know that union of nested compact sets is non-empty. So, we can have

$$a = \lim_{k \rightarrow \infty} a_k \quad \text{and} \quad b = \lim_{k \rightarrow \infty} b_k.$$

[Meaning: $n \geq k \implies a_k \leq x_n \leq b_k$. So, for large enough n , $a \leq x \leq b$.]

Example 1.5.1

- $x_n = (-1)^n$. Then,

$$a_k = \inf_{n \geq k} x_n = -1 \quad \text{and} \quad b_k = \sup_{n \geq k} x_n = 1.$$

So, $a = -1$ and $b = 1$.

- $x_n = \frac{(-1)^n}{n}$. We have

$$a = 0 \quad \text{and} \quad b = 0.$$

Definition 1.5.2 (lim inf and lim sup). Suppose x_n is a sequence, then

$$a = \liminf_{n \rightarrow \infty} x_n = \lim_{k \rightarrow \infty} \inf_{n \geq k} x_n = \sup_k \inf_{n \geq k} x_n, \quad \text{and}$$

$$b = \limsup_{n \rightarrow \infty} x_n = \lim_{k \rightarrow \infty} \sup_{n \geq k} x_n = \inf_k \sup_{n \geq k} x_n.$$

These encode where the sequence is localized when n gets large.

Theorem 1.5.3

$$\liminf_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} x_n.$$

Proof 1. Simply take $\lim_{k \rightarrow \infty}$ of $a_k \leq b_k$.

Q.E.D. ■

Remark 2. (Meaning of lim sup). If $\limsup_{n \rightarrow \infty} x_n = b$, then $\forall \varepsilon > 0$, we know that

$$b + \varepsilon > \limsup_{n \rightarrow \infty} x_n = \lim_{k \rightarrow \infty} b_k$$

- By definition of supremum: $\exists k$ s.t. $b_k < b + \varepsilon$.
- By definition of limit: $\exists k$ s.t. $\forall n \geq k, \quad x_n < b + \varepsilon$.

Theorem 1.5.4

$$\liminf_{n \rightarrow \infty} x_n = \limsup_{n \rightarrow \infty} x_n = x \iff \lim_{n \rightarrow \infty} x_n = x.$$

Proof 3. (\Rightarrow): $\forall \varepsilon > 0, \forall n$ large enough, $x_n < b + \varepsilon$ ($b = \limsup x_n$) and $a - \varepsilon < x_n$ ($a = \liminf x_n$).

If $a = b = x$, then $\forall \varepsilon > 0, \forall n$ large enough, $x - \varepsilon < x_n < x + \varepsilon \implies |x_n - x| < \varepsilon$. Hence, $\lim_{n \rightarrow \infty} x_n = x$.

(\Leftarrow): Trivial.

Q.E.D. ■

Lemma 1.5.5 If $x_n \leq y_n \quad \forall n$, then

$$\limsup_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} y_n.$$

Proof 4. $\forall k$, we have that

$$\begin{aligned} \sup_{n \geq k} x_n \leq \sup_{n \geq k} y_n &\implies \limsup_{k \rightarrow \infty} \sup_{n \geq k} x_n \leq \limsup_{k \rightarrow \infty} \sup_{n \geq k} y_n \\ &\implies \limsup_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} y_n. \end{aligned}$$

Q.E.D. ■

Corollary 1.5.6 If $x_n \leq y_n \leq z_n$, then

$$\liminf x_n \leq \liminf y_n \leq \limsup y_n \leq \limsup z_n.$$

If $\lim x_n = \lim z_n = c$, then

$$\liminf y_n = \limsup y_n = c, \quad \text{and} \quad \lim_{n \rightarrow \infty} y_n = c.$$

Theorem 1.5.7 Sub-additivity of \limsup

$$\limsup_{n \rightarrow \infty} (x_n + y_n) \leq \limsup_{n \rightarrow \infty} x_n + \limsup_{n \rightarrow \infty} y_n.$$

Example 1.5.8

$x_n = (-1)^n$ and $y_n = (-1)^{n+1}$. Then,

$$x_n + y_n = 0, \quad \limsup x_n = 1, \quad \limsup y_n = 1,$$

but $\limsup(x_n + y_n) = 0$. So, $\limsup(x_n + y_n) < \limsup x_n + \limsup y_n$.

Notation 1.5.9 (Unbounded Sequences). If $A \subset \mathbb{R}$ is not bounded above, we write

$$\sup A = \infty.$$

If sequence x_n is not bounded above, we write

$$\limsup_{n \rightarrow \infty} x_n = \infty.$$

Remark 5. (What does $\liminf x_n = \infty$ mean?).

$$\implies a_k \rightarrow a = \infty$$

$$\implies \limsup x_n = \infty$$

$$\iff \lim_{n \rightarrow \infty} x_n = \infty$$

Example 1.5.10

What is the logical negation of the statement $\lim_{n \rightarrow \infty} x_n = \infty$?

Solution 6.

$$\lim x_n = \infty \iff \liminf x_n = \infty \quad \text{and} \quad \neg(\lim x_n = \infty) \iff \liminf x_n < \infty.$$

□

2 Borel Sets and Functions

2.1 σ -Algebra

Definition 2.1.1 (σ -Algebra). Let X be a set and $\mathcal{P}(X)$ be the set of all subsets of X . A σ -algebra on X is a set $\mathcal{A} \subset \mathcal{P}(X)$ with the following properties:

- $\emptyset \in \mathcal{A}$
- $A \in \mathcal{A} \implies A^c \in \mathcal{A}$, where $A^c = X \setminus A$.
- For any sequence $A_j \in \mathcal{A}$, $j = 1, 2, 3, \dots$,

$$\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}.$$

Observations:

- $X \in \mathcal{A}$. [following from ① and ②]
- $A_j \in \mathcal{A} \implies \bigcap_{j=1}^{\infty} A_j \in \mathcal{A}$.

Proof 1.

$$\begin{aligned} A_j \in \mathcal{A} &\implies A_j^c \in \mathcal{A} \\ &\implies \bigcup_{j=1}^{\infty} A_j^c \in \mathcal{A} \quad \text{and} \quad \bigcup_{j=1}^{\infty} A_j^c = \left(\bigcap_{j=1}^{\infty} A_j \right)^c \\ &\implies \left(\bigcap_{j=1}^{\infty} A_j \right)^c \in \mathcal{A} \\ &\implies \bigcap_{j=1}^{\infty} A_j \in \mathcal{A}. \end{aligned}$$

Q.E.D. ■

- For any A_1, \dots, A_n , then $\bigcup_{j=1}^n A_j \in \mathcal{A}$ and $\bigcap_{j=1}^n A_j \in \mathcal{A}$.

Proof 2. By taking $A_{n+1}, A_{n+2}, \dots = \emptyset$. Then,

$$\bigcup_{j=1}^n A_j = \bigcup_{j=1}^{\infty} A_j.$$

Q.E.D. ■

- $A, B \in \mathcal{A} \implies A \setminus B \in \mathcal{A}$.

Proof 3.

$$A \setminus B = A \cap B^c.$$

Q.E.D. ■

Example 2.1.2

- Maximal σ -algebra on X : $\mathcal{A} = \mathcal{P}(X)$.
- Minimal σ -algebra on X : $\mathcal{A} = \{\emptyset, X\}$.

Proposition 2.1.3 The intersection of a non-empty family (set) of σ -algebras on X is a σ -algebra on X .**Proof 4.** Let \mathcal{F} be the set of σ -algebras on X :

$$\mathcal{F} = \{\mathcal{A}_\alpha \mid \alpha \in I\},$$

where \mathcal{A}_α 's are σ -algebras and I is the index set. Define their intersection

$$\mathcal{A} = \bigcap_{\alpha \in I} \mathcal{A}_\alpha.$$

[WTS: \mathcal{A} is a σ -algebra.]

- $\emptyset \in \mathcal{A}_\alpha \quad \forall \alpha \in I \implies \emptyset \in \mathcal{A}$.
- $A \in \mathcal{A} \implies A \in \mathcal{A}_\alpha \quad \forall \alpha \in I \implies A^c \in \mathcal{A}_\alpha \quad \forall \alpha \in I \implies A^c \in \bigcap_{\alpha \in I} \mathcal{A}_\alpha = \mathcal{A}$
- $A_n \in \mathcal{A} \quad \forall n \in \mathbb{N}$. Then, $A_n \in \mathcal{A}_\alpha \quad \forall \alpha \in I, n \in \mathbb{N}$. So,

$$\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}_\alpha \quad \forall \alpha \in I \implies \bigcup_{n \in \mathbb{N}} A_n \in \bigcap_{\alpha \in I} \mathcal{A}_\alpha \in \mathcal{A}.$$

Hence, \mathcal{A} is a σ -algebra.

Q.E.D. ■

Definition 2.1.4 (σ -Algebra Generated by \mathcal{F}). Let $\mathcal{F} \subset \mathcal{P}(X)$. The σ -algebra generated by \mathcal{F} is the intersection of all σ -algebras on X containing \mathcal{F} (which have \mathcal{F} as a subset).

Remark.

- By Proposition 2.3, this is indeed a σ -algebra.
- We also think of this as the smallest σ -algebra that contains \mathcal{F} .

2.2 Borel Sets and Functions

Definition 2.2.1 (Borel σ -Algebra and Borel Sets). Let X be a metric space and \mathcal{T} be the topology induced by the metric. Then, the *Borel σ -algebra* on X is the σ -algebra generated by \mathcal{T} , denoted as \mathcal{B}_X . Elements of \mathcal{B}_X are called *Borel sets*.

Example 2.2.2

- Discrete metric on X

$$d(x, y) = \begin{cases} 1, & x \neq y \\ 0, & x = y \end{cases}$$

gives topology $\mathcal{T} = \mathcal{P}(X)$, which generates $\mathcal{B}_X = \mathcal{P}(X)$.

- $X = \mathbb{R}$. Then,

$$\mathcal{B}_{\mathbb{R}} \subsetneq \mathcal{P}(\mathbb{R}).$$

Proof uses axiom of choice (omitted).

Definition 2.2.3 (Measurable Function). Let $f : X \rightarrow Y$. Suppose \mathcal{A} is a σ -algebra on X and \mathcal{B} is a σ -algebra on Y . Denote

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}.$$

f is *measurable* if $B \in \mathcal{B} \implies f^{-1}(B) \in \mathcal{A}$.

[*Connection: definition of continuity.*]

Definition 2.2.4 (Borel Function). Let $f : X \rightarrow Y$, where X and Y are metric spaces. f is a *Borel function* if $B \in \mathcal{B}_Y \implies f^{-1}(B) \in \mathcal{B}_X$.

Proposition 2.2.5 If $f : X \rightarrow Y, g : Y \rightarrow Z$ are Borel functions, then

$$g \circ f : X \rightarrow Z$$

is also a Borel function.

Proof 1. $C \in \mathcal{B}_Z$. Then, $g^{-1}(C) \in \mathcal{B}_Y \implies f^{-1}(g^{-1}(C)) \in \mathcal{B}_X$. That is,

$$(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C)) \in \mathcal{B}_X.$$

Q.E.D. ■

Definition 2.2.6 (Pushforward). Let $f : X \rightarrow Y$ and \mathcal{A} be a σ -algebra on X . Define

$$\mathcal{C} = \{C \subset Y \mid f^{-1}(C) \in \mathcal{A}\}.$$

Then, \mathcal{C} is a σ -algebra and called the *pushforward* of \mathcal{A} by f .

Proof2. Let's show that \mathcal{C} is a σ -algebra.

- $f^{-1}(\emptyset) = \emptyset \in \mathcal{A} \implies \emptyset \in \mathcal{C}$.
- Assume $C \in \mathcal{C}$, $f^{-1}(C) \in \mathcal{A}$.

$$f^{-1}(C^c) = \{x \mid f(x) \in C^c\} = \{x \mid f(x) \notin C\} = \{x \mid f(x) \in C\}^c = (f^{-1}(C))^c \in \mathcal{A}.$$

So, $C^c \in \mathcal{C}$.

- Assume $C_n \in \mathcal{C}$, $f^{-1}(C_n) \in \mathcal{A}$. Then,

$$f^{-1}\left(\bigcup_{n=1}^{\infty} C_n\right) = \bigcup_{n=1}^{\infty} f^{-1}(C_n) \in \mathcal{A}.$$

$$\text{So, } \bigcup_{n=1}^{\infty} C_n \in \mathcal{C}.$$

Q.E.D. ■

Proposition 2.2.7 Any continuous function $f : X \rightarrow Y$ is a Borel function.

Proof3. [WTS: $C \in \mathcal{B}_Y \implies f^{-1}(C) \in \mathcal{B}_X$]. By continuity, we know C is open $\implies f^{-1}(C)$ is open. Then, $f^{-1}(C) \in \mathcal{B}_X$. [WTS: C is Borel.]

Note that

$$\{C \mid f^{-1}(C) \in \mathcal{B}_X\} \text{ is a } \sigma\text{-algebra}$$

(because it is the pushforward of \mathcal{B}_X by f) contains all open sets. Hence, it contains all Borel sets.

Q.E.D. ■

Remark. \mathcal{A} contains all open sets and \mathcal{A} is a σ -algebra $\implies \mathcal{A}$ contains all Borel set (i.e., $\mathcal{B}_X \subset \mathcal{A}$).

2.3 Borel σ -Algebra on Euclidean Spaces

Goal: $\mathcal{B}_{\mathbb{R}}$ is generated by $\{(a, \infty) \mid a \in \mathbb{R}\}$.

Corollary 2.3.1 If $\mathcal{A} \subset \mathcal{P}(\mathbb{R})$, \mathcal{A} contains all half-lines (a, ∞) and \mathcal{A} is a σ -algebra, then \mathcal{A} contains all Borel sets of \mathbb{R} .

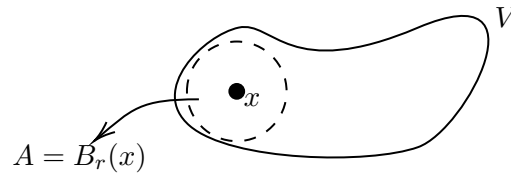
Definition 2.3.2 (Base). A base U of a topological space (X, \mathcal{T}) is a family $U \subset \mathcal{T}$ with the property: $\forall V$ open and $\forall x \in V, \exists A \in U$ s.t. $x \in A \subset V$.

Example 2.3.3

In a metric space,

$$U = \{B_r(x) \mid x \in X, r \in (0, \infty)\}$$

is a base.

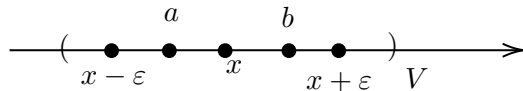


Definition 2.3.4 (Second-Countable). X is *second-countable* if it has a countable base.

Lemma 2.3.5 In $\mathbb{R}, U = \{(a, b) \mid a, b \in \mathbb{Q}, a < b\}$ is a countable base. Thus, \mathbb{R} is second-countable.

Proof 1.

- Countability: $a, b \in \mathbb{Q}$, countable. Pair of countable is also countable.
- Base: Let $V \subset \mathbb{R}$ be open. Suppose $x \in V$, then $\exists \varepsilon > 0$ s.t. $(x - \varepsilon, x + \varepsilon) \subset V$. Since \mathbb{Q} is dense in $\mathbb{R}, \exists a \in (x - \varepsilon, x) \cap \mathbb{Q}$ and $\exists b \in (x, x + \varepsilon) \cap \mathbb{Q}$. Then, $x \in (a, b) \subset (x - \varepsilon, x + \varepsilon) \subset V$.



Q.E.D. ■

Lemma 2.3.6 Let U be a base of X . Every open set V can be written as a union of some elements of the base.

Proof 2. (\subseteq): $\bigcup_{\substack{A \in U \\ A \subset V}} A \subset V$, trivially.

(\supseteq): For any $x \in V, \exists A \in U$ s.t. $x \in A \subset V$. Then, for any $x \in V, x \in \bigcup_{\substack{A \in U \\ A \subset V}} A$. Then, $V \subset \bigcup_{\substack{A \in U \\ A \subset V}} A$.

Hence, $\bigcup_{\substack{A \in U \\ A \subset V}} A = V$.

Q.E.D. ■

Theorem 2.3.7

$\mathcal{B}_{\mathbb{R}}$ is the σ -algebra generated by $(a, \infty), a \in \mathbb{R}$.

Proof3. Denote $\mathcal{A} = \sigma$ -algebra generated by (a, ∞) , $a \in \mathbb{R}$. [WTS: $\mathcal{A} = \mathcal{B}_{\mathbb{R}}$.]

(\subseteq): Note that (a, ∞) is open. So, $(a, \infty) \in \mathcal{B}_{\mathbb{R}} \implies \mathcal{A} \subset \mathcal{B}_{\mathbb{R}}$.

(\supseteq): [WTS: \mathcal{A} contains all open sets.]

$$\begin{aligned} (a, \infty) \in \mathcal{A} \implies & \textcircled{1} \quad (-\infty, a] = (a, \infty)^c \in \mathcal{A}. \\ & \textcircled{2} \quad (-\infty, b) = \bigcup_{n=1}^{\infty} \left(-\infty, b - \frac{1}{n} \right] \in \mathcal{A}. \\ & \textcircled{3} \quad (a, b) = (a, \infty) \cap (-\infty, b) \in \mathcal{A}. \end{aligned}$$

So, if V is open,

$$V = \bigcup_{\substack{a, b \in \mathbb{Q} \\ (a, b) \in V}} (a, b) \in \mathcal{A}.$$

That is, \mathcal{A} contains all open sets. Since $\mathcal{B}_{\mathbb{R}}$ is the smallest σ -algebra containing all the open sets, $\mathcal{B}_{\mathbb{R}} \subset \mathcal{A}$.

So, $\mathcal{B}_{\mathbb{R}} = \mathcal{A}$.

Q.E.D. ■

Corollary 2.3.8 $f : X \rightarrow \mathbb{R}$ is Borel $\iff f^{-1}((a, \infty))$ is Borel $\forall a \in \mathbb{R}$.

Proof4. (\implies): Suppose f is Borel and (a, ∞) is a Borel set, then

$$f^{-1}((a, \infty)) \in \mathcal{B}_X.$$

(\impliedby): [WTS: $\forall B \in \mathcal{B}_X$, $f^{-1}(B) \in \mathcal{B}_X$] Consider the pushforward of \mathcal{B}_X by f :

$$\mathcal{C} = \{C \subset \mathbb{R} \mid f^{-1}(C) \in \mathcal{B}_X\}$$

is a σ -algebra on \mathbb{R} . If $f^{-1}((a, \infty)) \in \mathcal{B}_X$, then $(a, \infty) \in \mathcal{C}$. So, $\mathcal{B}_{\mathbb{R}} \subset \mathcal{C}$.

Therefore, $\forall B \in \mathcal{B}_{\mathbb{R}}$, $f^{-1}(B) \in \mathcal{B}_X \implies f$ is a Borel function.

Q.E.D. ■

Example 2.3.9

If $f, g : X \rightarrow \mathbb{R}$ are Borel functions, then

$$h(x) = \max \{f(x), g(x)\}$$

is a Borel function.

Proof5. [WTS: If $B \in \mathcal{B}_X$, then $h^{-1}(B) \in \mathcal{B}_X$. $\implies h^{-1}((a, \infty)) \in \mathcal{B}_X$.]

$$\begin{aligned} h^{-1}((a, \infty)) &= \{x \mid h(x) \in (a, \infty)\} = \{x \mid h(x) > a\} \\ &= \{x \mid f(x) > a \text{ or } g(x) > a\} \\ &= \{x \mid f(x) > a\} \cup \{x \mid g(x) > a\} \\ &= f^{-1}((a, \infty)) \cup g^{-1}((a, \infty)) \in \mathcal{B}_X. \end{aligned}$$

Q.E.D. ■

We proved \mathbb{R} is second-countable with countable bases

$$\{(a, b) \mid a < b, a, b \in \mathbb{Q}\}.$$

Lemma 2.3.10 The space of \mathbb{R}^n has a countable base:

$$U = \left\{ \prod_{j=1}^n (a_j, b_j) \mid a_j, b_j \in \mathbb{Q}, a_j < b_j \right\}.$$

Corollary 2.3.11 $\mathcal{B}_{\mathbb{R}^n}$ is generated by rectangles $\prod_{j=1}^n (a_j, b_j)$, $a_j < b_j$.

Lemma 2.3.12 Let $h : X \rightarrow \mathbb{R}^n$ be defined as

$$h = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}. \quad \text{where } h_j : X \rightarrow \mathbb{R}.$$

Denote $\pi_j : \mathbb{R}^n \rightarrow \mathbb{R}$, the projection to the j -th entry. i.e.,

$$\pi_j \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_j, \quad \text{and} \quad h_j = \pi_j \circ h.$$

$h : X \rightarrow \mathbb{R}^n$ is Borel $\iff h_j = \pi_j \circ h : X \rightarrow \mathbb{R}$ are Borel $\forall j$.

Proof 6. (\implies): Suppose h is Borel. $\pi_j : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous, so it is Borel $\implies h_j = \pi_j \circ h$ is also Borel, as composition of Borel functions is Borel.

(\impliedby): Assume h_1, \dots, h_n are Borel functions.

$$\begin{aligned} h^{-1} \left(\prod_{j=1}^n (a_j, b_j) \right) &= \left\{ x \mid h(x) \in \prod_{j=1}^n (a_j, b_j) \right\} \\ &= \{x \mid h_j(x) \in (a_j, b_j) \quad \forall j\} \\ &= \bigcap_{j=1}^n \{x \mid h_j(x) \in (a_j, b_j)\} \\ &= \bigcap_{j=1}^n f^{-1}((a_j, b_j)) \in \mathcal{B}_X. \end{aligned}$$

So, h is Borel.

Q.E.D. ■

Theorem 2.3.13

If $f, g : X \rightarrow \mathbb{R}$ are Borel, then $f + g$ and fg are Borel functions.

Proof 7. Let $h = \begin{pmatrix} f \\ g \end{pmatrix}$, so $h : X \rightarrow \mathbb{R}^2$.

- $F : \mathbb{R}^2 \rightarrow \mathbb{R}$ s.t. $F(x, y) = x + y$ is continuous, so it is Borel. Hence, $F \circ h = f + g$ is Borel, as a composition of two Borel functions is Borel.
- $G : \mathbb{R}^2 \rightarrow \mathbb{R}$ s.t. $G(x, y) = xy$. Then, similarly, $G \circ h = fg$ is Borel.

Q.E.D. ■

Lemma 2.3.14 $f : X \rightarrow \mathbb{C}$ is Borel \iff $\operatorname{Re} f$ and $\operatorname{Im} f$ are Borel.

Proof 8. $\mathbb{C} = \mathbb{R}^2$ with $d(z, w) = |z - w|$.

Q.E.D. ■

Lemma 2.3.15 If $f, g : X \rightarrow \mathbb{C}$ are Borel, then $f + g$ and fg are Borel.

2.4 Extended Real Line and Limits

Goal: If $f_n : X \rightarrow \widehat{\mathbb{R}}$ are Borel, then $\sup_{n \in \mathbb{N}} f_n$, $\inf_{n \in \mathbb{N}} f_n$, $\limsup_{n \rightarrow \infty} f_n = \inf_k \sup_{n \geq k} f_n$, $\liminf_{n \rightarrow \infty} f_n = \sup_k \inf_{n \geq k} f_n$ are Borel.

Recall: $\widehat{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

Proof 1. [We will just prove $\sup_{n \in \mathbb{N}} f_n$ is Borel, and the other follows.]

[WTS: $f^{-1}(B)$ is Borel if B is Borel.]

Lemma 2.4.1 $U = \{(a, b) \mid a < b, a, b \in \mathbb{Q}\} \cup \{(a, \infty) \mid a \in \mathbb{Q}\} \cup \{[-\infty, b) \mid b \in \mathbb{Q}\}$ is a continuous base for $\widehat{\mathbb{R}}$.

Claim 2.4.2 If $V \subset \widehat{\mathbb{R}}$ is open, $\infty \in V$. Then, $\exists x_0 \in \mathbb{R}$ s.t. $(x_0, \infty) \subset V$.

Proof. By assumption, $\exists \varepsilon > 0$ s.t. $|\arctan x - \arctan \infty| < \varepsilon$, $x \in V$. That is,

$$\begin{aligned} \left| \arctan x - \frac{\pi}{2} \right| < \varepsilon &\implies x \in V \\ \arctan x > \frac{\pi}{2} - \varepsilon &\implies x \in V \\ x > \tan\left(\frac{\pi}{2} - \varepsilon\right) &\implies x \in V. \end{aligned}$$

Let $x_0 = \tan\left(\frac{\pi}{2} - \varepsilon\right)$. Pick $a > x_0$, $a \in \mathbb{Q}$, $(a, \infty) \in V$. □

Corollary 2.4.3 $\mathcal{B}_{\widehat{\mathbb{R}}} = \sigma$ -algebra generated by $(a, \infty]$, $a \in \mathbb{R}$.

It is hard to work on a general Borel set. Consider half lines

$$\begin{aligned} f^{-1}((a, \infty]) &= \{x \mid f(x) > a\} = \left\{x \mid \sup_n f_n(x) > a\right\} \\ &= \{x \mid \exists n, f_n(x) > a\} \\ &= \bigcup_n \{x \mid f_n(x) > a\} \\ &= \bigcup_n f_n^{-1}((a, \infty]). \end{aligned}$$

[What's still missing in this proof: $(a, \infty]$ is Borel to $\widehat{\mathbb{R}}$.] Denote $\mathcal{A} = \sigma$ -algebra generated by $(a, -\infty]$, $a \in \mathbb{R}$. Then,

$$\{-\infty\} = (-\infty, \infty)^c = (\mathbb{R} \cup \{\infty\})^c = \left(\bigcup_{j=1}^{\infty} (-j, \infty] \right)^c.$$

Q.E.D. ■

Corollary 2.4.4 $f_n : X \rightarrow \widehat{\mathbb{R}}$ are Borel. Then, $\{x \mid f_n(x) \text{ bounded below}\}$ is a Borel set.

Proof 2.

$$\begin{aligned} \{x \mid f_n(x) \text{ bounded below}\} &= \left\{ x \mid \sup_n f_n(x) = \infty \right\}^c \\ &= f^{-1}(\{\infty\})^c \text{ is Borel.} \end{aligned}$$

Q.E.D. ■

Lemma 2.4.5 For $A \subset \widehat{\mathbb{R}}$, $A \in \mathcal{B}_{\widehat{\mathbb{R}}} \iff A \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}$.

Proof 3. Simply ignore $\{-\infty, +\infty\}$, $A \cap \mathbb{R} \subset \mathbb{R}$.

Q.E.D. ■

Lemma 2.4.6 $f : X \rightarrow \widehat{\mathbb{R}}$ is Borel $\iff f^{-1}(\{\infty\}) \in \mathcal{B}_X$, $f^{-1}(\{-\infty\}) \in \mathcal{B}_X$, and $f|_{f^{-1}(\mathbb{R})} : f^{-1}(\mathbb{R}) \rightarrow \mathbb{R}$ is Borel.

3 Measure and Abstract Integration

3.1 Measures

Definition 3.1.1 (Measure). Let \mathcal{A} be a σ -algebra on X . A *measure* on \mathcal{A} is a map:

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

such that

- $\mu(\emptyset) = 0$, and
- for any pairwise disjoint $A_n \in \mathcal{A}$,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

This property is called σ -*additivity*. [*pairwise disjoint means that $A_k \cap A_\ell = \emptyset$ if $k \neq \ell$.*]

Definition 3.1.2 (Borel Measure). A *Borel measure* μ on X is a measure on \mathcal{B}_X .

Definition 3.1.3 (Finite Measure and Probability Measure).

- μ is *finite* if $\mu(X) < \infty$, and
- μ is a *probability measure* if $\mu(X) = 1$.

Example 3.1.4 Simple Measures

- Trivial measure: $\mu \equiv 0$.
- Fix $x_0 \in X$. *Dirac measure* at x_0 is

$$\delta_{x_0}(A) = \begin{cases} 1 & \text{if } x_0 \in A \\ 0 & \text{if } x_0 \notin A. \end{cases}$$

- Counting measure:

$$\mu(A) = \begin{cases} |A| = \#A, & \text{if } A \text{ is finite (cardinality of } A) \\ \infty & \text{if } A \text{ is infinite.} \end{cases}$$

Properties of Measures

For any measure μ on \mathcal{A} and sets in \mathcal{A} :

- For any finite n , if A_1, \dots, A_n pairwise disjoint, then

$$\mu\left(\bigcup_{j=1}^n A_j\right) = \sum_{j=1}^n \mu(A_j).$$

Proof 1. Take $A_{n+1}, A_{n+2}, \dots, = \emptyset$. Then, by σ -additivity, it is trivial. Q.E.D. ■

- If $S \subset T$, then $\mu(S) \leq \mu(T)$.

Proof 2. $T = S \cup (T \setminus S) \implies \mu(T) = \mu(S) + \mu(T \setminus S) \geq \mu(S)$. Q.E.D. ■

- If $S \subset T$ and $\mu(S)$ is finite, then

$$\mu(T \setminus S) = \mu(T) - \mu(S).$$

Proof 3. Subtract $\mu(S)$ from $\mu(T) = \mu(S) + \mu(T \setminus S)$.

[When subtracting, we need to worry about if $\mu(S) = \infty$.] Q.E.D. ■

- If $B_n \subset B_{n+1} \quad \forall n \in \mathbb{N}$ (nested). Then,

$$\mu\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n \rightarrow \infty} \mu(B_n).$$

Proof 4. Denote $A_1 = B_1, A_n = B_n \setminus B_{n-1}$ for $n \geq 2$. Then, $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$, and A_n 's are pairwise disjoint. Also,

$$B_n = \bigcup_{j=1}^n A_j.$$

By σ -additivity,

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} B_n\right) &= \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \\ &= \lim_{n \rightarrow \infty} \sum_{n=1}^{\infty} \mu(A_n) \\ &= \lim_{n \rightarrow \infty} \sum_{n=1}^{\infty} \mu(B_n). \end{aligned}$$

Q.E.D. ■

- If $C_n \supset C_{n+1} \quad \forall n \in \mathbb{N}$ (decreasing sequence), and $\exists k \in \mathbb{N}$ such that $\mu(C_k) < \infty$, then

$$\mu\left(\bigcap_{n=1}^{\infty} C_n\right) = \lim_{n \rightarrow \infty} \mu(C_n).$$

Proof 5. [$\mu(C_k) < \infty$ is required for doing subtraction.] WLOG, take $k = 1$. So, $\mu(C_1) < \infty$, and $C_1 \setminus C_n \subset C_1 \setminus C_{n+1}$. Then,

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} (C_1 \setminus C_n)\right) &= \lim_{n \rightarrow \infty} \mu(C_1 \setminus C_n) \\ \bigcup_{n=1}^{\infty} (C_1 \setminus C_n) &= C_1 \setminus \bigcap_{n=1}^{\infty} C_n \\ \mu\left(C_1 \setminus \bigcap_{n=1}^{\infty} C_n\right) &= \lim_{n \rightarrow \infty} \mu(C_1 \setminus C_n) \\ \mu(C_1) - \mu(C_n) &= \lim_{n \rightarrow \infty} (\mu(C_1) - \mu(C_n)) \\ \mu(C_n) &= \lim_{n \rightarrow \infty} \mu(C_n) \quad \text{[Subtract } C_1] \end{aligned}$$

Q.E.D. ■

Example 3.1.5 When this property fails

Let μ be some continuous measure on \mathbb{N} . Denote

$$C_n = \{n, n+1, n+2, \dots\} = \{j \in \mathbb{N} \mid j \geq n\}.$$

Then, $\mu(C_n) = \infty \quad \forall n$. However,

$$\bigcap_{n=1}^{\infty} C_n = \{j \in \mathbb{N} \mid j \geq N \quad \forall n\} = \emptyset.$$

- For any sequence A_n , σ -subadditivity:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Proof 6. Use $B_n = \bigcup_{j=1}^n A_j$ (increasing sequence). Then,

$$\begin{aligned} \bigcup_{j=1}^{\infty} A_j &= \bigcup_{j=1}^{\infty} B_j \\ \mu\left(\bigcup_{j=1}^{\infty} A_j\right) &= \mu\left(\bigcup_{j=1}^{\infty} B_j\right) = \lim_{n \rightarrow \infty} \mu(B_n). \end{aligned}$$

Denote $C_n = B_n \setminus B_{n-1}$. Then,

$$\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty} C_j$$

[Why? $x \in \bigcup_{j=1}^{\infty} A_j \iff \exists j \text{ s.t. } x \in A_j \iff \exists n \text{ smallest s.t. } x \in A_n \iff \exists n \text{ s.t. } x \in B_n \setminus B_{n-1}$.]

Therefore, since C_j are pairwise disjoint, by σ -additivity,

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \mu\left(\bigcup_{j=1}^{\infty} C_j\right) = \sum_{j=1}^{\infty} \mu(C_j).$$

Note that $C_n \subset A_n$. Then,

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(C_j) \leq \sum_{j=1}^{\infty} \mu(A_j).$$

Q.E.D. ■

3.2 lim sup of Sets

Definition 3.2.1 (lim sup of Sets). Let A_n be sets. Then,

$$\limsup_{n \rightarrow \infty} A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n.$$

In other words, $x_n \in \limsup_{n \rightarrow \infty} A_n \iff \exists$ infinitely many n s.t. $x \in A_n$.

Proof 1.

$$\begin{aligned} x_n \in \limsup_{n \rightarrow \infty} A_n &\iff \forall k, \exists n \geq k \text{ s.t. } x \in A_n \\ &\iff \{n \in \mathbb{N} \mid x \in A_n\} \text{ is unbounded} \\ &\iff \{n \in \mathbb{N} \mid x \in A_n\} \text{ is infinite} \\ &\iff \exists \text{ infinitely many } n \text{ s.t. } x \in A_n. \end{aligned}$$

Q.E.D. ■

Lemma 3.2.2 (Borel-Cantelli, I). If $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, then $\mu\left(\limsup_{n \rightarrow \infty} A_n\right) = 0$.

Proof 2. Fix k . Then,

$$\mu\left(\limsup_{n \rightarrow \infty} A_n\right) \leq \mu\left(\bigcup_{n=k}^{\infty} A_n\right) \leq \underbrace{\sum_{n=k}^{\infty} \mu(A_n)}_{\rightarrow 0 \text{ as } k \rightarrow \infty}$$

Let $k \rightarrow \infty$, we have $\mu\left(\limsup_{n \rightarrow \infty} A_n\right) = 0$.

Q.E.D. ■

Definition 3.2.3 (Almost Everywhere). A property P holds *almost everywhere* (a.e./ μ -a.e.) if

$$\{x \mid P(x) \text{ is false}\} \text{ has measure zero.}$$

Lemma 3.2.4 (Borel-Cantelli, II). If $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, then for μ -a.e. x , \exists finitely many $n \in \mathbb{N}$ s.t. $x \in A_n$.

3.3 Integration of Positive Functions

Set-Up: Let $f : X \rightarrow [0, \infty]$, \mathcal{A} is a σ -algebra on X , μ is a measure on \mathcal{A} .

Meanwhile, for $B \in \mathcal{B}_{\mathbb{R}}$, $f^{-1}(B) \in \mathcal{A} \implies \forall c, f^{-1}((c, \infty]) \in \mathcal{A}$. So, f is also measurable.

Definition 3.3.1 (Simple Function). A function $s : X \rightarrow [0, \infty]$ is *simple* if it takes finitely many values. If s takes values c_1, \dots, c_n , and define

$$A_j = s^{-1}(\{c_j\}),$$

then

$$\int s \, d\mu = \sum_{j=1}^n c_j \mu(A_j).$$

Example 3.3.2 Simple Functions

$$s(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

is a simple function. It is not Riemann integrable, but integrable under this framework.

Notation 3.3.3 If $c > 0$, then $c \cdot \infty = \infty$ and $0 \cdot \infty = 0$.

Proposition 3.3.4 (Properties of Integration of Simple Functions).

- If $s, t > 0$ are simple functions, then $\int (s + t) \, d\mu = \int s \, d\mu + \int t \, d\mu$.
- If $s \leq t$ (pointwisely, i.e., $s(x) \leq t(x) \quad \forall x$), then $\int s \, d\mu \leq \int t \, d\mu$.

Proof 1.

- $s + t$ is a simple, positive function. Suppose s takes values c_1, \dots, c_n on $A_j = s^{-1}(\{c_j\})$ and t takes

values d_1, \dots, d_m on $B_k = t^{-1}(\{d_k\})$. Then, $(s + t)$ takes values $c_j + d_k$ on $A_j \cap B_k$.

$$\begin{aligned}
\int (s + t) d\mu &= \sum_{j=1}^n \sum_{k=1}^m (c_j + d_k) \mu(A_j \cap B_k) \\
&= \sum_{j=1}^n \sum_{k=1}^m c_j \mu(A_j \cap B_k) + \sum_{j=1}^n \sum_{k=1}^m d_k \mu(A_j \cap B_k) \\
&= \sum_{j=1}^n c_j \sum_{k=1}^m \mu(A_j \cap B_k) + \sum_{k=1}^m d_k \sum_{j=1}^n \mu(A_j \cap B_k) \\
&= \sum_{j=1}^n c_j \mu\left(\bigcup_{k=1}^m (A_j \cap B_k)\right) + \sum_{k=1}^m d_k \mu\left(\bigcup_{j=1}^n (A_j \cap B_k)\right) \\
&= \sum_{j=1}^n c_j \mu(A_j) + \sum_{k=1}^m d_k \mu(B_k) \\
&= \int s d\mu + \int t d\mu.
\end{aligned}$$

- $t - s \geq 0$ and $t - s$ is a simple function. From ①:

$$\int t d\mu = \int s d\mu + \int (t - s) d\mu \geq \int s d\mu.$$

Q.E.D. ■

Definition 3.3.5 (Characteristic Functions). *Characteristic functions* of a set A is defined as

$$\chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A. \end{cases}$$

Take sets A_1, \dots, A_n , and $c_1, \dots, c_n \in [0, \infty)$. Then,

$$s = \sum_{j=1}^n c_j \chi_{A_j} \text{ is simple,}$$

and

$$\int s d\mu = \sum_{j=1}^n \int c_j \chi_{A_j} d\mu = \sum_{j=1}^n (c_j \mu(A_j) + 0 \cdot \mu(A_j^c)) = \sum_{j=1}^n c_j \mu(A_j).$$

Note that

$$c_j \chi_{A_j} \begin{cases} = c_j, & x \in A_j \\ 0, & x \in A_j^c. \end{cases}$$

Definition 3.3.6 (Integration of Positive Functions). Suppose $f : X \rightarrow [0, \infty]$ is measurable.

Define

$$\int f \, d\mu = \sup_{\substack{s \text{ simple} \\ 0 \leq s \leq f}} \int s \, d\mu \in [0, \infty].$$

This definition is analogy to “lower sum” definition of Riemann integration.

Remark. If t is simple, then

$$\int t \, d\mu = \sup_{\substack{s \text{ simple} \\ 0 \leq s \leq t}} \int s \, d\mu$$

because

$$s \leq t \implies \int s \, d\mu \leq \int t \, d\mu$$

$$s = t \implies \int s \, d\mu = \int t \, d\mu.$$

Theorem 3.3.7 Properties of Integration

- If $0 \leq f \leq g$, then $\int f \, d\mu \leq \int g \, d\mu$.
- If $f \geq 0$ and $c \in [0, \infty)$, then $\int cf \, d\mu = c \int f \, d\mu$.

Proof 2.

- If s is simple, and $0 \leq s \leq f$, then $0 \leq s \leq g$. Hence,

$$\int s \, d\mu \leq \int g \, d\mu.$$

Take sup over s :

$$\sup_s \int s \, d\mu = \int f \, d\mu \leq \int g \, d\mu.$$

- If $c = 0$, $0 = 0$, and the proof is trivial. Now, assume $c > 0$. Let s be simple such that $0 \leq s \leq f$. Then, $0 \leq cs \leq cf$ is also simple.

$$\begin{aligned} \int cs \, d\mu &= c \int s \, d\mu \\ \sup_s \int cs \, d\mu &= \sup_s c \int s \, d\mu = c \sup_s \int s \, d\mu \\ \int cf \, d\mu &= c \int f \, d\mu. \end{aligned}$$

Q.E.D. ■

Corollary 3.3.8 If $0 \leq f \leq g$ and $\int g \, d\mu < \infty \implies \int f \, d\mu < \infty$.

Notation 3.3.9 Integral over a subset $E \subset X$ defined and denoted by

$$\int E f \, d\mu = \int f \chi_E \, d\mu,$$

where χ_E is the characteristic function of E .

Theorem 3.3.10 Monotone Convergence Theorem/MCT

For any sequence of $f_n : X \rightarrow [0, \infty]$ s.t. $f_n \leq f_{n+1} \quad \forall n \in \mathbb{N}$ (monotone) (i.e., $f_n(x) \leq f_{n+1}(x) \quad \forall x, \forall n \in \mathbb{N}$, pointwise), we have that

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu = \int \lim_{n \rightarrow \infty} f_n \, d\mu.$$

Proof 3.

Lemma 3.3.11 (Special Case). If $s : X \rightarrow [0, \infty)$ simple, and $E_n \subset E_{n+1} \quad \forall n$ s.t.

$$\bigcup_{n \in \mathbb{N}} E_n = X,$$

then

$$\lim_{n \rightarrow \infty} \int_{E_n} s \, d\mu = \int s \, d\mu.$$

This is a special case because $s \chi_{E_n} \leq s \chi_{E_{n+1}}$, and $\lim_{n \rightarrow \infty} s \chi_{E_n} = s$.

Proof.

$$\begin{aligned} s = \sum_{j=1}^m c_j \chi_{A_j} &\implies s \chi_{E_n} = \sum_{j=1}^m c_j \chi_{A_j} \chi_{E_n} \\ &= \sum_{j=1}^m c_j \chi_{A_j \cap E_n}. \end{aligned}$$

Then,

$$\int s \chi_{E_n} \, d\mu = \sum_{j=1}^m c_j \mu(A_j \cap E_n).$$

As $n \rightarrow \infty$,

$$\sum_{j=1}^m c_j \mu(A_j \cap E_n) \rightarrow \sum_{j=1}^m c_j \mu(A_j) = \int s \, d\mu.$$

[Since E_n is increasing, $E_n \rightarrow X$. Then, $A_j \cap E_n \rightarrow A_j \cap X = A_j$. More formally, $A_j \cap E_n \subset A_j \cap E_{n+1}$:

$$\bigcup_{n=1}^{\infty} (A_j \cap E_n) = A_j \cap \bigcup_{n=1}^{\infty} E_n = A_j \cap X = A_j.$$

] \square

Denote $f = \lim_{n \rightarrow \infty} f_n$ (pointwise). [WTS: $\lim_{n \rightarrow \infty} \int f_n d\mu \leq \int f d\mu$ and $\lim_{n \rightarrow \infty} \int f_n d\mu \geq \int f d\mu$.]
 (\leq): Since $f_n(x) \leq f_{n+1}(x) \quad \forall x$, we have $f_n(x) \leq f(x)$. Then,

$$\int f_n d\mu \leq \int f d\mu \quad \forall n.$$

So,

$$\lim_{n \rightarrow \infty} \int f_n d\mu \leq \int f d\mu.$$

(\geq): Fix $c \in (0, 1)$. Fix simple function $0 \leq s \leq f$. Define

$$E_n = \{x \mid f_n(x) \geq cs(x)\}.$$

Then, $E_n \subset E_{n+1}$ because $f_n(x) \leq f_{n+1}(x)$.

Claim $\bigcup_{n=1}^{\infty} E_n = X$.

Proof. **Case I** If $f(x) = 0$, then $s(x) = 0$. So, $0 \leq f_n(x)$, $x \in E_n \quad \forall n$.

Case II If $0 < f(x) < \infty$, then

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) > cf(x) \geq cs(x).$$

Then, \forall large enough n , $f_n(x) > cs(x)$. So, $x \in E_n$.

Case III If $f(x) = \infty$, then

$$\lim_{n \rightarrow \infty} f_n(x) = \infty > cs(x).$$

Since $cs(x) < \infty$ (simple function), $f_n(x) \geq cs(x)$ for large enough n . So, $x \in E_n$. \square

Then, we have $f_n \geq f_n \chi_E \geq cs \chi_E$. So,

$$\int f_n d\mu \geq \int cs \chi_E d\mu.$$

Take $n \rightarrow \infty$,

$$\lim_{n \rightarrow \infty} \int f_n d\mu \geq \lim_{n \rightarrow \infty} \int cs \chi_E d\mu = \int cs d\mu = c \int s d\mu \quad [\text{by Lemma 3.11}]$$

Take $c \uparrow 1$ (c goes up to 1):

$$\lim_{n \rightarrow \infty} \int f_n d\mu \geq \int s d\mu \quad [\text{LHS} \perp c]$$

Take $\sup_{\substack{s \text{ simple} \\ 0 \leq s \leq f}}$:

$$\lim_{n \rightarrow \infty} \int f_n d\mu \geq \sup_{\substack{s \text{ simple} \\ 0 \leq s \leq f}} \int s d\mu = \int f d\mu.$$

Since $\lim_{n \rightarrow \infty} \int f_n d\mu \leq \int f d\mu$ and $\lim_{n \rightarrow \infty} \int f_n d\mu \geq \int f d\mu$, it must be that

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu.$$

Q.E.D. ■

Example 3.3.12

Let ν be the counting measure on \mathbb{N} . Let $f : \mathbb{N} \rightarrow [0, \infty]$. What is $\int f d\nu$?

Let $E_n = \{1, 2, \dots, n\}$. Then, $f\chi_{E_n}$ is simple, and

$$\int f\chi_{E_n} d\nu = \sum_{j=1}^n f(j)\nu(\{j\}) = \sum_{j=1}^n f(j).$$

Note that $E_n \subset E_{n+1}$ and $\bigcup_{n=1}^{\infty} E_n = \mathbb{N}$. So, $f\chi_{E_n} \leq f\chi_{E_{n+1}}$, and $f\chi_{E_n} \rightarrow f$. Then,

$$\begin{aligned} \int f d\nu &= \lim_{n \rightarrow \infty} \int f\chi_{E_n} d\nu && \text{[MCT]} \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^n f(j) \\ \int f d\nu &= \sum_{j=1}^{\infty} f(j). \end{aligned}$$

So, for positive functions on \mathbb{N} , series is integral over counting measure.

Proposition 3.3.13 (Re-arrangement). Let $a_n \geq 0$, $\pi : \mathbb{N} \rightarrow \mathbb{N}$ be a bijection. Then,

$$\sum_{n=1}^{\infty} a_n = \sum_{k=1}^{\infty} a_{\pi(k)}.$$

Proof 4. We know $\sum_{n=1}^{\infty} a_n = \int f d\nu$, where $f(n) = a_n$. Take $E_n = \{\pi(1), \pi(2), \dots, \pi(n)\}$. Note that

$$\bigcup_{n=1}^{\infty} E_n = \mathbb{N} \quad \text{and} \quad E_n \subset E_{n+1}.$$

So,

$$\begin{aligned} \int f d\nu &= \lim_{n \rightarrow \infty} \int_{E_n} f d\nu = \lim_{n \rightarrow \infty} \sum_{k=1}^n f(\pi(k)) \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n a_{\pi(k)} = \sum_{k=1}^{\infty} a_{\pi(k)}. \end{aligned}$$

That is,

$$\sum_{n=1}^{\infty} a_n = \sum_{k=1}^{\infty} a_{\pi(k)}.$$

Q.E.D. ■

Lemma 3.3.14 For $f : X \rightarrow [0, \infty]$, \exists simple functions s_n s.t. $0 \leq s_n \leq s_{n+1} \quad \forall n$ and

$$\lim_{n \rightarrow \infty} s_n = f.$$

Proof 5. (Sketch). Consider the floor function

$$\lfloor y \rfloor = \max \{k \in \mathbb{Z} \mid k \leq y\}.$$

[For example, $\lfloor \sqrt{2} \rfloor = 1$, $\lfloor 10\sqrt{2} \rfloor = 14$, $\lfloor \frac{1}{10} \rfloor 10\sqrt{2} \rfloor = 1.4$, $\frac{1}{10^2} \lfloor 10^2 \sqrt{2} \rfloor = 1.41$.]

Define

$$s_n(x) = \min \{2^{-n} \lfloor 2^n f(x) \rfloor, n\} \implies s_n \leq s_{n+1}.$$

So, s_n takes values $\{k2^{-n} \mid k \in \mathbb{Z}, 0 \leq k2^n \leq n\}$.

Q.E.D. ■

Corollary 3.3.15 By MCT,

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int s_n \, d\mu.$$

Lemma 3.3.16 (Additivity of Integrals). If $f, g : X \rightarrow [0, \infty]$, then

$$\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu.$$

Proof 6. \exists simple functions s_n, t_n s.t. $0 \leq s_n \leq s_{n+1}$ and $0 \leq t_n \leq t_{n+1}$ with

$$s_n \rightarrow f \quad \text{and} \quad t_n \rightarrow g.$$

Then, $s_n + t_n \leq s_{n+1} + t_{n+1}$ and $s_n + t_n \rightarrow f + g$. We know

$$\int (s_n + t_n) \, d\mu = \int s_n \, d\mu + \int t_n \, d\mu.$$

By MCT,

$$\int (f + g) \, d\mu = \int f \, d\mu + \int g \, d\mu.$$

Q.E.D. ■

Proposition 3.3.17 For any sequence of $g_n : X \rightarrow [0, \infty]$,

$$\int \left(\sum_{n=1}^{\infty} g_n \right) \, d\mu = \sum_{n=1}^{\infty} \int g_n \, d\mu.$$

Proof 7. Note that $f_n = \sum_{j=1}^n g_j$ is an increasing sequence, since g_j 's are positive. So, $0 \leq f_n \leq f_{n+1}$,

and by additivity,

$$\int f_n \, d\mu = \sum_{j=1}^n \int g_j \, d\mu.$$

Note that

$$\lim_{n \rightarrow \infty} f_n = \sum_{j=1}^{\infty} g_j.$$

So, by MCT,

$$\begin{aligned} \int \left(\sum_{n=1}^{\infty} g_n \right) \, d\mu &= \int \left(\lim_{n \rightarrow \infty} f_n \right) \, d\mu = \lim_{n \rightarrow \infty} \int f_n \, d\mu \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^{\infty} \int g_j \, d\mu \\ &= \sum_{j=1}^{\infty} \int g_j \, d\mu. \end{aligned}$$

Q.E.D. ■

Corollary 3.3.18 Suppose $a_{j,n} \geq 0$. Then,

$$\sum_{n=1}^{\infty} \sum_{j=1}^{\infty} a_{j,n} = \sum_{j=1}^{\infty} \sum_{n=1}^{\infty} a_{j,n}.$$

Proof 8. Recall ν is the counting measure on \mathbb{N} , and

$$\int f \, d\nu = \sum_{n=1}^{\infty} f(n).$$

Define $g(n) = \mathbb{N} \rightarrow [0, \infty]$, $g_n(j) = a_{j,n}$. Then,

$$\begin{aligned} \sum_{j=1}^{\infty} \sum_{n=1}^{\infty} g_n(j) &= \int \sum_{n=1}^{\infty} g_n(j) \, d\nu \\ &= \sum_{n=1}^{\infty} \int g_n(j) \, d\nu \\ &= \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} g_n(j). \end{aligned}$$

Q.E.D. ■

MCT is great, but it requires monotonicity. What if we don't have an increasing sequence, but still want to study the limit?

Theorem 3.3.19 Fatou's Lemma

For any $f_n : X \rightarrow [0, \infty]$,

$$\int \liminf_{n \rightarrow \infty} f_n \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n \, d\mu.$$

Example 3.3.20 Fatou's Lemma in Action

$$\begin{array}{r} 1 + 2 = 3 \\ 2 + 1 = 3 \\ \hline \min 1 + 1 < 3. \end{array}$$

$$\implies \sum_j \min_k a_{j,k} \leq \min_k \sum_j a_{j,k}.$$

Proof 9. Define $g_n = \inf_{k \geq n} f_k$. Then, $\liminf_{n \rightarrow \infty} f_n = \lim_{n \rightarrow \infty} g_n = \sup_n g_n$. Note that $0 \leq g_n \leq g_{n+1}$. Then, by MCT,

$$\int \liminf_{n \rightarrow \infty} f_n \, d\mu = \int \lim_{n \rightarrow \infty} g_n \, d\mu = \lim_{n \rightarrow \infty} \int g_n \, d\mu.$$

Since $g_n = \inf_{k \geq n} f_k$, $g_n \leq f_n$. Then,

$$\int g_n \, d\mu \leq \int f_n \, d\mu.$$

We don't know if the limit exists, but we can certainly take the \liminf on both sides:

$$\liminf_{n \rightarrow \infty} \int g_n \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n \, d\mu.$$

The LHS is indeed a limit because $\int g_n \, d\mu$ is convergent (monotone sequence converges on $\widehat{\mathbb{R}}$). So,

$$\lim_{n \rightarrow \infty} \int g_n \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n \, d\mu.$$

So,

$$\int \liminf_{n \rightarrow \infty} f_n \, d\mu = \lim_{n \rightarrow \infty} \int g_n \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n \, d\mu.$$

[Key of this proof: \liminf/\limsup always exists for any sequence on the extended real line $\widehat{\mathbb{R}}$.] Q.E.D. ■

Proposition 3.3.21 Let μ be a measure on σ -algebra \mathcal{A} on X , $f : X \rightarrow [0, \infty]$. Then, a measure ν on \mathcal{A} is defined by

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

Moreover, for any $g : X \rightarrow [0, 1]$,

$$\int g \, d\nu = \int fg \, d\mu.$$

We concisely state this as $d\nu = f d\mu$.

Proof 10.

- $\nu(\emptyset) = \int_{\emptyset} f d\mu = \int \chi_{\emptyset} f d\mu = \int 0 d\mu = 0.$
- Disjoint sets E_n :

$$\begin{aligned}\chi_{\bigcup E_n} &= \sum_{n=1}^{\infty} \chi_{E_n} \\ \chi_{\bigcup E_n} f &= \sum_{n=1}^{\infty} \chi_{E_n} f \\ \int \chi_{\bigcup E_n} f &= \int \sum_{n=1}^{\infty} \chi_{E_n} f = \sum_{n=1}^{\infty} \int \chi_{E_n} f && \text{[By Corollary 3.18]} \\ \nu\left(\bigcup E_n\right) &= \sum_{n=1}^{\infty} \nu(E_n).\end{aligned}$$

So, ν is a measure. Now, let's prove the second half of the statement. If g is simple, $g = \sum_{j=1}^m c_j \chi_{A_j}$. Then,

$$\begin{aligned}\int g d\nu &= \sum_{j=1}^m c_j \nu(A_j) \\ &= \sum_{j=1}^m c_j \int \chi_{A_j} f d\mu && \text{[Exchange } \int \text{ and } \sum \text{]} \\ &= \int \underbrace{\sum_{j=1}^m c_j \chi_{A_j}}_{=g} f d\mu \\ &= \int g f d\mu\end{aligned}$$

If g is positive, take simple functions s_n s.t. $0 \leq s_n \leq s_{n+1}$ with $s_n \rightarrow g$. Then, $0 \leq s_n f \leq s_{n+1} f$, and $s_n f \rightarrow g f$. Since s_n is simple,

$$\int s_n d\nu = \int s_n f d\mu.$$

By MCT,

$$\begin{aligned}\int \lim_{n \rightarrow \infty} s_n d\nu &= \int \lim_{n \rightarrow \infty} s_n f d\mu \\ \int g d\nu &= \int g f d\mu.\end{aligned}$$

Q.E.D. ■

Lemma 3.3.22 (Markov's Inequality). For $f : X \rightarrow [0, \infty]$ and $c > 0$,

$$\mu(\{x \mid f(x) \geq c\}) \leq \frac{1}{c} \int f \, d\mu.$$

Proof 11. Denote $A = \{x \mid f(x) \geq c\}$. Then,

$$\begin{aligned} c\chi_A &\leq f \\ c\chi_A(x) &\leq f(x) \\ \int c\chi_A(x) \, d\mu &\leq \int f(x) \, d\mu \\ c \int \chi_A(x) \, d\mu &\leq \int f(x) \, d\mu \\ c\mu(A) &\leq \int f \, d\mu \\ \mu(A) &= \frac{1}{c} \int f \, d\mu. \end{aligned}$$

Q.E.D. ■

Corollary 3.3.23 If $f : X \rightarrow [0, \infty]$ and $\int f \, d\mu < \infty$, then $f < \infty$ μ -a.e. i.e.,

$$\mu(\{x \mid f(x) = \infty\}) = 0.$$

Proof 12. By Markov's Inequality,

$$\mu(\{x \mid f(x) \geq c\}) \leq \frac{1}{c} \int f \, d\mu.$$

Let $c \in \mathbb{N}$ and $\varepsilon_c = \{x \mid f(x) \geq c\}$. Then,

$$\{x \mid f(x) = \infty\} = \bigcap_{c \in \mathbb{N}} \varepsilon_c.$$

So,

$$\begin{aligned} \lim_{c \rightarrow \infty} \mu(\{x \mid f(x) \geq c\}) &\leq \lim_{c \rightarrow \infty} \frac{1}{c} \int f \, d\mu = 0 \\ \mu(\{x \mid f(x) = \infty\}) &= 0. \end{aligned}$$

Q.E.D. ■

Proposition 3.3.24 For $f : X \rightarrow [0, \infty]$,

$$\int f \, d\mu = 0 \iff f = 0 \quad \mu - a.e.$$

Proof 13. (\Leftarrow): Assume $f = 0$ μ -a.e.. For simple s s.t. $0 \leq s \leq f$. Then, $s = 0$ μ -a.e.. That is,

$$s = \sum c_j \chi_{A_j} \begin{cases} c_j > 0 \\ c_j = 0 \end{cases} \implies \mu(A_j) = 0.$$

So,

$$\int s \, d\mu = 0 \quad \text{and} \quad \int f \, d\mu = \sup_s \int s \, d\mu = \sup 0 = 0.$$

(\Rightarrow): Assume $\int f \, d\mu = 0$. Then, by Markov's Inequality,

$$\mu(\{x \mid f(x) \geq c\}) = 0 \quad \text{for any } c > 0.$$

Take $c = \frac{1}{k}$, $k \in \mathbb{N}$. Then,

$$\bigcup \{x \mid f(x) \geq c\} = \{x \mid f(x) > 0\}.$$

So,

$$\mu(\{x \mid f(x) > 0\}) = \sum \mu(\{x \mid f(x) \geq c\}) = 0.$$

i.e.,

$$\mu(\{x \mid f(x) = 0\}^c) = 0 \implies f = 0, \quad \mu - a.e.$$

Q.E.D. ■

Corollary 3.3.25 If $f = g$ μ -a.e., then

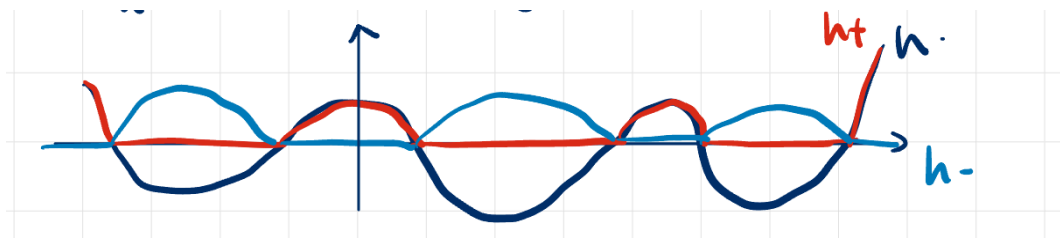
$$\int f \, d\mu = \int g \, d\mu.$$

3.4 Integration of Real and Complex Function

Definition 3.4.1 (h_+ and h_-). Given a real-valued function $h : X \rightarrow \mathbb{R}$, we define

$$h_+ = \max \{h, 0\} \geq 0$$

$$h_- = \max \{-h, 0\} \geq 0.$$



Remark. Note: $h = h_+ - h_-$ and $|h| = h_+ + h_-$.

Definition 3.4.2 (Integral of Real-Valued Functions). If at least one of the integrals $\int h_+ d\mu$ or $\int h_- d\mu$ is finite, we define

$$\int h d\mu = \int h_+ d\mu - \int h_- d\mu.$$

Example 3.4.3

Suppose $f : X \rightarrow (0, \infty]$ and $\int f d\mu < +\infty$, then

$$\log f \leq f$$

$$(\log f)_+ \leq f$$

$$\int (\log f)_+ d\mu \leq \int f d\mu < \infty$$

$\int \log f d\mu$ is defined, although its value can be $-\infty$.

Definition 3.4.4 (Integrability). In most cases, we'll be working with functions h for which both $\int h_{\pm} d\mu$ are finite, and such functions are called *integrable*.

Theorem 3.4.5

Let $h : X \rightarrow \mathbb{R}$. Then,

$$\int |h| d\mu < \infty \iff \int h_+ d\mu < \infty \quad \text{and} \quad \int h_- d\mu < \infty.$$

Proof 1. (\Rightarrow): Assume $\int |h| d\mu < \infty$. Since $0 \leq h_{\pm} \leq |h|$, we have

$$\int h_{\pm} d\mu < \infty.$$

(\Leftarrow): Assume $\int h_{\pm} d\mu < \infty$. Since $|h| = h_+ + h_-$, we have

$$\int |h| d\mu < \infty.$$

Q.E.D. ■

Proposition 3.4.6 (Integration of Real-Valued Functions is a Linear Operation).

Let μ be a measure on X . Then, the following is true:

- If $c \in \mathbb{R}$ and $f : X \rightarrow \mathbb{R}$ is integrable, then cf is integrable, and

$$\int cf \, d\mu = c \int f \, d\mu \quad (1)$$

- If $f, g : X \rightarrow \mathbb{R}$ are integrable, then $f + g$ is integrable, and

$$\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu \quad (2)$$

Proof2.

- cf is integrable because

$$\int |cf| \, d\mu = \int |c||f| \, d\mu = |c| \int |f| \, d\mu < \infty,$$

since $|c|$ and $|f|$ are positive and f is integrable. So, cf is integrable. Now, prove Equation (1).

Case I If $c \geq 0$, then

$$\begin{aligned} (cf)_+ &= \max\{cf, 0\} = c \max\{f, 0\} = cf_+ \\ (cf)_- &= \max\{-cf, 0\} = c \max\{-f, 0\} = cf_- \end{aligned}$$

So,

$$\begin{aligned} \int (cf) \, d\mu &= \int (cf)_+ \, d\mu - \int (cf)_- \, d\mu \\ &= \int cf_+ \, d\mu - \int cf_- \, d\mu \\ &= c \int f_+ \, d\mu - c \int f_- \, d\mu && \text{[Linearity]} \\ &= c \left(\int f_+ \, d\mu - \int f_- \, d\mu \right) \\ &= c \int f \, d\mu. \end{aligned}$$

Case II If $c < 0$, then

$$\begin{aligned} (cf)_+ &= \max\{cf, 0\} = \max\{(-c)(-f), 0\} \\ &= (-c) \max\{-f, 0\} = -cf_- \\ (cf)_- &= \max\{c(-f), 0\} = \max\{(-c)f, 0\} \\ &= (-c) \max\{f, 0\} = -cf_+. \end{aligned}$$

So,

$$\begin{aligned}
 \int (cf) \, d\mu &= \int (cf)_+ \, d\mu - \int (cf)_- \, d\mu \\
 &= \int -cf_- \, d\mu - \int -cf_+ \, d\mu \\
 &= -c \int f_- \, d\mu + c \int f_+ \, d\mu \\
 &= c \left(\int f_+ \, d\mu - \int f_- \, d\mu \right) \\
 &= c \int f \, d\mu.
 \end{aligned}$$

- $f + g$ is integrable because

$$\int |f + g| \, d\mu \leq \int (|f| + |g|) \, d\mu = \int |f| \, d\mu + \int |g| \, d\mu < \infty,$$

by triangle inequality. Denote $h = f + g$. Then,

$$\begin{aligned}
 h_+ - h_- &= h = (f_+ - f_-) + (g_+ - g_-) \\
 h_+ + f_- + g_- &= f_+ + g_+ + h_- && \text{[Rearrange]} \\
 \int h_+ + f_- + g_- \, d\mu &= \int f_+ + g_+ + h_- \, d\mu \\
 \int h_+ \, d\mu + \int f_- \, d\mu + \int g_- \, d\mu &= \int f_+ \, d\mu + \int g_+ \, d\mu + \int h_- \, d\mu && \text{[Additivity]} \\
 \int h_+ \, d\mu - \int h_- \, d\mu &= \int f_+ \, d\mu - \int f_- \, d\mu + \int g_+ \, d\mu - \int g_- \, d\mu \\
 \int h \, d\mu &= \int f \, d\mu + \int g \, d\mu.
 \end{aligned}$$

Q.E.D. ■

Definition 3.4.7 (\mathcal{L}^1 Space, Integrable Functions). We denote by $\mathcal{L}^1(X, d\mu) = \mathcal{L}^1(d\mu)$ the set of all measurable functions $f : X \rightarrow \mathbb{C}$ s.t.

$$\int_X |f| \, d\mu < \infty.$$

We call such f integrable.

Lemma 3.4.8

$$f \in \mathcal{L}^1(d\mu) \iff \int |\operatorname{Re} f| \, d\mu < \infty \quad \text{and} \quad \int |\operatorname{Im} f| \, d\mu < \infty.$$

Proof 3. (\Rightarrow): Suppose $f \in \mathcal{L}^1(d\mu)$. Then, $\int |f| \, d\mu < \infty$. Since $|\operatorname{Re} f| \leq |f| = \sqrt{|\operatorname{Re} f|^2 + |\operatorname{Im} f|^2}$ and

$$|\operatorname{Im} f| \leq |f|,$$

$$\int |\operatorname{Re} f| \, d\mu < \infty \quad \text{and} \quad \int |\operatorname{Im} f| \, d\mu < \infty.$$

(\Leftarrow): Suppose $\int |\operatorname{Re} f| \, d\mu < \infty$ and $\int |\operatorname{Im} f| \, d\mu < \infty$. Then,

$$|f| = \sqrt{|\operatorname{Re} f|^2 + |\operatorname{Im} f|^2} \leq |\operatorname{Re} f| + |\operatorname{Im} f|,$$

since $(\sqrt{a^2 + b^2})^2 = a^2 + b^2 \leq a^2 + b^2 + 2ab = (a + b)^2 \implies \sqrt{a^2 + b^2} \leq a + b$. Therefore,

$$\int |f| \, d\mu < \infty.$$

Q.E.D. ■

Definition 3.4.9 (Integral of Complex-Valued Functions). For $f \in \mathcal{L}^1(d\mu)$, we define

$$\int f \, d\mu = \int \operatorname{Re} f \, d\mu + i \int \operatorname{Im} f \, d\mu.$$

Theorem 3.4.10 Linearity of Integrals of Complex-Valued Functions

- If $c \in \mathbb{C}$ and $f \in \mathcal{L}^1(d\mu)$, then $cf \in \mathcal{L}^1(d\mu)$ and Equation (1) holds.
- If $f, g \in \mathcal{L}^1(d\mu)$, then $f + g \in \mathcal{L}^1(d\mu)$ and Equation (2) holds.

Proof 4.

- cf is integrable because

$$\int |cf| \, d\mu = \int |c||f| \, d\mu = |c| \int |f| \, d\mu < \infty.$$

Now, let's prove Equation (1):

$$\begin{aligned} \operatorname{Re}(cf) &= \operatorname{Re}[(\operatorname{Re} c + i \operatorname{Im} c)(\operatorname{Re} f + i \operatorname{Im} f)] \\ &= \operatorname{Re}[\operatorname{Re} c \operatorname{Re} f - \operatorname{Im} c \operatorname{Im} f + i(\operatorname{Re} c \operatorname{Im} f + \operatorname{Im} c \operatorname{Re} f)] \\ &= \operatorname{Re} c \cdot \operatorname{Re} f - \operatorname{Im} c \cdot \operatorname{Im} f \\ \operatorname{Im}(cf) &= \operatorname{Re} c \cdot \operatorname{Im} f + \operatorname{Im} c \cdot \operatorname{Re} f. \end{aligned}$$

Therefore,

$$\begin{aligned} \int cf \, d\mu &= \int \operatorname{Re}(cf) \, d\mu + i \int \operatorname{Im}(cf) \, d\mu \\ &= \int \operatorname{Re} c \cdot \operatorname{Re} f \, d\mu - \int \operatorname{Im} c \cdot \operatorname{Im} f \, d\mu + i \left(\int \operatorname{Re} c \cdot \operatorname{Im} f \, d\mu + \int \operatorname{Im} c \cdot \operatorname{Re} f \, d\mu \right) \end{aligned}$$

by additivity of real-valued functions. Then,

$$\begin{aligned} \int cf \, d\mu &= \operatorname{Re} c \left(\underbrace{\int \operatorname{Re} f \, d\mu + i \int \operatorname{Im} f \, d\mu}_{\int f \, d\mu} \right) + i \operatorname{Im} c \left(\underbrace{\int \operatorname{Re} f \, d\mu + i \int \operatorname{Im} f \, d\mu}_{\int f \, d\mu} \right) \\ &= \int f \, d\mu \underbrace{(\operatorname{Re} c + i \operatorname{Im} c)}_c \\ &= c \int f \, d\mu. \end{aligned}$$

- $f + g$ is integrable because

$$\int |f + g| \, d\mu \leq \int (|f| + |g|) \, d\mu = \int |f| \, d\mu + \int |g| \, d\mu < \infty.$$

Note that

$$\operatorname{Re}(f + g) = \operatorname{Re} f + \operatorname{Re} g$$

$$\operatorname{Im}(f + g) = \operatorname{Im} f + \operatorname{Im} g.$$

Then,

$$\begin{aligned} \int f + g \, d\mu &= \int \operatorname{Re}(f + g) \, d\mu + i \int \operatorname{Im}(f + g) \, d\mu \\ &= \int \operatorname{Re} f + \operatorname{Re} g \, d\mu + i \int \operatorname{Im} f + \operatorname{Im} g \, d\mu \\ &= \int \operatorname{Re} f \, d\mu + \int \operatorname{Re} g \, d\mu + i \int \operatorname{Im} f \, d\mu + i \int \operatorname{Im} g \, d\mu \\ &= \left(\int \operatorname{Re} f \, d\mu + i \int \operatorname{Im} f \, d\mu \right) + \left(\int \operatorname{Re} g \, d\mu + i \int \operatorname{Im} g \, d\mu \right) \\ &= \int f \, d\mu + \int g \, d\mu. \end{aligned}$$

Q.E.D. ■

Theorem 3.4.11

If $f \in \mathcal{L}^1(d\mu)$, then

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu.$$

Proof 5. Pick $\omega \in \mathcal{C}$ s.t.

$$\begin{cases} |\omega| = 1 \\ \omega \int f \, d\mu = \left| \int f \, d\mu \right|. \end{cases}$$

Then,

$$\begin{aligned}
 \left| \int f \, d\mu \right| &= \omega \int f \, d\mu = \operatorname{Re} \int \omega f \, d\mu \\
 &= \int \operatorname{Re}(\omega f) \, d\mu \\
 &\leq \int |\omega f| \, d\mu \\
 &= \int |\omega| |f| \, d\mu \\
 &= \int |f| \, d\mu.
 \end{aligned}$$

Q.E.D. ■

Theorem 3.4.12 Dominated Convergence Theorem, DCT

Let $(f_n)_{n=1}^{\infty}$ be a sequence of $\mathcal{L}^1(d\mu)$, “dominated” by some function $g \in \mathcal{L}^1(d\mu)$ in the sense that

$$|f_n(x)| \leq g(x) \quad \forall n \in \mathbb{N} \quad \mu - a.e. \, x \quad (3)$$

Assume that f_n converges pointwise μ -a.e. defining a function

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) \quad \mu - a.e. \quad (4)$$

Then, $f \in \mathcal{L}^1(d\mu)$, and

$$\lim_{n \rightarrow \infty} \int |f_n - f| \, d\mu = 0, \quad \text{and} \quad (5)$$

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu = \int f \, d\mu \quad (6)$$

Proof 6.

- From (3) and (4), we have

$$\begin{aligned}
 |f| &\leq g \quad \mu - a.e. \\
 \implies f &\in \mathcal{L}^1(d\mu) \quad \text{since } g \in \mathcal{L}^1(d\mu).
 \end{aligned}$$

- Let $h_n = 2g - \underbrace{|f_n - f|}_{\leq 2g} \geq 0$. Also note that

$$\lim_{n \rightarrow \infty} h_n = 2g \quad \mu - a.e.$$

By Fatou’s Lemma,

$$\int 2g \, d\mu \leq \liminf_{n \rightarrow \infty} \int (2g - |f_n - f|) \, d\mu$$

Since $2g$ and $|h_n - h|$ are integrable, we have

$$\begin{aligned} \int 2g \, d\mu &\leq \liminf_{n \rightarrow \infty} \left(\int 2g \, d\mu - \int |f_n - f| \, d\mu \right) \\ &= \int 2g \, d\mu + \liminf_{n \rightarrow \infty} \left(- \int |f_n - f| \, d\mu \right) \\ &= \int 2g \, d\mu - \limsup_{n \rightarrow \infty} \int |f_n - f| \, d\mu. \end{aligned}$$

So,

$$\limsup_{n \rightarrow \infty} \int |f_n - f| \, d\mu \leq 0.$$

But

$$\limsup_{n \rightarrow \infty} \int |f_n - f| \, d\mu \geq 0.$$

Then, Equation (5) holds.

- Note that

$$0 \leq \left| \int f_n \, d\mu - \int f \, d\mu \right| \leq \int |f_n - f| \, d\mu.$$

Since $|f_n - f| \rightarrow 0$, we have

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu = \int f \, d\mu.$$

Q.E.D. ■

Definition 3.4.13 (Summation Over a Set).

$$\sum_{j \in J} f(j) = \int f \, d\nu,$$

where ν is the counting measure. If $J = \mathbb{N}$, then,

$$\sum_{j \in \mathbb{N}} f(j) = \sum_{j=1}^{\infty} f(j)$$

when the series is absolutely convergent. That is, $\sum_{j \in \mathbb{N}} f(j)$ is only defined when $\sum_{j=1}^{\infty} |f(j)| < \infty$.

Example 3.4.14

If $z_n \rightarrow z$ in \mathbb{C} , then $\lim_{n \rightarrow \infty} \left(1 + \frac{z_n}{n}\right)^n = \underbrace{\sum_{j=0}^{\infty} \frac{z^j}{j!}}_{e^z}$.

Also note that $\left(1 + \frac{x}{n}\right)^n \rightarrow e^x$. This proof can be completed using DCT.

4 Lebesgue-Stieltjes Measures

4.1 Increasing Functions and Distribution Functions

Definition 4.1.1 (Increasing Functions). $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ is an increasing function if $x \leq y \implies \alpha(x) \leq \alpha(y)$.

Remark. This definition will be related to measure by

$$\mu((x, y]) = \alpha(y) - \alpha(x).$$

We would also have

$$\begin{aligned} \mu((x, y]) + \mu((y, z]) &= \alpha(x) - \alpha(y) + \alpha(y) - \alpha(z) \\ &= \alpha(x) - \alpha(z) = \mu((x, z]). \end{aligned}$$

This is the additivity of measures.

Lemma 4.1.2 Let $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ increasing, then,

- One-sided Limits:

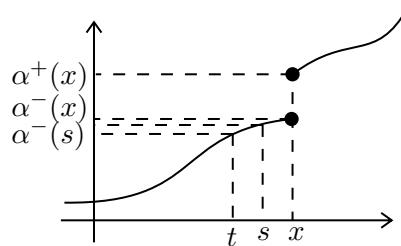
$$\alpha^+(x) = \lim_{t \downarrow x} \alpha(t) \quad \text{and} \quad \alpha^-(x) = \lim_{t \uparrow x} \alpha(t)$$

exist $\forall x \in \mathbb{R}$, and so do

$$\alpha^+(-\infty) = \lim_{t \downarrow -\infty} \alpha(t) \quad \text{and} \quad \alpha^- (+\infty) = \lim_{t \uparrow +\infty} \alpha(t).$$

Proof 1. Define $\alpha^-(x) = \sup_{t < x} \alpha(t)$ and $\alpha^+(x) = \inf_{t > x} \alpha(t)$. Then, $\lim_{t \uparrow x} \alpha(t) = \alpha^-(x)$ because $\forall \varepsilon > 0$, $\exists t < x$ s.t. $\alpha(t) < \alpha^-(x) - \varepsilon \implies \forall t < s < x$, $\alpha^-(x) - \varepsilon < \alpha(s) < \alpha^-(x)$. The other conclusions follow a similar argument. Note that

$$\alpha^- (+\infty) = \sup_{t \in \mathbb{R}} \alpha(t).$$



Q.E.D. ■

- $\alpha^-(x) \leq \alpha(x) \leq \alpha^+(x)$

Proof2. $\alpha(t) \leq \alpha(x) \quad \forall t < x$. Taking $\sup_{t < x}$, we have $\alpha^-(x) \leq \alpha(x)$. Q.E.D. ■

- If $x < y$, then $\alpha^+(x) \leq \alpha^-(y)$.

Proof3. Use some $t \in (x, y)$. Then,

$$\alpha^+(x) \leq \alpha(t) \leq \alpha^-(y).$$

Q.E.D. ■

- If $x < y$, then $\alpha^+(x) \leq \alpha^+(y) \implies \alpha^+$ is another increasing function, so is α^- .
- $(\alpha^-)^+ = \alpha^+$ and $(\alpha^+)^- = \alpha^-$.
- α can have countably many discontinuities (point x at which α is not continuous).

Proof4. For every x s.t. $\alpha^-(x) < \alpha^+(x)$ (i.e., α is not continuous), take the interval $(\alpha^-(x), \alpha^+(x))$. Different such x produce disjoint intervals. Each of the interval contains at least one rational number. So, we cannot have more intervals than rationals. That is, the discontinuity is countably many. Q.E.D. ■

Definition 4.1.3 (Distribution Function). Let μ be a Borel measure on \mathbb{R} . The function α is said to be a distribution function of μ if

$$\mu((x, y]) = \alpha(y) - \alpha(x) \quad \forall x, y \in \mathbb{R}, x < y.$$

Remark. Distribution functions are uniquely defined up to a constant.

Proposition 4.1.4 For a Borel measure μ on \mathbb{R} , the following are equivalent (TFAE):

- μ is finite on compact sets. i.e.,

$$\mu(K) < \infty \quad \forall \text{ compact } K.$$

- $\mu((x, y]) < \infty \quad \forall x, y \in \mathbb{R}, x < y$.
- μ has a distribution function.

Proof5. ③ \implies ② is trivial.

② \implies ①: Any compact K is bounded: $K \subset (-c, c]$. Then,

$$\mu(K) \leq \mu((-c, c]) < \infty.$$

① \implies ②: $K = [x, y]$ is compact, so,

$$\mu((x, y)) \leq \mu([x, y]) < \infty.$$

② \implies ③: Define a distribution function

$$\alpha(x) = \begin{cases} \mu((0, x]), & \text{if } x > 0 \\ 0, & \text{if } x = 0 \\ -\mu((x, 0]), & \text{if } x < 0. \end{cases}$$

Suppose $x < y$.

Case I $0 < x < y$:

$$\alpha(y) - \alpha(x) = \mu((0, y]) - \mu((0, x]) = \mu((x, y]).$$

Case II $x < 0 < y$:

$$\alpha(y) - \alpha(x) = \mu((0, y]) + \mu((x, 0]) = \mu((x, y]).$$

Case III $x < y < 0$:

$$\alpha(y) - \alpha(x) = -\mu((y, 0]) + \mu((x, 0]) = \mu((x, y]).$$

It is also easy to verify cases where $x = 0$ or $y = 0$.

Q.E.D. ■

Proposition 4.1.5 If α is a distribution function, then α is right-continuous. i.e.,

$$\alpha^+(x) = \alpha(x) \quad \forall x \in \mathbb{R}.$$

Proof 6.

$$\begin{aligned} \alpha^+(x) &= \lim_{t \downarrow x} \alpha(t) \\ \alpha^+(x) - \alpha(x) &= \lim_{t \downarrow x} (\alpha(t) - \alpha(x)) \\ &= \lim_{t \downarrow x} \mu((x, t]) \\ &= \lim_{n \rightarrow \infty} \mu\left(\left(x, x + \frac{1}{n}\right]\right) && \text{[Take a sequence]} \\ &= \mu\left(\bigcap_{n \in \mathbb{N}} \left(x, x + \frac{1}{n}\right]\right) \\ &= \mu(\emptyset) \\ &= 0. \end{aligned}$$

So, $\alpha^+(x) = \alpha(x)$.

Q.E.D. ■

Example 4.1.6

The same argument *does not* apply to $\alpha^-(x)$.

$$\alpha(x) - \alpha^-(x) = \lim_{n \rightarrow \infty} \mu\left(\left(x - \frac{1}{n}, x\right]\right) = \mu\left(\bigcap_{n \in \mathbb{N}} \left(x - \frac{1}{n}, x\right]\right) = \mu(\{x\})$$

Consider

$$\delta_0(A) = \begin{cases} 1, & 0 \in A \\ 0, & 0 \notin A. \end{cases}$$

Then, a distribution function is

$$\mu((x, y]) = \alpha(y) - \alpha(x) = \begin{cases} 1, & \text{if } x < 0 < y \\ 0, & \text{otherwise.} \end{cases}$$

So,

$$\alpha(x) = \begin{cases} 0, & x < 0 \\ 1, & x > 0. \end{cases}$$

Corollary 4.1.7 Distribution function α is continuous $\iff \mu(\{x\}) = 0 \quad \forall x$.

Lemma 4.1.8 If μ is a measure with distribution function α , then

$$\mu((x, y)) = \alpha^-(y) - \alpha(x).$$

Proof 7.

$$\begin{aligned} \mu((x, y)) &= \lim_{n \rightarrow \infty} \mu\left(\left(x, y - \frac{1}{n}\right]\right) && \text{[increasing sequence: } \bigcup_{n \in \mathbb{N}} \left(x, y - \frac{1}{n}\right] = (x, y)] \\ &= \lim_{n \rightarrow \infty} \left(\alpha\left(y - \frac{1}{n}\right) - \alpha(x)\right) \\ &= \alpha^-(y) - \alpha(x). \end{aligned}$$

Q.E.D. ■

4.2 Outer Measure

Definition 4.2.1 (Outer Measure). An *outer measure* on X is a map $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$ with the properties:

- $\mu^*(\emptyset) = 0$
- $A \subset B \implies \mu^*(A) \leq \mu^*(B)$
- σ -subadditivity:

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

[An outer measure is not a measure.]

Build an Outer Measure Fix a class of “elementary” sets $\mathcal{E} \subset \mathcal{P}(X)$. Fix a weight $\rho : \mathcal{E} \rightarrow [0, \infty]$. Assume that

- $\emptyset, X \in \mathcal{E}$
- $\rho(\emptyset) = 0$.

Example 4.2.2

$X = \mathbb{R}$, a valid elementary set is defined as

$$\mathcal{E} = \{(a, b) \mid a < b\} \cup \{\emptyset, \mathbb{R}\}$$

Then, $\rho((a, b)) = b - a$, $\rho(\emptyset) = 0$, and $\rho(\mathbb{R}) = \infty$.

Take infimum over all countable covers of A by elementary sets:

$$\mu^*(A) = \inf \left\{ \sum_{j=1}^{\infty} \rho(E_j) \mid E_j \in \mathcal{E}, A \subset \bigcup_{j=1}^{\infty} E_j \right\}.$$

Proposition 4.2.3 μ^* is an outer measure on X .

Proof 1.

- If $A = \emptyset$, then take all $E_j = \emptyset$.

$$\mu^*(\emptyset) = 0.$$

- Any countable cover for B is also a countable cover for A .
- Fix $\varepsilon > 0$. For each n , pick countable cover $A_n \subset \bigcup_{j=1}^{\infty} E_{n,j}$ s.t.

$$\sum_{j=1}^{\infty} \rho(E_{n,j}) \leq \mu^*(A_n) + \frac{\varepsilon}{2^n} \quad \text{[Pick a countable cover close to the optimal one]}$$

$$\bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} \bigcup_{j=1}^{\infty} E_{n,j} \quad \text{[countable cover for } \bigcup_{n=1}^{\infty} A_n \text{]}$$

$$\sum_{n=1}^{\infty} \sum_{j=1}^{\infty} \rho(E_{n,j}) \leq \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon.$$

So,

$$\mu^* \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \mu^*(A_n) + \varepsilon.$$

Since ε was chosen arbitrarily, take $\varepsilon \rightarrow 0$:

$$\mu^* \left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

Q.E.D. ■

Remark.

- Instead of countable cover, we can also take finite cover.
- For any elementary set, $\mu^*(E) \leq \rho(E)$. To see this, cover E by $E, \emptyset, \emptyset, \dots$

Definition 4.2.4 (Carathéodory). The set $A \subset X$ is *measurable with respect to μ^** if $\forall S \subset X$,

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

Remark.

- By σ -subadditivity,

$$\mu^*(S) \leq \mu^*(S \cap A) + \mu^*(S \cap A^c).$$

- Motivation: $\mu^*(X) < \infty$, $A \subset X$, and $\mu^*(A)$ can be viewed as an upper sum. The inner measure, which can be viewed as a lower sum, is defined as

$$\begin{aligned} \mu_*(A) &= \mu^*(X) - \mu^*(X \setminus A) \\ \implies \mu_*(A) &\leq \mu^*(A). \end{aligned}$$

That is, lower sum \leq upper sum, and

$$\mu_*(A) = \mu^*(A) \iff \mu^*(A) + \mu^*(X \setminus A) = \mu^*(X).$$

Theorem 4.2.5 Carathéodory

For any outer measure μ^* on X , the set

$$\mathcal{A} = \{A \mid A \text{ measurable w.r.t. } \mu^*\}$$

is a σ -algebra, and $\mu^*|_{\mathcal{A}}$ is a measure on \mathcal{A} .

Proof2.

$$\mathcal{A} = \{A \mid \forall S \text{ s.t. } \mu^*(S) = \mu^*(S \cap A) + \mu^*(S \cap A^c)\}.$$

- \mathcal{A} is a σ -algebra.

1. If $A = \emptyset$, $\mu^*(S) = \mu^*(\emptyset) + \mu^*(S) = \mu^*(S)$. So, $\emptyset \in \mathcal{A}$.
2. If $A \in \mathcal{A}$, $A^c \in \mathcal{A}$ because the definition is symmetric to A and A^c .
3. Let $A_n \in \mathcal{A}$. Let

$$B = \bigcup_{j=1}^{\infty} A_j \quad \text{and} \quad B_n = \bigcup_{j=1}^n A_j.$$

[WTS: $B \in \mathcal{A}$.] Let $S \subset X$, then Carathéodory definition, since $A_n \in \mathcal{A}$,

$$\begin{aligned}\mu^*(S \cap B_{n-1}^c) &= \mu^*(S \cap B_{n-1}^c \cap A_n) + \mu^*(S \cap B_{n-1}^c \cap A_n^c) \\ &= \mu^*(S \cap B_{n-1}^c \cap A_n) + \mu^*(S \cap \underbrace{(B_{n-1} \cup A_n)^c}_{=B_n^c}) \\ &= \mu^*(S \cap B_{n-1}^c \cap A_n) + \mu^*(S \cap B_n^c)\end{aligned}$$

By induction on n and “telescoping”:

$$\begin{aligned}\mu^*(S) &= \mu^*(S \cap B_n^c) + \sum_{j=1}^n \mu^*(S \cap B_{j-1}^c \cap A_j) \\ &\geq \underbrace{\mu^*(S \cap B^c)}_{\substack{\perp n \\ \text{By monotonicity,} \\ B^c \subset B_n^c}} + \sum_{j=1}^n \mu^*(S \cap B_{j-1}^c \cap A_j)\end{aligned}$$

Let $n \rightarrow \infty$, we have

$$\mu^*(S) \geq \mu^*(S \cap B^c) + \sum_{j=1}^{\infty} \mu^*(S \cap B_{j-1}^c \cap A_j).$$

Conversely, by σ -subadditivity,

$$\begin{aligned}\mu^*(S) &\leq \mu^*(S \cap B^c) + \mu^*(S \cap B) \\ &\leq \mu^*(S \cap B^c) + \sum_{j=1}^{\infty} \mu^*(S \cap \underbrace{B_{j-1}^c \cap A_j}_{B_j \setminus B_{j-1}}) \leq \mu^*(S).\end{aligned}$$

Hence, these are all equal. Since $\forall S$, $\mu^*(S) = \mu^*(S \cap B^c) + \mu^*(S \cap B)$, so we have $B \in \mathcal{A}$.

We have concluded that \mathcal{A} is a σ -algebra.

- $\mu^*|_{\mathcal{A}}$ is a measure on \mathcal{A} . [WTS: $\mu^*|_{\mathcal{A}}$ is σ -additive.]

Let $A_n \in \mathcal{A}$ with A_n disjoint. Use the same definition for B and B_n . By equality we proven earlier,

$$\begin{aligned}\mu^*(S) &= \mu^*(S \cap B^c) + \sum_{j=1}^{\infty} \mu^*(S \cap B_{j-1}^c \cap A_j) \\ &= \mu^*(S \cap B^c) + \sum_{j=1}^{\infty} \mu^*(S \cap A_j).\end{aligned}$$

Pick $S = B$, we have

$$\mu^*(B) = \underbrace{\mu^*(B \cap B^c)}_{=0} + \sum_{j=1}^{\infty} \underbrace{\mu^*(B \cap A_j)}_{=A_j} = \sum_{j=1}^{\infty} \mu^*(A_j).$$

So, $\mu^*|_{\mathcal{A}}$ has σ -additivity, and it is a measure on \mathcal{A} .

Q.E.D. ■

4.3 Borel Measures on \mathbb{R}

Application Develop the Lebesgue-Stieltje's measure

Set-Up Fix an increasing function $\alpha : \mathbb{R} \rightarrow \mathbb{R}$.

$$\mathcal{E} = \{\emptyset\} \cup \{(a, b) \mid a, b \in \mathbb{R}, a < b\}.$$

Define $\rho : \mathcal{E} \rightarrow [0, \infty]$ as

$$\rho(\emptyset) = 0, \quad \rho((a, b)) = \alpha^-(b) - \alpha^+(a).$$

Let μ^* be outer measure corresponding to this:

$$\mu^*(A) = \inf \left\{ \sum_i \rho(E_i) \mid E_i \subset \mathcal{E}, A \subset \bigcup_{j=1}^{\infty} E_j \right\}.$$

Lemma 4.3.1 For the μ^* defined above, for any $p, q \in \mathbb{R}$, $p \leq q$,

$$\mu^*([p, q]) = \alpha^+(q) - \alpha^-(p).$$

Proof 1. (\leq): $[p, q] \subset (p - \varepsilon, q + \varepsilon)$.

$$\mu^*([p, q]) \leq \rho((p - \varepsilon, q + \varepsilon))$$

$$\mu^*([p, q]) \leq \alpha^-(q + \varepsilon) - \alpha^+(p - \varepsilon)$$

Let $\varepsilon \rightarrow 0$:

$$\mu^*([p, q]) \leq (\alpha^-)^+(q) - (\alpha^+)^-(p)$$

$$\mu^*([p, q]) \leq \alpha^+(q) - \alpha^-(p).$$

(\geq): Consider any countable cover

$$[p, q] \subset \bigcup_{j=1}^{\infty} (a_j, b_j).$$

Then, \exists a finite subcover

$$[p, q] \subset \bigcup_{j=1}^n (a_j, b_j)$$

Take finite subcover with the smallest possible number of intervals. This minimality implies

$$a_j \neq a_k \quad \text{for } j \neq k \quad \text{and} \quad b_j \neq b_k \quad \text{for } j \neq k.$$

So, WLOG, assume

$$a_1 < a_2 < \cdots < a_n, \quad \text{and} \quad b_j < b_k \quad \text{for } j < k.$$

So, the cover will look like “telescoping,” and

$$\begin{aligned} \sum_{j=1}^{\infty} \rho((a_j, b_j)) &\geq \sum_{j=1}^n \rho((a_j, b_j)) \\ &= \sum_{j=1}^n (\alpha^-(b_j) - \alpha^+(a_j)) \\ &\geq \alpha^-(b_n) - \alpha^+(a_n) + \underbrace{\sum_{j=1}^{n-1} (\alpha^+(a_{j+1}) - \alpha^+(a_j))}_{\text{telescoping}} \\ &\quad [\text{use } \alpha^-(b_j) \geq \alpha^+(a_{j+1}) \text{ for } j < n.] \\ &= \alpha^-(b_n) - \alpha^+(a_n) + \alpha^+(a_n) - \alpha^+(a_1) \\ &= \alpha^-(b_n) - \alpha^+(a_1) \\ &\geq \alpha^+(q) - \alpha^-(p). \end{aligned}$$

Therefore, $\mu^*([p, q]) = \alpha^+(q) - \alpha^-(p)$.

Q.E.D. ■

Lemma 4.3.2 $\forall a < b, \quad \mu^*((a, b)) = \alpha^-(b) - \alpha^+(a) = \rho((a, b))$.

Proof 2. (\leq): $(a, b) \subset (a, b)$. So,

$$\mu^*((a, b)) \leq \rho((a, b)) = \alpha^-(b) - \alpha^+(a).$$

(\geq): Let $[p, q] \subset (a, b)$. Then,

$$\mu^*((a, b)) \geq \mu^*([p, q]) = \alpha^+(q) - \alpha^-(p).$$

Take $p \downarrow a$ and $q \uparrow b$, then $\alpha^+(q) \rightarrow \alpha^-(b)$ and $\alpha^-(p) \rightarrow \alpha^+(a)$. So,

$$\mu^*((a, b)) \geq \alpha^-(b) - \alpha^+(a).$$

Q.E.D. ■

Lemma 4.3.3

$$\mu^*((a, c]) = \alpha^+(c) - \alpha^+(a).$$

Goal Carathéodory Theorem gives a σ -algebra

$$\mathcal{A} = \{A \mid A \text{ measurable w.r.t. } \mu^*\} = \{A \mid \forall S, \mu^*(S) = \mu^*(S \cap A^c) + \mu^*(S \cap A)\}.$$

Also, $\mu^*|_{\mathcal{A}}$ is a measure.

We want to show:

- $\mathcal{B}_{\mathbb{R}} \subset \mathcal{A}$
- Denote $\mu^*|_{\mathcal{B}_{\mathbb{R}}} = \mu_{\alpha}$, then

$$\mu_{\alpha}([a, b]) = \alpha^+(b) - \alpha^+(a).$$

[To show $\mathcal{B}_{\mathbb{R}} \subset \mathcal{A}$, we need to show \mathcal{A} contains all the half-lines. So, we should work on half-lines first.]

Lemma 4.3.4 For any $I \in \mathcal{E}$ and any $c \in \mathbb{R}$,

$$\mu^*(I) = \mu^*(I \cap (-\infty, c]) + \mu^*(I \cap (c, \infty)).$$

Proof 3. Case I $I = \emptyset$ is trivial.

Case II $I = (a, b) \subset (-\infty, c]$, and $I = (a, b) \subset (c, \infty)$ are also trivial because

$$\mu^*(I) = \mu^*(I) + \mu^*(\emptyset).$$

Case III $a < c < b$. Note that

$$\alpha^-(b) - \alpha^+(a) = (\alpha^+(c) - \alpha^+(a)) + (\alpha^-(b) - \alpha^+(c))$$

So,

$$\begin{aligned} \mu^*((a, b)) &= \mu^*((a, c]) + \mu^*((b, c]) \\ &= \mu^*((a, b) \cap (-\infty, c]) + \mu^*((a, b) \cap (c, \infty)). \end{aligned}$$

Q.E.D. ■

Lemma 4.3.5 $\forall c, (c, \infty)$ is measurable with respect to μ^* . In other words, $\forall c \in \mathbb{R}, \forall S \subset \mathbb{R}$,

$$\mu^*(S) = \mu^*(S \cap (-\infty, c]) + \mu^*(S \cap (c, \infty)).$$

Proof 4. (\geq): Fix $S \subset \mathbb{R}$. Fix a countable cover of S by $I_j \in \mathcal{E}$. Then,

$$\begin{aligned} \underbrace{\sum_{j=1}^{\infty} \mu^*(I_j)}_{\text{countable cover for } S} &= \underbrace{\sum_{j=1}^{\infty} \mu^*(I_j \cap (-\infty, c])}_{\text{countable cover for } S \cap (-\infty, c]} + \underbrace{\sum_{j=1}^{\infty} \mu^*(I_j \cap (c, \infty))}_{\text{countable cover for } S \cap (c, \infty)} \\ &\geq \mu^*(S \cap (-\infty, c]) + \mu^*(S \cap (c, \infty)) \\ \sum_{j=1}^{\infty} \rho(I_j) &\geq \mu^*(S \cap (-\infty, c]) + \mu^*(S \cap (c, \infty)). \end{aligned}$$

Take inf over covers:

$$\mu^*(S) \geq \mu^*(S \cap (-\infty, c]) + \mu^*(S \cap (c, \infty)).$$

(\leq): $\mu^*(S) \leq \mu^*(S \cap (-\infty, c]) + \mu^*(S \cap (c, \infty))$ is trivial by σ -additivity. Q.E.D. ■

As a summary here,

$$(c, \infty) \in \mathcal{A} \implies \mathcal{B}_{\mathbb{R}} \in \mathcal{A}.$$

Therefore, $\mu^*|_{\mathcal{A}}$ is a measure, and $\mu^*|_{\mathcal{B}_{\mathbb{R}}}$ is also a measure.

$$\mu^*((a, b]) = \alpha^+(b) - \alpha^+(a).$$

If α is right continuous,

$$\mu^*((a, b]) = \alpha(b) - \alpha(a).$$

Theorem 4.3.6 Lebesgue-Stieltjes Measure

For any $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ increasing, right continuous, \exists a Borel measure

$$\mu_{\alpha} = \mu^*|_{\mathcal{B}_{\mathbb{R}}} \quad \text{on } \mathbb{R}$$

, called the Lebesgue-Stieltjes measure, such that $\forall, a < b$,

$$\mu_{\alpha}((a, b]) = \alpha(b) - \alpha(a).$$

Definition 4.3.7 (Complete Measure). A measure μ on \mathcal{A} is said to be *complete* if for every $B \in \mathcal{A}$ with $\mu(B) = 0$, then $\forall A \subset B, A \in \mathcal{A}$.

Definition 4.3.8 (Lebesgue Measurable Function and Lebesgue Measure). Specialize to $\mu^*|_{\mathcal{A}}$ and $\mu^*|_{\mathcal{B}_{\mathbb{R}}}$:

$$E \in \mathcal{A} \iff \exists A, B \in \mathcal{B}_{\mathbb{R}} \text{ s.t. } A \subset E \subset B \quad \text{and} \quad \mu_{\alpha}(B \setminus A) = 0.$$

- *Lebesgue Measurable Function:* $f : \mathbb{R} \rightarrow \mathbb{R}$.

f is measurable as a function form $(\mathbb{R}, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$:

$$f^{-1}((c, \infty)) \in \mathcal{A}.$$

- *Lebesgue Measure m :* take $\alpha(x) = x$.

$$m((a, b]) = b - a.$$

Theorem 4.3.9

If $f : [a, b] \rightarrow \mathbb{R}$ is Borel and Riemann integrable, then

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

Proof 5. (Sketch).

lower Darboux sum = integral of piecewise constant function, which is simple.

So,

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

Similarly,

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

Therefore,

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

Q.E.D. ■

Notation 4.3.10 We will use

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

even if f is not Riemann integrable. With respect to improper integrals, for us,

$$\int f \, dm$$

is only defined if $\int |f| \, dm < \infty$.

4.4 The Monotone Class Theorem

Motivation μ, ν are measures on \mathbb{R} that are finite on compact sets, with

$$\mu((a, b]) = \nu((a, b]) \quad \forall a < b.$$

Is it true that $\mu(B) = \nu(B) \quad \forall$ Borel set B ?

Attempt Define

$$\mathcal{A} = \{B \mid \mu(B) = \nu(B)\}.$$

Prove that \mathcal{A} is a σ -algebra.

- We know it contains intervals $\implies \mathcal{B}_{\mathbb{R}} \subset \mathcal{A}$.

- However, if sets $A_n \in \mathcal{A}$ with $\mu(A_n) = \nu(A_n) \quad \forall n$, we don't know if

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \nu\left(\bigcup_{n=1}^{\infty} A_n\right).$$

Definition 4.4.1 (Algebra). An *algebra* on X is $\mathcal{G} \subset \mathcal{P}(X)$ with properties:

- $\emptyset \in \mathcal{G}$,
- $A \in \mathcal{G} \implies A^c \in \mathcal{G}$, and
- $A_1, A_2 \in \mathcal{G} \implies A_1 \cap A_2 \in \mathcal{G}$.

Example 4.4.2

- Any σ -algebra is an algebra.
- The following example is an algebra, but not a σ -algebra:

$$\mathcal{G} = \{A \subset \mathbb{R} \mid A \text{ is finite or } A^c \text{ is finite}\}.$$

Definition 4.4.3 (Monotone Class). A *monotone class* on X is a subset $\mathcal{C} \subset \mathcal{P}(X)$ with properties:

- $A_n \in \mathcal{C}, A_n \subset A_{n+1} \quad \forall n \in \mathbb{N} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{C}$.
- $B_n \in \mathcal{C}, B_n \supset B_{n+1} \quad \forall n \in \mathbb{N} \implies \bigcap_{n=1}^{\infty} B_n \in \mathcal{C}$.

Example 4.4.4

$$\mathcal{C} = \{\emptyset, \mathbb{R}\} \cup \{(n, \infty) \mid n \in \mathbb{Z}\}$$

is a monotone class but not a σ -algebra.

Definition 4.4.5 (Monotone Class Generated by \mathcal{E}). The *monotone class generated by $\mathcal{E} \subset \mathcal{P}(X)$* is the smallest monotone class containing \mathcal{E} . It is also the intersection of all monotone class containing \mathcal{E} .

Theorem 4.4.6 Monotone Class Theorem

If $\mathcal{G} \subset \mathcal{P}(X)$ is an algebra, then the monotone class generated by \mathcal{G} equals the σ -algebra generated by \mathcal{G} .

Proof 1. Denote by \mathcal{C} the monotone class generated by \mathcal{G} , and denote by \mathcal{A} the σ -algebra generated by \mathcal{G} . [WTS: $\mathcal{C} = \mathcal{A}$].

(\subseteq): \mathcal{A} is a monotone class, so $\mathcal{C} \subset \mathcal{A}$.

(\supseteq): [To show $\mathcal{A} \subset \mathcal{C}$, we want to show \mathcal{C} is a σ -algebra].

- \mathcal{C} is an algebra. [Goal: $\forall E, F \in \mathcal{C}, E \setminus F, F \setminus E, E \cap F \in \mathcal{C}$.]

Fix E . Define

$$\mathcal{C}_E = \{F \in \mathcal{C} \mid E \setminus F, F \setminus E, E \cap F \in \mathcal{C}\}.$$

Then, \mathcal{C}_E is a monotone class. [$F_n \subset F_{n+1}, F_n \in \mathcal{C}_E \implies E \setminus F_n \supset E \setminus F_{n+1} \implies E \setminus F_{n+1} \in \mathcal{C}$. Hence, $E \setminus F_n \in \mathcal{C} \implies E \setminus (\bigcup F_n) = \bigcap (E \setminus F_n) \in \mathcal{C} \implies \bigcup F_n \in \mathcal{C}_E$.]

We know: $E, F \in \mathcal{G} \implies E \setminus F, F \setminus E, E \cap F \in \mathcal{G} \subset \mathcal{C}$ because \mathcal{G} is an algebra.

1. If $E \in \mathcal{G}, F \in \mathcal{G}$, then $F \in \mathcal{C}_E$.
2. If $E \in \mathcal{G}$, then $\mathcal{G} \subset \mathcal{C}_E$.

\mathcal{C}_E is a monotone class containing \mathcal{G} , so it contains \mathcal{C} . So, $\mathcal{C} \subset \mathcal{C}_E$. Then,

$$E \in \mathcal{G}, F \in \mathcal{C} \implies E \setminus F, F \setminus E, E \cap F \in \mathcal{C}.$$

We could also have, by symmetry for E and F , that

$$\begin{aligned} E \in \mathcal{C}, F \in \mathcal{G} &\implies E \setminus F, F \setminus E, E \cap F \in \mathcal{C} \\ \forall E \in \mathcal{C}, F \in \mathcal{G} &\implies F \in \mathcal{C}_E, \mathcal{G} \subset \mathcal{C}_E. \end{aligned}$$

Since \mathcal{C}_E is a monotone class, and $\mathcal{C} \subset \mathcal{C}_E$, we have

$$E, F \in \mathcal{C} \implies E \setminus F, F \setminus E, E \cap F \in \mathcal{C}.$$

So, \mathcal{C} is an algebra:

1. $\emptyset \in \mathcal{G} \subset \mathcal{C}$
 2. $X \in \mathcal{G}$, so $A \in \mathcal{C} \implies X \setminus A \in \mathcal{C}$.
 3. $A_1, A_2 \in \mathcal{C} \implies A_1 \cap A_2 \in \mathcal{C}$.
- \mathcal{C} is a σ -algebra.

\mathcal{C} is an algebra and a monotone class. If $A_n \in \mathcal{C}$, denote $B_n = A_1 \cup \dots \cup A_n \in \mathcal{C}$. So,

$$\bigcup A_n = \bigcup B_n \in \mathcal{C} \quad \text{since } B_n \subset B_{n+1}.$$

Therefore, \mathcal{C} is a σ -algebra.

Q.E.D. ■

Theorem 4.4.7 Motivating Example

μ, ν on \mathbb{R} are Borel measures finite on compact with

$$\mu((a, b]) = \nu((a, b]) \quad \forall a < b.$$

Then, $\mu(B) = \nu(B) \quad \forall$ Borel sets B .

Proof2.

- Show $\mu(B) = \nu(B)$ for B in algebra generated by half open intervals. [Challenges are

1. μ, ν may be infinite.

2. We don't know if

$$\mu(A_1) = \nu(A_1) \quad \text{and} \quad \mu(A_2) = \nu(A_2) \implies \mu(A_1 \cap A_2) = \nu(A_1 \cap A_2).$$

]

Lemma 4.4.8 Denote

$$\mathcal{I} = \{(a, b] \mid a < b, a \text{ can be } -\infty\} \cup \{(a, \infty) \mid a \in \mathbb{R} \cup \{-\infty\}\}.$$

So, $\mu = \nu$ on \mathcal{I} . Also, define

$$\mathcal{G} = \{\emptyset\} \cup \left\{ \bigcup_{j=1}^n I_j \mid n \in \mathbb{N}, I_j \in \mathcal{I}, I_j \cap I_k = \emptyset, j \neq k \right\}.$$

Then, \mathcal{G} is the algebra generated by \mathcal{I} .

Proof. [WTS: \mathcal{G} is an algebra] $E \in \mathcal{G} \iff \chi_E$ is left continuous and is continuous except at finitely many points.

1. If $A \in \mathcal{G}$, then χ_A has the properties. Note that

$$\chi_{A^c} = 1 - \chi_A$$

has the same property. So, $A^c \in \mathcal{G}$.

2. If $A_1, A_2 \in \mathcal{G}$, then χ_{A_1} and χ_{A_2} have the properties. Then,

$$\chi_{A_1 \cap A_2} = \chi_{A_1} \chi_{A_2}$$

is left continuous and is continuous except at finitely many points.

So, \mathcal{G} is an algebra. \square

For $E \in \mathcal{G}$, if $E = \bigcup_{j=1}^n I_j$, with $I_j \cap I_k = \emptyset$ for $j \neq k$, then

$$\mu(E) = \mu\left(\bigcup_{j=1}^n I_j\right) = \sum_{j=1}^n \mu(I_j) = \sum_{j=1}^n \nu(I_j) = \nu\left(\bigcup_{j=1}^n I_j\right) = \nu(E).$$

[This addresses challenge #2. Also, since only addition was involved, we don't need to worry about #1.]

- Monotone Class Theorem to show \forall Borel sets B . [

Attempt $\{E \in \mathcal{B}_{\mathbb{R}} \mid \mu(E) = \nu(E)\}$ is a monotone class.

Issue If $E_n \supset E_{n+1}$, then

$$\mu\left(\bigcap E_n\right) = \lim_{n \rightarrow \infty} \mu(E_n) = \lim_{n \rightarrow \infty} \nu(E_n) = \nu\left(\bigcap E_n\right)$$

only if $\mu(E_k) < \infty$ for some k .

]

Instead Define

$$\mathcal{C} = \{E \in \mathcal{B}_{\mathbb{R}} \mid \forall k \in \mathbb{N}, \mu(E \cap (-k, k]) = \nu(E \cap (-k, k])\}$$

[since μ, ν is finite on compact, so $E \cap (-k, k] \subset (-k, k]$, and $\mu((-k, k]) < \infty$.]

[WTS: \mathcal{C} is a monotone class.] If $A_n \subset A_{n+1}$, $A_n \cap (-k, k] \subset A_{n+1} \cap (-k, k]$. Since

$$\mu(A_n \cap (-k, k]) = \nu(A_n \cap (-k, k]),$$

by taking $n \rightarrow \infty$, we have

$$\mu\left(\bigcup A_n \cap (-k, k]\right) = \nu\left(\bigcup A_n \cap (-k, k]\right).$$

For $A_n \supset A_{n+1}$, these arguments work similarly.

For $E \in \mathcal{G}$, $E \cap (-k, k] \in \mathcal{G}$. So, $\mu(E \cap (-k, k]) = \nu(E \cap (-k, k])$. Then, $\mathcal{G} \subset \mathcal{C}$, and \mathcal{C} contains monotone class generated by $\mathcal{G} = \mathcal{B}_{\mathbb{R}}$ (by Monotone Class Theorem). Then, since

$$\mu(E \cap (-k, k]) = \nu(E \cap (-k, k]) \quad \forall k \in \mathbb{N}$$

by taking $k \rightarrow \infty$, $\mu(E) = \nu(E)$.

Q.E.D. \blacksquare

5 Product Measures

5.1 Product Measures

Goal We want to construct a measure κ such that

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

Definition 5.1.1 (Product σ -Algebra). If \mathcal{A} is a σ -algebra on X , \mathcal{B} is a σ -algebra on Y . The *product σ -algebra* on $X \times Y$ is the σ -algebra generated by “rectangles” $A \times B$, where $A \in \mathcal{A}$ and $B \in \mathcal{B}$. The product σ -algebra will be denoted $\mathcal{A} \otimes \mathcal{B}$.

We want to construct product measure by Carathéodory's Theorem. Let μ be a measure on \mathcal{A} , ν be a measure on \mathcal{B} . Take class of elementary sets

$$\mathcal{E} = \{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\},$$

with weight

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

We get outer measure

$$\kappa^*(S) = \inf \left\{ \sum_{j=1}^{\infty} \mu(A_j)\nu(B_j) \mid S \subset \bigcup_{j=1}^{\infty} A_j \times B_j \right\}.$$

Updated Goal This gives a measure on $\mathcal{A} \otimes \mathcal{B}$.

Lemma 5.1.2 If $A \in \mathcal{A}$, $B \in \mathcal{B}$, then

$$\kappa^*(A \times B) = \mu(A)\nu(B).$$

Proof 1. (\leq): $A \times B$ covered by $A \times B, \emptyset, \emptyset, \dots$. So,

$$\kappa^*(A \times B) \leq \rho(A \times B) = \mu(A)\nu(B).$$

(\geq): For the opposite inequality, let

$$\begin{aligned} A \times B &\subset \bigcup_{j=1}^{\infty} A_j \times B_j \\ \chi_{A \times B} &\leq \sum_{j=1}^{\infty} \chi_{A_j \times B_j} \end{aligned}$$

Let $x \in X$ and $y \in Y$, then

$$\chi_A(x)\chi_B(y) \leq \sum_{j=1}^{\infty} \chi_{A_j}(x)\chi_{B_j}(y).$$

Fix x , integrate in y :

$$\chi_A(x)\nu(B) \leq \sum_{j=1}^{\infty} \chi_{A_j}(x)\nu(B_j)$$

Integrate in x :

$$\begin{aligned} \mu(A)\nu(B) &\leq \sum_{j=1}^{\infty} \mu(A_j)\nu(B_j) \\ \mu(A)\nu(B) &\leq \kappa^*(A \times B). \end{aligned}$$

Q.E.D. ■

Lemma 5.1.3 For any $S, R \in \mathcal{E}$,

$$\kappa^*(S) = \kappa^*(S \cap R^c) + \kappa^*(S \cap R).$$

Proof 2. By sub-additivity of outer measure,

$$\begin{aligned} \kappa^*(S) &\leq \kappa^*(S \cap R^c) + \kappa^*(S \cap R) \\ &\leq \kappa^*(S \cap (A \times B^c)) + \kappa^*(S \cap (A^c \times B^c)) + \kappa^*(S \cap (A^c \times B)) + \kappa^*(S \cap (A \times B)). \end{aligned}$$

Defining $S = C \times D$, LHS = RHS = $\mu(C)\nu(D)$. So, we have an equality,

$$\kappa^*(S) = \kappa^*(S \cap R^c) + \kappa^*(S \cap R).$$

Q.E.D. ■

Lemma 5.1.4 Every rectangle $R \in \mathcal{E}$ is measurable with respect to κ^* .

Proof 3. (\geq): Let $S \subset X \times Y$ be an arbitrary set. Then, $S \subset \bigcup A_j \times B_j$. For any rectangle R , applying Lemma 5.3 to $A_j \times B_j$ and R ,

$$\begin{aligned} \sum \kappa^*(A_j \times B_j) &= \sum \kappa^*((A_j \times B_j) \cap R^c) + \sum \kappa^*((A_j \times B_j) \cap R) \\ &\geq \kappa^*(S \cap R^c) + \kappa^*(S \cap R) \end{aligned} \quad [\sigma\text{-additivity}]$$

Take inf over countable covers,

$$\kappa^*(S) \geq \kappa^*(S \cap R^c) + \kappa^*(S \cap R).$$

(\leq): By sub-additivity,

$$\kappa^*(S) \leq \kappa^*(S \cap R^c) + \kappa^*(S \cap R).$$

Q.E.D. ■

By Carathéodory's Theorem, restricted of κ^* to $\mathcal{A} \otimes \mathcal{B}$ is a measure, which we call the product measure of μ, ν and denote it $\mu \otimes \nu$.

Remark.

$$\mathcal{M} = \{S \subset X \times Y \mid S \text{ is measurable w.r.t. } \kappa^*\}.$$

By Carathéodory, \mathcal{M} is a σ -algebra and contains all rectangles, so $\mathcal{A} \otimes \mathcal{B} \subset \mathcal{M}$. Further, $\kappa^*|_{\mathcal{M}}$ is a measure $\implies \kappa^*|_{\mathcal{A} \otimes \mathcal{B}}$ is a measure called the product measure.

5.2 Tonelli's and Fubini's Theorem

Definition 5.2.1 (σ -Finite). μ is σ -finite if $\exists K_n \subset K_{n+1}$ s.t. $\mu(K_n) < \infty$ and $\bigcup_{n=1}^{\infty} K_n = X$.

Theorem 5.2.2 Tonelli

If μ, ν are σ -finite, then for every $f : X \times Y \rightarrow [0, \infty]$ measurable:

- For every $y \in Y$, $x \mapsto f(x, y)$ is measurable.
- The function $y \mapsto \int f(x, y) d\mu(x)$ is measurable.
- For every $x \in X$, $y \mapsto f(x, y)$ is measurable.
- The function $x \mapsto \int f(x, y) d\nu(y)$ is measurable.
- Iterated integrals $\perp\!\!\!\perp$ the order of integration.

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

Proof 1. Let $\mathcal{M} = \{f : X \times Y \rightarrow [0, \infty] \mid \textcircled{1}\text{-}\textcircled{5} \text{ hold}\}$. Note that

- $f, g \in \mathcal{M} \implies f + g \in \mathcal{M}$
- $f \in \mathcal{M}, c \in [0, \infty) \implies cf \in \mathcal{M}$
- If $f_n \in \mathcal{M}$, $f_n \leq f_{n+1}$, and $f_n \rightarrow f$ pointwise, then $f \in \mathcal{M}$.

Goal If $S \in \mathcal{A} \otimes \mathcal{B}$, then $\chi_S \in \mathcal{M}$. [For this, we will use Monotone Class Theorem.]

If $E = A \times B$, then $\chi_E \in \mathcal{M}$:

$$\chi_E(x, y) = \chi_A(x)\chi_B(y).$$

Note that fix x , $\chi_B(y)$ is measurable, and fix y , $\chi_A(x)$ is measurable. So,

$$\iint \chi_A(x)\chi_B(y) d\mu(x) d\nu(y) = \int \mu(A)\chi_B(y) d\nu(y) = \mu(A)\nu(B).$$

Likewise, for

$$\iint \chi_A(x)\chi_B(y) d\nu(y)d\mu(x).$$

So,

$$\int \chi_E(x, y) d(\mu \otimes \nu) = (\mu \otimes \nu)(E) = \mu(A)\nu(B).$$

Define an algebra (one can check this easily)

$$\mathcal{G} = \{\emptyset\} \cup \left\{ \bigcup_{j=1}^n (A_j \times B_j) \mid A_j \times B_j \text{ are pairwise disjoint rectangles} \right\}.$$

For $S \in \mathcal{G}$, $\chi_S \in \mathcal{M}$. If $S = A \times B$, then

$$\iint \chi_S(x, y) d\mu(x)d\nu(y) = \int \mu(A)\chi_B(y) d\nu(y) = \mu(A)\nu(B).$$

Since μ, ν are σ -finite, so,

$$\begin{aligned} \exists K_n, \quad \mu(K_n) < \infty, \quad X &= \bigcup_{n=1}^{\infty} K_n \\ \exists L_n, \quad \nu(L_n) < \infty, \quad Y &= \bigcup_{n=1}^{\infty} L_n. \end{aligned}$$

Assume, WLOG, $K_n \subset K_{n+1}$ and $L_n \subset L_{n+1}$. Define

$$C_n = \{S \in \mathcal{A} \otimes \mathcal{B} \mid \chi_S \chi_{K_n \times L_n} \in \mathcal{M}\}.$$

So,

$$\begin{aligned} S_j \subset S_{j+1}, \dots, \rightarrow S, \quad S_j \in C_n &\implies S \in C_n \\ S_j \supset S_{j+1}, \dots, \rightarrow S, \quad S_j \in C_n &\implies S \in C_n \quad [\text{use } K_n \times L_n \setminus S_j] \end{aligned}$$

Also, note that \mathcal{G} is an algebra of a set. Since

$$S \in \mathcal{G} \implies S \cap (K_n \times L_n) \in \mathcal{G} \implies S \in C_n.$$

So, C_n is a monotone class containing \mathcal{G} . So, by Monotone Class Theorem,

$$\mathcal{A} \otimes \mathcal{B} \subset C_n \implies \mathcal{A} \otimes \mathcal{B} = C_n.$$

We know for each $S \in \mathcal{A} \otimes \mathcal{B}$, $\chi_S \chi_{K_n \times L_n} \in \mathcal{M}$. When $n \rightarrow \infty$, $\chi_S \in \mathcal{M}$.

Q.E.D. ■

Corollary 5.2.3 If $\int f d(\mu \otimes \nu) < \infty$, then

$$\int f(x, y) d\mu(x) < \infty \quad \text{for } \nu\text{-a.e. } y \quad \text{and} \quad \int f(x, y) d\nu(y) < \infty \quad \text{for } \mu\text{-a.e. } x.$$

Theorem 5.2.4 Fubini's Theorem

Suppose $f : X \times Y \rightarrow \mathbb{C}$. If $\int |f| d(\mu \otimes \nu) < \infty$, then

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

Proof 2. If $\int |f| d(\mu \otimes \nu)$ implies

$$\iint |f(x, y)| d\mu(x) < \infty \quad \text{for } \nu\text{-a.e. } y \quad \text{and} \quad \iint |f(x, y)| d\nu(y) < \infty \quad \text{for } \mu\text{-a.e. } x.$$

Key to the proof:

$$f = (\operatorname{Re} f)_+ - (\operatorname{Re} f)_- + i[(\operatorname{Im} f)_+ - (\operatorname{Im} f)_-].$$

Q.E.D. ■

Corollary 5.2.5 If $g_n : X \rightarrow \mathbb{C}$ is a sequence of measurable functions such that

$$\sum_n \int |g_n| d\mu < \infty, \quad \text{then} \quad \sum_n \int g_n d\mu = \int \sum_n g_n d\mu.$$

Proof 3. Method I Fubini's Theorem.

This is a special case of Fubini for $\mu \otimes \nu$, where ν is the counting measuring on \mathbb{N} . Apply Fubini to $G : X \times \mathbb{N} \rightarrow \mathbb{C}$ defined by

$$G(x, n) = g_n(x).$$

Method II Dominated Convergence Theorem.

$$f_n = g_1 + \cdots + g_n \quad [\textit{partial sum}]$$

Then, $|f_n| \leq |g_1| + \cdots + |g_n| \leq h$. Dominating function:

$$h = \sum_{n=1}^{\infty} |g_n|.$$

Q.E.D. ■

Example 5.2.6

$f : X \rightarrow [0, \infty]$. Define $g(y) = \mu(\{x \mid f(x) > y\})$. Then,

$$\int f d\mu = \int_0^{\infty} g(y) dy.$$

Proof 4. Take $h(x, y) = f(x) - y$. Measurable as $h : X \times (0, \infty) \rightarrow \mathbb{R} \cup \{\infty\}$. Then, $E = h^{-1}((0, \infty))$

is measurable. By Tonelli's Theorem on χ_E :

$$\iint \chi_E(x, y) \, d\mu(x) dy = \iint \chi_E(x, y) \, dy d\mu(x).$$

- RHS: $\int \chi_E(x, y) \, dy = f(x)$
- LHS: Fix y , $(x, y) \in E \iff f(x) > y \iff x \in \{x \mid f(x) > y\}$. So,

$$\int \chi_E(x, y) \, d\mu(x) = g(y).$$

So,

$$\int g(y) \, dy = \int f(x) \, d\mu(x).$$

Q.E.D. ■

5.3 Lebesgue Measure on \mathbb{R}^n

Theorem 5.3.1

If X, Y are second countable, then $\mathcal{B}_X \otimes \mathcal{B}_Y = \mathcal{B}_{X \times Y}$

Proof 1.

Lemma 5.3.2 If X, Y are metric spaces, then $X \times Y$ is a metric space with

$$d((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2).$$

The rest of the proof is omitted.

Q.E.D. ■

Corollary 5.3.3 $\mathcal{B}_{\mathbb{R}^k} \otimes \mathcal{B}_{\mathbb{R}^\ell} = \mathcal{B}_{\mathbb{R}^{k+\ell}}$.

Definition 5.3.4 (n -Dimensional Lebesgue Measure). Let $m : \mathcal{B}_{\mathbb{R}} \rightarrow [0, \infty]$ Lebesgue measure on \mathbb{R} . Define, inductively,

$$m_1 = m, \quad \text{and} \quad m_n = m_{n-1} \otimes m$$

to be the n -dimensional Lebesgue measure. So,

$$m_n : \mathcal{B}_{\mathbb{R}^n} \rightarrow [0, \infty].$$

Proposition 5.3.5 (Translation Invariance). m_1 is translation-invariant because $m_1(A + t) = m_1(A)$, $A \in \mathcal{B}_{\mathbb{R}}, t \in \mathbb{R}$. Hence, m_n is also translation-invariant:

$$m_n(A + \mathbf{v}) = m_n(A), \quad A \in \mathcal{B}_{\mathbb{R}^n}, \mathbf{v} \in \mathbb{R}^n.$$

Proposition 5.3.6 (Uniqueness of Lebesgue Measure). Let μ be a Borel measure on \mathbb{R}^n . If μ is translation-

invariant and

$$\mu((0, 1]^n) = c < \infty,$$

then $\mu = cm_n$.

Proof2. By translation invariance,

$$\begin{aligned} \mu\left(\left(0, \frac{1}{q_1}\right] \times \left(0, \frac{1}{q_2}\right] \times \cdots \times \left(0, \frac{1}{q_n}\right]\right) &= \frac{c}{q_1 q_2 \cdots q_n} \\ \mu\left(\left(0, \frac{p_1}{q_1}\right] \times \cdots \times \left(0, \frac{p_n}{q_n}\right]\right) &= c \frac{p_1}{q_1} \frac{p_2}{q_2} \cdots \frac{p_n}{q_n} \\ \implies \mu((0, b_1] \times \cdots \times (0, b_n]) &= c b_1 b_2 \cdots b_n \quad \text{for } b_1, b_2, \dots, b_n > 0 \\ \implies \mu((a_1, b_1] \times \cdots \times (a_n, b_n]) &= c \prod_{j=1}^n (b_j - a_j) \\ &= cm_n((a_1, b_1] \times \cdots \times (a_n, b_n]). \end{aligned}$$

So, $\mu(B) = cm_n(B)$ for rectangle B . Define

$$C_k = \{B \mid \mu(B \cap (-k, k]^n) = cm_n(B \cap (-k, k]^n)\}.$$

Note that C_k is a monotone class. By Monotone Class Theorem,

$$\mu(B \cap (-k, k]^n) = cm_n(B \cap (-k, k]^n) \quad \forall \text{ Borel } B.$$

When $k \rightarrow \infty$, $\mu(B) = cm_n(B)$.

Q.E.D. ■

Theorem 5.3.7 Rotation Invariance

Let S be an invertible $n \times n$ real matrix. Define

$$SB = \{S\mathbf{x} \mid \mathbf{x} \in B\}.$$

Then, $m_n(SB) = |\det S| m_n(B)$. In particular, if S is a rotation matrix ($\det S = 1$), then

$$m_n(SB) = m_n(B).$$

This property is called *rotation invariance*.

Proof3. Fix S . Define $\mu_S(B) = m_n(SB)$. μ_S is translation-invariant because

$$\mu_S(\mathbf{v} + B) = m_n(S\mathbf{v} + SB) = m_n(SB) = \mu_S(B).$$

Now, consider

$$\mu_S((0, 1]^n) = m_n(\underbrace{S(0, 1]^n}_{\text{bounded set}}) = c(S) < \infty.$$

$$\mu_S = c(S)m_n, \quad \text{and} \quad m_n(SB) = c(S)m_n(B).$$

It remains to prove $c(S) = |\det S|$. Observe that $c(S_1 S_2) = c(S_1)c(S_2)$, and $|\det(S_1 S_2)| = |\det S_1| \cdot |\det S_2|$. It is sufficient to prove for elementary matrices (appeared in Gaussian Elimination).

- $S =$ permutation matrix. $\exists \sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ direction. $S_{j\sigma(j)} = 1$, and all other entries are 0. Permutation does not change the Lebesgue measure.
- $S =$ diagonal matrix.

$$S = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \implies \begin{array}{l} \text{Lebesgue measure is charged by} \\ |\lambda_1 \lambda_2 \cdots \lambda_n| = |\det(S)| \end{array}.$$

•

$$S = \left[\begin{array}{cc|c} 1 & 1 & 0 \\ & 1 & \\ \hline & & 1 \\ 0 & & \ddots \\ & & & 1 \end{array} \right]$$

By translation invariance, Lebesgue measure is unaffected.

We can multiply elementary matrices to S , to get an identity eventually. Hence, we have shown

$$m_n(SB) = |\det S| m_n(B).$$

Q.E.D. ■

Example 5.3.8

Definition 5.3.9 (Gamma Function).

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt, \quad \operatorname{Re} z > 0$$

Note that $\int_0^{\infty} |t^{z-1} e^{-t}| dt < \infty$. So, Gamma function is well-defined.

Definition 5.3.10 (Beta Function).

$$B(z, w) = \int_0^1 t^{z-1} (1-t)^{w-1} dt, \quad \operatorname{Re} z, \operatorname{Re} w > 0.$$

Lemma 5.3.11 (Relation Between Gamma and Beta Functions).

$$B(z, w)\Gamma(z + w) = \Gamma(z)\Gamma(w)$$

Proof 4.

$$\begin{aligned} \text{RHS} = \Gamma(z)\Gamma(w) &= \int_0^\infty x^{z-1}e^{-x} dx \int_0^\infty y^{w-1}e^{-y} dy \\ &= \int_0^\infty \int_0^\infty x^{z-1}e^{-x}y^{w-1}e^{-y} dy \quad [\text{Fubini}] \\ &= \int_{(0,\infty)^2} x^{z-1}e^{-x}y^{w-1}e^{-y} d(x, y) \quad [\text{Integral over product measure}] \end{aligned}$$

Change of variable: $s = x - y$ and $t = x + y$. Then,

$$\begin{pmatrix} s \\ t \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

where

$$\det \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = 2 \implies d(s, t) = 2d(x, y).$$

So,

$$\text{RHS} = \int_A \left(\frac{s+t}{2}\right)^{z-1} \left(\frac{t-s}{2}\right)^{w-1} e^{-t} \frac{1}{2} d(s, t),$$

where $A = \{(s, t) \mid |s| \leq t\}$. That is, by Fubini,

$$\text{RHS} = \int_0^\infty \int_{-t}^t \left(\frac{s+t}{2}\right)^{z-1} \left(\frac{t-s}{2}\right)^{w-1} e^{-t} \frac{1}{2} ds dt.$$

Another change of variable: $s = -t + 2tu$. So, $-t \leq s \leq t \implies 0 \leq u \leq 1$, and $ds = 2du$. Then,

$$\begin{aligned} \text{RHS} &= \int_0^\infty \int_0^1 u^{z-1}(1-u)^{w-1} t^{z+w-1} e^{-t} du dt \\ &= \int_0^1 u^{z-1}(1-u)^{w-1} du \int_0^\infty t^{z+w-1} e^{-t} dt \\ &= B(z, w)\Gamma(z + w) \\ &= \text{LHS}. \end{aligned}$$

Q.E.D. ■

6 L^p Spaces and Fourier Transforms

6.1 Norms, Banach, and L^p Spaces

Definition 6.1.1 (Seminorm and Norm). Let V be a vector space over \mathbb{C} . $\|\cdot\| : V \rightarrow [0, \infty)$ is a norm if

- $\|\lambda x\| = |\lambda| \cdot \|x\|$
- $\|x + y\| \leq \|x\| + \|y\|$
- $x \neq 0 \implies \|x\| \neq 0$.

Note that ① and ② defines a seminorm.

Definition 6.1.2 (Norm-Induced Metric). A norm induces a metric

$$d(x, y) = \|x - y\|.$$

Example 6.1.3 Space with Norm

- \mathbb{R}^n
- $C(K) = \{f : K \rightarrow \mathbb{C} \mid f \text{ continuous}\}$, where K is compact.

1. Norm:

$$\|f\|_K = \max_{x \in K} |f(x)|$$

2. $C(K)$ is a complete metric space.

3. Convergence: $\|f_n - f\|_K \rightarrow 0$ is uniform convergence.

Definition 6.1.4 (Banach Space). A Banach space is a vector space with a norm, which is complete with respect to the metric induced by the norm.

Definition 6.1.5 (L^p Norm (informal)). X is a space, \mathcal{A} is a σ -algebra on X , and μ is a measure on \mathcal{A} .

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

- For $p \in [1, \infty)$,

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

- For $p = \infty$,

$$\|f\|_\infty = \inf \{t \in [0, \infty] \mid |f| \leq t, \mu\text{-a.e.}\}$$

Disproof 1. Is $\|f\|_1$ a norm on $\mathcal{L}^1(X, d\mu)$? Recall

$$\mathcal{L}^1(X, d\mu) = \left\{ f : X \rightarrow \mathbb{C} \mid \int |f| d\mu < \infty \right\}$$

is a vector space over \mathbb{C} .

- $\int |\lambda f| d\mu = |\lambda| \int |f| d\mu \implies \|\lambda f\|_1 = |\lambda| \cdot \|f\|_1.$
- $\int |f + g| d\mu \leq \int (|f| + |g|) d\mu = \int |f| d\mu + \int |g| d\mu \implies \|f + g\|_1 \leq \|f\|_1 + \|g\|_1.$
- However,

$$\int |f| d\mu = 0 \iff f = 0 \quad \mu\text{-a.e.}$$

That is, $\|f\|_1$ is just a seminorm on $\mathcal{L}^1(X, d\mu)$. ■

How to make it a norm? Define an *equivalence relation*:

$$f \sim g \iff f = g \quad \mu\text{-a.e.}$$

An equivalence class is denoted $[f]$. Define

$$L^1(X, d\mu) = \{[f] \mid f \in \mathcal{L}^1(X, d\mu)\} = \underbrace{\mathcal{L}^1(X, d\mu)}_{\text{Quotient space}} / \sim$$

Now, $\|f\|_1$ is a norm on $L^1(X, d\mu)$:

$$\|[f]\|_1 = \|f\|_1.$$

Proof 2. This norm is well-defined.

- $\|[\lambda f]\|_1 = |\lambda| \cdot \|f\|_1 = |\lambda| \cdot \|[f]\|_1.$
- $\|[f + g]\|_1 \leq \|[f]\|_1 + \|[g]\|_1.$
-

$$\begin{aligned} \|[f]\|_1 = 0 &\iff f = 0 \quad \mu\text{-a.e.} \\ &\iff f \sim 0 \\ &\iff [f] = [0] \end{aligned}$$

So, $\|\cdot\|_1$ is a norm on $L^1(X, d\mu)$. Q.E.D. ■

Lemma 6.1.6 (Young's Inequality). Let $p, q \in (1, \infty)$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$. Then, $\forall x, y \in [0, \infty)$,

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

Proof 3. Use log and exponentials, we have

$$xy = e^{\log x + \log y} \quad \text{and} \quad \frac{x^p}{p} + \frac{y^q}{q} = \frac{1}{p}e^{p \log x} + \frac{1}{q}e^{q \log y}.$$

Denote $u = p \log x, v = q \log y, t = \frac{1}{p}$, and $1 - t = \frac{1}{q}$. Then,

$$xy = e^{\log x + \log y} = e^{tu + (1-t)v} \quad \text{and} \quad \frac{x^p}{p} + \frac{y^q}{q} = te^u + (1-t)e^v.$$

Since exponential is a convex function,

$$e^{tu + (1-t)v} \leq te^u + (1-t)e^v.$$

Q.E.D. ■

Lemma 6.1.7 (Hölder's Inequality). If $\frac{1}{p} + \frac{1}{q} = 1$, and

$$\int |f|^p d\mu < \infty \quad \text{and} \quad \int |g|^q d\mu < \infty,$$

then $\bar{g}f \in \mathcal{L}^1(X, d\mu)$ and

$$\left| \int \bar{g}f d\mu \right| \leq \|f\|_p \|g\|_q.$$

Remark 4. (Special Cases).

- Young's Inequality: $p = q = 2$.

$$xy \leq \frac{x^2}{2} + \frac{y^2}{2}$$

- Hölder's Inequality: $p = q = 2$ gives Cauchy-Schwarz Inequality.

$$\left| \int \bar{g}f d\mu \right| \leq \|f\|_2 \|g\|_2, \quad \text{and} \quad \left| \sum_{j=1}^n x_j y_j \right| \leq \sqrt{\sum_{j=1}^n |x_j|^2} \sqrt{\sum_{j=1}^n |y_j|^2}.$$

Proof 5.

- If $\|f\|_p = 0$, then $0 \leq 0$
- If $\|g\|_q = 0$, then $0 \leq 0$.
- Otherwise, WLOG, assume $\|f\|_p = 1$ and $\|g\|_q = 1$. [For the general case, if $0 \leq \|f\|_p < \infty$, let

$C = \frac{1}{\|f\|_p}$. Let $F = Cf$. Let $D = \frac{1}{\|g\|_q}$, $G = Dg$. Then, $\|F\|_p = \|G\|_q = 1$. Then,

$$\left| \int \overline{G}F \, d\mu \right| \leq \|F\|_p \|G\|_q.$$

Multiply by $\frac{1}{CD}$ to finish the proof. By Young's inequality, we have

$$\begin{aligned} |\overline{g}(x)f(x)| &= |g(x)||f(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q} \\ \int |\overline{g}f| \, d\mu &\leq \frac{1}{p} \int |f|^p \, d\mu + \frac{1}{q} \int |g|^q \, d\mu = \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p \|g\|_q \end{aligned}$$

Thus, $\overline{g}f \in \mathcal{L}^1$. Further,

$$\left| \int \overline{g}f \, d\mu \right| \leq \int |\overline{g}f| \, d\mu \leq \|f\|_p \|g\|_q.$$

Q.E.D. ■

Corollary 6.1.8 For fixed $f \in \mathcal{L}^p(X, d\mu)$. If we pick

$$g = \begin{cases} |f|^{p-2}f, & f \text{ non-zero} \\ 0, & f \text{ zero.} \end{cases}$$

We have $g \in \mathcal{L}^q$, and

$$\|g\|_q = \left(\int |f|^{(p-1)q} \, d\mu \right)^{1/q} = \left(\int |f|^p \, d\mu \right)^{1/q} = \|f\|_p^{p/q}.$$

Then, we have equality in Hölder's Inequality:

$$\int \overline{g}f \, d\mu = \|f\|_p \|g\|_q.$$

Lemma 6.1.9 (Minkowski's Inequality). For $f_1, f_2 \in \mathcal{L}^p(X, d\mu)$, $p \in (1, \infty)$. Then,

$$\|f_1 + f_2\|_p \leq \|f_1\|_p + \|f_2\|_p.$$

Proof 6. For every $g \in \mathcal{L}^q$,

$$\begin{aligned} \left| \int \overline{g}(f_1 + f_2) \, d\mu \right| &= \left| \int \overline{g}f_1 \, d\mu + \int \overline{g}f_2 \, d\mu \right| \\ &\leq \|g\|_q \|f_1\|_p + \|g\|_q \|f_2\|_p. \end{aligned}$$

Pick the g that achieves equality in Hölder's Inequality (by Corollary 6.8). That is, $g = |f_1 + f_2|^{p-2}(f_1 + f_2)$. Then,

$$\left| \int \overline{g}(f_1 + f_2) \, d\mu \right| = \|g\|_q \|f_1 + f_2\|_p.$$

So,

$$\begin{aligned}\|g\|_q \|f_1 + f_2\|_p &\leq \|g\|_q \|f_1\|_p + \|g\|_q \|f_2\|_p \\ \|f_1 + f_2\|_p &\leq \|f_1\|_p + \|f_2\|_p.\end{aligned}$$

Q.E.D. ■

Proof 7. (Rigorous) Let $f_1, f_2 \in \mathcal{L}^p$. Pick $g = |f_1 + f_2|^{p-2}(f_1 + f_2)$. Then,

$$\begin{aligned}\|g\|_q &= \left(\int |g|^q d\mu \right)^{1/q} = \left(\int |f_1 + f_2|^{p-1} d\mu \right)^{1/q} = \|f_1 + f_2\|_p^{p/q} \\ \left| \int \bar{g} f_1 d\mu \right| &\leq \|g\|_q \|f_1\|_p \\ \left| \int \bar{g} f_2 d\mu \right| &\leq \|g\|_q \|f_2\|_p \\ \int |f_1 + f_2|^p d\mu &= \left| \int \bar{g}(f_1 + f_2) d\mu \right| \leq \|g\|_q (\|f_1\|_p + \|f_2\|_p) \\ \|f_1 + f_2\|_p^p &\leq \|f_1 + f_2\|_p^{p/q} (\|f_1\|_p + \|f_2\|_p) \\ \|f_1 + f_2\|_p &\leq \|f_1\|_p + \|f_2\|_p.\end{aligned}$$

Q.E.D. ■

Definition 6.1.10 (L^p Space). Define $\|f\|_p = 0 \iff f = 0 \quad \mu$ -a.e.. Then,

$$L^p(X, d\mu) = \{[f] \mid f \in \mathcal{L}^p(X, d\mu)\},$$

where the equivalence class is defined as $f \sim g$ if $f = g \quad \mu$ -a.e. $L^p(X, d\mu)$ is a normed vector space.

Theorem 6.1.11 Riesz-Fischer

For $p \in [1, \infty)$, $L^p(X, d\mu)$ is complete.

Proof 8. Pick a Cauchy sequence $f_n \in L^p$. [WTS: it is convergent. i.e., $\exists f \in L^p$ s.t. $\|f_n - f\|_p \rightarrow 0$ as $n \rightarrow \infty$.] Pick a subsequence such that

$$\|f_{n_{k+1}} - f_{n_k}\|_p \leq \frac{1}{4^k}.$$

Notationally, denote $f_{n_0} = 0$. By geometric sum,

$$\sum_{k=1}^{\infty} \|f_{n_k} - f_{n_{k-1}}\|_p < \infty.$$

Define

$$h(x) = \sum_{k=1}^{\infty} |f_{n_k}(x) - f_{n_{k-1}}(x)|.$$

This is an increasing limit of functions

$$h_m(x) = \sum_{k=1}^{\infty} |f_{n_k}(x) - f_{n_{k-1}}(x)|.$$

By Minkowski,

$$\begin{aligned} \|h_m\|_p &\leq \sum_{k=1}^m \|f_{n_k} - f_{n_{k-1}}\|_p \\ \left(\int |h_m|^p d\mu \right)^{1/p} &\leq \sum_{k=1}^m \|f_{n_k} - f_{n_{k-1}}\|_p. \end{aligned}$$

Taking $m \rightarrow \infty$, by monotone convergence,

$$\begin{aligned} \left(\int |h|^p d\mu \right)^{1/p} &\leq \sum_{k=1}^{\infty} \|f_{n_k} - f_{n_{k-1}}\|_p \\ \|h\|_p &\leq \sum_{k=1}^{\infty} \|f_{n_k} - f_{n_{k-1}}\|_p < \infty. \end{aligned}$$

So, $h < \infty$ μ -a.e.

For any x s.t. $h(x) < \infty$,

$$\begin{aligned} \sum_{k=1}^{\infty} |f_{n_k}(x) - f_{n_{k-1}}(x)| &< \infty \\ \implies (f_{n_k}(x))_{k=1}^{\infty} &\text{ is Cauchy in } \mathbb{C} \\ \implies (f_{n_k}(x))_{k=1}^{\infty} &\text{ is convergent.} \end{aligned}$$

Define $f(x) = \lim f_{n_k}(x)$, the pointwise limit. So, $f_{n_k}(x) \rightarrow f(x)$. Then,

$$\begin{aligned} f_{n_m}(x) &= \sum_{k=1}^m (f_{n_k}(x) - f_{n_{k-1}}(x)) \\ |f_{n_m}(x)| &\leq \sum_{k=1}^m |f_{n_k}(x) - f_{n_{k-1}}(x)| \leq h(x) \end{aligned}$$

Then, taking $m \rightarrow \infty$, $|f(x)| \leq h(x)$. Note that

$$\begin{aligned} |f_{n_k} - f| &\leq |f_{n_k}| + |f| \leq 2h \\ |f_{n_k} - f|^p &\leq 2^p h^p. \end{aligned}$$

By Dominated Convergence Theorem, with dominating function $2^p h^p$, we have

$$\lim_{k \rightarrow \infty} \int |f_{n_k} - f|^p d\mu = \int \lim_{k \rightarrow \infty} |f_{n_k} - f|^p d\mu = \int 0 d\mu = 0.$$

So, $\|f_{n_k} - f\|_p \rightarrow 0$ as $k \rightarrow \infty$.

Given a Cauchy sequence, showing its subsequence converges implies that the original sequence converges. Q.E.D. ■

Theorem 6.1.12

Let $L^p(X, d\mu)$, where $p \in [1, \infty)$, then simple functions s such that

$$\mu(\{x \mid s(x) \neq 0\}) < \infty$$

are dense in $L^p(X, d\mu)$.

Proof 9. Let $s = \chi_A$.

$$\begin{aligned} \|s\|_p &= (\mu(A))^{1/p} < \infty \implies \mu(A) < \infty \\ s &= \sum_{j=1}^n c_j \chi_{A_j} \in L^p \iff \mu(A_j) < \infty \quad \forall j \text{ with } c_j \neq 0. \end{aligned}$$

[*M is dense in metric space X if $\overline{M} = X \iff \forall x \in X, \exists y_n \in M$ s.t. $\lim_{n \rightarrow \infty} y_n = x$.*]

First assume $f \in L^p, f \geq 0$. Since $f \geq 0, \exists s_n$ simple such that $0 \leq s_n \leq s_{n+1}$ with $s_n \rightarrow f$ pointwise, and by Dominated Convergence Theorem,

$$\lim_{n \rightarrow \infty} \int |s_n - f|^p d\mu = \int \lim_{n \rightarrow \infty} |s_n - f|^p d\mu = \int 0 d\mu = 0.$$

We can apply DCT because $|s_n - f| \leq f \implies |s_n - f|^p \leq f^p$. Use f^p as the dominating function.

Now, assume $f \in L^p$ and f is real-valued. Then, $f = f_+ - f_-$, where $f_{\pm} = \max\{\pm f(x), 0\} \geq 0 \implies f_{\pm}^p \leq |f|^p$. So,

$$\int f_{\pm}^p d\mu < \infty.$$

Hence, $f_{\pm} \in L^p$ are positive functions. Take simple functions $s_{n,+} \rightarrow f_+$ and $s_{n,-} \rightarrow f_-$. Then,

$$s_{n,+} - s_{n,-} \rightarrow f_+ - f_- \quad \text{in } L^p.$$

[*Alternative reasoning: $\|s_{n,+} - f_+\|_p \rightarrow 0$ and $\|s_{n,-} - f_-\|_p \rightarrow 0$ as $n \rightarrow \infty$. So, by triangle inequality,*

$$\begin{aligned} \|(s_{n,+} - s_{n,-}) - (f_+ - f_-)\|_p &= \|(s_{n,+} - f_+) - (s_{n,-} - f_-)\|_p \\ &\leq \|s_{n,+} - f_+\|_p + \|s_{n,-} - f_-\|_p \leq 0. \end{aligned}$$

]

Finally, for $f \in L^p$ arbitrary, $f = \operatorname{Re} f + i \operatorname{Im} f$. By linear combination, we complete this proof.

Q.E.D. ■

6.2 Approximation of Sets and Functions

Theorem 6.2.1

Let μ be a finite Borel measure on metric space X . \forall Borel set E and $\forall \varepsilon > 0$, \exists closed F and \exists open U such that

$$F \subset E \subset U \quad \text{and} \quad \mu(U \setminus F) < \varepsilon.$$

Proof 1.

$$\mathcal{A} = \{E \text{ Borel} \mid \forall \varepsilon > 0, \exists U \text{ open } \exists F \text{ closed, } F \subset E \subset U, \mu(U \setminus F) < \varepsilon\}$$

[WTS: \mathcal{A} is a σ -algebra and contains all closed sets.]

• \mathcal{A} is a σ -algebra.

1. $E = \emptyset, F = U = \emptyset$
2. $E \in \mathcal{A} \implies E^c \in \mathcal{A}$ because

$$F \subset E \subset U \implies U^c \subset E^c \subset F^c.$$

3. Assume sets $E_n \in \mathcal{A}, E = \bigcup_n E_n$. [WTS: $E \in \mathcal{A}$.]

Fix $\varepsilon > 0, \forall n, \exists F_n$ closed and U_n open such that $F_n \subset E_n \subset U_n$, and

$$\mu(U_n \setminus F_n) < \frac{\varepsilon}{2^n}.$$

Take $U = \bigcup U_n$ and $A = \bigcup F_n$. Then,

$$A \subset E \subset U \quad \text{and} \quad \mu(U \setminus A) < \sum_n \frac{\varepsilon}{2^n} = \varepsilon.$$

[The inequality because $x \in U \setminus A \implies x \in U, x \notin A \implies \exists n$ s.t. $x \in U_n$ and $\forall n, x \notin F_n$. So, $\exists n$ s.t. $x \in U_n \setminus F_n$. That is,

$$U \setminus A \subset \bigcup_{n=1}^{\infty} U_n \setminus F_n \implies \mu(U \setminus A) < \sum_n \mu(U_n \setminus F_n).$$

]

Hence, U is open, but A is not necessarily closed. [WTS: create a closed set.]

$$A_N = \bigcup_{n=1}^N f_n \quad \text{closed.}$$

$\mu(A_N) \rightarrow \mu(A)$ as $N \rightarrow \infty$. Then, $\mu(A \setminus A_N) \rightarrow 0$ as $N \rightarrow \infty$ since μ is a finite measure. Pick N so that $\mu(A \setminus A_N) < \varepsilon$. Denote $F = A_N$. Then,

$$F \subset A \subset E \subset U \quad \text{and} \quad \mu(U \setminus F) < \mu(U \setminus A) + \mu(A \setminus F) < \varepsilon + \varepsilon = 2\varepsilon.$$

So, $E \subset \mathcal{A}$.

So, \mathcal{A} is a σ -algebra.

- \mathcal{A} contains closed sets:

Assume E is closed. Take $F = E$. Define $U_n = \left\{ x \mid \text{dist}(x, E) < \frac{1}{n} \right\}$. Then, U_n are nested open sets: $U_{n+1} \subset U_n$.

$$\bigcap_{n=1}^{\infty} U_n = E \quad \text{because } E \text{ is closed.}$$

$$\mu(U_n) \rightarrow \mu(E) \implies \mu(U_n \setminus E) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

[Note that we can do subtraction since μ is finite.] Fix $\varepsilon > 0$. Pick n s.t. $\mu(U_n \setminus E) < \varepsilon$. Pick $F = E$ and $U = U_n$. Then, $F = E \subset U$ with $\mu(U \setminus F) < \varepsilon$.

Q.E.D. ■

Theorem 6.2.2

Let μ be finite Borel measure on a compact metric space X . Then, $C(X)$ are a dense subset of $L^p(X, d\mu)$ for $p \in [1, \infty)$.

Corollary 6.2.3 $C([a, b])$ are dense in $L^p([a, b], d\mu)$.

Proof 2. (of Theorem 6.2.2)

Observation $f \in C(X) \implies f$ is bounded $\implies |f| \leq C$. Then,

$$\begin{aligned} \int |f|^p d\mu &\leq C^p \mu(X) < \infty \\ \implies C(X) &\subset L^p(X, d\mu). \end{aligned}$$

The remaining needs to prove $C(X)$ is dense. Let $E \subset X$ be Borel. Fix $\varepsilon > 0$, F closed, and U open such that

$$F \subset E \subset U \quad \text{and} \quad \mu(U \setminus F) < \varepsilon.$$

Take $f(x) = (1 - \text{dist}(x, F)/\delta)_+$. The,

$$f \in C(x) \quad \text{and} \quad f : X \rightarrow [0, 1], \quad f|_F = 1.$$

Take $\delta = \text{dist}(F, U^c) > 0$. In other words, $x \in U^c \implies \text{dist}(x, F) \geq \delta$. Then, $f(x) = 0$. Hence, $f|_{U^c} = 0$. So,

$$\begin{aligned} \int |f - \chi_E|^p d\mu &\leq 1 \cdot \mu(U \setminus F) < \varepsilon \\ \|f - \chi_E\|_p &\leq \varepsilon^{1/p}. \end{aligned}$$

Now, take simple function

$$s = \sum_{j=1}^n c_j \chi_{E_j}.$$

Fix $\varepsilon > 0$. Find $f_j \in C(X)$ s.t. $\|f_j - \chi_{E_j}\|_p < \varepsilon$. Take $f = \sum_{j=1}^n c_j f_j$, then

$$\|s - f\|_p = \left\| \sum_{j=1}^n c_j (\chi_{E_j} - f_j) \right\|_p \leq \sum_{j=1}^n |c_j| \varepsilon.$$

Finally, $g \in L^p$. Fix $\varepsilon > 0$, $\exists s$ simple such that $\|s - g\|_p < \varepsilon$ and $\exists f \in C(X)$ such that $\|s - f\|_p < \varepsilon$. Then,

$$\|g - f\|_p \leq \|s - g\|_p + \|s - f\|_p < 2\varepsilon.$$

Q.E.D. ■

Theorem 6.2.4

Let μ be a Borel measure finite on compact sets. Then, $C(\mathbb{R}^d)$ is a dense subset of $L^p(\mathbb{R}^d, d\mu)$ for $p \in [1, \infty)$.

Proof 3. Note that $f \in C(\mathbb{R}^d)$ is not necessarily bounded and not necessarily in L^p . We will use $f \in C_c(\mathbb{R}^d)$, where

$$\text{supp } f = \overline{\{x \mid f(x) \neq 0\}}, \quad \text{and } C_c(x) = \{f \in C(X) \mid \text{supp } f \text{ compact}\}.$$

Define $K = \text{supp } f$, $C = \max_k |f|$. Then,

$$|f| \leq C \chi_K, \quad |f|^p \leq C^p \chi_K, \quad \int |f|^p d\mu \leq C^p \mu(K) < \infty.$$

Q.E.D. ■

Theorem 6.2.5

Let μ be a measure on \mathbb{R}^d which is finite on compacts. Then, $C_c(\mathbb{R}^d)$ is a dense subset of $L^p(\mathbb{R}^d, d\mu)$, $p \in [1, \infty)$.

Proof 4. Let $E \subset \mathbb{R}^d$ be a Borel set. Then, $\mu(E) < \infty$. Denote $K_n = [-n, n]^d$, then $E \setminus K_n$ is decreasing, and

$$\bigcap (E \setminus K_n) = \emptyset.$$

Since E has finite measure, $\mu(E \setminus K_n) \rightarrow 0$ when $n \rightarrow \infty$. Fix $\varepsilon > 0$. Pick n such that $\mu(E \setminus K_n) < \varepsilon$.

Let's work on K_n : $\exists F \subset E \cap K_n$ such that F is closed and

$$\mu(E \cap K_n \setminus F) < \varepsilon.$$

Define $f_k(x) = (1 - k \operatorname{dist}(x, F))_+$. Then, $f_k = 1$ on F . When $k \rightarrow \infty$, $f_k \rightarrow \chi_F$ pointwise. By Dominated Convergence Theorem,

$$\lim_{k \rightarrow \infty} \int |f_k - \chi_F|^p d\mu = \int \lim_{k \rightarrow \infty} |f_k - \chi_F|^p d\mu = 0.$$

To apply DCT, consider the dominating function $\chi_{K_{n+1}}$.

[We are not choosing 1 as our dominating function because

$$\int_{\mathbb{R}^d} 1 d\mu \quad \text{may be infinite (non-integrable).}$$

But,

$$\int_{\mathbb{R}^d} \chi_{K_{n+1}} d\mu = \mu(K_{n+1}) < \infty.$$

]

And $0 \leq f_k \leq 1$, and $f_k = 0$ on K_{n+1}^c . From DCT, we also know that $\exists k$ such that

$$\begin{aligned} \|f_k - \chi_F\|_p &< \varepsilon \\ \|\chi_F - \chi_{E \cap K_n}\|_p &< \varepsilon^{1/p} \\ \|\chi_E - \chi_{E \cap K_n}\|_p &< \varepsilon^{1/p} \\ \implies \|\chi_E - \chi_F\|_p &< \varepsilon + 2\varepsilon^{1/p}. \end{aligned}$$

Then, simple functions can be approximated. So, any function in L^p can be approximated. Q.E.D. ■

Definition 6.2.6 (Space \mathbb{T} /Torus). \mathbb{T} is the unit circle with length 2π . Note that \mathbb{T} is compact, and $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. So, addition is defined as $x + y \pmod{2\pi}$. Meanwhile, $\mathbb{T}^d = \mathbb{T} \times \mathbb{T} \times \dots \times \mathbb{T}$.

6.3 Continuity of Translation

Definition 6.3.1 (Translate of a Function). For a function $f : \mathbb{T}^d \rightarrow \mathbb{C}$ (for $f : \mathbb{R}^d \rightarrow \mathbb{C}$), its translate by y is the function

$$(\tau_y f)(x) = f(x - y).$$

Theorem 6.3.2

For any $f \in C_c(\mathbb{R}^d)$

$$\|f - \tau_u f\|_{\mathbb{R}^d} \rightarrow 0 \quad \text{as } y \rightarrow 0.$$

Proof 1. Let $\|y\| < 1$ and $\operatorname{supp} f \subset [-n, n]^d$.

Take $K = [-n-1, n+1]^d$. f is continuous on K , so it is uniformly continuous. That is, $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $x, y \in K$ and

$$\|x - y\| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Change the notation: $\forall \varepsilon > 0, \exists \delta > 0$ such that $\forall x, x - y \in K$, if $\|y\| < \delta$, then $|f(x) - f(x - y)| < \varepsilon$. Outside of K , $f(x) - f(x - y) = 0$.

So, $\forall \varepsilon > 0, \exists \delta > 0$ s.t. $\|y\| < \delta \implies \forall x \in \mathbb{R}^d, |f(x) - f(x - y)| < \varepsilon$. That is, $\|f - \tau_y f\|_{\mathbb{R}^d} \leq \varepsilon$.

Q.E.D. ■

Example 6.3.3 Why we Need Uniform Continuity?

$f(x) = x^2$ on \mathbb{R}^2 is not uniformly continuous.

$$(\tau_y f)(x) = f(x) - f(x - y) = 2xy - y^2.$$

$$\|f - \tau_y f\|_{\mathbb{R}^d} = \infty.$$

Extension 6.3.1 On $\mathbb{T}^d, \forall f \in C(\mathbb{T}^d), \|f - \tau_y f\|_{\mathbb{T}^d} \rightarrow 0$ as $y \rightarrow 0$.

Theorem 6.3.4

For any $f \in L^p(\mathbb{R}^d)$ (with Lebesgue measure), $p \in [1, \infty)$. Then,

$$\|f - \tau_y f\|_p \rightarrow 0 \quad \text{as } y \rightarrow 0.$$

Proof 2. [Key idea: use density.] Fix $\varepsilon > 0, \exists g \in C_c(\mathbb{R}^d)$ s.t. $\|f - g\|_p < \varepsilon \implies \|\tau_y f - \tau_y g\|_p < \varepsilon$. We have $\|g - \tau_y g\|_{\mathbb{R}^d} \rightarrow 0$ as $y \rightarrow 0$. If $\text{supp } g \subset [-n, n]^d$ and $\|y\| < 1$, then

$$g - \tau_y g \quad \text{is supported in } [-n - 1, n + 1]^d.$$

Then,

$$\int |g - \tau_y g|^p d\mu \leq m([-n - 1, n + 1]^d) \|g - \tau_y g\|_{\mathbb{R}^d}^p \rightarrow 0 \quad \text{as } y \rightarrow 0.$$

So, $\exists \delta > 0$ s.t. $\forall y$ with $\|y\| < \delta \implies \|g - \tau_y g\|_p < \varepsilon$. That is, $\exists \delta > 0$ such that $\forall y$ with $\|y\| < \delta$,

$$\|f - \tau_y f\|_p \leq \|f - g\|_p + \|\tau_y f - \tau_y g\|_p + \|g - \tau_y g\|_p < 3\varepsilon.$$

Q.E.D. ■

Extension 6.3.2 In $\mathbb{T}^d, \forall f \in L^p(\mathbb{T}^d), p \in [1, \infty)$,

$$\|f - \tau_y f\|_p \rightarrow 0 \quad \text{as } y \rightarrow 0.$$

[Note: implied measure:

$$d\mu = \frac{dm}{(2\pi)^d} \quad \text{and} \quad \mu(\mathbb{T}^d) = 1.$$

]

6.4 Fourier Series and the Fejér Kernel

Set-Up f : 2π -period function on \mathbb{R} .

Goal Approximate f by 2π -period functions. A natural attempt is to use trigonometric polynomials.

For example,

$$p(x) = \sum_{k=n_1}^{n_2} c_k e^{ikx}, \quad C_k \in \mathbb{C}.$$

If $f \approx p$, then

$$\int_0^{2\pi} p(x) e^{-ijx} \frac{dx}{2\pi} = c_j.$$

Definition 6.4.1 (Fourier Coefficients). For a function $f \in L^1(\mathbb{T})$, Fourier coefficients are

$$\hat{f}_j = \int_0^{2\pi} f(x) e^{-ijx} \frac{dx}{2\pi}.$$

Question

$$f \stackrel{?}{=} \sum_{j=-N}^N \hat{f}_j e^{ijx}$$

Definition 6.4.2 (Partial Fourier Sum).

$$\begin{aligned} f(x + 2\pi) &= f(x) \\ \hat{f}_n &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{inx} dx, \quad n \in \mathbb{Z} \\ (S_n f)(x) &= \sum_{n=-N}^N \hat{f}_n e^{inx}. \end{aligned}$$

Definition 6.4.3 (Dirichlet Kernel).

$$D_N(t) = \sum_{n=-N}^N e^{int}.$$

Proposition 6.4.4

$$D_N(t) = \begin{cases} \frac{\sin((2N+1)t/2)}{\sin(t/2)}, & t \neq 0 \\ 2N+1, & t = 0. \end{cases}$$

Proof 1. By geometric sum,

$$\begin{aligned}
 D_N(t) &= e^{-iNt} \sum_{j=0}^{2N} e^{ijt} = e^{-iNt} \frac{1 - e^{i(2N+1)t}}{1 - e^{it}} \\
 &= \frac{e^{-i\frac{(2N+1)t}{2}} - e^{i\frac{(2N+1)t}{2}}}{e^{-it/2} - e^{it/2}} \\
 &= \frac{-2i \sin\left(\frac{2N+1}{2}t\right)}{-2i \sin\left(\frac{t}{2}\right)} \quad [\text{Euler's formula}] \\
 &= \frac{\sin((2N+1)t/2)}{\sin(t/2)}.
 \end{aligned}$$

Q.E.D. ■

Attempt to answer the question

$$\begin{aligned}
 (S_N f)(x) &= \sum_{n=-N}^N \int_0^{2\pi} f(y) e^{iny} e^{inx} \frac{dy}{2\pi} \\
 &= \int_0^{2\pi} \sum_{n=-N}^N f(y) e^{iny} e^{inx} \frac{dy}{2\pi} \quad [\text{Exchange } \sum \text{ and } \int \text{ because we have finite sums}] \\
 &= \int_0^{2\pi} f(y) \sum_{n=-N}^N e^{in(x-y)} \frac{dy}{2\pi} \\
 &= \int_0^{2\pi} f(y) D_N(x-y) \frac{dy}{2\pi} \quad [D_N(t) = \sum_{n=-N}^N e^{int} \text{ is the Dirichlet kernel}]
 \end{aligned}$$

Updated Question

- $f \in C(\mathbb{T})$:

$$\|S_N f - f\|_{\mathbb{T}} \rightarrow 0 \quad \text{as } N \rightarrow \infty? \quad \text{Not in general.}$$

- $f \in L^1(\mathbb{T})$:

$$\|S_N f - f\|_1 \rightarrow 0 \quad \text{as } N \rightarrow \infty? \quad \text{Not in general.}$$

Why this Dirichlet Kernel Fails?

$$\int_0^{2\pi} D_N(t) \frac{dt}{2\pi} = \int_0^{2\pi} \sum_{k=-N}^N e^{ikt} \frac{dt}{2\pi} = 1,$$

but

$$\sup_N \int_0^{2\pi} |D_N(t)| \frac{dt}{2\pi} = \infty.$$

So, we have oscillations.

Definition 6.4.5 (Cesàrd Convergence). The sequence x_n is *cesard-convergent* if

$$\frac{1}{N} \sum_{n=0}^{N-1} x_n \text{ (Cesard-averages) is convergent.}$$

Example 6.4.6

The sequence $1, -1, 1, -1, \dots$ is not convergent, but their cesard-averages converge to 0.

Definition 6.4.7 (Cesàrd-Averages of Fourier Partial Sums).

$$\begin{aligned} (C_N f)(x) &= \frac{1}{N} \sum_{n=0}^{N-1} (S_N f)(x) \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=-n}^n \hat{f}_k e^{ikx} \\ &= \frac{1}{N} \sum_{k=-N+1}^{N-1} \sum_{n=|k|}^{N-1} \hat{f}_k e^{ikx} \\ &= \sum_{k=-N+1}^{N-1} \frac{N-|k|}{N} \hat{f}_k e^{ikx} \quad \text{[No dependency on } n; \text{ sum it up.]} \end{aligned}$$

Definition 6.4.8 (Fejér Kernel).

$$(C_N f)(x) = \int_0^{2\pi} f(y) F_N(x-y) \frac{dy}{2\pi},$$

where

$$F_N(t) = \frac{1}{N} \sum_{n=0}^{N-1} D_n(t) \text{ is the Fejér kernel.}$$

Note that

$$F_N(t) = \frac{1}{N} \left(\frac{\sin(Nt/2)}{\sin(t/2)} \right)^2,$$

and

$$\int_0^{2\pi} F_N(t) \frac{dt}{2\pi} = \frac{1}{N} \sum_{n=0}^{N-1} \int_0^{2\pi} D_n(t) \frac{dt}{2\pi} = 1.$$

Lemma 6.4.9 (Fejér Kernel Properties). Fejér kernel has the following properties:

- $F_N(t) \geq 0$
- $\int_0^{2\pi} F_N(t) \frac{dt}{2\pi} = 1$

- For any $\varepsilon > 0$,

$$\lim_{N \rightarrow \infty} \int_{\varepsilon}^{2\pi - \varepsilon} F_N(t) \frac{dt}{2\pi} = 0.$$

Proof2.

- ①-② are evident. To show ③, if $\varepsilon \leq t \leq 2\pi - \varepsilon$, then

$$\sin \frac{t}{2} \geq \sin \frac{\varepsilon}{2}.$$

So,

$$\begin{aligned} F_N(t) &= \frac{1}{N} \left(\frac{\sin Nt/2}{\sin t/2} \right)^2 \leq \frac{1}{N \sin^2(\varepsilon/2)}. \\ \int_{\varepsilon}^{2\pi - \varepsilon} F_N(t) \frac{dt}{2\pi} &\leq \int_{\varepsilon}^{2\pi - \varepsilon} \frac{1}{N \sin^2(\varepsilon/2)} \frac{dt}{2\pi} \\ &= \frac{2\pi - 2\varepsilon}{2\pi} \frac{1}{N \sin^2(\varepsilon/2)} \rightarrow 0 \quad \text{when } N \rightarrow \infty. \end{aligned}$$

Q.E.D. ■

Theorem 6.4.10 Fejér's Theorem

For any $f \in C(\mathbb{T})$,

$$\|C_N f - f\|_{\mathbb{T}} \rightarrow 0 \quad \text{as } N \rightarrow \infty.$$

Proof3.

$$\begin{aligned} (C_N f)(x) - f(x) &= \int_0^{2\pi} F_N(y) f(x-y) \frac{dy}{2\pi} - f(x) \int_0^{2\pi} F_N(y) \frac{dy}{2\pi} \\ &= \int_0^{2\pi} F_N(y) (f(x-y) - f(x)) \frac{dy}{2\pi}. \end{aligned}$$

[Idea: for small y , $f(x-y) - f(x)$ will be small. For large y , $F_N(y)$ will be small.]

$$\begin{aligned} |(C_N f)(x) - f(x)| &\leq \left| \int_{-\varepsilon}^{\varepsilon} F_N(y) (f(x-y) - f(x)) \frac{dy}{2\pi} \right| + \int_{\varepsilon}^{2\pi - \varepsilon} F_N(y) (f(x-y) - f(x)) \frac{dy}{2\pi} \\ \|C_N f - f\|_{\mathbb{T}} &\leq \int_{-\varepsilon}^{\varepsilon} F_N(y) \|\tau_y f - f\|_{\mathbb{T}} \frac{dy}{2\pi} + \int_{\varepsilon}^{2\pi - \varepsilon} F_N(y) \cdot 2\|f\|_{\mathbb{T}} \frac{dy}{2\pi}. \end{aligned}$$

For any $\delta > 0$, $\exists \varepsilon > 0$ such that $|y| < \varepsilon \implies \|\tau_y f - f\|_{\mathbb{T}} < \delta$. So,

$$\begin{aligned} \|C_N f - f\|_{\mathbb{T}} &\leq \delta \underbrace{\int_{-\varepsilon}^{\varepsilon} F_N(y) \frac{dy}{2\pi}}_{\leq 1 \text{ since total integral}=1} + \int_{\varepsilon}^{2\pi - \varepsilon} F_N(y) \cdot 2\|f\|_{\mathbb{T}} \frac{dy}{2\pi} \\ &\leq \delta + \underbrace{\int_{\varepsilon}^{2\pi - \varepsilon} F_N(y) \cdot 2\|f\|_{\mathbb{T}} \frac{dy}{2\pi}}_{\rightarrow 0 \text{ by Lemma 6.9}}. \end{aligned}$$

So, when $N \rightarrow \infty$,

$$\limsup_{N \rightarrow \infty} \|C_N f - f\|_{\mathbb{T}} \leq \delta.$$

Since $\delta > 0$ is arbitrary,

$$\begin{aligned} \limsup_{N \rightarrow \infty} \|C_N f - f\|_{\mathbb{T}} &= 0 \\ \lim_{N \rightarrow \infty} \|C_N f - f\|_{\mathbb{T}} &= 0. \end{aligned}$$

Q.E.D. ■

Corollary 6.4.11 (Weierstrass' Second Theorem). Trigonometric polynomials are dense in $C(\mathbb{T})$.

Corollary 6.4.12 For any $f \in L^1(\mathbb{T})$,

$$\lim_{n \rightarrow \pm\infty} \hat{f}_n = 0.$$

Proof 4. If p is a trigonometric polynomial,

$$p(x) = \sum_{k=a}^{\ell} c_k e^{ikx}.$$

Then, $\hat{p}_n = 0$ if $n > \ell$ or $n < a$.

For any $f \in L^1(\mathbb{T})$, any $\varepsilon > 0$, $\exists g \in C(\mathbb{T})$ such that $\|f - g\|_1 < \varepsilon$ and $\exists p$ trigonometric polynomial such that $\|g - p\|_{\mathbb{T}} < \varepsilon$. Then,

$$\begin{aligned} \|g - p\|_1 &= \int |g - p| dx \leq \|g - p\|_{\mathbb{T}} \int 1 dx < 2\pi\varepsilon \\ \|f - p\| + 1 &\leq \|f - g\|_1 + \|g - p\|_1 < \varepsilon + 2\pi\varepsilon = (2\pi + 1)\varepsilon \\ |\hat{f}_n - \hat{p}_n| &= \left| \int (f(x)e^{-inx} - p(x)e^{-inx}) \frac{dx}{2\pi} \right| \\ &\leq \int |f - p| \frac{dx}{2\pi} \\ &= \frac{1}{2\pi} \|f - p\|_1 < \frac{2\pi + 1}{2\pi} \varepsilon. \end{aligned}$$

Note that for any $|n|$ large enough, $|\hat{p}_n| = 0$. So,

$$|\hat{f}_n| < \frac{2\pi + 1}{2\pi} \varepsilon.$$

Q.E.D. ■

Definition 6.4.13 (Fourier Coefficients of Measures). Let μ be a finite measure on \mathbb{T} . Its Fourier coefficients are given by

$$\hat{\mu}_n = \int e^{-inx} d\mu(x).$$

If $\hat{\mu}_n = \hat{\nu}_n \quad \forall n \in \mathbb{Z}$, then $\mu = \nu$.

6.5 Convolutions

In this section, assume $G = \mathbb{T}^d$ or $G = \mathbb{R}^d$, $f, g \in L^1(G)$, and define

$$h(x) = \int f(x-y)g(y) \, dy.$$

Lemma 6.5.1 $h(x)$ is well-defined for a.e. x and $h \in L^1(G)$ with

$$\|h\|_1 \leq \|f\|_1 \|g\|_1.$$

Proof 1.

$$\begin{aligned} \iint |f(x-y)g(y)| \, dy dx &= \iint |f(x-y)g(y)| \, dx dy && \text{[Tonelli]} \\ &= \iint |f(x)g(y)| \, dx dy && \text{[Translation Invariance]} \\ &= \int \left(|g(y)| \int |f(x)| \, dx \right) \, dy \\ &= \int |f(x)| \, dx \cdot \int |g(y)| \, dy < \infty. \end{aligned}$$

Then,

$$\int |f(x-y)g(y)| \, dy < \infty \quad \text{for a.e. } x.$$

Hence,

$$h(x) = \int f(x-y)g(y) \, dy$$

is well-defined for a.e. x . Moreover,

$$\begin{aligned} \int |h(x)| \, dx &= \int \left| \int f(x-y)g(y) \, dy \right| \, dx \\ &\leq \iint |f(x-y)g(y)| \, dy dx \\ &= \|f\|_1 \|g\|_1. \end{aligned}$$

That is, $\|h\|_1 \leq \|f\|_1 \|g\|_1$. So, $h \in L^1(G)$.

Q.E.D. ■

Definition 6.5.2 (Convolution). h is called the convolution of f and g . Notationally,

$$h = f * g.$$

Remark.

- Convolution is commutative. Consider x fixed, y variable, and the substitution $u = x - y$.

Then,

$$\int f(x-y)g(y) dy = \int f(u)g(x-u) du$$

$$f * g = g * f.$$

- One can also show

$$(f * g) * h = f * (g * h)$$

- No identity element for convolution. Instead, we have an *approximate identity*, which is a family of functions (sequence, or parametrized by $\lambda > 0$), $K_\lambda \in L^1(G)$ satisfying

1. $K_\lambda \geq 0$

2. $\int K_\lambda(x) dx = 1$

3. \forall open $U \subset G$ with $0 \in U$,

$$\lim_{\lambda \rightarrow \infty} \int_{U^c} K_\lambda(x) dx = 0.$$

[This definition is similar to Fejér kernel's. A more general definition is that

1. $\sup_\lambda \int |K_\lambda(x)| dx < \infty$

2. $\int K_\lambda(x) dx = 1$

3. \forall open U with $0 \in U$,

$$\int_{U^c} |K_\lambda(x)| dx \rightarrow 0.$$

]

Theorem 6.5.3 Approximate Identity

For an approximate identity $K_\lambda, \forall f \in C(G), K_\lambda * f \rightarrow f$ in $C(G)$

Remark. $g \equiv 1$ is not an identity.

$$(f * g)(x) = \int f(x-y)g(y) dy = \int f(x-y) dy$$

$$= \int f(y) dy$$

$$= \text{constant} \perp\!\!\!\perp x.$$

Fourier Series on \mathbb{T}^d

- Fourier coefficients:

$$\hat{f}_n = \int f(x) e^{-inx} \frac{dx}{(2\pi)^d}, \quad n \in \mathbb{Z}^d$$

- Fourier partial sums

$$S_N f = \sum_{\substack{n \in \mathbb{Z}^d \\ |n_j| \leq N \forall j}} \widehat{f}_n e^{inx}$$

- Cesàrd-Averages:

$$C_N f = \frac{1}{N} \sum_{k=0}^{N-1} S_{N-k} f, \quad \text{and } C_N f = F_N * f$$

- Fejér kernel in \mathbb{T}^d :

$$F_N(x) = \prod_{j=1}^d F_N(x_j)$$

This Fejér kernel is an approximate identity in \mathbb{T}^d . So,

$$C_N f \rightarrow f \quad \text{in } C(\mathbb{T}^d).$$

Theorem 6.5.4

Let K_λ be an approximate identity. Then, $\forall f \in L^1(G)$,

$$K_\lambda * f \rightarrow f \quad \text{in } L^1(G).$$

Proof 2.

$$\begin{aligned} (K_\lambda * f)(x) - f(x) &= \int K_\lambda(y) f(x-y) dy - \int K_\lambda(y) f(x) dy && [\int K_\lambda(y) dy = 1] \\ &= \int K_\lambda(y) [f(x-y) - f(x)] dy \\ &= \int K_\lambda(y) (\tau_y f - f)(x) dy. \end{aligned}$$

So,

$$\begin{aligned} \|K_\lambda * f - f\|_1 &\leq \iint K_\lambda(y) |(\tau_y f - f)(x)| dy dx \\ &\leq \int K_\lambda(y) \|\tau_y f - f\|_1 dy \end{aligned}$$

Fix $\varepsilon > 0$. Let U open with $0 \in U$ such that

$$\|\tau_y f - f\| < \varepsilon \quad \text{whenever } y \in U.$$

[Why we can pick such a set? By continuity of translation:

$$\tau_y f \rightarrow f \quad \text{in } L^1 \quad \text{as } y \rightarrow 0 \implies \exists \delta > 0, \quad \|y\| < \delta \implies \|\tau_y f - f\|_1 < \varepsilon.$$

] Hence,

$$\begin{aligned} \|K_\lambda * f - f\|_1 &\leq \int_U K_\lambda(y) \|\tau_y f - f\|_1 \, dy + \int_{U^c} K_\lambda(y) \|\tau_y f - f\|_1 \, dy \\ &\leq \int_U \varepsilon K_\lambda(y) \, dy + \underbrace{\int_{U^c} 2\|f\|_1 K_\lambda(y) \, dy}_{\rightarrow 0 \text{ as } \lambda \rightarrow \infty} \end{aligned}$$

Hence,

$$\limsup_{\lambda \rightarrow \infty} \|K_\lambda * f - f\|_1 \leq \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary,

$$\lim_{\lambda \rightarrow \infty} \|K_\lambda * f - f\|_1 = 0.$$

Q.E.D. ■

Proposition 6.5.5 (Convolution Improves Differentiability).

$$\begin{aligned} (f * g)(x) &= \int f(x - y)g(y) \, dy \\ (f * g)'(x) &= \int f'(x - y)g(y) \, dy \\ (f * g)' &= f' * g. \end{aligned}$$

6.6 Fourier Integrals

Recall Fourier series: for $f \in L^1(\mathbb{T}^d)$,

$$\hat{f}_n = \frac{1}{(2\pi)^d} \int e^{-inx} f(x) \, dx \quad n \in \mathbb{Z}^d.$$

Definition 6.6.1 (Fourier Integral). Let $f \in L^1(\mathbb{R}^d)$ [*we lose periodicity here*],

$$\hat{f}(k) = \frac{1}{(2\pi)^{d/2}} \int e^{-ik \cdot x} f(x) \, dx \quad k \in \mathbb{R}^d.$$

Definition 6.6.2 (Pointwise Decaying Function). $f(x)$ is a *decaying function* if $\lim_{x \rightarrow \pm\infty} f(x) = 0$.

Remark. Note that $f \in L^1(\mathbb{R}^d) \not\Rightarrow f$ is decaying. However, $\hat{f}(k)$ is decaying. So, the Fourier integral produces a decaying function.

Proposition 6.6.3 (Some Properties of Fourier Transformation).

- $|e^{-ikx} f(x)| = |f(x)|$
- $\widehat{f * g} = (2\pi)^{d/2} \hat{f} \hat{g}$

- \widehat{f} is a bounded function

Proof 1.

$$\begin{aligned} |\widehat{f}(k)| &\leq \frac{1}{(2\pi)^{d/2}} \int |e^{-ikx} f(x)| dx \\ \implies |\widehat{f}(k)| &\leq \frac{1}{(2\pi)^{d/2}} \|f\|_1 \quad \forall k \\ \implies \widehat{f} &\text{ is a bounded function} \end{aligned}$$

Q.E.D. ■

- \widehat{f} is uniformly continuous.

Proof 2.

$$\begin{aligned} \widehat{f}(k+a) - \widehat{f}(k) &= \frac{1}{(2\pi)^{d/2}} \int (e^{-i(k+a)x} - e^{-ikx}) f(x) dx \\ &= \frac{1}{(2\pi)^{d/2}} \int (e^{-iax} - 1) e^{-ikx} f(x) dx \\ |\widehat{f}(k+a) - \widehat{f}(k)| &\leq \frac{1}{(2\pi)^{d/2}} \int |e^{-iax} - 1| \cdot |f(x)| dx \end{aligned}$$

It is enough to prove:

$$\lim_{a \rightarrow 0} \int |e^{-iax} - 1| \cdot |f(x)| dx = 0.$$

Note that $|e^{-iax} - 1| \cdot |f(x)| \leq 2|f(x)|$ and $e^{-iax} \rightarrow 1$ pointwise when $a \rightarrow \infty$. [Recall the fact that

$$\lim_{a \rightarrow 0} g(a) = 0 \iff \forall \text{ sequence } a_n, \quad a_n \rightarrow 0 \implies g(a_n) \rightarrow 0.$$

] So, we will do the same. Pick a sequence $a_n \rightarrow 0$. Apply dominated convergence theorem, and we complete the proof. Q.E.D. ■

- **Lemma Riemann-Lebesgue** If $f \in L^1(\mathbb{R}^d)$, then

$$\lim_{\|k\| \rightarrow \infty} \widehat{f}(k) = 0.$$

That is, \widehat{f} is decaying, and $\forall \varepsilon > 0, \exists R \quad \forall k \text{ s.t.}$

$$\|k\| > R \implies |\widehat{f}(k)| < \varepsilon.$$

Proof 3. We will just prove a special case here. Let $f \in C_C(\mathbb{R}^d)$. Then,

$$\begin{aligned} \widehat{f}(k) &= \frac{1}{(2\pi)^{d/2}} \int e^{-ik \cdot x} f(x) dx \\ \widehat{f}(k) &= \frac{1}{(2\pi)^{d/2}} \int -e^{-ik \cdot (x + \frac{\pi}{k})} f(x) dx \end{aligned} \tag{7}$$

Define $y = x + \frac{\pi}{k}$. Then,

$$\widehat{f}(k) = \frac{1}{(2\pi)^{d/2}} \int -e^{-ik \cdot y} f\left(x - \frac{\pi}{k}\right) dy \quad (8)$$

(7) + (8), we have

$$\begin{aligned} 2\widehat{f}(k) &= \frac{1}{(2\pi)^{d/2}} \int e^{-ik \cdot x} \left[f(x) - f\left(x - \frac{\pi}{k}\right) \right] dx \\ 2|\widehat{f}(k)| &\leq \frac{1}{(2\pi)^{d/2}} \int \left| f(x) - f\left(x - \frac{\pi}{k}\right) \right| dx \rightarrow 0 \quad \text{as } k \rightarrow \infty \end{aligned}$$

since f is uniform convergent ($f \in C_C(\mathbb{R}^d)$).

Q.E.D. ■

Lemma 6.6.4 If $f \in L^1(\mathbb{R})$ and $\int |xf(x)| dx < \infty$, then \widehat{f} is differentiable, and

$$\left(\widehat{f}\right)'(k) = \widehat{(-ixf(x))}(k).$$

Proof 4. Denote $g(x) = ixf(x)$. So, $g \in L^1(\mathbb{R})$. [WTS: $\widehat{f}' = \widehat{g}$. Instead of differentiation, let's show for $a < b$,

$$\int_a^b \widehat{g}(k) dk = \widehat{f}(b) - \widehat{f}(a).$$

We already know that \widehat{g} is continuous. So, this will imply $\widehat{f}' = \widehat{g}$ by Fundamental Theorem of Calculus.]

Fix $a < b$. Then,

$$\int_a^b \widehat{g}(k) dk = \frac{1}{\sqrt{2\pi}} \int_a^b \int g(x) e^{-ik \cdot x} dx dk.$$

To justify exchanging intergrals, let's check:

$$\begin{aligned} \int_a^b \int |g(x) e^{-ik \cdot x}| dx dk &= \int_a^b \int |g(x)| dx dk \\ &= \int_a^b \|g\|_1 dk \\ &= \|g\|_1 (b - a) < \infty. \end{aligned}$$

So, by Fubini's Theorem,

$$\begin{aligned} \boxed{\int_a^b \widehat{g}(k) dk} &= \frac{1}{\sqrt{2\pi}} \int \int_a^b g(x) e^{-ik \cdot x} dk dx \\ &= \frac{1}{\sqrt{2\pi}} \int \int_a^b -ixf(x) e^{-ik \cdot x} dk dx && [\partial_k (e^{-ikx}) = -ixe^{-ikx}] \\ &= \frac{1}{\sqrt{2\pi}} \int f(x) e^{-ik \cdot x} \Big|_{k=a}^{k=b} dx \\ &= \frac{1}{\sqrt{2\pi}} \int f(x) e^{-ib \cdot x} - f(x) e^{-ia \cdot x} dx = \boxed{\widehat{f}(b) - \widehat{f}(a)}. \end{aligned}$$

Q.E.D. ■

For $f \in L^1(\mathbb{R})$, does \widehat{f} uniquely determine f ? The answer is Yes. To prove it, we need Fejér Kernel.

Definition 6.6.5 (Fejér Kernel).

$$K_\lambda(x) = \frac{\lambda}{2\pi} \left(\frac{\sin(\lambda x/2)}{(\lambda x/2)} \right)^2 \geq 0.$$

Note that

- $\int K_\lambda(x) dx = 1$
- For any $\delta > 0$, as $\lambda \rightarrow \infty$.

$$\int_{\mathbb{R} \setminus (-\delta, \delta)} K_\lambda(x) dx \rightarrow 0.$$

- So, $K_\lambda * f \rightarrow f$ in $L^1(\mathbb{R})$.

Lemma 6.6.6 (Fejér Kernel as an Integral).

$$K_\lambda(x) = \frac{1}{2\pi} \int_{-\lambda}^{\lambda} \left(1 - \frac{|k|}{\lambda} \right) e^{ik \cdot x} dk.$$

Corollary 6.6.7 For all $f \in L^1(\mathbb{R})$,

$$(K_\lambda * f)(x) = \frac{1}{\sqrt{2\pi}} \int_{-\lambda}^{\lambda} \left(1 - \frac{|k|}{\lambda} \right) \widehat{f}(k) e^{-ik \cdot x} dk.$$

Corollary 6.6.8 Fourier transforms uniquely determine the function. i.e., if $\widehat{f} = \widehat{g} \quad \forall k$, then $f = g$.

Proof 5. Fourier integrals uniquely determine the convolutions, which converges to the function.

Q.E.D. ■

7 Hilbert Space

7.1 Inner Product

Definition 7.1.1 (Dot Product).

- $\forall x, y \in \mathbb{R}^n$, the dot product is defined as $x \cdot y = \sum_{j=1}^n x_j y_j$, and $x \cdot x = \|x\|^2$.
- $\forall x, y \in \mathbb{C}^n$, the dot product is defined as $\langle x, y \rangle = \sum_{j=1}^n \bar{x}_j y_j$, and $\langle x, x \rangle = \sum_{j=1}^n |x_j|^2$.

Definition 7.1.2 (Inner Product). Let V be a vector space over the field \mathbb{C} . An *inner product* on V is a map $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ with the following properties:

- Linearity in the second parameter: If $\lambda \in \mathbb{C}$, $x, y, \tilde{y} \in V$, then

$$\langle x, \lambda y \rangle = \lambda \langle x, y \rangle \quad \text{and} \quad \langle x, y + \tilde{y} \rangle = \langle x, y \rangle + \langle x, \tilde{y} \rangle.$$

- Skew Symmetry: $\overline{\langle x, y \rangle} = \langle y, x \rangle$.
- Strict Positivity Away from 0: $\forall x \in V, \quad x \neq 0 \implies \langle x, x \rangle > 0$.

Proposition 7.1.3 (Conjugate Linearity in the First Parameter). Combining properties ① and ② in Definition 7.1.2, we have that

$$\begin{aligned} \langle \lambda x, y \rangle &= \overline{\langle y, \lambda x \rangle} = \overline{\lambda \langle y, x \rangle} = \bar{\lambda} \langle y, x \rangle = \bar{\lambda} \langle x, y \rangle \quad \lambda \in \mathbb{C} \\ \langle x + \tilde{x}, y \rangle &= \overline{\langle y, x + \tilde{x} \rangle} = \overline{\langle y, x \rangle + \langle y, \tilde{x} \rangle} = \langle x, y \rangle + \langle \tilde{x}, y \rangle. \end{aligned}$$

Definition 7.1.4 (Induced Norm). [We have yet to show that this is a proper norm.]

$$\|x\| = \sqrt{\langle x, x \rangle} \quad \text{or} \quad \|x\|^2 = \langle x, x \rangle.$$

Proposition 7.1.5 (Properties of the Induced Norm). The inner product induced norm has positive homogeneity and positivity away from 0.

Proof 1.

- $\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x \rangle} = \sqrt{\bar{\lambda} \lambda \langle x, x \rangle} = \sqrt{|\lambda|^2 \langle x, x \rangle} = |\lambda| \cdot \|x\|$
- $x \geq 0 \implies \langle x, x \rangle > 0 \implies \|x\| \geq 0$.

Q.E.D. ■

[To show the induced norm is a norm, we still need the triangle inequalities. We will develop some tools.]

Definition 7.1.6 (Orthogonality). We say that x_1, \dots, x_n are (pairwise) orthogonal if $\langle x_j, x_k \rangle = 0$ whenever $j \neq k$.

Theorem 7.1.7 Pythagorean Theorem

If x_1, \dots, x_n are orthogonal, then

$$\left\| \sum_{j=1}^n x_j \right\|^2 = \sum_{j=1}^n \|x_j\|^2.$$

Proof 2.

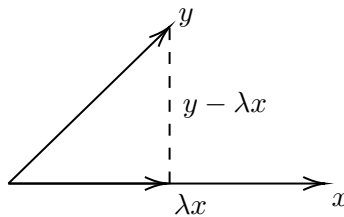
$$\begin{aligned} \left\| \sum_{j=1}^n x_j \right\|^2 &= \left\langle \sum_{j=1}^n x_j, \sum_{k=1}^n x_k \right\rangle \\ &= \sum_{j=1}^n \sum_{k=1}^n \langle x_j, x_k \rangle \\ &= \sum_{j=1}^n \langle x_j, x_j \rangle = \sum_{j=1}^n \|x_j\|^2. \end{aligned}$$

Q.E.D. ■

Lemma 7.1.8 (Cauchy-Schwarz Inequality). $\forall x, y \in V$, we have

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|.$$

Proof 3. If $x = 0$, then both sides are zero. Now, assume $x \neq 0$. Look for $\lambda \in \mathcal{C}$ s.t. $y - \lambda x$ is orthogonal to x .



So, we want

$$\begin{aligned} \langle x, y - \lambda x \rangle &= 0 \\ \langle x, y \rangle - \langle x, \lambda x \rangle &= 0 \\ \langle x, y \rangle &= \lambda \langle x, x \rangle \\ \lambda &= \frac{\langle x, y \rangle}{\langle x, x \rangle} = \frac{\langle x, y \rangle}{\|x\|^2} \quad [x \neq 0] \end{aligned}$$

So, $\langle \lambda x, y - \lambda x \rangle = 0$. By Pythagorean Theorem,

$$\begin{aligned}\|y\|^2 &= \|\lambda x\|^2 + \|y - \lambda x\|^2 \geq \|\lambda x\|^2 \\ \|y\| &\geq \|\lambda x\| = |\lambda| \cdot \|x\| \\ \|y\| &\geq \frac{|\langle x, y \rangle|}{\|x\|^2} \|x\| \\ |\langle x, y \rangle| &\leq \|x\| \cdot \|y\|\end{aligned}$$

The equality holds when $y = \lambda x$.

Q.E.D. ■

Theorem 7.1.9 Triangle Inequality

$\forall x, y \in V$,

$$\|x + y\| \leq \|x\| + \|y\|.$$

Proof 4. This is equivalent to proving

$$\begin{aligned}\|x + y\|^2 &\leq (\|x\| + \|y\|)^2 \\ \langle x + y, x + y \rangle &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| \\ \cancel{\langle x, x \rangle} + \langle x, y \rangle + \langle y, x \rangle + \cancel{\langle y, y \rangle} &\leq \cancel{\langle x, x \rangle} + \cancel{\langle y, y \rangle} + 2\|x\| \cdot \|y\| \\ \langle x, y \rangle + \overline{\langle x, y \rangle} &\leq 2\|x\| \cdot \|y\| \\ 2 \operatorname{Re} \langle x, y \rangle &\leq 2\|x\| \cdot \|y\|.\end{aligned}$$

This is then correct, because,

$$\operatorname{Re} \langle x, y \rangle \leq |\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

by Cauchy-Schwarz Inequality.

Q.E.D. ■

Definition 7.1.10 (Hilbert Space). For V with inner product $\langle \cdot, \cdot \rangle$, if V is complete with respect to the induced norm, then V is called a *Hilbert space*, and usually denoted \mathcal{H} .

So, every Hilbert space is a Banach space, but not conversely.

Example 7.1.11

$L^2(X, d\mu)$ with inner product $\langle f, g \rangle = \int \bar{f}g \, d\mu$ is a Hilbert space.

Proof 5. If $f, g \in L^2$, then by Hölder's Inequality,

$$\begin{aligned}\int |fg| \, d\mu &\left(\int |f|^2 \, d\mu \right)^{1/2} \left(\int |g|^2 \, d\mu \right)^{1/2} \\ \int |\bar{f}g| \, d\mu &= \left(\int |f|^2 \, d\mu \right)^{1/2} \left(\int |g|^2 \, d\mu \right)^{1/2}\end{aligned}$$

So, $\bar{f}g$ is integrable, and $\langle f, g \rangle$ is well-defined.

- Linear in the second parameter: $\langle f, g \rangle = \overline{\langle g, f \rangle}$.
- $\langle f, f \rangle = \int |f|^2 d\mu \geq 0$, and $\int |f|^2 d\mu = 0 \iff f = 0 \quad \mu - a.e. \iff f = 0$ as element of L^2 .
- $\langle f, f \rangle = \|f\|_2^2$.

The induced norm is exactly the norm we used in L^p spaces. So, we have triangle inequality.

Moreover, we know L^2 is complete with respect to $\|\cdot\|_2$. So, L^2 is a Hilbert space.

Q.E.D. ■

Proposition 7.1.12 If vectors $x_j, \quad j \in \mathbb{N}$ are orthogonal, and if

$$\sum_{j=1}^{\infty} \|x_j\|^2 < \infty,$$

then $\sum_{j=1}^{\infty} x_j$ converges in Hilbert space \mathcal{H} , and

$$\left\| \sum_{j=1}^{\infty} x_j \right\|^2 = \sum_{j=1}^{\infty} \|x_j\|^2.$$

Proof 6. Let $y_n = \sum_{j=1}^n x_j$ and $y = \sum_{j=1}^{\infty} x_j$. [WTS: y_n is Cauchy.]

Suppose $m < n$. Then,

$$\|y_n - y_m\|^2 = \left\| \sum_{j=m+1}^n x_j \right\|^2 = \sum_{j=m+1}^n \|x_j\|^2$$

by Paythagorean Theorem. Since $\sum_{j=1}^{\infty} \|x_j\|^2 < \infty, \forall \varepsilon > 0, \exists n_0$ s.t. $m > n_0 \implies \sum_{j=m+1}^{\infty} \|x_j\|^2 < \varepsilon$. So, for $n > m > n_0$,

$$\|y_n - y_m\|^2 = \sum_{j=m+1}^n \|x_j\|^2 < \varepsilon.$$

So, $(y_n)_{n=1}^{\infty}$ is a Cauchy sequence. By completeness, $y_n \rightarrow y \in \mathcal{H}$.

Lemma In Banach Spaces,

$$y_n \rightarrow y \implies \|y_n\| \rightarrow \|y\|.$$

Proof.

$$\begin{aligned} \|y_n\| &\leq \|y\| + \|y_n - y\| \\ \|y\| &\leq \|y_n\| = \|y - y_n\| \end{aligned}$$

Combine, we get

$$-\|y_n - y\| \leq \|y_n\| - \|y\| \leq \|y_n - y\|.$$

When $y_n \rightarrow 0$, $\|y_n - y\| \rightarrow 0$. So,

$$\begin{aligned} \|y_n\| - \|y\| &\rightarrow 0 \\ \|y_n\| &\rightarrow \|y\|. \end{aligned}$$

We have $\|y_n\|^2 = \sum_{j=1}^n \|x_j\|^2$. So,

$$\|y\|^2 = \sum_{j=1}^{\infty} \|x_j\|^2.$$

□

Q.E.D. ■

Example 7.1.13

In $L^2(\mathbb{T})$, consider functions e^{inx} for $n \in \mathbb{Z}$. The inner product is defined as

$$\langle f, g \rangle = \int_0^{2\pi} \overline{f(x)}g(x) \frac{dx}{2\pi}.$$

Then,

$$\int_0^{2\pi} \overline{e^{inx}} e^{ikx} \frac{dx}{2\pi} = \begin{cases} 1, & k = n \\ 0, & k \neq n \end{cases}$$

Start with some sequence $c_n \in \mathbb{C}$, with $\sum |c_n|^2 < \infty$. Then, $\sum_{n \in \mathbb{Z}} c_n e^{inx}$ converges in Hilbert space $L^2(\mathbb{T})$. By mutual orthogonality, $\|c_n e^{inx}\| = |c_n|$.

Conclusion: \forall sequence $c_n \in \ell^2(\mathbb{Z})$, $\exists f = \sum_{n \in \mathbb{Z}} c_n e^{inx} \in L^2(\mathbb{T})$ such that $\hat{f}_n = c_n$.

Proof 7. If $\sum x_j$ converges, then $\forall v$,

$$\left\langle v, \sum x_j \right\rangle = \sum \langle v, x_j \rangle.$$

Equivalently, if $y_n \rightarrow y$, then $\forall v \in \mathcal{H}$,

$$\langle v, y_n \rangle \rightarrow \langle v, y \rangle.$$

This is a continuity statement about inner product:

$$\begin{aligned} |\langle v, y_n \rangle - \langle v, y \rangle| &= |\langle v, y_n - y \rangle| \\ &\leq \|v\| \cdot \|y_n - y\| && \text{[Cauchy-Schwarz]} \\ &\rightarrow 0 \quad \text{as } \|y_n - y\| \rightarrow 0. \end{aligned}$$

So, $\langle v, y_n \rangle \rightarrow \langle v, y \rangle$. Hence,

$$\begin{aligned}\langle e^{ikx}, f(x) \rangle &= \left\langle e^{ikx}, \sum_{n \in \mathbb{Z}} c_n e^{inx} \right\rangle \\ &= \sum_{n \in \mathbb{Z}} c_n \langle e^{ikx}, e^{inx} \rangle \\ &= c_k.\end{aligned}$$

Q.E.D. ■

7.2 Orthogonal Complement and Projection

Remark. A vector space S of a Hilbert space \mathcal{H} isn't always closed. For example, in $\ell^2(\mathbb{N})$, let

$$\begin{aligned}S &= \{x \in \ell^2(\mathbb{N}) \mid x_n = 0 \quad \forall \text{ large enough } n\} \\ &= \{x \in \ell^2(\mathbb{N}) \mid \exists n_0, \quad \forall n > n_0, x_n = 0\} \\ &= \text{span} \{\delta_n \mid n \in \mathbb{N}\},\end{aligned}$$

where

$$(\delta_n)_j = \begin{cases} 1, & j = n \\ 0, & j \neq n. \end{cases}$$

$S \neq \ell^2(\mathbb{N})$ because

$$x_n = \frac{1}{n}, \quad x \in \ell^2(\mathbb{N}), \quad \text{but } x \notin S.$$

But S is dense in $\ell^2(\mathbb{N})$.

Proof 1. $\forall x \in \ell^2(\mathbb{N}), \forall \varepsilon > 0, \exists n_0,$

$$\sum_{n=n_0+1}^{\infty} |x_n|^2 < \varepsilon.$$

Take $y = (x_1, x_2, \dots, x_{n_0}, 0, 0, \dots) \in S$. So,

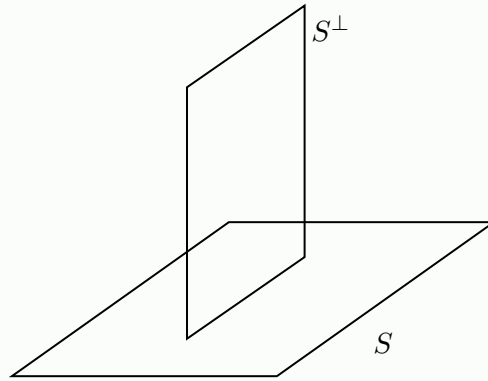
$$\|x - y\|^2 = \sum_{n=n_0+1}^{\infty} |x_n|^2 < \varepsilon.$$

Q.E.D. ■

Definition 7.2.1 (Orthogonal Complement). The *orthogonal complement* of S is

$$S^\perp = \{y \in \mathcal{H} \mid \langle x, y \rangle = 0 \quad \forall x \in S\}.$$

Remark. If S is closed, then $\mathcal{H} = S \oplus S^\perp$.



Lemma 7.2.2 S^\perp is a closed subspace of \mathcal{H} .

Proof2. Let $y, z \in S^\perp$. Then, $\forall x \in S$,

$$\langle x, y \rangle = 0 \quad \text{and} \quad \langle x, z \rangle = 0.$$

So,

$$\langle x, y \rangle + \langle x, z \rangle = \langle x, y + z \rangle = 0.$$

Hence, $y + z \in S^\perp$.

Similarly, for $c \in \mathbb{C}$,

$$c \langle x, y \rangle = \langle x, cy \rangle = 0.$$

So, $cy \in S^\perp$. Therefore, S^\perp is a subspace of \mathcal{H} .

For closedness, if $y_n \in S^\perp$, $y_n \rightarrow y$. Then, $\forall x \in S$,

$$\langle x, y_n \rangle = 0 \implies \langle x, y \rangle = 0.$$

So, $y \in S^\perp$, and S^\perp is closed. Q.E.D. ■

Lemma 7.2.3 (Projection Lemma). Let S be a closed subspace of \mathcal{H} . $\forall x \in \mathcal{H}$, \exists unique $y \in S$ such that $x - y \in S^\perp$.

Proof3.

- Uniqueness is easy:

If $y, z \in S$ such that $x - y \in S^\perp$ and $x - z \in S^\perp$, then $y - z \in S$ and $y - z \in S^\perp$. [Since both S and S^\perp are closed.] So,

$$\langle y - z, y - z \rangle = 0$$

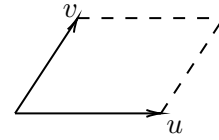
$$\|y - z\|^2 = 0$$

$$y = z.$$

- Existence: Denote $c = \inf_{z \in S} \|x - z\|$. Then, $\forall n \in \mathbb{N}$, $\exists y_n \in S$ such that $c^2 \leq \|x - y_n\|^2 \leq c^2 + \frac{1}{n}$.

Lemma (Parallelogram Identity)

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2$$



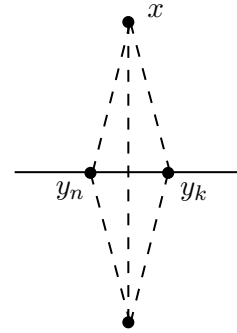
Proof.

$$\begin{aligned} \|u \pm v\|^2 &= \langle u \pm v, u \pm v \rangle \\ &= \langle u, u \rangle + \langle v, v \rangle \pm \langle u, v \rangle \pm \langle v, u \rangle \\ \|u + v\|^2 + \|u - v\|^2 &= 2\langle u, u \rangle + 2\langle v, v \rangle \\ &= 2\|u\|^2 + 2\|v\|^2 \end{aligned}$$

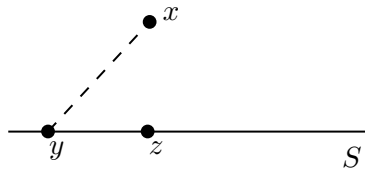
□

We have

$$\begin{aligned} \underbrace{2\|x - y_n\|^2}_{\leq 2(c^2 + \frac{1}{n})} + \underbrace{2\|x - y_k\|^2}_{\leq 2(c^2 + \frac{1}{k})} &= \|y_n - y_k\|^2 + \underbrace{\left\| 2\left(x - \frac{y_n + y_k}{2}\right) \right\|^2}_{\geq 4c^2} \\ 2c^2 + \frac{2}{n} + 2c^2 + \frac{2}{k} &\leq \|y_n - y_k\|^2 + 4c^2 \\ \|y_n - y_k\|^2 &\leq \frac{2}{n} + \frac{2}{k}. \end{aligned}$$



So, y_n is a Cauchy sequence. Then, y_n is convergent by completeness. So $\lim y_n = y \in S$ since S is closed.



$$\|x - y\| = \inf_{z \in S} \|x - z\|$$

Fix $z \in S$. Assume $\langle x - y, z \rangle \neq 0$. Then,

$$\begin{aligned} \|x - y - \lambda z\|^2 &= \langle x - y - \lambda z, x - y - \lambda z \rangle \\ &= \|x - y\|^2 - \langle \lambda z, x - y \rangle - \langle x - y, \lambda z \rangle + \|\lambda z\|^2 \\ &= \|x - y\|^2 - 2\operatorname{Re}(\lambda \langle x - y, z \rangle) + |\lambda|^2 \|z\|^2 \\ &< \|x - y\|^2 \quad \text{for suitable small } \lambda \end{aligned}$$

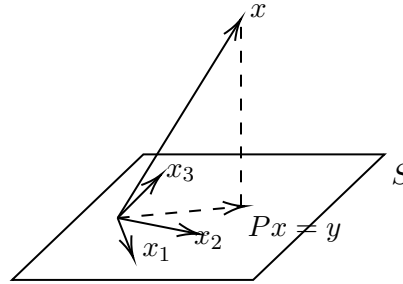
* This is a contradiction because y should be the minimizer.

Q.E.D. ■

Corollary 7.2.4 If S is a closed subspace of \mathcal{H} , and $S \neq \mathcal{H}$, then $S^\perp \neq \{0\}$.

Proof 4. Take $x \in \mathcal{H} \setminus S$. Then, $\exists y \in S$ s.t. $x - y \in S^\perp$. Since $x \neq y$, $x - y \neq 0$, and $x - y \in S^\perp$, the proof is complete. Q.E.D. ■

Proposition 7.2.5 Let x_1, \dots, x_n be orthogonal.



Then,

- $S = \text{span} \{x_1, \dots, x_n\}$ is a closed subspace and orthogonal projection to S is

$$Px = \sum_{j=1}^n \langle x_j, x \rangle x_j = y.$$

Proof 5. $x \in \mathcal{H}, y \in S$. Then,

$$\begin{aligned} \langle x_k, x \rangle &= \langle x_k, y \rangle \quad \forall k \\ \langle x_k, x - y \rangle &= 0 \quad \forall k \\ \implies \left\langle \sum \lambda_k x_k, x - y \right\rangle &= 0 \\ x - y &\in S^\perp. \end{aligned}$$

So, we proved orthogonal projection part. It remains to prove the closedness: $S\bar{S}$. $S \subset \bar{S}$ is trivial. If $x \in \bar{S}$, then $x - y \in \bar{S}$, and $x - y \in S^\perp = \bar{S}^\perp$ [proved in HW]. So, $x - y = x - Px = 0$. So, $x = Px = y \in S$. Then, $\bar{S} \subset S$. So, $S = \bar{S}$. Q.E.D. ■

- **Lemma Bessel's Inequality** For any $x \in \mathcal{H}$,

$$\sum_{j=1}^n |\langle x_j, x \rangle|^2 \leq \|x\|^2.$$

Proof 6.

$$\begin{aligned} \|x\|^2 &= \|y\|^2 + \|x - y\|^2 \geq \|y\|^2 \\ &= \sum_{j=1}^n \|\langle x_j, x \rangle x_j\|^2 \\ &= \sum_{j=1}^n |\langle x_j, x \rangle|^2. \end{aligned}$$

Q.E.D. ■

Example 7.2.6 Application of Bessel's Inequality

On $L^2(\mathbb{T})$, consider e^{inx} for $n \in \mathbb{Z}$. Let $f \in L^2(\mathbb{T})$. Then,

$$\begin{aligned}\widehat{f}_n &= \int e^{-inx} f(x) \frac{dx}{2\pi} \\ \widehat{f}_n &= \langle e^{inx}, f \rangle \\ \sum_{k=-n}^n |\widehat{f}_k|^2 &\leq \|f\|^2 \quad \text{[Bessel's Inequality]}\end{aligned}$$

Take the limit $n \rightarrow \infty$:

$$\sum_{k \in \mathbb{Z}} |\widehat{f}_k|^2 \leq \|f\|^2.$$

Let $g = \sum_{k \in \mathbb{Z}} \widehat{f}_k e^{ikx}$ [Note: $\widehat{f}_k e^{ikx}$ are mutually orthogonal], then we have

$$\sum \|\widehat{f}_k e^{ikx}\|^2 = \sum |\widehat{f}_k|^2 < \infty.$$

Then, g is well-defined, and $g \in L^2(\mathbb{T})$,

$$\|g\|_2^2 = \sum |\widehat{f}_k|^2 \leq \|f\|^2.$$

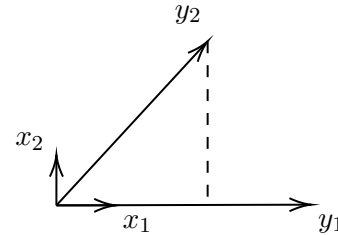
But can we say $\|g\|_2^2 = \|f\|^2$? See Corollary 7.3.6 in the next section.

Proposition 7.2.7 (Gram-Schmidt Process). Given y_1, y_2, \dots, y_n .

- Remove any y_k which is a linear combination of y_1, \dots, y_{k-1} . We get $y_k, \dots, y_{k+1}, \dots$, which is linearly independent and has the same span as the original sequence.
- Reindex the linearly independent sequence y_1, y_2, y_3, \dots

- $x_1 = \frac{y_1}{\|y_1\|}$
- $x_2 = \frac{y_2 - \langle x_1, y_2 \rangle x_1}{\|y_2 - \langle x_1, y_2 \rangle x_1\|}$.
- In general,

$$x_n = \frac{y_n - \sum_{k=1}^{n-1} \langle x_k, y_n \rangle x_k}{\left\| y_n - \sum_{k=1}^{n-1} \langle x_k, y_n \rangle x_k \right\|}$$



- Then, x_1, x_2, \dots is an orthonormal sequence with the same span as y_1, y_2, \dots

7.3 Orthonormal Sets and Bases

Definition 7.3.1 (Orthonormal Set). The *orthonormal family of vectors* $x \in \mathcal{H}$ satisfy:

- If $x_1, x_2 \in X$ and $x_1 \neq x_2$, then $\langle x_1, x_2 \rangle = 0$.
- If $x \in X$, then $\|x\| = 1$.

Definition 7.3.2 (Orthonormal Basis). An *orthonormal basis* is an orthonormal family X such that $\text{span } X$ is dense in \mathcal{H} .

Example 7.3.3

$$X = \{e^{ikx} \mid k \in \mathbb{Z}\} \quad \text{is an orthonormal family.}$$

$\text{span } X = \text{span of all trigonometric polynomials. It is dense in } C(\mathbb{T}). \text{ Further, } C(\mathbb{T}) \text{ is dense in } L^2(\mathbb{T}). \text{ So, } \text{span } X \text{ is dense in } L^2(\mathbb{T}). \text{ So,}$

$$X = \{e^{ikx} \mid k \in \mathbb{Z}\} \quad \text{is an orthonormal basis of } L^2(\mathbb{T}).$$

Theorem 7.3.4

For an orthonormal family X , the following are equivalent:

- There is no $y \notin X$ s.t. $X \cup \{y\}$ is an orthonormal family.
- X is an orthonormal basis.
- If $\langle z, x \rangle = 0 \quad \forall x \in X$, then $z = 0$.

Proof 1. (① \implies ②): We will prove the contrapositive: $\neg \text{①} \implies \neg \text{②}$. Let $S = \text{span } X$. Assume S is not dense. Then, $\overline{S} \neq \mathcal{H}$. Since $\overline{S}^\perp \neq \{0\}$, $\exists z \in \overline{S}^\perp$, $z \neq 0$ such that

$$\langle x, z \rangle = 0 \quad \forall x \in X.$$

Let $y = \frac{z}{\|z\|}$. Then, $X \cup \{y\}$ is an orthonormal family.

(② \implies ③): If $\langle z, x \rangle = 0 \quad \forall x \in X$. Then,

$$\langle z, v \rangle = 0 \quad \forall v \in \text{span } X = S.$$

Since $v_n \in S$, then $v_n \rightarrow v \in \overline{S}$. So,

$$\langle z, v_n \rangle = 0, \quad \text{and} \quad \langle z, v_n \rangle \rightarrow \langle z, v \rangle.$$

Hence, $\forall v \in \bar{S}, \langle z, v \rangle = 0$. That is, $z \in \bar{S}^\perp$. Since S is dense in \mathcal{H} , $\bar{S} = \mathcal{H}$. So, $z = 0$.

(③ \implies ①): Assume ③ is true. Assume y such that $X \cup \{y\}$ is an orthonormal family. Then,

$$\langle x, y \rangle = 0 \quad \forall x \in X.$$

So, $y = 0$. * This contradicts that $\|y\| = 1$. Then, such a y must not exist. Q.E.D. ■

Theorem 7.3.5 Parseval's Identity

Assume $(x_n)_{n=1}^\infty$ is an orthonormal basis of \mathcal{H} . Let $y \in \mathcal{H}$. Then,

$$y = \sum_{n=1}^{\infty} \langle x_n, y \rangle x_n \quad \text{and} \quad \boxed{\|y\|^2 = \sum_{n=1}^{\infty} |\langle x_n, y \rangle|^2}.$$

[Note: the boxed equality is the equality of Bessel's Inequality and is referred as Parseval's Identity.]

Proof2.

$$y_n = \sum_{k=1}^n \langle x_k, y \rangle x_k.$$

By Bessels' Inequality,

$$\sum_{k=1}^n |\langle x_k, y \rangle|^2 \leq \|y\|^2 \quad \forall n.$$

Take $n \rightarrow \infty$, then

$$\sum_{k=1}^{\infty} |\langle x_k, y \rangle|^2 \leq \|y\|^2.$$

Suppose $m < n$. By Pythagorean Theorem,

$$\begin{aligned} \|y_n - y_m\|^2 &= \left\| \sum_{k=m+1}^n \langle x_k, y \rangle x_k \right\|^2 \\ &= \sum_{k=m+1}^n |\langle x_k, y \rangle|^2. \end{aligned}$$

So, $(y_n)_{n=1}^\infty$ is a Cauchy sequence. Since \mathcal{H} is complete,

$$y_n \rightarrow z \quad \text{for some } z \in \mathcal{H}.$$

[WTS: $z = y$.] Note that

$$\langle x_j, z \rangle = \lim_{n \rightarrow \infty} \langle x_j, y_n \rangle.$$

For $n \geq j$,

$$\langle x_j, y_n \rangle = \langle x_j, y \rangle.$$

So,

$$\begin{aligned}\langle x_j, z \rangle &= \lim_{n \rightarrow \infty} \langle x_j, y_n \rangle = \langle x_j, y \rangle \\ \langle x_j, y - z \rangle &= 0 \quad \forall j.\end{aligned}$$

Hence, by Theorem 7.3.4, since $\{x_j \mid j \in \mathbb{N}\}$ is an orthonormal basis,

$$y - z = 0 \implies y = z.$$

Finally, for Parseval's Equality,

$$\begin{aligned}\|y\|^2 &= \lim_{n \rightarrow \infty} \|y_n\|^2 = \lim_{n \rightarrow \infty} \sum_{k=1}^n |\langle x_k, y \rangle|^2 \\ &= \sum_{k=1}^{\infty} |\langle x_k, y \rangle|^2.\end{aligned}$$

Q.E.D. ■

Corollary 7.3.6 In $L^2(\mathbb{T})$. Suppose $f \in L^2(\mathbb{T})$.

$$\begin{aligned}\hat{f}_n &= \langle e^{inx}, f \rangle \\ f &= \lim_{n \rightarrow \infty} \sum_{k=-n}^n \hat{f}_k e^{ikx}.\end{aligned}$$

By Parseval's Identity,

$$\|f\|_2^2 = \sum_{n \in \mathbb{Z}} |\hat{f}_n|^2.$$

Proposition 7.3.7 (Parseval's Identity in $L^2(\mathbb{R})$). In $L^2(\mathbb{R})$,

$$\int |f(x)|^2 dx = \int |\hat{f}(k)|^2 dk.$$

8 Bonus Topics

8.1 Cantor Set and Cantor Function

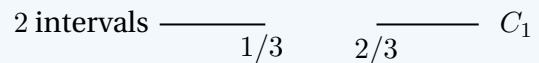
Definition 8.1.1 (Cantor Set). The Cantor Set is constructed iteratively:

- $C_0 = [0, 1]$

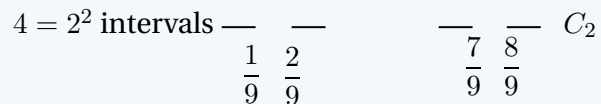


- $C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$

We removed the middle third.



- C_2 : remove the middle third of each sub-interval.



- In the n -th step, we remove 2^{n-1} open intervals from the set C_{n-1} to obtain C_n , which consists of 2^n closed intervals.

- The “middle third Cantor set” is the limit of this process.

Remark 1. (Generalization of this process). Let $\gamma_n \in (0, 1)$. They will tell us how much we remove in the n -th step.

Precise Construction and Properties of Cantor Sets (General Case)

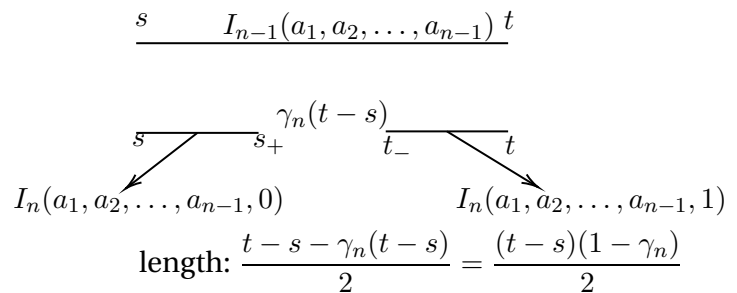
- Fix some constants $0 < \gamma_n < 1$ for $n \in \mathbb{N}$ (note that the middle third Cantor set corresponds to $\gamma_n = \frac{1}{3} \quad \forall n \in \mathbb{N}$).
- Inductively, we define set C_n so that C_n is a union of 2^n closed interval:

$$C_n = \bigcup_{a_1, \dots, a_n \in \{0,1\}} I_n(a_1, a_2, \dots, a_n).$$

- We start with

$$C_0 = [0, 1] = I_0.$$

We define I_n inductively.



If

$$I_{n-1}(a_1, a_2, \dots, a_{n-1}) = [s, t],$$

then we define intervals $I_n(a_1, a_2, \dots, a_n)$ for $a_1, \dots, a_n \in \{0, 1\}$ by

$$\begin{aligned}\min I_n(a_1, a_2, \dots, a_{n-1}, 0) &= s \\ \max I_n(a_1, a_2, \dots, a_{n-1}, 0) &= s_+ = s + \frac{1 - \gamma_n}{2}(t - s) \\ \min I_n(a_1, a_2, \dots, a_{n-1}, 1) &= t_- = t - \frac{1 - \gamma_n}{2}(t - s) \\ \max I_n(a_1, a_2, \dots, a_{n-1}, 1) &= t.\end{aligned}$$

- Note that

$$I_n(a_1, a_2, \dots, a_n) \subset I_{n-1}(a_1, a_2, \dots, a_{n-1}) \quad (9)$$

$$\max I_n(a_1, a_2, \dots, a_{n-1}, 0) < \min I_n(a_1, a_2, \dots, a_{n-1}, 1) \quad (10)$$

- So, $C_n \subset C_{n-1}$ and each C_n is closed.
- We can define

$$C = \bigcap_{n \in \mathbb{N}} C_n.$$

Definition 8.1.2 (Lexicographic Order). The *lexicographic order* on the set $S = \{0, 1\}^n$ for some $n \in \mathbb{N}$ or on the set $S = \{0, 1\}^{\mathbb{N}}$ is defined by the following: for any $a, b \in S$ [they are sequence of 0's and 1's], $a \prec b$ if $\exists k$ s.t.

- $a_j = b_j \quad j < k$, and
- $a_k < b_k$.

Moreover, we define $a \preceq b$ if $a \prec b$ or $a = b$.

Example 8.1.3

$a, b \in \{0, 1\}^5$, and

$$a = (1, 0, 1, \boxed{1}, 0)$$

$$b = (1, 0, 1, \boxed{0}, 1)$$

So, $b \prec a$, and the k in definition is 4.

Lemma 8.1.4 The lexicographic order is a total order on S . i.e., it is reflexive, anti-symmetric, and transitive, and for any $a, b \in S$, $a \preceq b$ or $b \preceq a$.

Lemma 8.1.5 For any $n \in \mathbb{N}$,

$$(a_1, \dots, a_n) \prec (b_1, \dots, b_n) \implies \max I_n(a_1, \dots, a_n) < \min I_n(b_1, \dots, b_n).$$

In particular, if $(a_1, \dots, a_n) \neq (b_1, \dots, b_n)$, then

$$I_n(a_1, \dots, a_n) \cap I_n(b_1, \dots, b_n) = \emptyset.$$

Proof2. Assume $(a_1, \dots, a_n) \prec (b_1, \dots, b_n)$. Then, $\exists k$ s.t. $a_j = b_j \quad \forall j < k$, and $a_k < b_k$. Thus, $a_k = 0$ and $b_k = 1$. By the nesting property (9), we have

- $I_n(a_1, \dots, a_n) \subset I_k(a_1, \dots, a_{k-1}, 0)$

$$\implies \max I_n(a_1, \dots, a_n) < \max I_k(a_1, \dots, a_{k-1}, 0).$$

- $I_n(b_1, \dots, b_n) \subset I_k(a_1, \dots, a_{k-1}, 1)$

$$\implies \min I_n(b_1, \dots, b_n) > \min I_k(a_1, \dots, a_{k-1}, 1).$$

From (10), we then have

$$\max I_k(a_1, \dots, a_{k-1}, 0) < \min I_k(a_1, \dots, a_{k-1}, 1).$$

Therefore,

$$\begin{aligned} \max I_n(a_1, \dots, a_n) &< \max I_k(a_1, \dots, a_{k-1}, 0) \\ &< \min I_k(a_1, \dots, a_{k-1}, 1) \\ &< \min I_n(b_1, \dots, b_n). \end{aligned}$$

Q.E.D. ■

Lemma 8.1.6 Any sequence $a \in \{0, 1\}^{\mathbb{N}}$ corresponds to some element $x \in C$ in the sense that

$$\bigcap_{n \in \mathbb{N}} I_n(a_1, \dots, a_n) = \{x\}.$$

Proof3. By construction, (9) holds. i.e., these closed intervals are nested, we also know that this intersection is non-empty. By construction, the diameter

$$\text{diam } I_n(a_1, \dots, a_n) < \frac{1}{2} \text{diam } I_{n-1}(a_1, \dots, a_{n-1}).$$

So, diameters of the sets above go to zero as $n \rightarrow \infty$. Therefore, the intersection is a single point $\{x\}$. Since $\{x\} \in I_n(a_1, \dots, a_n) \quad \forall n \in \mathbb{N}$, we know that

$$x \in \bigcap_{n \in \mathbb{N}} C_n = C.$$

Q.E.D. ■

The previous Lemma allows us to define the following mapping

$$\tau : \{0, 1\}^{\mathbb{N}} \rightarrow C \quad \text{s.t.} \quad \{\tau(a)\} = \bigcap_{n \in \mathbb{N}} I_n(a_1, \dots, a_n).$$

We think of $(a_n)_{n=1}^{\infty}$ as coordinates of the point $\tau(a) \in C$.

Lemma 8.1.7 The mapping $\tau : \{0, 1\}^{\mathbb{N}} \rightarrow C$ is an order-preserving bijection. That is, it is a bijection, and

$$a \prec b \implies \tau(a) < \tau(b).$$

Proposition 8.1.8 The Cantor set is an uncountable closed subset of \mathbb{R} without isolated points and with empty interior.

Proof 4.

- C is closed because it's an intersection of closed sets.
- Since $\tau : \{0, 1\}^{\mathbb{N}} \rightarrow C$ is a bijection, and $\{0, 1\}^{\mathbb{N}}$ is uncountable, we have C is also uncountable.
- To see that C has no isolated points, let $x = \tau(a) \in C$. For any $n \in \mathbb{N}$, choose $b_k = a_k$ for $k \neq n$ and $b_n \neq a_n$. Then, $x = \tau(a) \neq \tau(b)$, and

$$|x - \tau(b)| < \frac{1}{2^n}$$

- C has empty interior: for any $n \in \mathbb{N}$, if $(\alpha, \beta) \subset C_n$, then $\beta - \alpha < \frac{1}{2^n}$.

$$\implies (\alpha, \beta) \subset C \implies \beta - \alpha < \frac{1}{2^n} \quad \forall n \in \mathbb{N}$$

$$\implies \beta - \alpha \leq 0$$

$\implies C$ does not contain any interval

$\implies C$ has empty interior.

Q.E.D. ■

Remark.

- $\bar{C} = C$, and $\partial C = \bar{C} \setminus \text{int } C = C$.
- ∂C_n is finite, but $\bigcup \partial C_n \neq \partial C$.

Proposition 8.1.9 The Lebesgue measure of the middle third Cantor set is zero.

Proof 5. If $\gamma_n = \frac{1}{3}$ $\forall n \in \mathbb{N}$, C_n consists of 2^n intervals of length $\frac{1}{3^n}$. Then,

$$m(C_n) = \frac{2^n}{3^n},$$

and

$$\lim_{n \rightarrow \infty} m(C_n) = 0 \implies m(C) = 0.$$

Q.E.D. ■

Proposition 8.1.10 In the special case $\left(\gamma_n = \frac{1}{3} \quad \forall n \in \mathbb{N}\right)$,

$$\tau(a) = x \iff x = \sum_{k=1}^{\infty} \frac{2a_k}{3^k}.$$

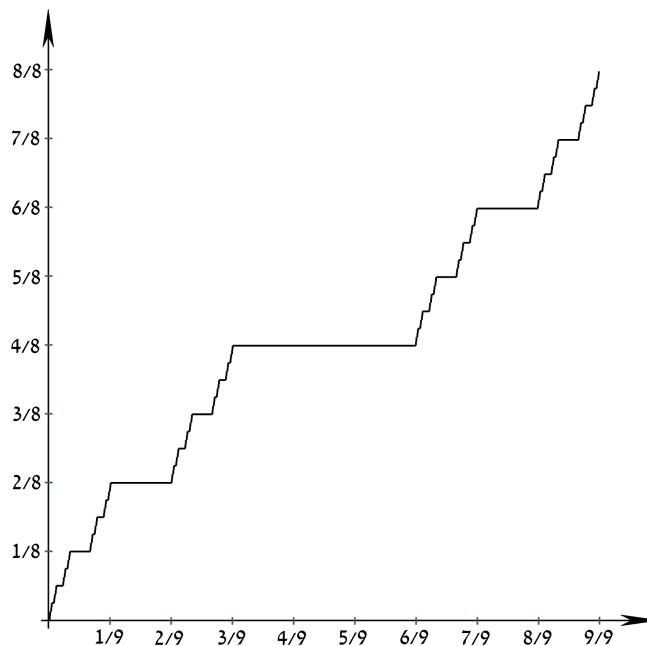
Definition 8.1.11 (Cantor Function). Define $\beta : C \rightarrow [0, 1]$ by

$$\beta(x) = \sum_{k=1}^{\infty} \frac{a_k}{2^k} \quad \text{if } \tau(a) = x.$$

That is,

$$x = \sum_{k=1}^{\infty} \frac{2a_k}{3^k} \implies \beta(x) = \sum_{k=1}^{\infty} \frac{a_k}{2^k}.$$

β can be extended to a continuous function on $[0, 1]$ by asking that β is constant on the removed intervals.



Proposition 8.1.12 $\beta : [0, 1] \rightarrow [0, 1]$ has the following properties:

- increasing,
- onto,
- continuous,

- $\beta'(x) = 0$ on $[0, 1] \setminus C$, and
- for all $x \in C$,

$$\lim_{\varepsilon \downarrow 0} \frac{\beta(x + \varepsilon) - \beta(x - \varepsilon)}{2\varepsilon} = \infty.$$

Proof 6. ①-④ are trivial to see.

⑤: [Goal: produce a lower bound that is ∞ .] Start with

$$x = \sum_{k=1}^{\infty} \frac{2a_k}{3^k}.$$

Let

$$x_n = \sum_{k=1}^n \frac{2a_k}{3^k}$$

$$y_n = \sum_{k=1}^n \frac{2a_k}{3^k} + \sum_{k=n+1}^{\infty} \frac{2}{3^k}.$$

Then, $x_n, y_n \in C$, and

$$\beta(y_n) - \beta(x_n) = \sum_{k=n+1}^{\infty} \frac{1}{2^k} = \frac{1}{2^n}.$$

Note that

$$y_n - x_n = \sum_{k=n+1}^{\infty} \frac{2}{3^k} = \frac{1}{3^n}.$$

If $3^{-n} \leq \varepsilon \leq 3^{-n+1}$,

$$x - \varepsilon \leq x_n \leq x \leq y_n \leq x + \varepsilon$$

$$\beta(x + \varepsilon) - \beta(x - \varepsilon) \geq \beta(y_n) - \beta(x_n) = \frac{1}{2^n}.$$

So,

$$\frac{\beta(x + \varepsilon) - \beta(x - \varepsilon)}{2\varepsilon} \geq \frac{\frac{1}{2^n}}{2 \cdot 3^{-n+1}} = \frac{1}{6} \left(\frac{3}{2} \right)^n \rightarrow \infty \text{ as } n \rightarrow \infty.$$

Hence,

$$\lim_{\varepsilon \downarrow 0} \frac{\beta(x + \varepsilon) - \beta(x - \varepsilon)}{2\varepsilon} = \infty.$$

Q.E.D. ■

Corollary 8.1.13 The length of the graph of $\beta(x)$ is 2.

Remark 7. (Cantor Function as a Counterexample of Fundamental Theorem of Calculus).

For $f \in C^1([0, 1])$, by Fundamental Theorem of Calculus,

$$f(1) - f(0) = \int_0^1 f'(x) dx.$$

For Lebesgue *a.e.* $x \in [0, 1]$, we know that $x \in [0, 1] \setminus C$ (because $m(C) = 0$), and $\beta'(x) = 0$. However,

$$\beta(1) - \beta(0) \neq \int_0^1 \beta'(x) \, dx.$$

How to fix this? *Absolute continuity*.

Definition 8.1.14 (Absolute Continuity). $f : [a, b] \rightarrow \mathbb{C}$ is called *absolutely continuous* if $\forall \varepsilon > 0, \exists \delta > 0$ s.t. $\forall n, \forall (c_j, d_j)$,

$$\sum_{j=1}^n |d_j - c_j| < \delta \implies \sum_{j=1}^n |f(d_j) - f(c_j)| < \varepsilon.$$

Theorem 8.1.15

If f is absolute continuous, then $f'(x)$ exists for Lebesgue *a.e.* x , $f' \in L^1([a, b])$, and

$$f(b) - f(a) = \int_a^b f'(x) \, dx.$$

Unfortunately, $\beta(x)$ is not absolutely continuous.

Definition 8.1.16 (Cantor Measure).

$$\mu_\beta((x, y]) = \beta(y) - \beta(x)$$

Definition 8.1.17 (Mutually Singular). Measures μ, ν on the same space are said to be *mutually singular* if \exists a set A such that $\mu(A) = 0$ and $\nu(A^c) = 0$.

Proposition 8.1.18 Cantor Measure and Lebesgue measure are mutually singular.

Proof.

$$\begin{aligned} \mu_\beta((-\infty, 0]) &= 0 \\ \mu_\beta((1, \infty)) &= 0 \\ \mu_\beta\left(\left(\frac{1}{3}, \frac{2}{3}\right)\right) &= 0 \end{aligned}$$

So, we have that

$$\mu_\beta(\mathbb{R} \setminus C) = 0 \quad \text{and} \quad m(C) = 0.$$

Q.E.D. ■