Real Analysis III

Abstract Measure

Author: Kumiko $\in$ Maki's Lab

Institute: Maki's Lab

Date: 2022 summer

Powered by ElegantBook: `https://github.com/ElegantLaTeX/ElegantBook`.

Contents

Chapter 1 A Brief History of Analysis

1.1 Fourier Series

1.2 Riemann's Habilitationsschrift[1]

1.3 Some Deficiencies in Riemann Integration

1.4 Jordan Content

1.5 The Men Who Changed Integration

[1] habilitation thesis (qualification to become an instructor)

Chapter 2 Lebesgue Measure on $\mathbb{R}$

2.1 Building Blocks

The most natural objects to which can be assigned the notion of a measure are intervals in $\mathbb{R}$, squares in $\mathbb{R}^2$, and cubes in $\mathbb{R}^3$.

Definition 2.1 (rectangles)

A (closed) ***rectangle*** *in* $\mathbb{R}^d$ *is a product of closed intervals* $[a_1, b_1] \times \cdots [a_d, b_d]$. ♣

If $-\infty < a < b < \infty$,

$$(a, b), [a, b], (a, b], [a, b)$$

2.2 Lebesgue Outer Measure

2.3 Measurability and Lebesgue Measure

2.4 Lebesgue σ-Algebra

2.5 Cantor Set

2.6 Nonmeasurable Sets

Chapter 3 Abstract Measures

3.1 Measures and Set-Algebraic Structures

Introduction

- *lebesgue measure review*
- *generating a σ-algebra*
- *axioms of a measure*
- *properties of measures*
- *Dynkin's system*

The more general definition of a measure should look like:

Definition 3.1 (a possible definition)

A set function μ is called a measure if it satisfies:

1. $\cdots$
2. $\cdots$
3. $\cdots$

♣

This is called the "**axiomatic fashion**", the items $1, 2, 3$ are called axioms. The core idea is when we study a mathematical object, we care about what it does rather what it is.

Example 3.1 In linear algebra, the identity element with respect to the addition in a vector space V is an element a such that $a + v = v + a = v$ for all $v \in V$. We usually denote the identity element by 0.

When I was learning linear algebra for the first time, I could not help recognize this 0 as the real number "0". It turns out that any element which has no contributions in addition is the identity element, and 0 is just a notation!

Now let's focus on what should a measure μ do (or what properties it should satisfy).

1. μ should be able to "measure" the empty set, and give a result of 0.
2. μ should possess the countable additivity, as shown in the above two examples.
3. What else?

We cannot go through every property of the Lebesgue measure m and simply change the letter from m to μ to get measure axioms, which leads to the loss of generality. The measure μ will be defined on a collection of subsets of X, where X is just a set (without any structure[1] a priori!). Hence, we can delete the regularity properties[2] from the candidates. Another important property which is helpful in computation is the lower and upper continuity. However, the proof of this property only uses the countable additivity of the Lebesgue measure: it is a corollary of the countable additivity, so it need not be a measure axiom.

Finally, a function needs a domain. The collection of Lebesgue measurable sets satisfies some closure condition:

1. If $\{E_n : n \in \mathbb{N}\}$ is a family of Lebesgue measurable sets, then $\bigcup_{n=1}^{\infty} E_n$ is Lebesgue measurable.

[1] such as topology, metric, norm, etc.

[2] A property related to measurable sets and open, closed, compact sets

2. If E is Lebesgue measurable, then so is E^c.
3. If $\{E_n : n \in \mathbb{N}\}$ is a family of Lebesgue measurable sets, then $\bigcap_{n=1}^{\infty} E_n$ is Lebesgue measurable.

Since taking complements on union yields intersection, we define a collection of subsets of X that is closed under some set operations.

Definition 3.2 (σ-algebra)

Let X be a set. A σ-algebraof sets on X is a nonempty collection $\mathcal{M}$ of subsets of X such that

1. *If $\{A_n : n \in \mathbb{N}\} \subset \mathcal{M}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$.*
2. *If $E \in \mathcal{M}$, then $E^c \in \mathcal{M}$.*

♣

Once we have a reasonable domain, we can define a set function on it.

Definition 3.3 (measure)

*Let X be a set equipped with a σ-algebra$\mathcal{M}$. A **measure** on $\mathcal{M}$ is a function $\mu : \mathcal{M} \to [0, \infty]$ such that*

1. $\mu(\varnothing) = 0$,
2. *For any sequence of disjoint sets $\{E_n : n \in \mathbb{N}\} \subset \mathcal{M}$, $\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n)$.*

♣

The next two subsections deal with some properties of σ-algebras and measures.

3.1.1 σ-Algebras

We begin by looking some examples of σ-algebras.

Example 3.2 If X is any set, then $\mathcal{P}(X)$ and $\{\varnothing, X\}$ are σ-algebras.

Example 3.3 If X is uncountable, then $\mathcal{A} = \{E \subset X : E \text{ is countable or } E^c \text{ is countable}\}$ is a σ-algebra, called the **σ-algebra of countable or co-countable sets**.

Proof Let $E \in \mathcal{A}$, then E^c is countable or $(E^c)^c$ is countable, so $E^c \in \mathcal{A}$. Let $\{E_n\}_{n\in\mathbb{N}} \subset \mathcal{A}$, and let $\mathcal{I} = \{n \in \mathbb{N} : E_n \text{ is countable}\}$, $\mathcal{J} = \{n \in \mathbb{N} : E_n^c \text{ is countable}\}$. Then

$$\bigcup_{n=1}^{\infty} E_n = \bigcup_{i\in\mathcal{I}} E_i \cup \bigcup_{j\in\mathcal{J}} E_j,$$

$$\left(\bigcup_{n=1}^{\infty}\right)^c = \left(\bigcup_{i\in\mathcal{I}} E_i\right)^c \cup \left(\bigcup_{j\in\mathcal{J}} E_j\right)^c.$$

Since $\bigcap_{j\in\mathcal{J}} E_j^c$ is countable, it follows that $\left(\bigcup_{n=1}^{\infty} E_n\right)^c$ is countable.

Next we introduce a concept that will be frequently used in the future. It is difficult or even impossible to give a complete description of a σ-algebra. In $\mathbb{R}^3$ we only need 3 vectors to describe every vector by taking linear combinations. The idea is to use something simpler to represent the complex structure. We make an analogy:

	simple set	operations
linear algebra	basis	linear combination
real analysis	easy-to-describe sets	countable union and complement

Recall that the span of vectors $v_1, \cdots, v_n$ is the set of all linear combinations of them, or the smallest vector space containing $v_1, \cdots, v_n$. Here, "smallest" means the intersection of all vector spaces containing $v_1, \cdots, v_n$. Using the same idea, we can represent a σ-algebra through simpler sets.

Definition 3.4

*Let $\mathcal{E}$ be a collection of subsets of X. The intersection of all σ-algebras containing $\mathcal{E}$ is called the σ-algebra **generated by** $\mathcal{E}$, and denoted $\sigma(\mathcal{E})$(or $\mathcal{M}(\mathcal{E})$).* ♣

Example 3.4 The **Borel σ-algebra** on $\mathbb{R}$ is the σ-algebragenerated by open sets of $\mathbb{R}$.

Remark It is a bit hard to imagine what a countable union of open sets looks like. In fact, the generating sets can be made even much simpler: open intervals $\{(a, b) : a < b\}$. To see this, we will need the following useful result.

Proposition 3.1

If $\mathcal{E} \subset \mathcal{M}$, where $\mathcal{M}$ is a σ-algebra, then $\sigma(\mathcal{E}) \subset \mathcal{M}$. ♠

Proof $\sigma(\mathcal{E})$ is the smallest σ-algebracontaining $\mathcal{E}$, so it is contained in $\mathcal{M}$.

Definition 3.5

*If X is any topological space, the σ-algebra generated by the family of open sets in X is called the **Borel σ-algebra** on X and is denoted by $\mathcal{B}_X$. Its members are called **Borel sets**.* ♣

The Borel σ-algebraon $\mathbb{R}$ is of vital importance. By definition it is generated by the family of open sets in $\mathbb{R}$, but we can find simpler generating families.

Proposition 3.2

$\mathcal{B}_{\mathbb{R}}$ is generated by each of the following:

1. *the open intervals: $\mathcal{E}_1 = \{(a, b) : a < b\}$,*
2. *the closed intervals: $\mathcal{E}_2 = \{[a, b] : a < b\}$,*
3. *the half-open intervals: $\mathcal{E}_3 = \{(a, b] : a < b\}$ or $\mathcal{E}_4 = \{[a, b) : a < b\}$,*
4. *the open rays: $\mathcal{E}_5 = \{(a, \infty) : a \in \mathbb{R}\}$ or $\mathcal{E}_6 = \{(-\infty, a) : a \in \mathbb{R}\}$,*
5. *the closed rays: $\mathcal{E}_7 = \{[a, \infty) : a \in \mathbb{R}\}$ or $\mathcal{E}_8 = \{(-\infty, a] : a \in \mathbb{R}\}$.* ♠

Proof

1. Let $\mathcal{G}$ be the family of open sets in $\mathbb{R}$. (a, b) is clearly in $\mathcal{G}$, so $\mathcal{E}_1 \subset \mathcal{G}$. Then $\sigma(\mathcal{E}_1) \subset \sigma(\mathcal{G}) = \mathcal{B}_{\mathbb{R}}$. Conversely, if G is an open set, then $G = \bigcup_{n=1}^{\infty}(a_n, b_n) \in \sigma(\mathcal{E}_1)$, so $\mathcal{G} \subset \sigma(\mathcal{E}_1)$, hence $\mathcal{B}_{\mathbb{R}} = \sigma(\mathcal{G}) \subset \mathcal{E}_1$.
2. Observe that
$$[a, b] = \bigcap_{n=1}^{\infty}\left(a - \frac{1}{n}, b + \frac{1}{n}\right),$$
$$(a, b) = \bigcup_{n=1}^{\infty}\left[a + \frac{1}{n}, b - \frac{1}{n}\right].$$
Then $\mathcal{E}_2 \subset \sigma(\mathcal{E}_1)$ and $\mathcal{E}_1 \subset \sigma(\mathcal{E}_2)$, so $\sigma(\mathcal{E}_2) = \sigma(\mathcal{E}_1) = \mathcal{B}_{\mathbb{R}}$.
The left are exercises.

3.1.2 Measures

If X is a set and $\mathcal{M} \subset \mathcal{P}(X)$ is a σ-algebra, $(X, \mathcal{M})$ is called a **measurable space**. The sets in $\mathcal{M}$ are called **measurable sets**. In Real analysis I, Lebesgue measurable sets are those behave well in geometric sense. Now, if a set is in a σ-algebra, then it is measurable! Once we have a σ-algebra, we can define a measure on it, then we get a triple $(X, \mathcal{M}, \mu)$ which is called a **measure space**.

Example 3.5 (counting measure) Let $X = \mathbb{N}$ be any nonempty set, let $\mathcal{M} = \mathcal{P}(\mathbb{N})$, define $\mu : \mathcal{M} \to [0, \infty]$ by $\mu(E) = |E|$(the cardinality of E) if E is finite, and $\mu(E) = \infty$ if E is infinite. For example,

$$\mu(\{1\}) = 1, \quad \mu(\{1,2\}) = 2, \quad \mu(\{1,2,\cdots\}) = \infty.$$

More generally, let X be any nonempty set, $\mathcal{M} = \mathcal{P}(X)$. Define μ on $\mathcal{M}$ by $\mu(\{x\}) = 1$ for each $x \in X$. μ is called the **counting measure**.

Example 3.6 (Lebesgue measure) You are very familiar with this!

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

Theorem 3.1 (monotonicity and subadditivity)

1. *If $E, F \in \mathcal{M}$ and $E \subset F$, then $\mu(E) \leq \mu(F)$.*
2. *If $\{E_n : n \in \mathbb{N}\} \subset \mathcal{M}$, then $\mu(\bigcup_{n=1}^\infty E_n) \leq \sum_{n=1}^\infty \mu(E_n)$.*

♡

Proof

1. $\mu(F) = \mu(E \cup (F \setminus E)) \geq \mu(E)$.
2. Trivial.

Theorem 3.2 (continuity of measures)

1. *If $\{E_n : n \in \mathbb{N}\}$ and $E_1 \subset E_2 \subset \cdots$, then $\mu(\bigcup_{n=1}^\infty E_n) = \lim_{n\to\infty} \mu(E_n)$.*
2. *If $\{E_n : n \in \mathbb{N}\}, E_1 \supset E_2 \supset \cdots$, and $\mu(E_1) < \infty$, then $\mu(\bigcap_{n=1}^\infty E_n) = \lim_{n\to\infty} \mu(E_n)$.*

♡

Proof Let $F_n = F_{n+1} \setminus F_n$, then $\{F_n\}_{n\in\mathbb{N}}$ is disjoint, and $\bigcup_{n=1}^\infty E_n = \bigcup_{k=1}^\infty F_k$, so

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^\infty E_n\right) &= \mu\left(\bigcup_{k=1}^\infty F_k\right) \\
&= \sum_{k=1}^\infty \mu(F_k) \\
&= \lim_{N\to\infty} \sum_{k=1}^N \mu(F_k) \\
&= \lim_{N\to\infty} \mu\left(\bigcup_{k=1}^N F_k\right) \\
&= \lim_{N\to\infty} \mu(E_N).
\end{aligned}$$

Next, let $F_j = E_1 \setminus E_j$, then $F_1 \subset F_2 \subset \cdots$, $\mu(E_1) = \mu(F_j) + \mu(E_j)$, and $\bigcup_{j=1}^\infty F_j =$

$E_1 \setminus (\bigcap_{j=1}^{\infty} E_j)$. Then,

$$\mu(E_1) = \mu\left((\bigcap_{j=1}^{\infty} E_j)\right) + \lim_{j\to\infty} \mu(F_j) = \mu\left((\bigcap_{j=1}^{\infty} E_j)\right) + \lim_{j\to\infty} (\mu(E_1) - \mu(E_j)).$$

Since $\mu(E_1) < \infty$, we have the desired result.

Definition 3.6

*A set $E \in \mathcal{M}$ with $\mu(E) = 0$ is called a null set. If $\mu(E) = 0$ and $F \subset E$, then $\mu(F)$ should equal to 0, but F is not necessarily in $\mathcal{M}$. A measure whose domain includes all subsets of null sets is called **complete**.* ♣

Theorem 3.3

Suppose that $(X, \mathcal{M}, \mu)$ is a measure space. Let $\mathcal{N} = \{N \in \mathcal{M} : \mu(N) = 0\}$ and $\overline{\mathcal{M}} = \{E \cup F : E \in \mathcal{M}, F \subset N \text{ for some } N \in \mathcal{N}\}$. Then $\overline{\mathcal{M}}$ is a σ-algebra, and there is a unique extension $\overline{\mu}$ of μ to a complete measure on $\overline{\mathcal{M}}$. ♡

Definition 3.7

*$\overline{\mu}$ is called the **completion** of μ, and $\overline{\mathcal{M}}$ is called the **completion** w.r.t. μ.* ♣

3.1.3 More Set-Algebraic Structures; Dynkin System

We have learned certain algebraic structures in linear algebra and modern algebra courses. Typical algebraic structures include vector spaces, groups, rings, fields, and modules. An algebraic structure is a set with some operations on it satisfying some axioms.

Example 3.7 A vector space V is a set V with addition and scalar multiplication which satisfy 8 properties.

Example 3.8 A group is a set G with a binary operation $\cdot$ such that

1. $a \cdot (b \cdot c) = (a \cdot b) \cdot c, \forall a, b, c \in G$.
2. There is an identity element 1 such that $1 \cdot a = a \cdot 1 \; \forall a \in G$.
3. For each $a \in G$ there exists $b \in G$ such that $a \cdot b = b \cdot a = 1$, and b is called an inverse of a.

A set-algebraic structure is a family of subsets of X(a set) that is closed under some set operations. For example, a σ-algebra is a family $\mathcal{M}$ of subsets of X that is closed under complement and countable union. There are a lot of set-algebraic structures, and we will meet them in the future. You can even create your own structures by arranging closure conditions on some set operations.

Example 3.9 A family of sets $\mathcal{R} \subset \mathcal{P}(X)$ is called a **ring** if it is closed under finite unions and differences:

1. If $E_1, \cdots, E_n \in \mathcal{R}$, then $\bigcup_{j=1}^{n} E_j \in \mathcal{R}$.
2. If $E, F \in \mathcal{R}$, then $E \setminus F \in \mathcal{R}$.

A ring that is closed under countable unions is called a σ-**ring**.

Now we study a structure that has many applications in measure theory and probability theory, called the Dynkin system (or λ-system)[3].

[3] The following part is from Homework 1 of Math 721 Real Analysis I, Fall 2022, UW-Madison, taught by Andreas

Definition 3.8 (Dynkin system)

A Dynkin-system [1]*$\mathcal{D}$ on X is a collection of subsets of X which has the following properties.*
(i) $X \in \mathcal{D}$.
(ii) If $A \in \mathcal{D}$ then its complement $A^c := X \backslash A$ belong to $\mathcal{D}$.
(iii) If A_n is a sequence of mutually disjoint sets in $\mathcal{D}$ then $\cup_{n=1}^{\infty} A_n \in \mathcal{D}$. ♣

Observe that every σ-algebra is a Dynkin system.

Historical Notes [4] Eugene Borisovich Dynkin (11 May 1924 – 14 November 2014) was a Soviet and American mathematician. He made contributions to the fields of probability and algebra, especially semisimple Lie groups, Lie algebras, and Markov processes. The Dynkin diagram, the Dynkin system, and Dynkin's lemma are named after him.

A1. In the literature one can also find a definition with alternative axioms (i), (ii)* , (iii)* where again
(i) $X \in \mathcal{D}$, and
(ii)* If A, B are in $\mathcal{D}$ and $A \subset B$ then $B \backslash A \in \mathcal{D}$.
(iii)* If $A_n \in \mathcal{D}$, $A_n \subset A_{n+1}$ for all $n = 1, 2, 3, \ldots$ then also $\cup_{n=1}^{\infty} A_n \in \mathcal{D}$.
Prove that the definition with (i), (ii), (iii) is equivalent with the definition with (i), (ii)*, (iii)*

Proof **Original axioms imply new axioms**: Let $A, B \in \mathcal{D}$ with $A \subset B$, then $(B \setminus A)^c = (B \cap A^c)^c = B^c \cup A \in \mathcal{D}$ since the union is disjoint. By (ii), $B \setminus A \in \mathcal{D}$. Let A_n be an increasing sequence of sets in $\mathcal{D}$. The ideal is to "disjointify" $\{A_n\}_{n \in \mathbb{N}}$ and preserve the countable union. Let $B_{n+1} = A_{n+1} \setminus A_n$ and $B_1 = A_1$, then $\{B_n\}_{n \in \mathbb{N}}$ is disjoint and $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. By (iii), $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.
New axioms imply original axioms: Let $A \in \mathcal{D}$, then $X \setminus A \in \mathcal{D}$, so (ii) holds. Let $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{D}$ be disjoint, then $B_n = \bigcup_{k=1}^{n} A_n$ is increasing, so $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathcal{D}$. □

A2. Verify: If $\mathcal{E}$ is any collection of subsets of X then the intersection of all Dynkin-systems containing $\mathcal{E}$ is a Dynkin system containing $\mathcal{E}$. It is the smallest Dynkin system containing $\mathcal{E}$. We call it the Dynkin-system generated by $\mathcal{E}$, and denote it by $\mathcal{D}(\mathcal{E})$.

Proof Write $\mathcal{D}(\mathcal{E}) = \bigcap_{i \in \mathcal{I}} \mathcal{D}_i$, where each $\mathcal{D}_i$ is a Dynkin system containing $\mathcal{E}$.

1. $X \in \mathcal{D}_i \ \forall i \in \mathcal{I} \implies X \in \bigcap_{i \in \mathcal{I}} D_i$.
2. Let $A \in \bigcap_{i \in \mathcal{I}} D_i$, then $A \in \mathcal{D}_i \ \forall i \in \mathcal{I}$, then $A^c \in \mathcal{D}_i \ \forall i \in \mathcal{I}$, hence $A^c \in \bigcap_{i \in \mathcal{I}} D_i$.
3. Let $\{A_n\}_{n \in \mathbb{N}} \subset \bigcap_{i \in \mathcal{I}} D_i$ be disjoint, then $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{D}_i \ \forall i \in \mathcal{I}$, hence $\bigcup_{n=1}^{\infty} A_n \subset \mathcal{D}_i \ \forall i \in \mathcal{I}$, so $\bigcup_{n=1}^{\infty} A_n \in \bigcap_{i \in \mathcal{I}} D_i$.

If $\mathcal{M}$ is any Dynkin system containing $\mathcal{E}$, then $\mathcal{M} \supset \bigcap_{i \in \mathcal{I}} D_i = \mathcal{D}(\mathcal{E})$. This is what "smallest" means. □

Definition: A collection $\mathcal{A}$ of subsets of X is $\cap$-stable if for $A \in \mathcal{A}$ and $B \in \mathcal{A}$ we also have $A \cap B \in \mathcal{A}$. Observe that a $\cap$-stable system is stable under finite intersections.

A3. (i) Show that if $\mathcal{D}$ is a $\cap$-stable Dynkin system, then the union of two sets in $\mathcal{D}$ is again

Seeger

[4] https://en.wikipedia.org/wiki/Eugene_Dynkin

in $\mathcal{D}$.
(ii) Prove: A Dynkin-system is a σ-algebra if and only if it is $\cap$-stable.

Proof (i) Let $A, B \in \mathcal{D}$. Observe that $(A \cup B)^c = A^c \cap B^c \in \mathcal{D}$ since $\mathcal{D}$ is $\cap$-stable.
(ii) A σ-algebra is clearly $\cap$-stable. Conversely, let the Dynkin system $\mathcal{D}$ be $\cap$-stable and let $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{D}$. Let $B_n = \bigcup_{k=1}^n A_n$, then $B_n \in \mathcal{D}$ since $\mathcal{D}$ is closed under finite union. Since B_n is increasing, $\bigcup_{n=1}^\infty B_n \in \mathcal{D}$, so $\mathcal{D}$ is a σ-algebra. □

The following theorem turns out to be very useful for the construction of σ-algebras (we may use it later in the proof of Fubini's theorem).

Theorem 3.4

Let $\mathcal{E}$ be any collection of subsets of X which is stable under finite intersections. Then the Dynkin-system $\mathcal{D}(\mathcal{E})$ generated by $\mathcal{E}$ is equal to the σ-algebra $\sigma(\mathcal{E})$ generated by $\mathcal{E}$. ♡

We will work out the following steps:
(i) Argue that it suffices to show that $\mathcal{D}(\mathcal{E})$ is a σ-algebra. By A3 it suffices to show that $\mathcal{D}(\mathcal{E})$ is $\cap$-stable.
(ii) Fix a set $B \in \mathcal{D}(\mathcal{E})$. Prove that the system

$$\Gamma_B = \{A \subset X : A \cap B \in \mathcal{D}(\mathcal{E})\}$$

is a Dynkin system.
(iii) Prove that $\mathcal{E} \subset \Gamma_B$ for all $B \in \mathcal{E}$, and hence $\mathcal{D}(\mathcal{E}) \subset \Gamma_B$ for all $B \in \mathcal{E}$.
(iv) Prove that $\mathcal{E} \subset \Gamma_B$ even for all $B \in \mathcal{D}(\mathcal{E})$, and hence $\mathcal{D}(\mathcal{E}) \subset \Gamma_B$ for all $B \in \mathcal{D}(\mathcal{E})$. Conclude.

Proof (i) If $\mathcal{D}(\mathcal{E})$ is a σ-algebra, then $\mathcal{D}(\mathcal{E}) \supset \sigma(\mathcal{E})$. Since $\sigma(\mathcal{E})$ itself is a Dynkin system containing $\mathcal{E}$, $\mathcal{D}(\mathcal{E}) \subset \sigma(\mathcal{E})$.
(ii) $X \cap B = B \in \mathcal{D}(\mathcal{E}) \implies X \in \Gamma_B$. Let $A \in \Gamma_B$, then $A^c \cap B = B \setminus A = B \setminus (A \cap B) \in \mathcal{D}(\mathcal{E})$, hence $A^c \in \Gamma_B$. Let $\{A_n\}_{n\in\mathbb{N}} \subset \Gamma_B$ be disjoint, then $\bigcup_{n=1}^\infty A_n \cap B = \bigcup_{n=1}^\infty (A_n \cap B) \in \mathcal{D}(\mathcal{E})$. Therefore, Γ_B is a Dynkin system.
(iii) If $A \in \mathcal{E}$, then $A \cap B \in \mathcal{E}$ for all $B \in \mathcal{E}$ since $\mathcal{E}$ is $\cap$-stable, hence $A \cap B \in \mathcal{D}(\mathcal{E})$, that is, $A \in \Gamma_B$. Since $\mathcal{D}(\mathcal{E})$ is minimal, it is contained in Γ_B for all $B \in \mathcal{E}$.
(iv) $\mathcal{D}(\mathcal{E}) \subset \Gamma_B$ implies $A \cap B \in \mathcal{D}(\mathcal{E})$ for all $A \in \mathcal{D}(\mathcal{E})$ and $B \in \mathcal{E}$. In other words, if $B \in \mathcal{E}$, then $B \in \Gamma_A$ for all $A \in \mathcal{D}(\mathcal{E})$. Hence $\mathcal{E} \subset \Gamma_A$ for all $A \in \mathcal{D}(\mathcal{E})$, then$\mathcal{D}(\mathcal{E}) \subset \Gamma_A$ for all $A \in \mathcal{D}(\mathcal{E})$.

3.2 Construction of a Measure

In this section we introduce a universal way of constructing a measure that can be widely applied in analysis and probability theory. As we have seen in Real Analysis I, the construction of the Lebesgue measure on $\mathbb{R}^d$ depends on a geometric observation: every open set in $\mathbb{R}^d$ is a countable union of (almost disjoint) cubes, to which we can assign a natural measure: volume (product of side lengths). We make the following abstraction:

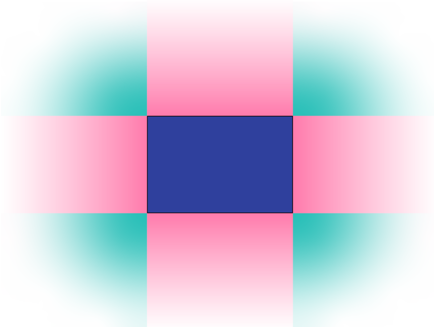
$\mathbb{R}^d$ case	abstraction
cube	elementary figures
volume	a volume function

Here an elementary figure is a set which is easy to deal with and can naturally be assigned a value (think of cubes in $\mathbb{R}^d$). Now let's investigate some set-algebraic properties of rectangles in $\mathbb{R}^d$. A rectangle is the product of intervals:

$$[a_1, b_1] \times \cdots \times [a_n, b_n],$$

where $-\infty \leq a_j, b_j \leq \infty$ for each j. Denote the collection of rectangles in $\mathbb{R}^d$ by $\mathcal{R}$.

1. The intersection of two rectangles is still a rectangle,
2. The complement of a rectangle is a finite disjoint union of rectangles.



We introduce a structure that is more elmentary than σ-algebra.

Definition 3.9

*A collection $\mathcal{S}$ of subsets of X is called a **semialgebra** if*

1. *$\varnothing \in \mathcal{S}$,*
2. *if $E, F \in \mathcal{S}$, then $E \cap F \in \mathcal{S}$,*
3. *if $E \in \mathcal{S}$, then E^c is a finite disjoint union of members of $\mathcal{S}$.*

*A set-valued function ρ on $\mathcal{S}$ is called a **volume** if $\rho(\varnothing) = 0$.* ♣

Let's go back to the Lebesgue measure. After declaring a volume, we assign an arbitrary subset E of $\mathbb{R}^d$ a value called the Lebesgue **outer measure**:

$$m^*(E) = \inf \left\{ \sum_{n=1}^{\infty} |Q_n| \right\},$$

where the infimum is taken over all $\{Q_n\}$ with $E \subset \bigcup_{n=1}^{\infty} Q_n$. Thus we extend the volume to a set-valued function on $\mathcal{P}(\mathbb{R}^d)$. However, this is too large, since not every set in $\mathbb{R}^d$ is measurable, so the last step is to restrict m^*, as we have seen in Real Analysis I. In general, the construction of a measure follows the following steps:

(volume, semialgebra) $\to$ (premeasure, algebra) $\to$ (outer measure, power set) $\to$ (measure, σ-algebra)

3.2.1 (volume, semialgebra)

In practice, we will often have a priori knowledge of elementary sets and a volume.

Example 3.10 In $\mathbb{R} \cup \{+\infty\}$, $\{(a, b] : -\infty \leq a, b \leq \infty\}$ is a semialgebra and the length of an interval is a volume.

3.2.2 (premeasure, algebra)

We want to extend the volume to a larger class. Before reaching countable unions, we would consider finite unions. We slightly weaken an assumption in the definition of σ-algebra.

Definition 3.10 (algebra)

*A nonempty collection $\mathcal{A}$ of subsets of X is called an **algebra** if*

1. *If $E_1, \cdots, E_n \in \mathcal{A}$, then $\bigcup_{j=1}^n E_j \in \mathcal{A}$,*
2. *if $E \in \mathcal{A}$, then $E^c \in \mathcal{A}$.*

♣

A semialgebra can be extended to an algebra by taking finite disjoint unions:

Proposition 3.3

If $\mathcal{S}$ is a semialgebra, then the collection $\mathcal{A}$ of finite disjoint unions of members of $\mathcal{S}$ is an algebra. That is,

$$\mathcal{A} = \left\{ \bigcup_{i\in\mathcal{I}} E_i : \mathcal{I} \text{ is finite and } \{E_i\}_{i\in\mathcal{I}} \subset \mathcal{S} \text{ is disjoint} \right\}.$$

♠

Proof Let $A, B \in \mathcal{A}$. We write $A \cup B$ as a disjoint union $A \cup B = (A \setminus B) \cup B$. We have $B^c = \bigcup_{i=1}^n B_i$, so $A \cap B^c = \bigcup_{i=1}^n A \cap B_i \in \mathcal{A}$, hence $A \cup B$ is a finite disjoint union of sets in $\mathcal{S}$. Now let $A_1, \cdots, A_n \in \mathcal{A}$ and suppose that $\bigcup_{i=1}^{n-1} A_i \in \mathcal{A}$, thus it is a disjoint union of sets in $\mathcal{S}$. Then

$$\bigcup_{i=1}^n A_i = \bigcup_{i=1}^{n-1} A_i \cup \left(A_n \setminus \bigcup_{i=1}^{n-1} A_i \right),$$

which is also a disjoint union of sets in $\mathcal{S}$.
We can deduce that $\mathcal{A}$ is closed under finite intersection. Let $A = \bigcup_{i=1}^n A_i, B = \bigcup_{j=1}^m B_j$, then $A \cap B = \bigcup_{j=1}^m (A \cap B_j) = \bigcup_{j=1}^m (\bigcup_{i=1}^n A_i \cap B_j) \in \mathcal{A}$. A similar induction shows the finite intersection is closed in $\mathcal{A}$. Now let $E \in \mathcal{A}$, then $E = \bigcup_{i=1}^n E_i$ with each $E_i \in \mathcal{S} \subset \mathcal{A}$, so $E^c = \bigcap_{i=1}^n E_i^c \in \mathcal{A}$. □

Then, we can extend the volume ρ from $\mathcal{S}$ to $\mathcal{A}$ by setting

$$\rho\left(\bigcup_{j=1}^n E_j \right) = \sum_{j=1}^n \rho(E_j),$$

where $E_j \in \mathcal{S}$ being disjoint.

Definition 3.11

*If $\mathcal{A}$ is an algebra, a function $\mu_0 : \mathcal{A} \to [0, \infty]$ is called a **premeasure** if*

1. *$\mu_0(\varnothing) = 0$,*
2. *if $\{A_j : j \in \mathbb{N}\} \subset \mathcal{A}$ with $\bigcup_{j=1}^\infty A_j \in \mathcal{A}$,(disjoint) then $\mu_0(\bigcup_{j=1}^\infty A_j) = \sum_{j=1}^\infty \mu_0(A_j)$.*

♣

3.2.3 (outer measure, power set)

First we make a generalization of the concept of an outer measure.

Definition 3.12 (outer measure axioms)

*An **outer measure** on a nonempty set X is a function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ that satisfies*

1. $\mu^*(\varnothing) = 0$,
2. $\mu^*(A) \leq \mu^*(B)$ *if* $A \subset B$,
3. $\mu^*(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

♣

Let $\mathcal{E}$ be a collection of sets on which a volume ρ has been defined (at this point we do not require and structure on $\mathcal{E}$) and let $\varnothing, X \in \mathcal{E}$. We cover any set $A \subset X$ by a countable union of "elementary sets" in $\mathcal{E}$, and define

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

where μ is a measure[5]

Proposition 3.4

μ^* is an outer measure.

♠

Proof If $A \subset B$, then any covering of B covers A. This shows the monotonicity.
For each $n \in \mathbb{N}$ let $A_n \subset \bigcup_{k=1}^{\infty} E_{n,k}$ with $\sum_{k=1}^{\infty} \mu(E_{n,k}) \leq \mu^*(A_n) + 2^{-k}\varepsilon$. Then

$$\bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} E_{n,k},$$

and

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

Since ε is arbitrary, we have $\mu^*(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$. □

Since μ^* is not necessarily a measure, we need to exclude some sets from X to obtain a σ-algebra, on which the restriction of μ^* will be a measure. Now we introduce a convenient way to rule out "bad" sets: **Carathéodory's criterion**.

Definition 3.13

*Let μ^* be an outer measure on X. A set $A \subset X$ is called μ^*-**measurable** if*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \subset X.$$

♣

If μ_0 is a premeasure on an algebra $\mathcal{A}$, it induces an outer measure

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu_0(A_j) : A_j \in \mathcal{A}, E \subset \bigcup_{j=1}^{\infty} A_j \right\}.$$

Exercise 3.1 Prove this. Compare the definition between a measure and a premeasure.

From now on we denote the family of μ^*-measurable sets by $\mathcal{M}^*$.

[5] Here we use measure just for convenience. In practice, the outer measure is usually induced by a more elementary set function.

Proposition 3.5

Let $\mathcal{A}$ be an algebra and μ^* given as above.

1. $\mu^*|\mathcal{A} = \mu_0$;
2. *every set in $\mathcal{A}$ is μ^*-measurable.*

Proof Let $A \in \mathcal{A}$, then A covers A, and it is the smallest covering, so $\mu^*(A) = \mu_0(A)$. Let $E \subset X$ be arbitrary and $\varepsilon > 0$. There is a covering $\{B_j\} \subset \mathcal{A}$ of E with $\sum_{j=1}^{\infty} \mu_0(B_j) \leq \mu^*(E)+\varepsilon$. Note that $B_j = (B_j \cap A) \cup (B_j \cap A^c)$, so $\mu_0(B_j) = \mu_0(B_j \cap A) + \mu_0(B_j \cap A^c)$. Then,

$$\mu^*(E) + \varepsilon \geq \sum_{j=1}^{\infty} \mu_0(B_j \cap A) + \sum_{j=1}^{\infty} \mu_0(B_j \cap A) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Since ε is arbitrary, A is μ^*-measurable. □

Remark The family of μ^*-measurable sets contains an algebra. That is, $\mathcal{A} \subset \mathcal{M}^*$, then $\sigma(\mathcal{A}) \subset \sigma(\mathcal{M}^*)$. What do you find?

3.2.4 (measure, σ-algebra)

Theorem 3.5 (Carathéodory's Theorem)

If μ^* is an outer measure on X, the collection $\mathcal{M}^*$ of μ^*-measurable sets is a σ-algebra, and the restriction of μ^* to $\mathcal{M}$ is a complete measure.

Proof

1. First we show that $\mathcal{M}^*$ is an algebra. $\mathcal{M}^*$ is clearly closed under complements. Let A, B be μ^*-measurable sets, we need to show that
$$\mu^*(E) = \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c).$$
The basic idea is to write $E \cap (A \cup B)$ as a disjoint union, and this can be done by observing that
$$A \cup B = (A \setminus B) \cup (A \cap B) \cup (B \setminus A),$$
thus
$$E \cap (A \cup B) = (E \cap A \cap B^c) \cup (E \cap A \cap B) \cup (E \cap B \cap A^c).$$
Then
$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap A) + \mu^*(E \cap A^c)\\ &= \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c)\\ &\geq \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)\end{aligned}$$
by subadditivity. Therefore $\mathcal{M}^*$ is an algebra. Moreover, if $A \cap B = \varnothing$, then
$$\mu^*(A \cup B) = \mu^*((A \cup B) \cap A) + \mu^*((A \cup B) \cap A^c) = \mu^*(A) + \mu^*(B),$$
so μ^* is finitely additive on $\mathcal{M}^*$.
2. We show that $\mathcal{M}^*$ is a σ-algebra. Let $\{A_n : n \in \mathbb{N}\} \subset \mathcal{M}^*$, we make it into a disjoint

sequence:

$$\begin{aligned}
B_1 &= A_1, \\
B_2 &= A_2 \setminus A_1, \\
B_3 &= A_3 \setminus (A_1 \cup A_2), \\
&\dots \\
B_n &= A_n \setminus \bigcup_{i=1}^{n=1} A_i, \\
&\dots
\end{aligned}$$

For each $N \in \mathbb{N}$, $\bigcup_{n=1}^{N} B_n \in \mathcal{M}^*$, so

$$\mu^*(E) = \mu^*(E \cap \bigcup_{n=1}^{N} B_n) + \mu^*(E \cap \left(\bigcup_{n=1}^{N} B_n\right)^c).$$

If we can show that $\mu^*(E \cap \bigcup_{n=1}^{N} B_n) = \sum_{n=1}^{N} \mu^*(E \cap B_n)$, then we would have

$$\begin{aligned}
\mu^*(E) &= \sum_{n=1}^{N} \mu^*(E \cap B_n) + \mu^*(E \cap \left(\bigcup_{n=1}^{N} B_n\right)^c) \\
&\geq \sum_{n=1}^{N} \mu^*(E \cap B_n) + \mu^*(E \cap \left(\bigcup_{n=1}^{\infty} B_n\right)^c).
\end{aligned}$$

Letting $N \to \infty$ leads to

$$\mu^*(E) \geq \sum_{n=1}^{\infty} \mu^*(E \cap B_n) + \mu^*(E \cap \left(\bigcup_{n=1}^{\infty} B_n\right)^c) \geq \mu^*(E \cap \bigcup_{n=1}^{\infty} B_n) + \mu^*(E \cap \left(\bigcup_{n=1}^{\infty} B_n\right)^c).$$

3. We justify the finite additivity. Let $C_N = \bigcup_{n=1}^{N} B_n$ and $C = \bigcup_{n=1}^{\infty} B_n$, then

$$\begin{aligned}
\mu^*(E \cap C_N) &= \mu^*(E \cap C_N \cap B_N) + \mu^*(E \cap C_N \cap B_N^c) \\
&= \mu^*(E \cap B_N) + \mu^*(E \cap C_{N-1}) \\
&= \mu^*(E \cap B_N) + \mu^*(E \cap B_{N-1}) + \mu^*(E \cap C_{N-2}) \\
&= \cdots \\
&= \mu^*(E \cap B_N) + \cdots + \mu^*(E \cap B_1).
\end{aligned}$$

Hence $\bigcup_{n=1}^{\infty} B_n \in \mathcal{M}^*$.

4. μ^* is a measure on $\mathcal{M}^*$. Let $E = \bigcup_{n=1}^{\infty} B_n$ in the step 2, we have $\mu^*(\bigcup_{n=1}^{\infty} B_n) \geq \sum_{n=1}^{\infty} \mu^*(B_n)$, hence $\mu^*(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu^*(B_n)$.
5. μ^* is complete. If $\mu^*(A) = 0$, then $\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*(E \cap A^c) \leq \mu^*(E)$, so $A \in \mathcal{M}^*$. □

Here comes the definitive stage of the construction.

Theorem 3.6

Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra, μ_0 a premeasure on $\mathcal{A}$, and $\mathcal{M}$ the σ-algebra generated by $\mathcal{A}$. We have the outer measure induced by μ_0:

$$\mu^*(E) = \inf\left\{\sum_{j=1}^{\infty} \mu_0(A_j) : A_j \in \mathcal{A}, E \subset \bigcup_{j=1}^{\infty} A_j\right\},$$

and $\mu = \mu^*|_{\mathcal{M}}$ is a measure on $\mathcal{M}$.

If μ_0 is σ-finite[a], then μ is the unique extension of μ_0 to a measure on $\mathcal{M}$.

[a]This means we can write $X = \bigcup_{j=1}^{\infty} E_j$ with $E_j \in \mathcal{A}$ and $\mu(E_j) < \infty$.

♡

Proof By Carathéodory's theorem, μ^* is a measure on $\mathcal{M}^*$. From section 3.2.3 we have $\mathcal{M}^* \supset \mathcal{M}$, hence $\mu^*|_{\mathcal{M}}$ is a measure on $\mathcal{M}$.

Uniqueness. Let ν be another extension of μ. We first show that μ and ν agree on sets of finite measure. Suppose that $F \in \mathcal{M}$ has finite measure. If $F \subset \bigcup_{j=1}^{\infty} E_j$ with $E_j \in \mathcal{A}$ then

$$\nu(F) \leq \nu\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} \nu(E_j) = \sum_{j=1}^{\infty} \mu_0(E_j).$$

Since the inequality holds for all $\{E_j\}$ covering F, taking infimum gives $\nu(F) \leq \mu^*(F) = \mu(F)$.

To prove the reverse inequality, note that if $E = \bigcup_{j=1}^{\infty} E_j$, then

$$\nu(E) = \lim_{n\to\infty} \nu\left(\bigcup_{j=1}^{n} E_j\right) = \lim_{n\to\infty} \mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \mu(E).$$

By the definition of an outer measure, we can choose $\{E_j\}$ so that $\mu(E) \leq \mu(F) + \varepsilon$, then $\mu(E \setminus F) \leq \varepsilon$ since $\mu(F) < \infty$, and therefore

$$\begin{aligned}
\mu(F) \leq \mu(E) = \nu(E) &= \nu(F) + \nu(E \setminus F) \\
&\leq \nu(F) + \mu(E \setminus F) \\
&\leq \nu(F) + \varepsilon.
\end{aligned}$$

Since ε is arbitrary, $\mu(F) = \nu(F)$.

Finally, we use this result that if μ is σ-finite, then $\mu = \nu$. We can write $X = \bigcup_{j=1}^{\infty} E_j$, where $E_j \in \mathcal{A}$ and are disjoint with $\mu(E_j) < \infty$. Then for any $F \in \mathcal{M}$ we have

$$\mu(F) = \sum_{j=1}^{\infty} \mu(F \cap E_j) = \sum_{j=1}^{\infty} \nu(F \cap E_j) = \nu(F).$$

□

A natural question raises: start from an algebra $\mathcal{A}$, we can define an outer measure on $\mathcal{P}(X)$, and we have two σ-algebras:

$\mathcal{M}^*$: collection of μ^*-measurable sets;
$\mathcal{M}$: the σ-algebragenerated by $\mathcal{A}$.

What is the difference between these two σ-algebras? In section 3.2.3 we see that every set in $\mathcal{A}$ is μ^*-measurable, so $\mathcal{M}^*$ is a σ-algebracontaining $\mathcal{A}$, hence $\mathcal{M}^* \supset \sigma(\mathcal{A}) = \mathcal{M}$. We can at least conclude that $\mathcal{M} \subset \mathcal{M}^*$. As we will see later, the Borel σ-algebrais generated by the collection of all open sets, while the Lebesgue σ-algebrais the collection of μ^*-measurable sets. The Borel σ-algebrais not complete, but the Lebesgue σ-algebrais complete.

Historical Notes [6] Constantin Carathéodory (13 September 1873 – 2 February 1950) was a Greek mathematician who spent most of his professional career in Germany. He made

[6]https://en.wikipedia.org/wiki/Constantin_Carath%C3%A9odory

significant contributions to real and complex analysis, the calculus of variations, and measure theory. He also created an axiomatic formulation of thermodynamics. Carathéodory is considered one of the greatest mathematicians of his era and the most renowned Greek mathematician since antiquity.

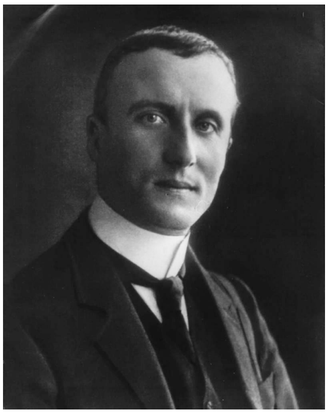


Figure 3.1: Constantin Carathéodory

3.3 Borel and Lebesgue-Stieltjes Measures

In this section we will apply the construction process to obtain a measure on $\mathcal{R}$. The primitive idea is to measure the length of an interval. The Borel σ-algebraon $\mathbb{R}$ can be generated by many types of intervals. We have already seen the following proposition:

Proposition 3.6

$\mathcal{B}_{\mathbb{R}}$ is generated by each of the following:

1. *the open intervals:* $\mathcal{E}_1 = \{(a, b) : a < b\}$,
2. *the closed intervals:* $\mathcal{E}_2 = \{[a, b] : a < b\}$,
3. *the half-open intervals:* $\mathcal{E}_3 = \{(a, b] : a < b\}$ *or* $\mathcal{E}_4 = \{[a, b) : a < b\}$,
4. *the open rays:* $\mathcal{E}_5 = \{(a, \infty) : a \in \mathbb{R}\}$ *or* $\mathcal{E}_6 = \{(-\infty, a) : a \in \mathbb{R}\}$,
5. *the closed rays:* $\mathcal{E}_7 = \{[a, \infty) : a \in \mathbb{R}\}$ *or* $\mathcal{E}_8 = \{(-\infty, a] : a \in \mathbb{R}\}$.

In particular, we will choose our build block as the elementary family $\mathcal{E} = \{(a, b] : -\infty \leq a < b < \infty\}$, which is called **h-intervals**.

Exercise 3.2 Show that the intersections of two h-intervals is an h-interval, and the complement of an h-interval is an h-interval or the disjoint union of two h-intervals.

3.3.1 Borel Measures On $\mathbb{R}$

The above exercise tells us that $\mathcal{S}$ is a semialgebra, which is in the first stage of the general construction process. We choose the volume function ρ to be the length of an interval:

$$\rho((a, b]) = b - a.$$

Recall that the collection of finite disjoint unions in $\mathcal{E}$ is an algebra $\mathcal{A}$, and we can extend ρ from $\mathcal{E}$ to $\mathcal{A}$.

Proposition 3.7

If $(a_j, b_j](j = 1, \cdots, n)$ *are disjoint h-intervals, let*

$$\mu_0\left(\bigcup_{j=1}^n (a_j, b_j]\right) = \sum_{j=1}^n \rho(b_j - a_j) = \sum_{j=1}^n (b_j - a_j)$$

and let $\mu_0(\varnothing) = 0$. *Then* μ_0 *is a premeasure on the algebra* $\mathcal{A}$.

We will postpone the proof to the end of this section. Then we can define an outer measure μ^* on $\mathcal{P}(\mathbb{R})$. Since $\mathcal{B}_{\mathbb{R}}$ is generated by $\mathcal{A}$, μ^* restricted to $\mathcal{B}_{\mathbb{R}}$ is a measure by Carathéodory's theorem.[7]

In general, if the domain of a measure is $\mathcal{B}_{\mathbb{R}}$, then it is a **Borel measure**. A large family of Borel measures that are extremely useful in probability theory is closely related to **distribution functions**. We will return to this concept later. We first show the motivation of what properties should a distribution function possess.

Proposition 3.8

Suppose that μ *is a finite Borel measure on* $\mathbb{R}$ *and let* $F = \mu((-\infty, x])$, *then* F *is increasing and right continuous. Such an* F *is called a distribution function.*

Proof Use continuity and monotonicity of a measure. □

This is how the condition "increasing and right continuous" comes. Now we construct a measure μ starting from an increasing and right continuous function F. Here μ is not necessarily finite! Let $\mathcal{S}$ and $\mathcal{A}$ be the same as above.

Proposition 3.9

Let $F : \mathbb{R} \to \mathbb{R}$ *be increasing and right continuous. If* $(a_j, b_j]$ *are disjoint, let*

$$\mu_0\left(\bigcup_{j=1}^n (a_j, b_j]\right) = \sum_{j=1}^n \rho(b_j - a_j) = \sum_{j=1}^n (F(b_j) - F(a_j))$$

and let $\mu_0(\varnothing) = 0$. *Then* μ_0 *is a premeasure on the algebra* $\mathcal{A}$.

The most difficult part is done, then we obtain a correspondence between Borel measures and distribution functions.

Theorem 3.7

If $F : \mathbb{R} \to \mathbb{R}$ *is any increasing and right continuous functions, there is a unique Borel measure* μ_F *on* $\mathbb{R}$ *such that* $\mu_F((a, b]) = F(b) - F(a)\ \forall a, b$. *If* G *is another such function, we have* $\mu_F = \mu_G \iff F - G$ *is a constant.*

Conversely, if μ *is a Borel measure on* $\mathbb{R}$ *that is finite on all bounded Borel sets and*

[7] Here we actually use Theorem 3.6, but we still refer to it as Carathéodory's theorem. Notice that $\mu^*|_{\mathcal{B}_{\mathbb{R}}}$ is not necessarily complete.

we define

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

then F is increasing and right continuous, and $\mu = \mu_F$. ♡

Proof Start from $\mu_F((a,b]) = F(b) - F(a)$, using proposition 3.9 we get a premeasure (w.r.t. the distribution function F). The premeasure induces an outer measure on $\mathcal{P}(\mathbb{R})$, by restricting the outer measure to $\sigma(\mathcal{A})$, we get a Borel measure μ_F. Since $\mathbb{R} = \bigcup_{j=-\infty}^{\infty}(j, j+1]$, $\mathbb{R}$ is σ-finite, thus μ_F is unique. If $\mu_F = \mu_G$, then $F(b) - F(a) = G(b) - G(a)$ for all $a, b \in \mathbb{R}$, so

$$F(x) - G(x) = F(0) - G(0)\ \forall x \in \mathbb{R}.$$

Conversely, if $F - G$ is a constant, then $\mu_F((a,b]) = \mu_G((a,b])$, hence μ_F and μ_G induces the same premeasure, by uniqueness of extension, $\mu_F = \mu_G$.

For the second assertion,

- If $0 < x < y$, then
$$\begin{aligned} F(y) - F(x) &= \mu((0,y]) - \mu((0,x]) \\ &= \mu((0,y] \setminus (0,x]) \\ &= \mu((x,y]) = y - x > 0. \end{aligned}$$
- If $x < 0 < y$, then
$$\begin{aligned} F(y) - F(x) &= \mu((0,y]) + \mu((x,0]) \\ &= \mu((0,y] \cup (x,0]) \\ &= \mu((x,y]) = y - x > 0. \end{aligned}$$
- If $x < y < 0$, then
$$\begin{aligned} F(y) - F(x) &= -\mu((y,0]) + \mu((x,0]) \\ &= \mu((x,0] \setminus (y,0]) \\ &= \mu((x,y]) = y - x > 0. \end{aligned}$$

Therefore, F is increasing. Let $x \geq 0$ and $x_n \to x$ with $x_n > x$, then

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} \mu((0,x_n]) = \mu\left(\bigcap_{n=1}^{\infty}(0,x_n]\right) = \mu((0,x]) = F(x).$$

The case of $x < 0$ is similar.

Now we prove proposition 3.9.

Proof [proof of proposition 3.7]

Step 1. We check that μ_0 is well-defined, since elements of $\mathcal{A}$ can be represented in more than one way as disjoint unions of h-intervals. If $\{(a_j, b_j]\}_{j=1}^n$ are disjoint and $\bigcup_{j=1}^n (a_j, b_j] = (a,b]$, then after relabeling we have

$$a = a_1 < b_1 = a_2 < b_2 = \cdots < b_n = b,$$

so

$$\sum_{j=1}^{n}[F(b_j) - F(a_j)] = F(b) - F(a).$$

If $\{I_i\}_{i=1}^n$ and $\{J_j\}_{j=1}^m$ are disjoint h-intervals such that $\bigcup_{i=1}^m I_i = \bigcup_{j=1}^n J_j$, then we decompose each I_i and J_j:

$$I_1 = I_1 \cap \bigcup_{j=1}^n J_j = \bigcup_{j=1}^n I_1 \cap J_j,$$

$$\cdots$$

$$I_m = I_m \cap \bigcup_{j=1}^n J_j = \bigcup_{j=1}^n I_m \cap J_j,$$

$$J_1 = J_1 \cap \bigcup_{i=1}^m I_i = \bigcup_{i=1}^m J_1 \cap I_i$$

$$\cdots$$

$$J_n = J_n \cap \bigcup_{i=1}^m I_i = \bigcup_{i=1}^m J_n \cap I_i.$$

Now each I_i is a finite disjoint union of h-intervals (it is easy to see that each $I_i \cap J_j$ is an h-interval), so we can apply the above reasoning to get

$$\sum_{i=1}^m \mu_0(I_i) = \sum_{i=1}^m \sum_{j=1}^n \mu_0(I_i \cap J_j) = \sum_{j=1}^n \mu_0(J_j).$$

Thus μ_0 is well-defined.

Step 2. It remains to show that if $\{I_j\}_{j=1}^\infty$ are disjoint h-intervals with $\bigcup_{j=1}^\infty I_j \in \mathcal{A}$, then $\mu_0\left(\bigcup_{j=1}^\infty I_j\right) = \sum_{j=1}^\infty \mu_0(I_j)$. Since $\bigcup_{j=1}^\infty I_j \in \mathcal{A}$, and recall that $\mathcal{A}$ is the collection of all finite disjoint unions of h-intervals, we have $\bigcup_{j=1}^\infty I_j$ is a finite disjoint union of h-intervals. This is a crucial observation, since the word "finite" in analysis is almost equivalent to "one". Then we can consider each h-interval component of $\bigcup_{j=1}^\infty I_j$. Each component is a disjoint union of some subsequence of $\{I_j\}_{j=1}^\infty$. By consider each subsequence separately and using the finite additivity of μ_0, we may assume that $\bigcup_{j=1}^\infty I_j$ is an h-interval $I = (a, b]$.

Step 3. We begin the estimate. Since $(a, b] = \bigcup_{j=1}^\infty I_j$, it follows that $(a, b] \supset \bigcup_{j=1}^N I_j$ for any $N \in \mathbb{N}$. Thus

$$\mu_0(I) = \mu_0\left(\bigcup_{j=1}^\infty I_j\right) = \mu_0\left(\bigcup_{j=1}^n I_j\right) + \mu_0\left(I \setminus \bigcup_{j=1}^\infty I_j\right) \geq \mu_0\left(\bigcup_{j=1}^n I_j\right) = \sum_{j=1}^n \mu_0(I_j).$$

Letting $n \to \infty$, we obtain $\mu_0(I) \geq \sum_{j=1}^\infty \mu(I_j)$.

Step 4. We prove the reverse inequality. First suppose that $a, b < \infty$. Fix $\varepsilon > 0$. Since F is right continuous, there exists $\delta > 0$ such that $F(a+\delta) - F(a) < \varepsilon$. Write $I_j = (a_j, b_j]$. For each j there exists $\delta > 0$ such that

$$F(b_j + \delta_j) - F(b_j) < \varepsilon 2^{-j}.$$

Now $\bigcup_{j=1}^\infty (a_j, b_j + \delta_j) \supset [a + \delta, b]$, so there is a finite subcover. By relabeling we may assume that

- $(a_1, b_1 + \delta_1), \cdots, (a_N, b_N + \delta_N)$ cover $[a + \delta, b]$,

- $b_j + \delta_j \in (a_{j+1}, b_{j+1} + \delta_{j+1})$ for $j = 1, \cdots, N-1$.

Then

$$\begin{aligned}
\mu_0(I) &= F(b) - F(a) \\
&< F(b) - F(a+\delta) + \varepsilon \\
&\leq F(b_N + \delta_N) - F(a_1) + \varepsilon \\
&= F(b_N + \delta_N) - F(a_N) + \sum_{j=1}^{N-1} [F(a_{j+1}) - F(a_j)] + \varepsilon \\
&\leq F(b_N + \delta_N) - F(a_N) + \sum_{j=1}^{N-1} [F(b_j + \delta_j) - F(a_j)] + \varepsilon \\
&< \sum_{j=1}^{N-1} [F(b_j) + \varepsilon 2^{-j} - F(a_j)] + \varepsilon \\
&< \sum_{j=1}^{\infty} \mu(I_j) + 2\varepsilon.
\end{aligned}$$

Since ε is arbitrary, the reverse inequality is proved. If $a = -\infty$, for any $M < \infty$ the intervals $(a_j, b_j + \delta_j)$ cover $[-M, b]$, so the same reasoning gives $F(b) - F(-M) \leq \sum_{j=1}^{\infty} \mu_0(I_j) + 2\varepsilon$ (note that RHS is independent of M!). If $b = \infty$, for any $M < \infty$ we obtain $F(M) - F(a) \leq \sum_{j=1}^{\infty} \mu_0(I_j) + 2\varepsilon$. □

3.3.2 From Borel to Lebesgue: Two Approaches of Completion

By Carathéodory's theorem, we could get a complete σ-algebra containing $\mathcal{B}_{\mathbb{R}}$ and a complete measure $\overline{\mu}_F$. Some questions raise:

- What do we call this complete σ-algebra?
- Is $\overline{\mu}_F$ the completion of μ_F?
- Is $\mathcal{B}_{\mathbb{R}}$ strictly contained in this complete σ-algebra?

Definition 3.14 (Lebesgue σ-algebra)

The completion of $\mathcal{B}_{\mathbb{R}}$ is called the Lebesgue σ-algebra, denoted $\mathcal{L}$. ♣

In Real Analysis I, we derived the notion of Lebesgue measurability in a purely geometric way and verified that the family of Lebesgue measurable sets is a σ-algebra. Here we obtain the same concept through another approach. To answer the third question, we need measurable functions; to answer the second question, we have two approaches: by studying regularity properties of Lebesgue measurable sets and through a measure-theoretic way. The first method has wide applications in the future.

Topological approach: regularity

In Real Analysis I, we have seen that Lebesgue measurable sets differ only by a set of measure 0 with some Borel sets. This is called the **regularity properties** of Lebesgue measurable sets:

Proposition 3.10 (regularity properties)

Let $E \subset \mathbb{R}$ be a Lebesgue measurable set.

1. *For each $\varepsilon > 0$ there is an open set $U \supset E$ with $\mu(U \setminus E) < \varepsilon$, and there is a closed set $F \subset E$ with $\mu(E \setminus F) < \varepsilon$.*
2. *$E = A \setminus N$, where A is a G_δ set and $m(N) = 0$.*
3. *$E = B \cup M$, where B is an F_σ set and $m(M) = 0$.*

Proof See Real Analysis I.

Measure-theoretic approach

We can also derive the Lebesgue σ-algebra via a more abstract way.

Exercise 3.3 Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra, $\mathcal{A}_\sigma$ the collection of countable unions of sets in $\mathcal{A}$, and $\mathcal{A}_{\sigma\delta}$ the collection of countable intersections of sets in $\mathcal{A}_\sigma$. Let μ_0 be a premeasure on $\mathcal{A}$ and μ^* the induced outer measure.

1. For any $E \subset X$ and $\epsilon > 0$ there exists $A \in \mathcal{A}_\sigma$ with $E \subset A$ and $\mu^*(A) \le \mu^*(E) + \epsilon$.
2. If $\mu^*(E) < \infty$, then E is μ^*-measurable implies that there exists $B \in \mathcal{A}_{\sigma\delta}$ with $E \subset B$ and $\mu^*(B \backslash E) = 0$.
3. If μ_0 is σ-finite, the restriction $\mu^*(E) < \infty$ in (b) is superfluous.

Proof By the definition of an outer measure, there exists $Q_n \in \mathcal{A}$ with $E \subset \bigcup_{n=1}^{\infty} Q_n$ such that

$$\mu^*\left(\bigcup_{n=1}^{\infty} Q_n\right) \le \mu^*(E) + \varepsilon,$$

set $A = \bigcup_{n=1}^{\infty} Q_n$ completes part (1).

For each $n \in \mathbb{N}$ there exists an $A_n \in \mathcal{A}_\sigma \subset \sigma(\mathcal{A})$ such that $\mu^*(A_n) \le \mu^*(E) + 1/n$, then

$$\mu^*\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu^*(A_n) \le \mu^*(E).$$

The reverse inequality is obvious. Set $B = \bigcap_{n=1}^{\infty} A_n$ and thus $\mu^*(B \setminus E) = 0$. If μ_0 is σ-finite, then $X = \bigcup_{n=1}^{\infty} X_n$ with $\mu^*(X_n) < \infty$, so we can write $E = \bigcup_{n=1}^{\infty} X_n \cap E$, where $E_n = X_n \cap E$. For each E_n we have $B_n \supset E$ with $\mu^*(B_n \setminus E_n) = 0$, hence

$$\mu^*\left(\bigcup_{n=1}^{\infty} B_n \setminus \bigcup_{n=1}^{\infty} E_n\right) = \mu^*\left(\bigcup_{n=1}^{\infty} (B_n \setminus E_n)\right) = 0.$$

Exercise 3.4 Let $(X, \mathcal{M}, \mu)$ be a measure space, μ^* the outer measure induced by μ, $\mathcal{M}^*$ the σ-algebra of μ^*-measurable sets, and $\bar{\mu} = \mu^* \mid \mathcal{M}^*$. If μ is σ-finite, then $\bar{\mu}$ is the completion of μ.

Proof Let $E \in \mathcal{M}^*$, then there exists $B \in \sigma(\mathcal{A})$ with $\mu^*(B \setminus E) = 0$, so $E = B \cup (B \setminus E)$

Definition 3.15

Let $F : \mathbb{R} \to \mathbb{R}$ be any increasing and right continuous function. We call $\overline{\mu}_F$ the Lebesgue-Stieltjes measure asscoiated to F, and usually denote this complete measure also by μ_F.

3.3.3 Junction: Carathéodory and Lebesgue

In Real Analysis I, we derive Lebesgue measure by restricting the outer measure to a smaller family of sets by defining E to be Lebesgue measurable if

$$\text{for every } \varepsilon > 0 \text{ there exists an open set } \mathcal{O} \supset E \text{ with } m^*(\mathcal{O} \setminus E) < \varepsilon.$$

Then we showed that under this condition, the family of Lebesgue measurable sets forms a σ-algebra.

In Real Analysis III we use the Carathéodory-style approach to obtain the Lebesgue σ-algebra by declaring E is Lebesgue measurable if

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c) \quad \forall A \subset \mathbb{R},$$

which is elegant and easy to manipulate.

Now we show that the above two conditions are equivalent. First we assume E satisfies the Carathéodory condition ($E \in \mathcal{M}_\mu$) and derive the first regularity property. We begin by a lemma modifying h-intervals to open intervals.

Lemma 3.1

Let μ be a fixed Lebesgue-Stieltjes measure with domain $\mathcal{M}_\mu$. For any $E \in \mathcal{M}_\mu$,

$$\mu(E) = \inf\left\{\sum_{j=1}^{\infty} \mu((a_j, b_j)) : E \subset \bigcup_{j=1}^{\infty} (a_j, b_j)\right\}.$$

♡

Proof Let us call the quantity on the right $\nu(E)$. Suppose $E \subset \bigcup_1^\infty (a_j, b_j)$. Each (a_j, b_j) is a countable disjoint union of h-intervals $I_j^k (k = 1, 2, \ldots)$; specifically, $I_j^k = c_j^k, c_j^{k+1}]$ where $\{c_j\}$ is any sequence such that $c_j^1 = a_j$ and c_j^k increases to b_j as $k \to \infty$. Thus $E \subset \bigcup_{j,k=1}^\infty I_j^k$, so

$$\sum_1^\infty \mu\left((a_j, b_j)\right) = \sum_{j,k=1}^\infty \mu(I_j^k) \geq \mu(E),$$

and hence $\nu(E) \geq \mu(E)$. On the other hand, given $\epsilon > 0$ there exists $\{(a_j, b_j]\}_1^\infty$ with $E \subset \bigcup_1^\infty (a_j, b_j]$ and $\sum_1^\infty \mu\left((a_j, b_j]\right) \leq \mu(E) + \epsilon$, and for each j there exists $\delta_j > 0$ such that $F\left(b_j + \delta_j\right) - F\left(b_j\right) < \epsilon 2^{-j}$. Then $E \subset \bigcup_1^\infty \left(a_j, b_j + \delta_j\right)$ and

$$\sum_1^\infty \mu\left((a_j, b_j + \delta_j)\right) \leq \sum_1^\infty \mu\left((a_j, b_j]\right) + \epsilon \leq \mu(E) + 2\epsilon,$$

so that $\nu(E) \leq \mu(E)$. □

Theorem 3.8

Let μ be a fixed Lebesgue-Stieltjes measure with domain $\mathcal{M}_\mu$. If $E \in \mathcal{M}_\mu$, then

$$\mu(E) = \inf\{\mu(U) : U \supset E, U \text{ is open}\}.$$

♡

Proof For any $\epsilon > 0$ there exist intervals $\left(a_j, b_j\right)$ such that $E \subset \bigcup_1^\infty \left(a_j, b_j\right)$ and $\mu(E) \leq \sum_1^\infty \mu\left((a_j, b_j)\right) + \epsilon$. If $U = \bigcup_1^\infty \left(a_j, b_j\right)$ then U is open, $U \supset E$, and $\mu(U) \leq \mu(E) + \epsilon$. On the other hand, $\mu(U) \geq \mu(E)$ whenever $U \supset E$, so the first equality is valid. □

The Lebesgue measure is a special case of Lebesgue-Stieltjes measure with $F(x) = x$, so we can apply the above results. Conversely, suppose that $E \subset X$ and for every $\varepsilon >$

0 there exists an open set $U \supset E$ with $m^*(U \setminus E) < \varepsilon$. Then by a limiting argument we can find a G_δ set G with $E = G \setminus N$ and $m(N) = 0$, thus $E = G \cup N \in \mathcal{M}_\mu$. The proof is complete.

3.3.4 Cantor Set and Cantor Function

3.4 Measurable Functions

3.4.1 Definitions

Definition 3.16

Let $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ be measurable spaces, a function $f : X \to Y$ is called $(\mathcal{M}, \mathcal{N})$-measurable, or just measurable if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$. ♣

We do not need to check the measurability of f on every set in $\mathcal{N}$, instead it is enough to consider generating sets.

Proposition 3.11

If $\mathcal{N}$ is generated by $\mathcal{E}$, then $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$ measurable if and only if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$. ♠

Proof If f is measurable, then clearly $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$. Conversely, consider $\mathcal{A} = \{E \subset Y : f^{-1}(E) \in \mathcal{M}\}$, then $\mathcal{A} \supset \mathcal{E}$, and $\mathcal{A}$ is a σ-algebra. This can be seen from

- If $E_n \in \mathcal{A}$, then $f^{-1}\left(\bigcup_{n=1}^\infty E_n\right) = \bigcup_{n=1}^\infty f^{-1}(E_n) \in \mathcal{M}$ since $\mathcal{M}$ is a σ-algebra.
- If $E \in \mathcal{A}$, then $f^{-1}(E^c) = f^{-1}(E)^c \in \mathcal{M}$.

Therefore, $\sigma(\mathcal{E}) = \mathcal{N} \subset \mathcal{A}$, thus $E \in \mathcal{N}$ implies that $f^{-1}(E) \in \mathcal{M}$. □

Most of the measurable functions we will use in the future are real-valued.

Definition 3.17

Let $(X, \mathcal{M})$ be a measurable space and $f : X \to \mathbb{R}$. f is called $\mathcal{M}$-measurable, or just measurable, if it is $(\mathcal{M}, \mathcal{B}_\mathbb{R})$-measurable. Likewise, $f : X \to \mathbb{C}$ is called measurable if it is $(\mathcal{M}, \mathcal{B}_\mathbb{C})$-measurable.

In particular, $f : \mathbb{R} \to \mathbb{R}$ is

- *Lebesgue measurable if it is $(\mathcal{L}, \mathcal{B}_\mathbb{R})$-measurable.*
- *Borel measurable if it is $(\mathcal{B}_\mathbb{R}, \mathcal{B}_\mathbb{R})$-measurable.* ♣

Proposition 3.12

If $(X, \mathcal{M})$ is a measurable space and $f : X \to \mathbb{R}$, the following are equivalent.

1. *f is $\mathcal{M}$-measurable.*
2. *$f^{-1}((a, \infty)) \in \mathcal{M}\ \forall a \in \mathbb{R}$.*
3. *$f^{-1}([a, \infty)) \in \mathcal{M}\ \forall a \in \mathbb{R}$.*
4. *$f^{-1}((\infty, a)) \in \mathcal{M}\ \forall a \in \mathbb{R}$.*
5. *$f^{-1}((\infty, a]) \in \mathcal{M}\ \forall a \in \mathbb{R}$.* ♠

Proof Use the generating sets for $\mathcal{B}_\mathbb{R}$ and Proposition 3.11. □

This coincides with the definition of a measurable function from $\mathbb{R} \to \mathbb{R}$ in Real Analysis I.

Definition 3.18

If $(X, \mathcal{M})$ is a measurable space, $f : X \to \mathbb{R}$ and $E \in \mathcal{M}$, we say that f is measurable on E if $f^{-1}(B) \cap E \in \mathcal{M}$ for all Borel sets B. ♣

The following properties of measurable functions is completely analogous to the $\mathbb{R} \to \mathbb{R}$ case, and the proof are the same as shown in Real Analysis I.

Properties

1. *If $f, g : X \to \mathbb{R}$ are measurable, then so are $f + g$ and fg.*
2. *If $\{f_j\}$ is a sequence of $\overline{R}$-valued measurables on $(X, \mathcal{M})$, then*

$$\sup_j f_j(x), \quad \limsup_{j\to\infty} f_j(x),$$
$$\inf_j f_j(x), \quad \liminf_{j\to\infty} f_j(x)$$

   are all measurable.
3. *If $f(x) = \lim_{j\to\infty} f(x)$ exists for every $x \in X$, then f is measurable.*

It is convenient to include adjoin $\pm\infty$ in $\mathbb{R}$ so we can say a limit converges to infinity. The definition of measurability of $f : X \to \overline{\mathbb{R}}$ admits a slight modification.

Exercise 3.5 Let $f : X \to \overline{\mathbb{R}}$ and $Y = f^{-1}(\mathbb{R})$. Then f is measurable if and only if $f^{-1}(\{-\infty\}) \in \mathcal{M}$, $f^{-1}(\{\infty\}) \in \mathcal{M}$, and f is measurable on Y.

Proof Suppose f is measurable. Since $\{-\infty\}, \{\infty\}$ are Borel sets, $f^{-1}(\{-\infty\}) \in \mathcal{M}$, $f^{-1}(\{\infty\}) \in \mathcal{M}$. Let B be a Borel set in $\overline{R}$. If B does not contain $\{-\infty, \infty\}$, then $Y \cap f^{-1}(B) = f^{-1}(B) \in \mathcal{M}$. If B contain ∞ or $-\infty$, then $f^{-1}(B) = f^{-1}(B \setminus \{\pm\infty\}) \cup f^{-1}(\{\pm\infty\}) \in \mathcal{M}$.

Conversely, Let B be a Borel set in $\overline{\mathbb{R}}$. Then consdier two cases: B contains $\pm\infty$ or not, and we are done. □

The following criterion is more applicable.

Exercise 3.6 If $f : X \to \overline{\mathbb{R}}$ and $f^{-1}((r, \infty]) \in \mathcal{M}$ for each $r \in \mathbb{Q}$, then f is measurable.

Proof Let $Y = f^{-1}(\mathbb{R})$, we fisrt show that f is measurable on Y, so the hypothesis can be rewritten as $f^{-1}((r, \infty)) \in \mathcal{M}$ for all $r \in \mathbb{Q}$. Let $a \in \mathbb{R}$, then there is a sequence $\{r_n\}$ increasing to a so that $\bigcup_{n=1}^{\infty}(r_n, \infty) = [a, \infty)$, thus $f^{-1}([a, \infty)) \in \mathcal{M}$ for all $a \in \mathbb{R}$. This shows that f is measurable on Y.

$\bigcap_{n=1}^{\infty} f^{-1}((n, \infty]) = f^{-1}(\{\infty\})$ and $\bigcap_{n=1}^{\infty} f^{-1}([-\infty, -n]) = f^{-1}(\{\infty\})$ implies that $f^{-1}(\{\infty\}) \in \mathcal{M}$ and $f^{-1}(\{-\infty\}) \in \mathcal{M}$. □

3.4.2 Random Variables

In this section we introduce the some basic concepts of probability theory. We will adopt the convention of notation in probability. Let Ω be a set and $\mathcal{F}$ be a σ-algebra on Ω. Let $P : \mathcal{F} \to [0, \infty]$ be a measure such that $P(\Omega) = 1$ (which is called a **probability measure**). The triple $(\Omega, \mathcal{F}, P)$ is called a **probability space** (just another name for measure space!).

Example 3.11 (discrete probability spaces) Let Ω be a at most countable set. Let $\mathcal{F} = \mathcal{P}(\Omega)$, let

$$P(A) = \sum_{\omega \in A} p(\omega) \text{ where } p(\omega) \geq 0, \sum_{\omega \in \Omega} p(\omega) = 1.$$

In many cases when Ω is a finite set, we have $p(\omega) = 1/|\Omega|$.

Definition 3.19

A real valued function $X : \Omega \to \mathbb{R}$ is called a random variable if for every Borel set $B \subset \mathbb{R}$, $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}$. ♣

3.4.3 Distributions

Definition 3.20

*If X is a random variable, then X induces a probability measure on $\mathbb{R}$ called its **distribution** by setting $\mu(A) = P(X \in A)$ for Borel sets A. The notation $P(X \in A)$ is equivalent as $P(X^{-1}(A))$.* ♣

Exercise 3.7 Verify that μ is a probability measure.

Proof

- $\mu(\mathbb{R}) = P(X \in \mathbb{R}) = P(X^{-1}(\mathbb{R})) = P(\Omega) = 1$.
- Let $\{A_n\}_{n=1}^{\infty} \subset \mathcal{B}_{\mathbb{R}}$ be disjoint. Then

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= P\left(X^{-1}\left(\bigcup_{n=1}^{\infty} A_n\right)\right) \\ &= P\left(\bigcup_{n=1}^{\infty} A_n X^{-1}(A_n)\right) \\ &= \sum_{n=1}^{\infty} P(X^{-1}(A_n)) \\ &= \sum_{n=1}^{\infty} \mu(A_n). \end{aligned}$$

□

Definition 3.21

The distribution function of a random variable X is given by

$$F(x) = P(X \leq x) = P(X^{-1}(-\infty, x]).$$

♣

Now $F(x) = P(X \leq x) = \mu((-\infty, x])$ is just the motivation we mentioned in the section of Borel measures, so F is increasing and right continuous. With P being a probability, F has some other properties.

Properties

1. $\lim_{x\to\infty} F(x) = 1, \lim_{x\to-\infty} F(x) = 0$.
2. $P(X = x) = F(x) - F(x^-)$.

Proof

1. $\lim_{x\to\infty} F(x) = \lim_{n\to\infty} F(n) = \lim_{n\to\infty} P(X^{-1}(-\infty, n]) = \mu(\mathbb{R}) = 1$, and likewise $\lim_{x\to-\infty} F(x) = 0$.

2. The interval $(-\infty, x)$ can be approximated by any sequence $x_n \to x$ with $x_n < x$.

$$\begin{aligned}
P(X = x) &= \mu(\{x\}) \\
&= \mu((-\infty, x] \setminus (-\infty, x)) \\
&= F(x) - \mu((-\infty, x)) \\
&= F(x) - \mu\left(\bigcup_{n=1}^{\infty} (-\infty, x_n]\right) \\
&= F(x) - \lim_{n\to\infty} F(x_n) \\
&= F(x) - F(x^-).
\end{aligned}$$

□

Theorem 3.9

If $F : \mathbb{R} \to \mathbb{R}$ is increasing, right continuous, and satisfies

$$\lim_{x\to\infty} F(x) = 1, \qquad \lim_{x\to-\infty} F(x) = 0,$$

then F is the distribution function of some random variable.

Proof Let $\Omega = (0, 1)$, $\mathcal{F} =$ the Borel sets, and P be the Borel measure. If $\omega \in (0, 1)$, let

$$X(\omega) = \sup\{y : F(y) < \omega\}.$$

The supremum exists since $y \in (0, 1)$. If we show that

$$\{\omega : X(\omega) \le x\} = \{\omega : \omega \le F(x)\},$$

then $F(x) = P(\{\omega : X(\omega) \le x\}) = P(\{\omega : \omega \le F(x)\})$ (since P is the Borel measure, the RHS is just the length of $(0, F(x))$).

If $\omega \le F(x)$, then $x \notin \{y : F(y) < \omega\}$. Notice that $\sup\{y : F(y) < \omega\}$ is the least upper bound of the set $(-\infty, \sup\{y : F(y) < \omega\}]$, so $x \ge X(\omega)$.

Conversely, if $\omega > F(x)$, then there exists $\varepsilon > 0$ so that $F(x+\varepsilon) < \omega$ since F is right continuous. By the construction of X we see that $X(\omega) \ge x + \varepsilon > x$, so $\{\omega : X(\omega) \le x\} \subset \{\omega : \omega \le F(x)\}$.

We conclude this section with a dictionary of probabilists' terms.

Analysis	Probability
measure space $(X, \mathcal{M}, \mu)$ $(\mu(X) = 1)$	sample space $(\Omega, \mathcal{F}, P)$
measurable set	event
measurable real-valued function f	random variable X

Chapter 4 Integration on Measure Spaces

Integration on measure spaces is essentially the same as we have seen in Real analysis I, what we need to do here is just change the letter m to μ.

4.1 Abstract Integration: 3 Stages and Convergence Theorems

Fix a measure space $(X, \mathcal{M}, \mu)$, denote M^+ to be the space of all measurable functions from X to $[0, \infty]$.

4.1.1 Stage 1: Simple Functions

If $E \subset X$, the **characteristic function** χ_E of E is defined by

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E, \\ 1 & \text{if } x \notin E. \end{cases}$$

Exercise 4.1 χ_E is measurable if and only if E is measurable.

Definition 4.1

A simple function on X is a finite $\mathbb{C}$-linear combination of characteristic functions of sets in $\mathcal{M}$. (We do not allow simple functions to assume the values $\pm\infty$.) ♣

Example 4.1 (equivalent definition) $f : X \to \mathbb{C}$ is simple if and only if f is measurable and the range of f is a finite subset of $\mathbb{C}$.

Proof If $\operatorname{range}(f) = \{c_1, \cdots, c_N\}$, we can set $E_n = f^{-1}(\{c_n\})$ so that $f = \sum_{n=1}^N c_n \chi_{E_n}$. We call this the **standard representation** of f. The other direction is obvious.

Theorem 4.1 (simple approximation)

Let $(X, \mathcal{M})$ be a measurable space. If $f : X \to [0, \infty]$ is measurable, then there is a sequence $\{\phi_n\}$ of simple functions such that $0 \le \phi_1(x) \le \phi_2(x) \le \cdots \le f(x)$ for all x and $\phi_n(x) \to f(x)$ for every x. Moreover, $\phi_n \to f$ uniformly on any set on which f is bounded. ♡

Proof For $n = 0, 1, 2, \cdots$ and $0 \le k \le 2^{2n} - 1$ let

$$E_n^k = f^{-1}((k2^{-n}, (k+1)2^{-n}]), \quad F_n = f^{-1}((2^n, \infty]),$$

and define

$$\phi_n = \sum_{k=1}^{2^{2n}-1} k2^{-n}\chi_{E_n^k} + 2^n \chi_{F_n}.$$

Then $\phi_n \le \phi_{n+1}$ for all n, and $0 \le f - \phi_n \le 2^{-n}$ on $f^{-1}((0, 2^n])$. □

Theorem 4.2 (complex simple approximation)

Let $(X, \mathcal{M})$ be a measurable space. If $f : X \to \mathbb{C}$ is measurable, then there is a sequence $\{\phi_n\}$ of simple functions such that $0 \le |\phi_1(x)| \le |\phi_2(x)| \le \cdots \le |f(x)|$

for all x and $\phi_n(x) \to f(x)$ for every x. Moreover, $\phi_n \to f$ uniformly on any set on which f is bounded. ♡

Proof Write $f = g + ih$, applying real simple approximation theorem to g^+, g^-, h^+, h^- completes the proof. □

Now we begin the construction of an integral starting from simple functions.

Definition 4.2

If ϕ is a simple function in L^+ with standard representation $\phi = \sum_1^n a_j \chi_{E_j}$, define the integral of ϕ w.r.t. μ by

$$\int \phi \, d\mu = \sum_{j=1}^{n} a_j \mu(E_j).$$

♣

Other notations:

1. $\int \phi(x) \, d\mu(x)$
2. $\int \phi(x) \, \mu(dx)$

Remark If $A \in \mathcal{M}$, then ϕ_{χ_A} is also simple, and we define $\int_A \phi \, d\mu$ to be $\int \phi \chi_A \, d\mu$.

Proposition 4.1 (properties of integration)

Let ϕ and ψ be simple functions in M^+.

1. *If $c \geq 0$, $\int c\phi = c \int \phi$.*
2. *$\int (\phi + \psi) = \int \phi + \int \psi$.*
3. *If $\phi \leq \psi$, then $\int \phi \leq \int \psi$.*
4. *The map $A \mapsto \int_A \phi \, d\mu$ is a measure on $\mathcal{M}$.*

The last property says that every simple function induces a measure on $\mathcal{M}$. ♠

Proof For (2), express the sum of two simple functions in terms of their "common intersections".

For (4), let $\{A_n\} \subset \mathcal{M}$ be disjoint and $A = \bigcup_{n=1}^{\infty} A_n$ then

$$\begin{aligned}
\int_A &= \sum_{n=1}^{\infty} a_n \mu(A \cap E_n) \\
&= \sum_{n,k} a_j \mu(A_k \cap E_n) \\
&= \sum_{k=1}^{\infty} \int_{A_k} \phi.
\end{aligned}$$

4.1.2 Stage 2: Nonnegative Functions

Definition 4.3

If $f \in L^+$, define

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

♣

Proposition 4.2

Let $f, g \in L^+$,

1. $\int f \le \int g$ *whenever* $f \le g$,
2. $\int cf = c \int f$ *for all* $c \ge 0$,
3. $\int (f+g) = \int f + \int g$. *(use MCT to prove)*

Theorem 4.3 (MCT)

If $\{f_n\}$ is a sequence in L^+ with $f_j \le f_{j+1}$ and $f = \lim_{n\to\infty} f_n$, then

$$\int \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int f_n,$$

i.e.,

$$\int f = \lim_{n\to\infty} \int f_n.$$

Proof Idea: use another definition of supremum.

$\int f_n \le \int f$ for all n implies $\lim_{n\to\infty} \int f_n \le \int f$. Conversely, Fix $\alpha \in (0,1)$, let ϕ be simple with $0 \le \phi \le f$ and let $E_n = \{x : f_n(x) \ge \alpha\phi(x)\}$. Then $\cup_{n=1}^{\infty} = X$ and $\int f_n \ge \int_{E_n} f_n \ge \alpha \int_{E_n} \phi$. Since $\lim_{n\to\infty} \int_{E_n} \phi = \int \phi$, $\lim_{n\to\infty} \int f_n \ge \alpha \int \phi$. Letting $\alpha \to 1^-$ and taking supremum over ϕ completes the proof.

The partial sums of nonnegative functions form an increasing sequence.

Theorem 4.4

Let $\{f_n\}$ be a sequence in L^+, then

$$\int \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int f_n.$$

Proof Let $F_N = \sum_{n=1}^{N} f_n$, then $F_N \nearrow \sum_{n=1}^{\infty} f_n$. By MCT,

$$\int \lim_N F_N = \lim_N \int F_N = \lim_N \sum_{n=1}^{N} \int f_n.$$

Proposition 4.3

If $f \in L^+$, then $\int f = 0$ iff $f = 0$ a.e.

Proof Let $E_n = \{x : f(x) > 1/n\}$.

Theorem 4.5 (Fatou's Lemma)

If $\{f_n\}$ is any sequence in L^+, then

$$\int \liminf f_n \le \liminf \int f_n.$$

Proof

$$\int \liminf f_n = \lim_{k\to\infty} \int \inf_{n\ge k} f_n \le \liminf \int f_n.$$

The last inequality follows from $\inf_{n\ge k} f_n \le f_j \ \forall j \ge k$, then

$$\int \inf_{n\ge k} f_n \le \int f_j \ \forall j \ge k,$$

hence
$$\int \inf_{n \geq k} f_n \leq \inf_{j \geq k} \int f_j.$$
□

Proposition 4.4

If $f \in M^+$ and $\int f < \infty$, then $\{x : f(x) = \infty\}$ is a null set and $\{x : f(x) > 0\}$ is σ-finite.

Proof exercise.

4.1.3 Stage 3: Complex Functions

Define $f^+(x) = \max(f(x), 0)$ and $f^-(x) = \max(-f(x), 0)$. Then $f = f^+ - f^-$ and $|f| = f^+ + f^-$.

Definition 4.4

If at least one of $\int f^+$ and $\int f^-$ is finite, we define
$$\int f = \int f^+ - \int f^-.$$
($\infty - \infty$ is undefined)
If $\int f^+$ and $\int f^-$ are both finite, we then say that f is integrable.

Proposition 4.5

f is integrable iff $\int |f| < \infty$.

4.1.4 Connections Between Measure and Integration

Theorem 4.6

Suppose $f : X \to [0, \infty]$ is integrable, and
$$\nu(E) = \int_E f d\mu, \quad (E \in \mathcal{M}).$$
Then ν is a measure on $\mathcal{M}$, and
$$\int g \, d\nu = \int gf \, d\mu.$$

Proof Begin with characteristic functions, then use MCT to complete the proof.

If $g = \chi_E$ for some $E \in \mathcal{M}$, then
$$\int g \, d\nu = \nu(E) = \int_E f \, d\mu = \int \chi_E f \, d\mu.$$
If g is a simple function, then by linearity we have
$$\int g \, d\nu = \int gf \, d\mu.$$
If g is a nonnegative measurable function, then use a sequence of simple functions to approximate g and by the monotone convergence theorem, $\int g \, d\nu = \int gf \, d\mu$. Finally, if $g \in L^1$, then $g = g^+ + g^-$ and applying the previous step to g^+ and g^- yields the desired result. □

Remark We often write

$$d\varphi = f\, d\mu,$$

which is only a notation. The converse is the Radon-Nikodym theorem.

4.2 L^1 Space

Denote L^1 the space of complex-valued integrable functions.

Proposition 4.6

If $f \in L^1$, then $|\int f| \le \int |f|$.

Proposition 4.7

1. *If $f \in L^1$, then $\{x : f(x) \neq 0\}$ is σ-finite.*
2. *If $f, g \in L^1$, then $\int_E f = \int_E g$ for all $E \in \mathcal{M}$ iff $f = g$ a.e. iff $\int |f - g| = 0$.*

$\rho(f, g) = \int |f - g|$ is a metric on L^1, thus $f_n \to f$ in L^1 iff $\int |f_n - f| \to 0$.

Theorem 4.7 (dominated convergence theorem)

Let $\{f_n\}$ be a sequence in L^1 such that

1. *$f_n \to f$ a.e.,*
2. *there exists a nonnegative $g \in L^1$ such that $|f_n| \le g$ a.e. for all n,*

then $f \in L^1$ and $\int f = \lim_{n\to\infty} \int f_n$.

Proof Apply Fatou's lemma to $g + f_n$ and $g - f_n$ (both are nonnegative),

$$\int g + \int f \le \liminf \int (g + f_n) = \int g + \liminf \int f_n,$$
$$\int g - \int f \le \liminf \int (g - f_n) = \int g - \limsup \int f_n$$

Theorem 4.8

Suppose $\{f_j\} \subset L^1$ with $\sum_{j=1}^{\infty} \int |f_j| < \infty$. Then $\sum_{j=1}^{\infty} f_j$ converges a.e. to a function in L^1, and

$$\int \sum_{j=1}^{\infty} f_j = \sum_{n=1}^{\infty} \int f_j.$$

Proof First consider the nonnegative case and we can apply MCT:

$$\sum_{j=1}^{\infty} \int |f_j| = \int \sum_{j=1}^{\infty} |f_j|,$$

hence $\sum_{j=1}^{\infty} |f_j|$ is integrable, so it is finite for a.e. x. Then $\sum_{j=1}^{\infty} f_j$ converges (absolute convergence $\implies$ convergence). Let $F_N = \sum_{j=1}^{N} f_j$, then

1. $F_N \to \sum_{j=1}^{\infty} f_j$,
2. $|F_N| \le \sum_{j=1}^{\infty} |f_j| \in L^1$,

so

$$\lim_{N\to\infty} \int F_N = \int \lim_{N\to\infty} F_N,$$

which is

$$\lim_{N\to\infty}\sum_{j=1}^{N}\int f_j=\int\sum_{j=1}^{\infty}f_j.$$

Theorem 4.9 (denseness)

If $f\in L^1(\mu)$ and $\varepsilon>0$, there is an integrable simple function $\phi=\sum a_j\chi_{E_j}$ such that $\int|f-\phi|\,d\mu<\varepsilon$.

If μ is a Lebesgue-Stieltjes measure on $\mathbb{R}$, the sets E_j can be taken to be finite unions of open intervals;

thre is a continuous function g with bounded support such that $\int|f-g|\,d\mu<\varepsilon$.

Summary:

1. *$\{$simple functions$\}$ is dense in L^1,*
2. *$\{$functions of bounded support$\}$ is dense in $L^1(\mathbb{R},\mu)$, where μ is a Lebesgue-Stieltjes measure.*

♡

Proof Urysohn's lemma and regularity of Lebesgue measure.

Theorem 4.10 (Egorov)

Suppose that $\mu(X)<\infty$, and $\{f_n\}$ is a sequence of measurable complex-valued functions on X such that $f_n\to f$ a.e. Then for every $\varepsilon>0$ there exists $E\subset X$ such that $\mu(E)<\varepsilon$ and $f_n\to f$ uniformly on E^c.

♡

We conclude this section with a fascinating example, which is a simple version of Vitali's convergence theorem. The proof of this example involves almost every result we have learned so far. First, we discuss the decaying property of an integral. To motivate, think f as a function on $\mathbb{R}$. f is integrable does not imply that $f(x)$ tends to 0 as $x\to\infty$, but we have $\int_N^\infty f(x)dx\to 0$ as $N\to\infty$, so the truncated integral cannot be too large. In an arbitrary measure space, the integral of an integrable function is controlled in a similar way.

Proposition 4.8 (absolute continuity)

If $f\in L^1$, then for every $\varepsilon>0$ there is a $\delta>0$ such that

$$\left|\int_E f\,d\mu\right|<\varepsilon \text{ whenever } \mu(E)<\delta.$$

♠

Proof Let $A_N=\{x:|f(x)|>N\}$, then $f\chi_{A_N}\to 0$ a.e. since $f\in L^1$. Also, $|f\chi_{A_N}|\le|f|\in L^1$, so by the dominated convergence theorem

$$\int_{A_N}f\,d\mu\to 0\quad(N\to 0).$$

Now

$$\begin{aligned}\int_E|f|d\mu&=\int_{E\cap A_N}|f|d\mu+\int_{E\cap A_N^c}|f|d\mu\\&\le\int_{A_N}|f|d\mu+\int_{E\cap A_N^c}N\,d\mu\\&\le\frac{\varepsilon}{2}+N\mu(E),\end{aligned}$$

choosing $\delta = \varepsilon/(2N)$ completes the proof. $\square$

Proposition 4.9 (principal part)

Let $f \in L^1$, then for every $\varepsilon > 0$ there exists a set E of finite measure such that

$$\int_{E^c} |f| < \varepsilon.$$

Proof Let $E_n = \{x \in X : |f(x)| > 1/n\}$, then

$$\int_X f \, d\mu = \int_{E_n} f \, d\mu + \int_{E_n^c} f \, d\mu$$

and $\bigcap_{n=1}^{\infty} X_n^c = \varnothing$. Hence $\mu\left(\bigcap_{n=1}^{\infty} E_n^c\right) = 0$, so for some large N we have $\mu(E_N^c) < \varepsilon$, then

$$\int_{E_N^c} f \, d\mu \leq \frac{\varepsilon}{N} < \varepsilon,$$

and clearly $\mu(E_N) < \infty$. $\square$

Example 4.2 [1] A sequence $\{f_n\}_{n\in\mathbb{N}}$ of real-valued measurable functions defined on a measure space (X, Σ, μ) is uniformly integrable if for every $\epsilon > 0$ there is a $\delta > 0$ so that $\sup_{n\in\mathbb{N}} \left|\int_E f_n d\mu\right| < \epsilon$ for all measurable subsets $E \subset X$ with measure at most δ.
(i) Suppose that $\mu(X) < \infty$, $f_n : X \to \mathbb{R}$ is a uniformly integrable sequence and $f_n(x)$ converges to $f(x)$ almost everywhere. Assume that $|f(x)| < \infty$ almost everywhere. Then show

$$\lim_{n\to\infty} \int |f_n - f| \, d\mu = 0$$

and $\lim_{n\to\infty} \int f_n d\mu = \int f d\mu$.
(ii) Show how the dominated convergence theorem can be deduced from part (i).

Proof Let $\varepsilon > 0$. Since $\{f_n\}$ is uniformly integrable, there exists $\delta > 0$ such that $\mu(A) < \delta$ implies

$$\int_E |f_n| d\mu < \varepsilon \ \forall n \in \mathbb{N}.$$

We can take A to be the "Egorov set". For this chosen δ there is a set E with $\mu(E) < \delta$ and $f_n \to f$ uniformly on E^c, thus

$$\int_{E^c} |f_n - f| d\mu < \varepsilon \text{ for all large } n.$$

We want to estimate

$$\begin{aligned}\int_X |f_n - f| d\mu &= \int_E |f_n - f| d\mu + \int_{E^c} |f_n - f| d\mu \\ &\leq \int_E |f_n| d\mu + \int_E |f| d\mu + \int_{E^c} |f_n - f| d\mu.\end{aligned}$$

The remaining part is $\int_E |f| d\mu$. From basic analysis we know if a sequence of numbers $|c_n| \leq M$ and $c_n \to c$, then its limit c must also be bounded by M. Now we have $\int_E |f_n| d\mu < \varepsilon$ and $f_n \to f$ a.e., how can we detour around this integral sign Here comes

[1] This is a Homework problem in Math 721 at UW-Madison in Fall 2022, taught by Andreas Seeger.

Fatou's lemma!

$$\begin{aligned}\int_E |f| &= \int_E \liminf_{n\to\infty} |f_n| d\mu \\ &= \liminf_{n\to\infty} \int_E |f_n| d\mu \\ &< \varepsilon.\end{aligned}$$

Combining these estimates together we have

$$\int_X |f_n - f| d\mu \to 0 \quad (n \to \infty).$$

For part (ii), we present two proofs.

1. This method use the principal-part property to obtain a set of finite measure. Let $\varepsilon > 0$ and g be the dominating function, then there is a set E of finite measure such that

$$\int_{E^c} |f_n| d\mu \leq \int_{E^c} |g| d\mu < \varepsilon \quad \forall n \in \mathbb{N}.$$

The condition $|f_n| \leq g \in L^1$ implies that $\{f_n\}$ is uniformly integrable, so we apply part (i) on the set E to get

$$\int_E |f_n - f| d\mu \to 0.$$

Hence,

$$\int_X |f_n - f| d\mu \leq \int_E |f_n - f| d\mu + 2 \int_{E^c} |g| d\mu,$$

completing the proof.

2. This method constructs a new finite measure. Define ν on Σ by $\nu(E) = \int_E g d\mu$. Then $\nu(X) < \infty$ since $g \in L^1(\mu)$. Applying part (i) to functions $\{f_n/g\}$ yields

$$\int_X \frac{|f_n - f|}{g} d\nu = \int_X |f_n - f| d\mu \to 0 \quad (n \to \infty).$$

□

4.3 Some Applications of the Dominated Convergence Theorem

The dominated convergence theorem (we will refer to it as DCT) gives a sufficient condition for interchanging the limit and integration. In fact, we can get a stronger result. $f_n \to f$ a.e. implies that $|f_n - f| \to 0$ a.e., and $|f_n| \leq g$ a.e. implies that $|f_n - f| \leq 2g$, so apply DCT to $|f_n - f|$ leads to

$$\lim_{n\to\infty} \int |f_n - f| d\mu = 0,$$

which is $\lim_{n\to\infty} \|f_n - f\|_{L^1} = 0$. In this section we mainly discuss the applications of DCT on showing some analytic properties of a function.

Definition 4.5 (Fourier Transform)

Let $f \in L^1(\mathbb{R}^d)$, define the Fourier transform $\widehat{f}$ of f by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i x\cdot\xi} dx,$$

where $x \in \mathbb{R}^d, \xi \in \mathbb{R}^d, x \cdot \xi = x_1\xi_1 + \cdots + x_d\xi_d$. ♣

Example 4.3 If $f \in L^1(\mathbb{R}^d)$, then $\widehat{f}$ is continuous on $\mathbb{R}^d$.

Proof We need to estimate

$$\begin{aligned}|\widehat{f}(\xi+h)-\widehat{f}(\xi)| &= \left|\int_{\mathbb{R}^d} f(x)e^{-2\pi ix\xi}(e^{-2\pi ixh}-1)dx\right| \\ &\leq \int_{\mathbb{R}^d} |f(x)||e^{-2\pi ixh}-1|dx.\end{aligned}$$

It suffices to show the above integral tends to 0 as $h \to 0$. We already have

1. $f(x)(e^{-2\pi ix\cdot h}-1) \to 0$ as $h \to 0$,
2. $|f(x)(e^{-2\pi ix\cdot h}-1)| \leq 2|f(x)| \in L^1$.

Let $\{h_n\}$ be any sequence with $h_n \to 0$, then by DCT

$$\lim_{n\to\infty}\int_{\mathbb{R}^d}|f(x)||e^{-2\pi ixh_n}-1|dx = \int_{\mathbb{R}^d}\lim_{n\to\infty}|f(x)||e^{-2\pi ixh_n}-1|dx = 0.$$

Since $\{h_n\}$ is arbitrary,

$$\lim_{h\to 0}|\widehat{f}(\xi+h)-\widehat{f}(\xi)| = 0.$$

□

Remark Using sequential continuity (or Heine's theorem) we can use sequences converging to 0, because DCT deals only with sequences of functions. From now on, in such situations we shall usually just say let $h \to 0$.

In basic analysis, integration depending on a parameter deals mainly with the problems of interchanging a limit or a derivative with an integral. We begin with a simple example, let $f : \mathbb{R}^2 \to \mathbb{R}$ and $f \in C^1 \cap L^1$, then we can define a function $F(x) = \int_{\mathbb{R}^2} f(x,y)dy$. We are interested in the continuity and differentiability of F, which turns out to be solving the following problems:

- When do we have
$$\lim_{h\to 0}\int f(x+h,y)-f(x,y)dy = \int \lim_{h\to 0}(f(x+h,y)-f(x,y))dy,$$
and
- $\displaystyle\lim_{h\to 0}\frac{F(x+h)-F(x)}{h} = \int \lim_{h\to 0}\frac{f(x+h,y)-f(x,y)}{h}dy$, that is,
$$\frac{dF}{dx}(x) = \int \frac{\partial}{\partial x}f(x,y)dy?$$

We have the following theorem:

Theorem 4.11

Suppose that $f : X \times [a,b] \to \mathbb{C}(-\infty < a < b < \infty)$ and that $f(\cdot,t) : X \to \mathbb{C}$ is integrable for each $t \in [a,b]$. Let $F(t) = \int_X f(x,t)d\mu(x)$.

1. *Suppose that there exists $g \in L^1(\mu)$ such that $|f(x,t)| \leq g(x)$ for all x, t. If $\lim_{t\to t_0} f(x,t) = f(x,t_0)$ for every x, then $\lim_{t\to t_0} F(t) = F(t_0)$; in particular, if $f(x,\cdot)$ is continuous for each x, then F is continuous.*
2. *Suppose that $\partial f/\partial t$ exists and there is a $g \in L^1(\mu)$ such that $|(\partial f/\partial t)(x,t)| \leq g(x)$ for all x, t. Then F is differentiable and $F'(x) = \int(\partial f/\partial t)(x,t)d\mu(x)$* ♡

Proof The proof of part (1) shares essentially the same idea with the above example, and we leave it as an exercise. For part (2), let $t_0 \in [a,b]$ and consider the difference quotient

$$\frac{F(t)-F(t_0)}{t-t_0} = \int \frac{f(x,t)-f(x,t_0)}{t-t_0}.$$

Let $\{t_n\} \subset [a, b]$ with $t_n \to t_0$ and observe that

$$\frac{\partial}{\partial t} f(x, t_0) = \lim_{n\to\infty} \frac{f(x, t_n) - f(x, t_0)}{t_n - t_0} := h_n(x),$$

then $\partial f / \partial t$ is measurable, and by the mean value theorem,

$$|h_n(x)| \leq \sup_{t\in[a,b]} \left| \frac{\partial}{\partial t} f(x, t) \right| \frac{|t_n - t_0|}{|t_n - t_0|} \leq g(x),$$

so by DCT we have

$$F'(t_0) = \lim \frac{F(t_n) - F(t_0)}{t_n - t_0} = \lim \int h_n(x) d\mu(x) = \int \frac{\partial f}{\partial t}(x, t) d\mu(x)$$

We use this criterion to prove a property of the Fourier transform.

Example 4.4 Let f be a smooth and integrable function on $\mathbb{R}$ such that

- $f \in C_0(\mathbb{R})$(f vanishes at infinity),
- $xf \in L^1$.

Then,

$$\frac{\mathrm{d}\widehat{f}}{\mathrm{d}\xi} = [(-2\pi i x) f]^{\wedge},$$

and

$$\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)^{\wedge}(\xi) = (2\pi i\xi)\widehat{f}(\xi).$$

Proof For the first one,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\xi}\widehat{f}(\xi) &= \frac{\mathrm{d}}{\mathrm{d}\xi} \int_{\mathbb{R}} f(x) e^{-2\pi i x\xi} dx \\
&= \int_{\mathbb{R}} f(x) \frac{\mathrm{d}}{\mathrm{d}\xi} e^{-2\pi i x\xi} dx \\
&= \int_{\mathbb{R}} (-2\pi i x) f(x) e^{-2\pi i x\xi} dx,
\end{aligned}$$

which is the Fourier transform of $(-2\pi i x) f(x)$.

For the second one,

$$\begin{aligned}
\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)^{\wedge}(\xi) &= \int_{\mathbb{R}} f'(x) e^{-2\pi i x\xi} dx \\
&= f(x) e^{-2\pi i x\xi}|_{-\infty}^{\infty} - \int_{\mathbb{R}} f(x)(-2\pi i\xi) e^{-2\pi i x\xi} dx \\
&= 2\pi i\xi \widehat{f}(\xi).
\end{aligned}$$

□

4.4 Product σ-Algebras

Technically, product σ-algebras are closely related to product of collections of sets. We are familiar with the product of sets. Let $\{X_n : n \in \mathbb{N}\}$ be a collection of nonempty sets and let $X = \prod_{n=1}^{\infty} X_n$. Let $\pi_n : X \to X_n$ be the coordinate maps. That is,

$$\pi_n(x_1, \cdots, x_n, x_{n+1}, \cdots) = x_n \in X_n.$$

This section requires some familiarity with the properties of cartesian products, especially with respect to coordinate maps and set operations. If you find difficulties in some arguments, refer to the appendix.

Definition 4.6

Suppose $\mathcal{M}_n$ is a σ-algebra on X_n for each n. We define the product σ-algebra on X to be the σ-algbra generated by

$$\{\pi_n^{-1}(E_n) : E_n \in \mathcal{M}_n, n \in \mathbb{N}\}.$$

Denote this σ-algebra by $\bigotimes_{n\in\mathbb{N}} \mathcal{M}_n$. ♣

The above definition is rarely used in practice. Intuitively, we expect a set in the product σ-algebra to be a product of sets from each component. Alternatively, we have

Proposition 4.10

$\bigotimes_{n\in\mathbb{N}} \mathcal{M}_n$ is the σ-algebra generated by

$$\mathcal{A} = \{\prod_{n=1}^{\infty} E_n : E_n \in \mathcal{M}_n\}.$$

♠

Proof Denote $\mathcal{F} = \{\pi_n^{-1}(E_n) : E_n \in \mathcal{M}_n, n \in \mathbb{N}\}$. From the original definition we know $\sigma(\mathcal{F}) = \bigotimes_{n\in\mathbb{N}} \mathcal{M}_n$. It suffices to show that $\mathcal{A} \subset \sigma(\mathcal{F})$ and $\mathcal{F} \subset \sigma(\mathcal{A})$. We first observe that

$$\pi_n^{-1}(E_n) = X_1 \times \cdots \times X_{n-1} \times E_n \times X_{n+1} \times \cdots,$$

thus $\prod_{n=1}^{\infty} E_n = \bigcap_{n=1}^{\infty} \pi_n^{-1}(E_n) \in \sigma(\mathcal{F})$. This shows that $\mathcal{A} \subset \sigma(\mathcal{F})$.
Conversely, $\pi_n^{-1}(E_n)$ is clearly in $\mathcal{A}$ (with $E_k = X_k$ if $k \neq n$), so $\mathcal{F} \subset \mathcal{A}$, hence $\mathcal{F} \subset \sigma(\mathcal{A})$. □

If we take into consideration each generating family $\mathcal{E}_n \subset \mathcal{M}_n$, a simpler original definition comes:

Proposition 4.11

Suppose that $\mathcal{M}_n$ is generated by $\mathcal{E}_n$, then $\bigotimes_{n=1}^{\infty} \mathcal{M}_n$ is generated by

$$\mathcal{F}_1 = \{\pi_n^{-1}(E_n) : E_n \in \mathcal{E}_n, n \in \mathbb{N}\}.$$

♠

Proof Let $\mathcal{A} = \{\pi_n^{-1}(E_n) : E_n \in \mathcal{M}_n, n \in \mathbb{N}\}$. It suffices to prove $\mathcal{F}_1 \subset \sigma(\mathcal{A})$ and $\mathcal{A} \subset \sigma(\mathcal{F}_1)$. The first assertion is obvious. The collection $\mathcal{B}_n = \{E \subset X_n : \pi_n^{-1}(E) \in \sigma(\mathcal{F}_1)\}$ is a σ-algebra on X_n that contains $\mathcal{E}_n$:

1. $\pi_n^{-1}(\cup_{k=1}^{\infty} E_k) = \cup_{k=1}^{\infty} \pi_n^{-1}(E_k) \in \sigma(\mathcal{F}_1)$;
2. $\pi_n^{-1}(E^c) = \pi_n^{-1}(E)^c \in \sigma(\mathcal{F}_1)$.

Hence $\mathcal{B}_n \supset \mathcal{M}_n$. In other words, if $E \in \mathcal{M}_n$, then $\pi_n^{-1}(E) \in \sigma(\mathcal{F}_1)$. Let n run through $\mathbb{N}$, we have $\mathcal{A} \subset \sigma(\mathcal{F}_1)$.

Corollary 4.1

Suppose in addition that $X_n \in \mathcal{E}_n$ for each n, then $\bigotimes_{n=1}^{\infty} \mathcal{M}_n$ is generated by

$$\mathcal{F}_2 = \{\prod_{n=1}^{\infty} E_n : E_n \in \mathcal{E}_n\}.$$

♡

Proof The idea is to compare $\mathcal{F}_1$ and $\mathcal{F}_2$. $\pi_n^{-1}(E_n) = X_1 \times \cdots \times E_n \times \cdots \in \mathcal{F}_2$ since

$X_n \in \mathcal{E}_n$. Thus we have $\mathcal{F}_1 \subset \mathcal{F}_2$. Conversely,

$$\prod_{n=1}^{\infty} E_n = \bigcap_{n=1}^{\infty} \pi_n^{-1}(E_n) \in \sigma(\mathcal{F}_1). \quad \square$$

Remark It is convenient to view $\mathcal{A}$ and $\mathcal{F}_2$ as **product of collections of sets**. We can write

$$\mathcal{A} = \prod_{n=1}^{\infty} \mathcal{M}_n, \quad \mathcal{F}_2 = \prod_{n=1}^{\infty} \mathcal{E}_n.$$

Then, the above conclusion can be rewritten as

If $\mathcal{M}_n = \sigma(\mathcal{E}_n)$ and $X_n \in \mathcal{E}_n$ for each n, then

$$\bigotimes_{n=1}^{\infty} \mathcal{M}_n = \sigma\left(\prod_{n=1}^{\infty} \mathcal{E}_n.\right)$$

The next proposition covers the most cases we will encounter.

Proposition 4.12

Let $X_1, \cdots, X_n$ be metric spaces and let $X = \prod_{j=1}^n X_j$, equipped with the product metric. Then $\bigotimes_{j=1}^n \mathcal{B}_{X_j} \subset \mathcal{B}_X$. If X_j's are separable, then $\bigotimes_{j=1}^n \mathcal{B}_{X_j} = \mathcal{B}_X$.

Proof Let $\mathcal{O}_n$ be the collection of open sets in X_n, then $\bigotimes_{j=1}^n \mathcal{B}_{X_j}$ is generated by $\prod_{j=1}^n \mathcal{O}_j$. Let $O_1 \times \cdots \times O_n \in \prod_{j=1}^n \mathcal{O}_j$, then each O_j is open in X_j, hence $O_1 \times \cdots \times O_n$ is open in X. This shows $\prod_{j=1}^n \mathcal{O}_j \subset \mathcal{B}_X$.

Let C_j be a countable dense subset in X_j, and let $\mathcal{R}_j$ be the collection of open balls in X_j with rational radius and center in C_j, then each open set in X_j is a countable union of elements of $\mathcal{R}_j$, hence $\sigma(\mathcal{R}_j) = \mathcal{B}_{X_j}$. Moreover, an open ball of radius r in X is the product of open balls in X_j of radius r (recall that we are using the product metric!), then $\prod_{j=1}^n \mathcal{R}_j$ generates $\mathcal{B}_X$. Meanwhile, $\bigotimes_{j=1} n\mathcal{B}_{X_j} = \sigma\left(\prod_{j=1}^n \mathcal{R}_j\right)$, completing the proof. $\square$

4.5 Product Measures

4.5.1 Construction

Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be measure spaces. To construct a measure on the product space $X \times Y$, we follow the standard process:

(volume, semialgebra) $\to$ (premeasure, algebra) $\to$ (outer measure, power set) $\to$ (measure, σ-algebra)

4.5.1.1 Step 1: (volume, semialgebra)

Define a *rectangle* to be a set of the form $A \times B$, where $A \in \mathcal{M}, B \in \mathcal{N}$. Then

$$(A \times B) \cap (E \times F) = (A \cap E) \times (B \cap F), \quad (A \times B)^c = (X \times B^c) \cup (A^c \times B).$$

Hence $\{\text{rectangles}\}$ is a semialgebra, and we define the volume π of the rectangle $E = A \times B$ to be $\pi(E) = \mu(A)\nu(B)$.

4.5.1.2 Step 2: (premeasure, algebra)

The collection $\mathcal{A}$ of finite disjoint union of rectangles is an algebra. Suppose $A \times B = \bigcup_{j=1}^{n}(A_j \times B_j)$(in general a finite union of rectangles may not be a rectangle). Then for $x \in X$ and $y \in Y$,

$$\chi_A(x)\chi_B(y) = \chi_{A\times B}(x,y) = \sum_{j=1}^{n} \chi_{A_j\times B_j}(x,y) = \sum_{j=1}^{n} \chi_{A_j}(x)\chi_{B_j}(y).$$

Integrating w.r.t. x,

$$\begin{aligned}\int \chi_A(x)\chi_B(y)d\mu(x) &= \sum_{j=1}^{n} \int \chi_{A_j}(x)\chi_{B_j}(y)d\mu(x) \\ &= \sum_{j=1}^{n} \mu(A_j)\chi_{B_j}(y)\end{aligned}$$

Integrating in y,

$$\begin{aligned}\int \sum_{j=1}^{n} \mu(A_j)\chi_{B_j}(y)d\nu(y) &= \sum_{j=1}^{n} \mu(A_j) \int \chi_{B_j}(y)d\nu(y) \\ &= \sum_{j=1}^{n} \mu(A_j)\nu(B_j) \\ &= \sum_{j=1}^{n} \pi(A_j \times B_j). \\ \mu(A)\nu(B) &= \sum_{j=1}^{n} \mu(A_j)\nu(B_j).\end{aligned}$$

Based on this observation, we define a premeasure π on $\mathcal{A}$. If $E = \bigcup_{j=1}^{n}(A_j \times B_j) \in \mathcal{A}$(not necessarily a rectangle), then we set

$$\pi(E) = \sum_{j=1}^{n} \mu(A_j)\nu(B_j).$$

4.5.1.3 Step 3: (outer measure, power set)

Now π generates an outer measure π^* on $\mathcal{P}(X \times Y)$ by

$$\pi^*(E) = \inf\left\{\sum_{j=1}^{\infty} \pi(A_j \times B_j) : E \subset \bigcup_{j=1}^{\infty} A_j \times B_j, A_j \times B_j \text{ rectangle}\right\}.$$

4.5.1.4 Step 4: (measure, σ-algebra)

Let's copy the proposition 4.10: $\mathcal{M} \otimes \mathcal{N}$ is the σ-algebra generated by

$$\mathcal{S} = \{E_1 \times E_2 : E_1 \in \mathcal{M}, E_2 \in \mathcal{N}\},$$

which is exactly our semialgebra. Then the algebra $\mathcal{A} = \sigma(\mathcal{S})$ definitely generates $\mathcal{M} \otimes \mathcal{N}$. The restriction $\pi^*|_{\mathcal{M}\otimes\mathcal{N}} := \pi$ is a measure on $\mathcal{M} \times \mathcal{N}$, called the **product** of μ and ν, denoted by

$$\mu \times \nu.$$

Remark If μ and ν are σ-finite: $X = \bigcup_{j=1}^{\infty} A_j, Y = \bigcup_{k=1}^{\infty} B_j$ with $\mu(A_j) < \infty, \nu(B_k) < \infty$, then $X \times Y = \bigcup_{j,k=1}^{\infty} A_j \times B_k$, and $\mu \times \nu(A_j \times B_k) = \mu(A_j)\nu(B_k) < \infty$, so $\mu \times \nu$ is also σ-finite. In this case, $\mu \times \nu$ is the unique measure on $\mathcal{M} \otimes \mathcal{N}$ such that $\mu \times \nu(A \times B) = \mu(A)\nu(B)$ for all rectangles $A \times B$.

4.5.2 Fubini's Theorem

Fubini's theorem is somewhat a converse of the construction of product measure: given a measure on a product space but we are asked to recover the "component measure".

Definition 4.7

*If $E \subset X \times Y$, for $x \in X$ and $y \in Y$ we define the x-**section** and y-**section** by*

$$E_x = \{y \in Y : (x, y) \in E\}, \quad E^y = \{x \in X : (x, y) \in E\}.$$

If f is a function on $X \times Y$ we define its sections by

$$f_x(y) = f(x, y) \ (x \text{ fixed}),$$
$$f^y(x) = f(x, y) \ (y \text{ fixed}).$$

♣

Proposition 4.13 (section of measurable sets and functions)

1. *If $E \in \mathcal{M} \otimes \mathcal{N}$, then $E_x \in \mathcal{N}$ for all $x \in X$ and $E^y \in \mathcal{M}$ for all $y \in Y$.*
2. *If f is $\mathcal{M} \otimes \mathcal{N}$-measurable, then f_x is $\mathcal{N}$-measurable for all $x \in X$ and f^y is $\mathcal{M}$-measurable for all $y \in Y$.*

♠

Proof Define $\mathcal{R} = \{R \in \mathcal{M} \otimes \mathcal{N} : E_x \in \mathcal{N} \text{ for all } x \in X \text{ and } E^y \in \mathcal{M} \text{ for all } y \in Y\}$, then $\mathcal{R}$ is a σ-algebracontaining $\mathcal{M} \otimes \mathcal{N}$.
For the second part, let B be a Borel set in $\mathbb{R}$. Then

$$(f_x)^{-1}(B) = \{y \in Y : f(x, y) \in B\} = \{y \in Y : (x, y) \in f^{-1}(B)\} = (f^{-1}(B))_x$$

and similarly $(f^y)^{-1}(B) = (f^{-1}(B))^y$.

Definition 4.8 (Monotone Class)

*Define a **monotone class** on a space X to be a subset $\mathcal{C} \subset \mathcal{P}(X)$ which is closed under countable increasing unions and countable decreasing intersections.*

♣

Lemma 4.1 (The Monotone Class Lemma)

If $\mathcal{A}$ is an algebra of subsets of X, then the monotone class $\mathcal{C}$ generated by $\mathcal{A}$ coincides with the σ-algebra $\mathcal{M}$ generated by $\mathcal{A}$.

♡

Proof Idea: construct a set-algebraic structure
Obviously $\mathcal{C} \subset \mathcal{M}$. If we show that $\mathcal{C}$ is a σ-algebra, we will have $\mathcal{M} \subset \mathcal{C}$. For $E \in \mathcal{C}$ we define

$$\mathcal{C}(E) = \{F \in \mathcal{C} : E \setminus F, F \setminus E, E \cap F \in \mathcal{C}\}.$$

Clearly $\varnothing, E \in \mathcal{C}(E)$, and $E \in \mathcal{C}(F)$ iff $F \in \mathcal{C}(E)$. Let $\{F_n\} \subset \mathcal{C}(E)$ be an increasing sequence, then

$$E \setminus \bigcup_{n=1}^{\infty} F_n = \bigcap_{n=1}^{\infty} E \setminus F_n \in \mathcal{C}$$

since $\mathcal{C}$ is a monotone class. Similarly $\mathcal{C}(E)$ is closed under countable decreasing intersections. This shows that $\mathcal{C}(E)$ is a monotone class.

If E and E are in $\mathcal{A}$, then $E \setminus F, F \setminus E, E \cap F$ are all in $\mathcal{A} \subset \mathcal{C}$, hence $F \in \mathcal{C}(E)$ for all $F \in \mathcal{A}$. That is, $\mathcal{A} \subset \mathcal{C}(E)$, and hence $\mathcal{C} \subset \mathcal{C}(E)$. Therefore, if $F \in \mathcal{C}$, then $F \in \mathcal{C}(E)$ for all $E \in \mathcal{A}$. By symmetry, $E \in \mathcal{C}(F)$ for all $E \in \mathcal{A}$, so $\mathcal{A} \subset \mathcal{C}(F)$ and hence $\mathcal{C} \subset \mathcal{C}(F)$.

If $E, F \in \mathcal{C}$, then $E \setminus F$ and $E \cap F$ are in $\mathcal{C}$, $\mathcal{C}$ is therefore an algebra. If $\{E_j\} \subset \mathcal{C}$, we have $\bigcup_{j=1}^n E_j \in \mathcal{C}$ for all , and since $\mathcal{C}$ is closed under countable increasing unions it follows that $\bigcup_{j=1}^\infty E_j \in \mathcal{C}$. In short, $\mathcal{C}$ is a σ-algebra. □

Remark One can skip this lemma because we have seen Dynkin system in Chapter 1. This lemma can be immediately deduced by Theorem 3.4.

The next theorem is de facto the Fubini-Tonelli theorem applied to characteristic functions, as we always start with characteristic functions when dealing with an integration formula.

Theorem 4.12

Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. If $E \in \mathcal{M} \otimes \mathcal{N}$, then the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable on X and Y, and

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

♡

Proof First assume that μ and ν are finite. Let

$$\mathcal{C} = \{E \in \mathcal{M} \otimes \mathcal{N} : x \mapsto \nu(E_x) \text{ and } y \mapsto \mu(E^y) \text{ are measurable on } X \text{ and } Y,$$
$$\mu \times \nu(E) = \int \nu(E_x) d\mu(x) = \int \mu(E^y) d\nu(y)\}.$$

If $E = A \times B$, then

$$\nu(E_x) = \int \chi_{E_x}(y) dy = \int \chi_E(x, y) dy = \int \chi_A(x) \chi_B(y) dy = \chi_A(x) \nu(B)$$

and $\mu(E^y) = \mu(A) \chi_B(y)$, so clearly $E \in \mathcal{C}$. It follows that finite disjoint unions of rectangles are in $\mathcal{C}$, so by the monotone class lemma it will suffice to show that $\mathcal{C}$ is a monotone class.

If $\{E_n\}$ is an increasing sequence in $\mathcal{C}$ and $E = \bigcup_{n=1}^\infty E_n$, then the functions $f_n(y) = \mu((E_n)^y)$ are measurable and increase pointwise to $f(y) = \mu(E^y)$. Hence f is measurable. By MCT,

$$\int \mu(E^y) d\nu(y) = \lim_{n \to \infty} \int \mu((E_n)^y) d\nu(y) = \lim_{n \to \infty} \mu \times \nu(E_n) = \mu \times \nu(E).$$

We say a few words about the second equality above. If $E = A \times B$ is a rectangle, then we already have $\int \mu(E^y) d\nu(y) = \mu(A)\nu(B) = \mu \times \nu(E)$. If E is a finite disjoint union of rectangles, say $E = \bigcup_{j=1}^n A_j \times B_j$, then

$$\int \mu\left(\bigcup_{j=1}^n (A_j \times B_j)^y\right) d\nu(y) = \sum_{j=1}^n \int \mu((A_j \times B_j)^y) d\nu(y) = \sum_{j=1}^n \mu \times \nu(A_j \times B_j) = \mu \times \nu(E).$$

It suffices to take E to be a finite disjoint union of rectangles (i.e., E belongs to the algebra $\mathcal{A}$ generated by rectangles) because the monotone class generated by $\mathcal{A}$ equals $\sigma(\mathcal{A})$.

Similarly $\mu \times \nu(E) = \int \nu(E_x) d\mu(x)$, so $E \in \mathcal{C}$. Let $\{E_n\}$ be a decreasing sequence in $\mathcal{C}$ and $F = \bigcap_{n=1}^\infty E_n$. Then

1. $y \mapsto \mu((E_1)^y)$ is measurable.
2. $\mu((E_n)^y) \to \mu(F^y) a.e.$
3. $\int \mu((E_1)^y)d\nu(y) \leq \mu(X)\nu(Y) < \infty$, that is, $\mu((E_1)^y) \in L^1(\nu)$, and $|\mu((E_n)^y)| \leq \mu((E_1)^y)$.

Now by DCT we have

$$\mu \times \nu(F) = \lim_{n\to\infty} \mu \times \nu(E_n) = \lim_{n\to\infty} \int \mu((E_n)^y)d\nu(y) = \int \mu(F^y)d\nu(y),$$

hence $F \in \mathcal{C}$. Thus $\mathcal{C}$ is a monotone class.

Finally, if μ and are σ-finite, $X \times Y = \bigcup_{j=1}^{\infty} X_j \times Y_j$, where $\{X_j \times Y_j\}$ is increasing. If $E \in \mathcal{M} \otimes N$, apply preceding argument to each $E \cap (X_j \times Y_j)$. Since

$$(E \cap (X_j \times Y_j))_x = E_x \cap (X_j \times Y_j)_x = \begin{cases} \varnothing & x \notin X_j \\ E_x \cap Y_j & x \in X_j \end{cases},$$

$\nu(E_x \cap (X_j \times Y_j)_x) = \chi_{X_j}(x)\nu(E_x \cap Y_j)$. Then,

$$\mu \times \nu(E \cap (X_j \times Y_j)) = \int \chi_{X_j}(x)\nu(E_x \cap Y_j)d\mu(x) = \int \chi_{Y_j}(y)\mu(E^y \cap X_j)d\nu(y).$$

By MCT $\mu \times \nu(E \cap (X_j \times Y_j)) \to \mu \times \nu(E \cap (X \times Y)) = \mu \times \nu(E)$.

Theorem 4.13 (Fubini-Tonelli)

Suppose that $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces.

1. *(Tonelli) If $f \in L^+(X \times Y)$, then the functions $g(x) = \int f_x \, d\nu$ and $h(y) = \int f^y \, d\mu$ are in $L^+(X)$ and $L^+(Y)$, and*
$$\int f d(\mu\times\nu) = \int \left(\int f(x,y)d\nu(y)\right) d\mu(x) = \int \left(\int f(x,y)d\mu(x)\right) d\nu(y).$$
2. *(Fubini) If $f \in L^1(\mu \times \nu)$, then $f_x \in L^1(\nu)$ for a.e. $x \in X$, $f^y \in L^1(\mu)$ for a.e. $y \in Y$, the a.e.-defined functions $g(x) = \int f_x \, d\nu$ and $h(y) = \int f^y \, d\nu$ are in $L^1(\mu)$ and $L^1(\nu)$ and (1) holds.*

♡

Proof characteristic functions $\to$ nonnegative simple functions $\to$ nonnegative measurable functions.

By Theorem 4.12, when f is a characteristic function, Tonelli's theorem holds, and it therefore holds for nonnegative simple functions by linearity. Let f be a nonnegative measurable function on $X \times Y$, and let $\{f_n\}$ be a sequence of nonnegative simple functions with $f_n(p) \nearrow f(p)$ for all $p \in X \times Y$. Then for the sections we have

$$(f_n)_x(y) \nearrow f_x(y) \; \forall y \in Y \text{ and } (f_n)^y(x) \nearrow f^y(x) \; \forall x \in X.$$

Denote

$$g_n(x) = \int (f_n)_x d\nu, \quad h_n(y) = \int (f_n)^y d\mu,$$

MCT implies

$$\int g d\mu = \lim_{n\to\infty} \int g_n d\mu = \lim_{n\to\infty} \int f_n d(\mu \times \nu) = \int f d(\mu \times \nu),$$
$$\int h d\nu = \lim_{n\to\infty} \int h_n d\nu = \lim_{n\to\infty} \int f_n d(\mu \times \nu) = \int f d(\mu \times \nu),$$

which establishes Tonelli's theorem.

4.5.3 Examples of Fubini's Theorem

First, we look at the "σ-finite" hypothesis.

Example 4.5 Let $X = Y = [0,1]$, $\mathcal{M} = \mathcal{N} = \mathcal{B}_{[0,1]}$, μ be the Lebesgue measure and ν be the counting measure. If $D = \{(x,x) : x \in [0,1]\}$, then $\iint \chi_D d\mu d\nu$, $\iint \chi_D d\nu d\mu$, $\int \chi_D d(\mu\times\nu)$ are all unequal.

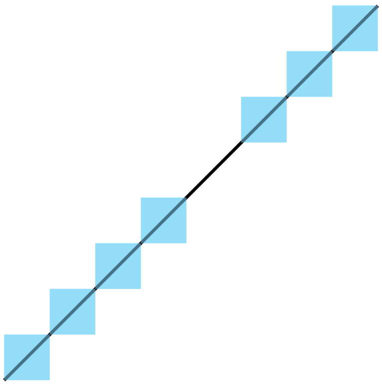


Figure 4.1: diagonal of $[0,1] \times [0,1]$

Proof The section $(\chi_D)^y(x) = \chi_D(x,y) = 0$ if $x \neq y$ and $= 1$ if $x = y$. For a fixed y, we have

$$\chi_D(x,y) = \chi_{\{y\}}(x).$$

Then,

$$\begin{aligned}\iint \chi_D(x,y)d\mu(x)d\nu(y) &= \int_Y \left(\int_X \chi_D(x,y)d\mu(x)\right) d\nu(y)\\ &= \int_Y \left(\int_X \chi_{\{y\}}(x)d\mu(x)\right) d\nu(y)\\ &= \int_Y \mu(\{y\})d\nu(y) = 0.\end{aligned}$$

Similarly,

$$\begin{aligned}\iint \chi_D d\nu d\mu &= \int_X \left(\int_Y \chi_D(x,y)d\nu(y)\right) d\mu(x)\\ &= \int_X \nu(\{x\})d\mu(x)\\ &= \int_X 1d\mu(x)\\ &= \mu(X) = 1.\end{aligned}$$

For the last one, it is impossible to write D as a product of rectangles. Our last hope is to use the outer measure, which is applied to all subsets of $X \times Y$. By definition,

$$(\mu\times\nu)^*(D) = \inf\left\{\sum_{n=1}^{\infty}(\mu\times\nu)(R_n) : D \subset \bigcup_{n=1}^{\infty} R_n, R_n = A_n \times B_n, A_n, B_n \in \mathcal{B}_{[0,1]}\ \forall n \in \mathbb{N}\right\}.$$

We observe that there must be an $R_N = A_N \times B_N$ with B_N uncountable (otherwise $\{B_n\}$ cannot cover $[0,1]$). Moreover, we may assume that $\mu(A_N) > 0$. Hence $(\mu\times\nu)(A_N \times B_N) = \mu(A_N)\mu(B_N) = \infty$, so $(\mu\times\nu)^*(D) = \infty$. It is not so obvious that D is measurable. To see this, let $D_n = \bigcup_{j=1}^{n}[(j-1)/n, j/n] \times [(j-1)/n, j/n]$, then $D = \bigcap_{n=1}^{\infty} D_n$. Now we can write $\mu\times\nu(D) = \infty$.

Example 4.6 Let $X=Y=\mathbb{N}, \mathcal{M}=\mathcal{N}=\mathcal{P}(\mathbb{N}), \mu=\nu=$ counting measure. Define $f(m,n)=1$ if $m=n$, $f(m,n)=-1$ if $m=n+1$, and $f(m,n)=0$ otherwise. Then $\int |f| d(\mu\times\nu)=\infty$, and $\iint f d\mu d\nu$ and $\iint f d\nu d\mu$ exist and are unequal.

Proof $\int |f| d(\mu\times\nu)=\sum_{m,n\in\mathbb{N}}|f(m,n)|$ is clearly ∞. For the second one,

$$\begin{aligned}\int_{\mathbb{N}}\int_{\mathbb{N}} f(m,n)d\mu(m)d\nu(n) &= \sum_{n\in\mathbb{N}}\sum_{m\in\mathbb{N}} f(m,n)\\ &= \sum_{n\in\mathbb{N}}(f(n,n)+f(n+1,n))\\ &= \sum_{n\in\mathbb{N}}(1-1)=0.\end{aligned}$$

For the last one,

$$\begin{aligned}\int_{\mathbb{N}}\int_{\mathbb{N}} f(m,n)d\nu(n)d\mu(m) &= \sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}} f(m,n)\\ &= \sum_{m\in\mathbb{N}}(f(m,m)+f(m,m-1))\\ &= f(1,1)+f(1,0)+\sum_{m\geq 2}(1-1)=1.\end{aligned}$$

Now we do some calculations using Fubini's theorem. The first example comes from a formula useful in proving the Fourier inversion formula. Recall that the Fourier transform $\widehat{f}$ of $f\in L^1(\mathbb{R}^d)$ is given by $\widehat{f}(\xi)=\int_{\mathbb{R}^d} f(x)e^{-2\pi i x\cdot\xi}dx$.

Example 4.7 (a multiplication formula) Suppose $f,g\in L^1(\mathbb{R}^d)$, then

$$\int_{\mathbb{R}^d}\widehat{f}(\xi)g(\xi)d\xi=\int_{\mathbb{R}^d} f(y)\widehat{g}(y)dy.$$

Proof By Fubini's theorem,

$$\begin{aligned}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} f(y)e^{-2\pi i y\xi}dy\right)g(\xi)d\xi &= \int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d} g(\xi)e^{-2\pi i y\xi}d\xi\right)f(y)dy\\ &= \int_{\mathbb{R}^d} f(y)\widehat{g}(y)dy.\end{aligned}$$

Example 4.8 (distribution function) Let f be a measurable function on X, the *distribution function* of f is the function d_f on $[0,\infty)$ defined by

$$d_f(\alpha)=\mu(\{x\in X:|f(x)|>\alpha\}).$$

Let (X,μ) be a σ-finite measure space, and suppose that $|f|^p\in L^1, 0<p<\infty$. Prove

$$\|f\|_{L^p}^p:=\int_X |f|^p d\mu=p\int_0^\infty \alpha^{p-1}d_f(\alpha)d\alpha.$$

Proof By passing to a characteristic function,

$$p\int_0^\infty \alpha^{p-1}d_f(\alpha)d\alpha=p\int_0^\infty \alpha^{p-1}\int_X \chi_{\{x:|f(x)|>\alpha\}}d\mu(x)d\alpha.$$

Now we need to change the order of integration. Observe that

$$\chi_{\{x:|f(x)|>\alpha\}}(x,\alpha)=\begin{cases}0 & \text{if } |f(x)|\leq\alpha,\\ 1 & \text{if } |f(x)|>\alpha.\end{cases}$$

Fix x, then the x-section of χ is

$$\chi_{\{x:|f(x)|>\alpha\}}(\alpha)=\begin{cases}0 & \text{if } \alpha\geq|f(x)|,\\ 1 & \text{if } \alpha<|f(x)|.\end{cases}$$

Hence,

$$
\begin{aligned}
p\int_0^\infty \alpha^{p-1}\int_X \chi_{\{x:|f(x)|>\alpha\}}d\mu(x)d\alpha &= p\int_X\int_0^\infty \alpha^{p-1}\chi_{\{x:\alpha<|f(x)|\}}(\alpha)d\alpha d\mu(x)\\
&= \int_X\int_0^{|f(x)|} p a^{p-1}d\alpha d\mu(x)\\
&= \int_X |f(x)|^p d\mu(x).
\end{aligned}
$$

4.6 Lebesgue Integral

In this section we study the change of variable formula of Lebesgue integral. This is a useful tool and readers are expected to know how to apply those formulas, but the proof can be skipped at the first time. We begin by reviewing basic properties of Lebesgue measure on $\mathbb{R}^d$.

Definition 4.9 (Lebesgue measure)

Lebesgue measure *m^d on $\mathbb{R}^d$ is the completion of the product of Lebesgue measure on $\mathbb{R}$ with itself, that is, the completion of $m\times\cdots\times m$ on $\mathcal{B}_\mathbb{R}\otimes\cdots\otimes\mathcal{B}_\mathbb{R}=\mathcal{B}_{\mathbb{R}^d}$, or equivalently the completion of $m\times\cdots\times m$ on $\mathcal{L}\otimes\cdots\otimes\mathcal{L}$. The domain $\mathcal{L}^d$ of m^d is called the class of* ***Lebesgue measurable sets*** *in $\mathbb{R}^d$, or just the* ***Lebesgue σ-algebra****. We shall usually omit the superscript d and write m for the d-dimensional Lebesgue measure. In the case $d=1$, the integral is usally written as $\int f(x)dx$ in place of $\int f dm$.* ♣

4.6.1 Translation-Invariance

Theorem 4.14 (translation-invariance)

Lebesgue measure is translation-invariant. For $a\in\mathbb{R}^d$ define

$$\tau_a:\mathbb{R}^d\to\mathbb{R}^d$$

$$\tau_a(x)=x+a.$$

- *If $E\in\mathcal{L}^d$, then $\tau_a(E)\in\mathcal{L}^d$, and $m(\tau_a(E))=m(E)$.*
- *If $f:\mathbb{R}^d\to\mathbb{C}$ is Lebesgue measurable, then so is $f\circ\tau_a$. Moreover, if either $f\geq 0$ or $f\in L^1(m)$, then*

$$\int (f\circ\tau_a)dm=\int f dm.$$

♡

Proof We first show the $d=1$ case, where we will invoke the measure-construction theorem (Theorem 3.6). Denote $m_a(E)=m(E+a)$, then for any intervals we have

$$m_a\left(\bigcup_{j=1}^N I_j\right)=m\left(\bigcup_{j=1}^N (I_j+a)\right),$$

hence m_a and m agree on the algebra $\mathcal{A}$ (the algebra in section 3.3), then they induces the same measure on $\mathcal{B}_\mathbb{R}$ by Theorem 3.6. Therefore, $m_a=m$.

Step 1: Rectangle. Let $E = E_1 \times \cdots \times E_d$ be a rectangle, where each $E_j \in \mathcal{B}_{\mathbb{R}}$. Then

$$\tau_a(E) = \{(x_1, \cdots, x_d) + (a_1, \cdots, a_d) : x_j \in E_j\} = (E_1 + a_1) \times \cdots \times (E_d + a_d).$$

Hence

$$\begin{aligned} m(\tau_a(E)) &= m((E_1 + a_1) \times \cdots \times (E_d + a_d)) \\ &= m(E_1 + a_1)m(E_2 + a_2) \cdots m(E_d + a_d) \\ &= m(E_1)m(E_2) \cdots m(E_d) \\ &= m(E_1 \times \cdots \times E_d) = m(E). \end{aligned}$$

Step 2: Finite union of rectangles. Suppose E is a finite disjoint union of rectangles, then

$$\tau\left(\bigcup_{n=1}^{N} E_n\right) = \bigcup_{n=1}^{N} E_n + a = \bigcup_{n=1}^{N} (E_n + a). \quad \text{(check this!)}$$

Now

$$\begin{aligned} m\left(\bigcup_{n=1}^{N} (E_n + a)\right) &= \sum_{n=1}^{N} m(E_n + a) \\ &= \sum_{n=1}^{N} m(E_n) \\ &= m\left(\bigcup_{n=1}^{N} E_n\right) = m(E). \end{aligned}$$

Step 3: Invoke Theorem 3.6. Similar to the $d = 1$ case, we view $m \circ \tau_a$ as another measure, and by the previous steps we see that m and $m \circ \tau_a$ agree on the algebra generated by rectangles[2], by the measure-construction theorem (3.6), they induces the same measure on $\mathcal{B}_{\mathbb{R}^d}$. Clearly m itself is induced by m(restricted to the algebra), hence $m = m \circ \tau_a$ on $\mathcal{B}_{\mathbb{R}^d}$. We are not done yet!

Step 4: Passing from Borel to Lebesgue. Since each Lebesgue measurable set is a union of a Borel set and a set of measure 0, the proof is complete once we solve the case of null set. First suppose that N is a Borel set with $m(N) = 0$, then for any $\varepsilon > 0$ there is an open set $\mathcal{O} \supset N$ with[3]

$$m(\mathcal{O} \setminus N) = m((\mathcal{O} \setminus N) + a) = m((\mathcal{O} + a) \setminus (N + a)) < \varepsilon,$$

hence $N + a$ is a Borel null set. Now we copy the definition of a complete measure:

> If $\mu(E) = 0$ and $F \subset E$, then $\mu(F)$ should equal to 0, but F is not necessarily in $\mathcal{M}$.
> A measure whose domain includes all subsets of null sets is called **complete**.

The Lebesgue σ-algebra thus contains all subsets of Borel null sets, that is, $m(N_0) = 0$ if $N_0 \subset N \in \mathcal{B}_{\mathbb{R}^d}$ with $m(N) = 0$. It is then clear that $m(N_0 + a) \leq m(N + a) = 0$. Now Suppose $E \in \mathcal{L}^d$, then $E = G \cup N$, where G is a Borel set and $m(N) = 0$, then

$$m(E + a) = m((G + a) \cup (N + a)) \leq m(G + a) + m(N + a) = m(G) + 0 = m(E),$$

[2]If you cannot understand this sentence, see section of product measures (4.5) or watch my video.

[3]$\mathcal{O} + a$ is open because the translation τ_a is a homeomorphism on $\mathbb{R}^d$.

completing the proof of the first assertion.

For the second assertion, let f be Lebesgue measurable and B be a Borel set in $\mathbb{C}$. Then $f^{-1}(B) \in \mathcal{L}^d$, hence $f^{-1}(B) = E \cup N$ where E is Borel and $m(N) = 0$. Since $\tau_a^{-1}(E)$ is Borel and $m(\tau_a^{-1}(N)) = m(\tau_{-a}(N)) = 0$, it follows that

$$(f \circ \tau_a)^{-1}(B) = \tau_a^{-1}(f^{-1}(B)) = \tau_a^{-1}(E \cup N) = \tau_a^{-1}(E) \cup \tau_a^{-1}(N) \in \mathcal{L}^d.$$

Hence $f \circ \tau_a$ is Lebesgue measurable. When $f = \chi_E$, the equality $\int (f \circ \tau_a) dm = \int f dm$ reduces to $m(\tau_{-a}(E)) = m(E)$. Then this is true for simple functions, and by the monotone convergence theorem we extends to nonnegative measurable functions. Taking positive and negative parts of real and imaginary parts then yields the result for $f \in L^1(m)$. $\square$

4.6.2 Linear Change of Variable

Let $e_1, \cdots, e_d$ be the standard basis of $\mathbb{R}^d$, and let T be a linear map on $\mathbb{R}^d$. Denote $GL(d, \mathbb{R})$ the group of invertible linear maps of $\mathbb{R}^d$. There are 3 types of elementary linear maps:

$$\begin{aligned} T_1(x_1, \cdots, x_j, \cdots, x_d) &= (x_1, \cdots, cx_j, \cdots, x_d); \\ T_2(x_1, \cdots, x_j, \cdots, x_d) &= (x_1, \cdots, x_j + cx_k, \cdots, x_d) \\ T_3(x_1, \cdots, x_j, \cdots, x_k, \cdots, x_d) &= (x_1, \cdots, x_k, \cdots, x_j, \cdots, x_d). \end{aligned}$$

From linear algebra, any $T \in GL(d, \mathbb{R})$ can be written as the product of finitely many elementary linear maps. (Every nonsingular matrix can be row-reduced to the identity matrix).

Theorem 4.15 (change of variable formula)

Suppose $T \in GL(d, \mathbb{R})$ and f is a Lebesgue measurable function on $\mathbb{R}^d$. Then $f \circ T$ is Lebesgue measurable. If $f \geq 0$ or $f \in L^1(m)$, then

$$\int f(x) dx = |\det T| \int f \circ T(x) dx. \tag{4.1}$$

♡

Proof **Step 1: Simplification.** Suppose that f is Borel measurale. Since T is a linear map, T is continuous. Hence $f \circ T$ is Borel measurable.[4] Observation: if the change of variable formula is true for the maps T and S, it is also true for $T \circ S$, because

$$\begin{aligned} \int f(x) dx &= |\det T| \int f \circ T(x) dx \\ &= |\det T| \int (f \circ T)(x) dx \quad \text{(apply the change of variable formula to } f \circ T) \\ &= |\det T| |\det S| \int (f \circ T) \circ S(x) dx \\ &= |\det(T \circ S)| \int f \circ (T \circ S)(x) dx. \end{aligned}$$

With this observation, it suffices to prove (4.1) when T is of the types T_1, T_2, T_3 described above.

Step 2: Elementary linear maps. We apply Fubini's theorem.

[4] Warning: If f is Lebesgue measurable and g is continuous, it does not follow that $f \circ g$ is Lebesgue measurable.

- For T_3, we have

$$\begin{aligned}
&\int_{\mathbb{R}^d} (f \circ T_3)(x_1, \cdots, x_j, \cdots, x_k, \cdots, x_d) dx_1 \cdots dx_j \cdots dx_k \cdots dx_d \\
&= \int_{\mathbb{R}^d} f(x_1, \cdots, x_k, \cdots, x_j, \cdots, x_d) dx_1 \cdots dx_j \cdots dx_k \cdots dx_d \\
&= \int_{\mathbb{R}^d} f(x_1, \cdots, x_k, \cdots, x_j, \cdots, x_d) dx_1 \cdots dx_k \cdots dx_j \cdots dx_d \\
&= \int_{\mathbb{R}^d} f(x) dx = |-1| \int_{\mathbb{R}^d} f(x) dx
\end{aligned}$$

since $\det T_3 = -1$.

- For T_1, we use the one-dimensional dilation formula (learned in Real Analysis I): $\int f(t)dt = |c| \int f(ct)dt$. Then,

$$\begin{aligned}
\int (f \circ T_1)(x) dx &= \int \cdots \left(\int f(x_1, \cdots, cx_j, \cdots, x_d) dx_j \right) dx_1 \cdots dx_d \\
&= \frac{1}{|c|} \int \cdots \int f(x_1, \cdots, x_j, \cdots, x_d) dx_j dx_1 \cdots dx_d \\
&= \frac{1}{|c|} \int f(x) dx \\
&= \frac{1}{|\det T_1|} \int f(x) dx.
\end{aligned}$$

- For T_3, by translation-invariance we have

$$\int f(x_1, \cdots, x_j + cx_k, \cdots, x_d) dx_j = \int f(x_1, \cdots, x_j, \cdots, x_d) dx_j.$$

The result follows by Fubini's theorem and $|\det T_2| = 1$.

Step 3: Passing from Borel to Lebesgue. Take $f = \chi_E$ where E is Borel, we have

$$m(E) = |\det T| \int \chi_E(Tx) dx = |\det T| m(T(E)).$$

Since T is invertible, T^{-1} is a linear map, hence is continuous. Now $T(E)$ is Borel whenever E is Borel. In particular, if N is a Borel null set, then there is an open set $\mathcal{O}$ such that $m(\mathcal{O} \setminus N) < \varepsilon$. Then $T(\mathcal{O} \setminus N) = T(\mathcal{O}) \setminus T(N)$, so

$$m(T(\mathcal{O}) \setminus T(N)) = m(T(\mathcal{O} \setminus N)) = |\det T| m(\mathcal{O} \setminus N) < |\det T| \varepsilon,$$

it follows that $T(N)$ is also a Borel null set. The same is true for T^{-1}. Since every Lebesgue null set is a subset of some Borel null set, we have, In summary,

the class of Lebesgue null sets is invariant under T and T^{-1}.

Now suppose f is a Lebesgue measurable function. Then for any Borel set E,

$$\begin{aligned}
(f \circ T)^{-1}(E) &= T^{-1}(f^{-1}(E)) \quad f^{-1}(E) \text{ is a Lebesgue measurable set} \\
&= T^{-1}(G \cup N) \\
&= T^{-1}(G) \cup T^{-1}(N),
\end{aligned}$$

which is Lebesgue measurable since $T^{-1}(G)$ is Borel and $T^{-1}(N)$ is null. This shows that $f \circ T$ is Lebesgue measurable, and by the same computation made above, the change of variable formula is proved.

Corollary 4.2

Suppose $T \in GL(d, \mathbb{R})$. If $E \in \mathcal{L}^d$, then $T(E) \in \mathcal{L}^d$ and $m(T(E)) = |\det T| m(E)$. ♡

Corollary 4.3

Lebesgue measure is invariant under rotations. ♡

Proof Rotations are orthogonal linear maps satisfying $TT^* = I$, where T^* is the transpose of T. Since $\det T = \det T^*$, $|\det T^*| = 1$. □

4.6.3 Differentiable Change of Variable

Let $G = (g_1, \cdots, g_d)$ be a map from an open set $\Omega \subset \mathbb{R}^d$ into $\mathbb{R}^d$ whose components $g_j \in C^1$. Denote $D_x G$ to be the Jacobian of G at x. i.e.,

$$D_x G = \begin{pmatrix} (\partial g_1/\partial x_1)(x) & \cdots & (\partial g_1/\partial x_d)(x) \\ \vdots & & \vdots \\ (\partial g_d/\partial x_1)(x) & \cdots & (\partial g_d/\partial x_d)(x) \end{pmatrix}.$$

Definition 4.10 (diffeomorphism)

G is called a C^1 ***diffeomorphism*** *if G is injective and $D_x G$ is invertible for all $x \in \Omega$.* ♣

If G is a C^1 diffeomorphism, by the inverse function theorme, $G^{-1} : G(\Omega) \to G$ is also a C^1 diffeomorphism and

$$D_x(G^{-1}) = [D_{G^{-1}(x)} G]^{-1} \quad \forall x \in G(\Omega).$$

Theorem 4.16

Suppose Ω is an open set in $\mathbb{R}^d$ and $G : \Omega \to \mathbb{R}^d$ is a C^1 diffeomorphism.

1. *If f is a Lebesgue measurable function on $G(\Omega)$, then $f \circ G$ is Lebesgue measurable on Ω. If $f \geq 0$ or $f \in L^1(G(\Omega, m))$, then*
$$\int_{G(\Omega)} f(x) dx = \int_\Omega f \circ G(x) |\det D_x G| dx.$$
2. *If $E \subset \Omega$ and $E \in \mathcal{L}^d$, then $G(E) \in \mathcal{L}^d$ and*
$$m(G(E)) = \int_E |\det D_x G| dx.$$

♡

4.7 Integration in Polar Coordinates

4.7.1 Homeomorphism: $\mathbb{R}^n \setminus \{0\} \cong (0, \infty) \times S^{n-1}$

For any nonzero $x \in \mathbb{R}^2$, we can express it in the polar coordinate: $x = (r \cos\theta, r \sin\theta)$, where $r = |x| = (x_1^2 + x_2^2)^{1/2}$, $(\cos\theta, \sin\theta) = x/|x| := x'$. Conversely, every element $(r, x') \in (0, \infty) \times S_1$ corresponds to a unique element $rx' \in \mathbb{R}^2 \setminus \{0\}$. Thus we obtain a homeomorphism (continuous bijection)

$$\begin{aligned} \Phi : \mathbb{R}^2 \setminus \{0\} &\to (0, \infty) \times S^1 \\ x &\mapsto (|x|, x/|x|). \end{aligned}$$

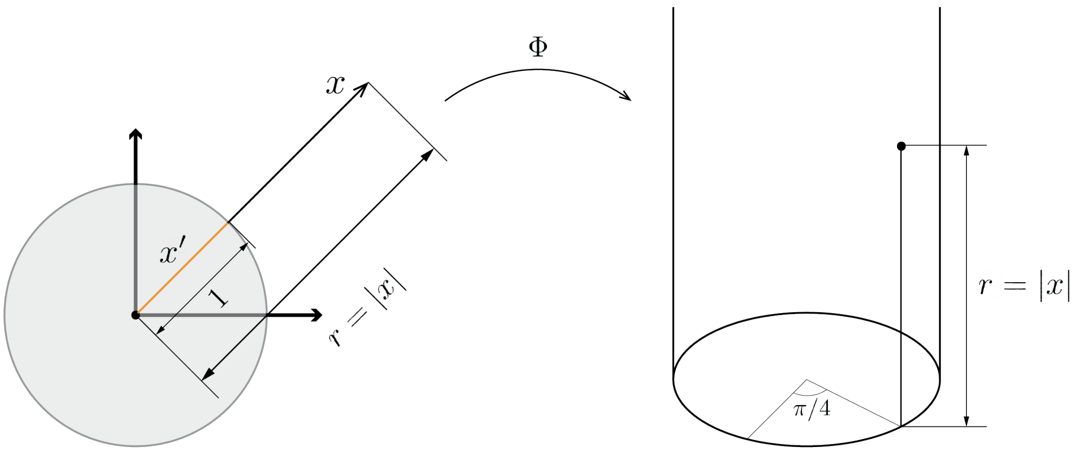


Figure 4.2: the cylinder homeomorphism

The idea is to "decompose" a point into its direction(represented by a unit vector) and norm. We now proceed to n-dimensional case. Let the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$.

Definition 4.11
If $x \in \mathbb{R}^n$ and $x \neq 0$, the polar coordinates of x are

$$r = |x| \in (0, \infty), \quad x' = \frac{x}{|x|} \in S^{n-1}.$$

♣

We define a map $\Phi : \mathbb{R}^n \setminus \{0\} \to (0, \infty) \times S^{n-1}$ by

$$\Phi(x) = (r, x') = \left(|x|, \frac{x}{|x|}\right).$$

This maps gives a homeomorphism of $\mathbb{R}^n \setminus \{0\}$ and the cylinder $(0, \infty) \times S^{n-1}$.

Example 4.9 If $n = 2$, then $\Phi((\sqrt{3}, 1)) = \left(2, (\sqrt{3}/2, 1/2)\right)$, $\Phi^{-1}(2, (\sqrt{3}/2, 1/2)) = 2(\sqrt{3}/2, 1/2) = (\sqrt{3}, 1)$.

4.7.2 Motivation

How is this homeomorphism connected with Lebesgue measure? We have all learnt change of variable in polar coordinates in calculus. For example,

$$\iint_{B(0,1)} f(x, y) dxdy = \int_0^{2\pi} \int_0^1 f(r\cos\theta, r\sin\theta) r dr d\theta.$$

Since an integration formula always starts with characteristic functions, we can let $f = 1$ for convenience. Then,

$$\iint_{B(0,1)} dxdy = \int_0^{2\pi} \int_0^1 r dr d\theta = \left(\int_0^\infty r\chi_{(0,1)} dr\right)\left(\int_0^{2\pi} d\theta\right). \tag{4.2}$$

For the $n = 3$ case, we have the integration in spherical coordinate:

$$\iiint_{B(0,1)} dxdydz = \int_0^\pi \int_0^{2\pi} \int_0^1 r^2 \sin\phi \, dr d\theta d\phi = \left(\int_0^\infty r^2 \chi_{(0,1)} \, dr\right)\left(\int_0^\pi \int_0^{2\pi} \sin\phi \, d\theta d\phi\right). \tag{4.3}$$

It seems that the Lebesgue measure in $\mathbb{R}^n \setminus \{0\}$ can be viewed as a product measure on $(0, \infty) \times S^{n-1}$:

- For $n = 2$, the measure (area) of unit disc is the product of $\rho(0,1) = \int_0^\infty r\chi_{(0,1)}dr$ and $\sigma(S^1) = \int_0^{2\pi} d\theta$, hence m may be decomposed into ρ and σ.
- For $n = 3$, the measure (volume) of unit ball is the product of $\rho(0,1) = \int_0^\infty r^2\chi_{(0,1)}\, dr$ and $\sigma(S^2) = \int_0^\pi \int_0^{2\pi} \sin\phi \; d\theta d\phi = 4\pi$, which is exactly the surface area of the unit ball.

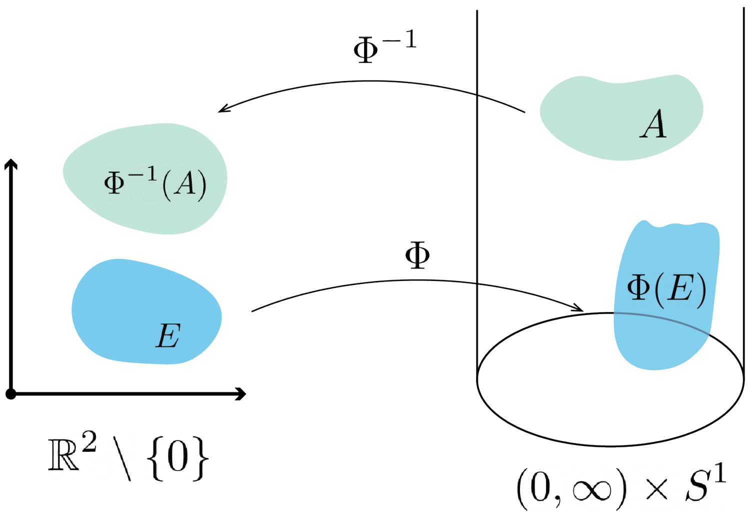
4.7.3 Surface Measure and Polar Integration Formula

Now we induce a measure via the homeomorphism. If $E \subset (0,\infty) \times S^{n-1}$ is a Borel set, then $\Phi^{-1}(E)$ is a Borel subset of $\mathbb{R}^n \setminus \{0\}$ since Φ is continuous. It is reasonable to require the measure is invariant under the homeomorphism Φ. Our new Borel measure m_* on $(0,\infty) \times S^{n-1}$ should satisfy

$$m(E) = m_*(\Phi(E)).$$

Or equivalently,

$$m_*(A) = m_*(\Phi^{-1}(A)) \text{ for all Borel sets } A \subset (0,\infty) \times S^{n-1}.$$



Definition 4.12

Define m_* to be the Borel measure on $(0,\infty) \times S^{n-1}$ induced by Φ:

$$m_*(A) = m(\Phi^{-1}(A)) \textit{ for all Borel sets } A \subset (0,\infty) \times S^{n-1}.$$

Define the measure ρ on $(0,\infty)$ by

$$\rho(I) = \int_I r^{n-1} dr.$$

♣

Our next goal is to decompose this measure into a product of two measures.

Theorem 4.17

There is a unique Borel measure $\sigma = \sigma_{n-1}$ on S^{n-1} (usually called the surface measure) such that $m_* = \rho \times \sigma$.

If f is Borel measurable on $\mathbb{R}^n$ and $f \geq 0$ or $f \in L^1(m)$, then

$$\int_{\mathbb{R}^n} f(x)dx = \int_{S^{n-1}} \int_0^\infty f(rx')r^{n-1}drd\sigma(x') \tag{4.4}$$

♡

Proof **Step 1: Contruct σ.** Let E be a Borel set in S^{n-1}, for $a > 0$ let

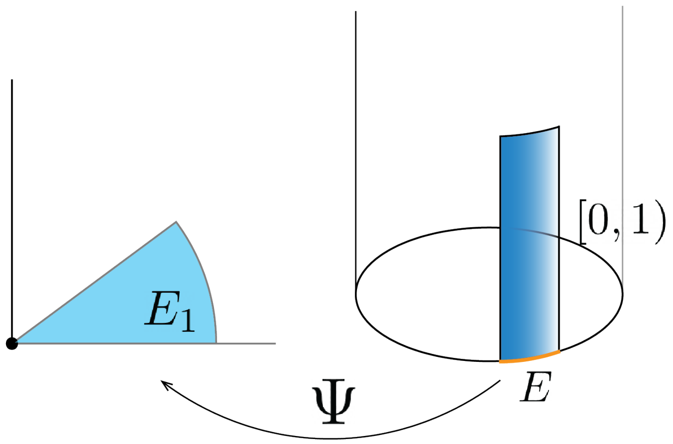
$$E_a = \Phi^{-1}((0,a] \times E) = \{rx' : 0 < r \leq a, x' \in E\}.$$

Since σ must satisfy the formula (4.4), we assume it holds when $f = \chi_{E_1}$, then

$$m(E_1) = \int_0^1 \int_E r^{n-1}d\sigma(')dr = \sigma(E)\int_0^1 r^{n-1}dr = \frac{\sigma(E)}{n}.$$

We therefore define $\sigma(E)$ to be $n \cdot m(E_1)$.

Step 2: Verify that σ is a measure. Denote the map $E \mapsto E_1$ by $\Psi(E) = \Phi^{-1}((0,1] \times E)$,



and let $\{E_k\}_{k=1}^\infty \subset \mathcal{B}_{S^{n-1}}$. Then Ψ satisfies

- $\Psi(\bigcup_{k=1}^\infty E_k) = \Phi^{-1}((0,1] \times \bigcup_{k=1}^\infty E_k) = \bigcup_{k=1}^\infty \Phi^{-1}((0,1] \times E_k) = \bigcup_{k=1}^\infty \Psi(E_k)$
- $\Psi(\bigcap_{k=1}^\infty E_k) = \bigcap_{k=1}^\infty \Psi(E_k)$.
- $\Psi(E^c) = \Psi(E)^c$.

Hence

$$\begin{aligned}
\sigma(\bigcup_{k=1}^\infty E_k) &= n \cdot m(\Psi(\bigcup_{k=1}^\infty E_k)) \\
&= n \cdot m(\bigcup_{k=1}^\infty \Psi(E_k)) \\
&= n \cdot \sum_{k=1}^\infty m(\Psi(E_k)) \\
&= \sum_{k=1}^\infty \sigma(E_k).
\end{aligned}$$

This shows that σ is a measure on S^{n-1}.

Step 3: $m_* = \rho \times \sigma$ on every rectangle. It is clear that $E_a = aE_1$, hence by the change of variable formula, $m(E_a) = a^n m(E_1)$, and hence, if $0 < a < b$,

$$\begin{aligned}
m_*((a,b] \times E) &= m_*((0,b] \times E \setminus (0,a] \times E) \\
&= m(\Phi^{-1}((0,b] \times E \setminus (0,a] \times E)) \\
&= m(\Phi^{-1}((0,b] \times E) \setminus \Phi^{-1}((0,a] \times E))) \\
&= m(E_b \setminus E_a) = m(E_b) - m(E_a) \\
&= (b^n - a^n)\frac{\sigma(E)}{n} \\
&= \sigma(E) \int_a^b r^{n-1} dr \\
&= \rho((a,b])\sigma(E) \\
&= (\rho \times \sigma)((a,b] \times E).
\end{aligned}$$

Remember that m_* and $\rho \times \sigma$ are defined independently from each other! And we have just shown that they agree on every rectangle $(a,b] \times E$, where $(a,b] \in \mathcal{B}_{(0,\infty)}, E \in \mathcal{B}_{S^{n-1}}$.

Step 4: $m_* = \rho \times \sigma$ on $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{n-1}}$. Fix $E \in \mathcal{B}_{S^{n-1}}$, let $\mathcal{A}_E$ be the collection of all finite disjoint unions of $(a,b] \times E$. Then $\mathcal{A}_E$ is an algebra on $(0,\infty) \times E$, which generates the σ-algebra $\mathcal{M}_E = \{A \times E : A \in \mathcal{B}_{(0,\infty)}$. It is easy to verify that $m_* = \rho \times \sigma$ on $\mathcal{A}_E$. By the measure-construction theorem, $m_* = \rho \times \sigma$ on $\mathcal{M}_E$, which holds for every $E \in \mathcal{B}_{S^{n-1}}$, hence $m_* = \rho \times \sigma$ on $\bigcup_{E \in \mathcal{B}_{S^{n-1}}} \mathcal{M}_E$. But $\bigcup_{E \in \mathcal{B}_{S^{n-1}}} \mathcal{M}_E$ is exactly the collection of all Borel rectangles in $(0,\infty) \times S^{n-1}$, which generates $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{n-1}}$. Invoking the measure-construction theorem again leads to

$$m_* = \rho \times \sigma \text{ on } \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{n-1}}.$$

Step 5: Integration formula. Let f be a characteristic function χ_{E_1}, where $E_1 = \Phi^{-1}((0,1] \times E\}$ as given in Step 1. Then

$$\begin{aligned}
\int_0^\infty \int_{S^{n-1}} \chi_{E_1}(rx') r^{n-1} d\sigma(x') dr &= \int_0^a \sigma(E) r^{n-1} dr \\
&= \rho((0,a])\sigma(E) \\
&= (\rho \times \sigma)((0,a] \times E) \\
&= m(\Phi^{-1}((0,a] \times E) \\
&= m(E_1),
\end{aligned}$$

where

$$\chi_{E_1}(rx') = \begin{cases} 1 & rx' \in E_1 \iff x' \in E \text{ and } r \in (0,a] \\ 0 & \text{otherwise.} \end{cases}$$

And from Step 3 we have

$$m_*((a,b] \times E) = m(E_b \setminus E_a) = \sigma(E) \int_a^b r^{n-1} dr = \int_{S^{n-1}} \int_0^\infty \chi_{\Phi^{-1}((a,b] \times E)}(rx') dr d\sigma(x').$$

This shows the measure m_* can also be given in the form of the integral as above! We can introduce a new function ν on $(0,\infty) \times S^{n-1}$ by setting

$$\nu((a,b] \times E) = \int_{S^{n-1}} \int_0^\infty \chi_{\Phi^{-1}((a,b] \times E)}(rx') dr d\sigma(x'),$$

then one can easily check that ν is countably additive. Since $m_* = \nu$ on all rectangles, they induce the same measure: $m_* = \nu$ on $\mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{n-1}}$. Thus we obtain

$$m_*(A) = m(\Phi^{-1}(A)) = \int_{S^{n-1}} \int_0^\infty \chi_{\Phi^{-1}(A)}(rx')d\sigma(x')dr, \quad A \in \mathcal{B}_{(0,\infty)} \otimes \mathcal{B}_{S^{n-1}}.$$

But Φ^{-1} is a homeomorphism, so $\Phi^{-1}(A)$ runs through all the Borel sets of $\mathbb{R}^n \setminus \{0\}$. Thus we conclude

$$\int \chi_B(x)dx = m(B) = \int_{S^{n-1}} \int_0^\infty \chi_{\Phi^{-1}(A)}(rx')d\sigma(x')dr$$

for every Borel set B in $\mathbb{R}^n \setminus \{0\}$.

It then follows for general f by the usual linearity and approximation arguments. The proof is complete. □

4.7.4 Applications

Proposition 4.14 (radial functions)

If f is a measurable function on $\mathbb{R}^d$, nonnegative or integrable, such that $f(x) = g(|x|)$ for some function g on $(0,\infty)$, then

$$\int f(x)dx = \sigma(S^{n-1}) \int_0^\infty g(r)r^{n-1}dr.$$

Proof Applying the integration formula and recall that $r = |x|$ in the definition of a polar coordinate. Then

$$\begin{aligned}\int f(x)dx &= \int_0^\infty \int_{S^{n-1}} g(r)r^{n-1}d\sigma(x')dr \\ &= \sigma(S^{n-1}) \int_0^\infty g(r)r^{n-1}dr.\end{aligned}$$

□

Proposition 4.15 (integrability of $|x|^{-a}$)

Let c and C be positive constants, $B = \{x \in \mathbb{R}^n : |x| < c\}$. Suppose f is a measurable function on $\mathbb{R}^n$.

1. *If $|f(x)| \le C|x|^{-a}$ for some $a < n$, then $f \in L^1(B)$.*
2. *If $|f(x)| \ge C|x|^{-n}$ on B, then $f \notin L^1(B)$.*
3. *If $|f(x)| \le C|x|^{-a}$ on B^c for some $a > n$, then $f \in L^1(B^c)$.*
4. *If $|f(x)| \ge C|x|^{-n}$ on B^c, then $f \notin L^1(B^c)$.*

Proof We write $\chi_B(x)$ in the form of a radial function. Observe that $x \in B$ if and only if $|x| < c$, hence $\chi_B(x) = \chi_{[0,c)}(|x|)$. Let $g(|x|) = |x|^{-a}\chi_{[0,c)}(|x|)$ and apply (4.14), then

$$\int |x|^{-a}\chi_B(x)dx = \sigma(S^{n-1}) \int_0^c r^{-a}r^{n-1}dr = \sigma(S^{n-1})\frac{r^{n-a}}{n-a}\Big|_0^c,$$

hence (1),(2) are proved. For (3),(4), observe that

$$\int |x|^{-a}\chi_{B^c}(x)dx = \sigma(S^{n-1}) \int_c^\infty r^{-a}r^{n-1}dr = \sigma(S^{n-1})\frac{r^{n-a}}{n-a}\Big|_c^\infty.$$

□

Example 4.10 (Gaussian) If $a > 0$, then

$$I_n = \int_{\mathbb{R}^n} e^{-a|x|^2}dx = \left(\frac{\pi}{a}\right)^{n/2}.$$

Proof

$$I_2 = \int_{S^1} d\sigma \int_0^\infty e^{-ar^2} r dr = 2\pi \int_0^\infty e^{-ar^2} r dr = \frac{\pi}{a}.$$

By Tonelli's theorem, $I_n = I_1^n$, and $I_1 = \sqrt{I_2} = \left(\frac{\pi}{a}\right)^{1/2}$, so $I_n = (\pi/a)^{n/2}$.

Definition 4.13 (Gamma Function)

Define

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt \quad (\operatorname{Re} z > 0).$$

Properties $\Gamma(z+1) = z\Gamma(z)$.

Proposition 4.16

$$\sigma(S^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}.$$

Proof By polar integration formula,

$$\int_{\mathbb{R}^n} e^{-|x|^2} dx = \int_{S^{n-1}} d\sigma \int_0^\infty r^{n-1} e^{-r^2} dr.$$

Substitute $s = r^2$, we have

$$\pi^{n/2} = \frac{\sigma(S^{n-1})}{2} \int_0^\infty s^{\frac{n}{2}-1} e^{-s} ds = \frac{\sigma(S^{n-1})}{2} \Gamma\left(\frac{n}{2}\right).$$

□

We now return to the intuitive observation of the surface measure: it measures the surface area of a sphere. This fact can be described in terms of **weak convergence**.

Definition 4.14 (weak convergence)

*The sequence $\{\mu_j\}$ of Borel measures on $\mathbb{R}^n$ **converges weakly** to a Borel measure μ if for all $\varphi \in C_c(\mathbb{R}^n)$,*

$$\int \varphi d\mu_j \to \int \varphi d\mu.$$

Proposition 4.17

σ^{n-1} is the weak limit of the measures $\delta^{-1} m|_{B(0,1+\delta)\setminus B(0,1)}$ as $\delta \to 0$.

Proof

Chapter 5 Complex Measures

Suppose $f : X \to [0, \infty]$ is integrable, then we can define a measure ν by setting

$$\nu(E) = \int_E f d\mu, \quad (E \in \mathcal{M}).$$

Under this notation, we have

$$\int g \, d\nu = \int gf \, d\mu,$$

and this is why we often write $d\nu = f \, d\mu$. It is natural to let f be a complex function on X so that ν will be a complex-valued function, which behaves like a measure. This is what we will learn in this chapter: the complex measure.

5.1 The Vector Space of Complex Measures

5.1.1 Definition and Properties

Let $f \in L^1(\mathbb{R}^d)$, then the function $\nu : \mathcal{L} \to \mathbb{R}$ defined by

$$\nu(E) = \int_E f(x)dx$$

satisfies countable additivity: $\nu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \nu(E_n)$ for any disjoint sequence of sets $\{E_n\}_{n\in\mathbb{N}}$. However, it is not necessary that $\nu(E) \geq 0$ for all measurable sets E. (for example, take f a negative function).

Definition 5.1 (complex measure)

Let $(X, \mathcal{M})$ be a measurable space.

- *A function $\nu : \mathcal{M} \to \mathbb{F}$ is called countably additive if*

$$\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \nu(E_n)$$

for every disjoint sequence $\{E_n\}_{n\in\mathbb{N}} \subset \mathcal{M}$.

- *A **complex measure** on $\mathcal{M}$ is a countably additive function $\nu : \mathcal{M} \to \mathbb{C}$.*
- *A countably additive function $\nu : \mathcal{M} \to \mathbb{R}$ is sometimes called a **real measure** or **signed measure**.*

♣

Example 5.1 Define ν on the Borel subsets of $[-1, 1]$ by

$$\nu(E) = m(E \cap [0, 1]) - m(E \cap [-1, 0)).$$

Then ν is a signed measure.

Proposition 5.1

Let ν be a complex measure on the measurable space $(X, \mathcal{M})$, then

1. *$\nu(\varnothing) = 0$.*
2. *$\sum_{k=1}^{\infty} |\nu(E_k)| < \infty$ for every disjoint sequence $\{E_k\}_{k=1}^{\infty} \subset \mathcal{M}$.*

♠

Proof (1) Since $\varnothing = \varnothing \cup \varnothing \cup \cdots$, $\nu(\varnothing) = \sum_{k=1}^{\infty} \nu(\varnothing)$. This series of constant terms converges only if $\nu(\varnothing) = 0$.

(2) First suppose ν is real. Write

$$\begin{aligned}\sum_{k=1}^{\infty}|\nu(E_k)| &= \sum_{\{k:\nu(E_k)\geq 0\}} \nu(E_k) - \sum_{\{k:\nu(E_k)<0\}} \nu(E_k) \\ &= \nu(\bigcup_{\{k:\nu(E_k)\geq 0\}} E_k) - \nu(\bigcup_{\{k:\nu(E_k)<0\}} E_k) \in \mathbb{R}\end{aligned}$$

since ν is real-valued.

Suppose that ν is complex, then

$$\sum_{k=1}^{\infty}|\nu(E_k)| = \sum_{k=1}^{\infty}|\operatorname{Re}\nu(E_k)+\operatorname{Im}\nu(E_k)| \leq \sum_{k=1}^{\infty}|\operatorname{Re}\nu(E_k)|+|\operatorname{Im}\nu(E_k)| < \infty.$$

□

Proposition 5.2

Suppose ν is a complex measure on a measurable space $(X, \mathcal{M})$. Then

1. $\nu(E \setminus D) = \nu(E) - \nu(D)$ *for all* $D, E \in \mathcal{S}$ *with* $D \subset E$*;*
2. $\nu(D \cup E) = \nu(D) + \nu(E) - \nu(D \cap E)$ *for all* $D, E \in \mathcal{M}$*;*
3. $\nu\left(\bigcup_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} \nu(E_k)$ *for all increasing sequences* $E_1 \subset E_2 \subset \cdots$ *of sets in* $\mathcal{M}$*;*
4. $\nu\left(\bigcap_{k=1}^{\infty} E_k\right) = \lim_{k\to\infty} \nu(E_k)$ *for all decreasing sequences* $E_1 \supset E_2 \supset \cdots$ *of sets in* $\mathcal{M}$*.*

♠

Proof The proof is similar to the case of a positive measure, we leave it as an exercise. □

5.1.2 Vector Space Structure on Measures

Definition 5.2 (addition and scalar multiplication)

Suppose $(X, \mathcal{S})$ is a measurable space and μ, ν are complex measures on $\mathcal{S}$, $\alpha \in \mathbb{F}$. Define complex measures $\mu + \nu$ and $\alpha\mu$ on $\mathcal{S}$ by

$$(\mu + \nu)(E) = \mu(E) + \nu(E), \quad (\alpha\mu)(E) = \alpha(\mu(E)).$$

♣

We show that $\mu + \nu$ is indeed a complex measure on $\mathcal{S}$, and the reader should verify that $\alpha\mu$ is a complex measure. Let E be a countable disjoint union: $E = \bigcup_{n=1}^{\infty} E_n$, then

$$\begin{aligned}(\mu+\nu)(E) &= \mu(E)+\nu(E) \\ &= \mu\left(\bigcup_{n=1}^{\infty} E_n\right) + \nu\left(\bigcup_{n=1}^{\infty} E_n\right) \\ &= \sum_{n=1}^{\infty}\mu(E_n) + \sum_{n=1}^{\infty}\nu(E_n) \\ &= \sum_{n=1}^{\infty}\mu(E_n)+\nu(E_n) \\ &= \sum_{n=1}^{\infty}(\mu+\nu)(E_n),\end{aligned}$$

and we are done. Now it is clear that if μ, ν are complex measures, then $\alpha\mu+\beta\nu$ $(\alpha, \beta \in \mathbb{F})$ is also a complex measure, so the addition and multiplication make the set of complex measures on $\mathcal{S}$ into a vector space.

Definition 5.3

Let $(X, \mathcal{S})$ be a measurable space, we denote $\mathcal{M}_{\mathbb{C}}(\mathcal{S})$ the vector space of complex measures on $\mathcal{S}$. Likewise, $\mathcal{M}_{\mathbb{R}}(\mathcal{S})$ denotes the vector space of real measures on $\mathcal{S}$. Finally, the notation $\mathcal{M}_{\mathbb{F}}(\mathcal{S})$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ generalize both conditions. ♣

Our goal is to make $\mathcal{M}_{\mathbb{F}}(\mathcal{S})$ into a complete normed vector space. We need some results from decomposition theorems.

5.2 Decomposition of Measures

If ν is a signed measure on $(X, \mathcal{M})$, a set $E \in \mathcal{M}$ is called

- **positive** for ν if $\nu(F) \geq 0$ for all $F \in \mathcal{M}$ with $F \subset E$;
- **negative** for ν if $\nu(F) \leq 0$ for all $F \in \mathcal{M}$ with $F \subset E$;
- **null** for ν if $\nu(F) = 0$ for all $F \in \mathcal{M}$ with $F \subset E$.

Example 5.2 In the motivation $\nu(E) = \int_E f d\mu$, E is positive, negative, or null precisely when $f \geq 0$, $f \leq 0$ or $f = 0$ a.e. on E.

Example 5.3 We write the negation of "positive": a set $E \in \mathcal{M}$ is *not positive* if there exists $F \in \mathcal{M}$ with $F \subset E$ such that $\nu(F) < 0$. Hence "not positive" does not imply "negative".

Lemma 5.1

Any measurable subset of a positive set is positive, and the union of any countable family of positive sets is positive. ♡

Proof The first assertion is obvious from the definition of positivity. Let $P_1, P_2, \cdots$ be positive sets and let $Q_n = P_n \setminus \bigcup_{j=1}^{n-1} P_j$. Then $Q_n \subset P_n$, so Q_n is positive. Hence if $E \subset \bigcup_{j=1}^{\infty} P_j$, then

$$\nu(E) = \nu(E \cap \bigcup_{j=1}^{\infty} P_j) = \nu(E \cap \bigcup_{j=1}^{\infty} Q_j) = \sum_{j=1}^{\infty} \nu(E \cap Q_j) \geq 0,$$

as desired. □

Theorem 5.1 (Hahn decomposition)

If ν is a signed measure on $(X, \mathcal{M})$, there exist a positive set P and a negative set N for ν such that $P \cup N = X$ and $P \cap N = \varnothing$. If P', N' is another such pair, then $P \Delta P' (= N \Delta N')$ is null for ν. ♡

Proof Let m be the supremum of $\nu(E)$ as E ranges over all positive sets, then there is a sequence $\{P_j\}$ of positive sets such that $\nu(P_j) \to m$ $(j \to \infty)$. Let $P = \bigcup_{j=1}^{\infty}$, interchanging the limit and countable union, we obtain $\nu(P) = \lim_{j\to\infty} \nu(P_j) = m < \infty$. We claim that $N = X \setminus P$ is negative. To show this, we assume that N is not negative and derive a contradiction. From the construction $N = X \setminus P$ we have two results:

- N cannot contain any nonnull positive sets. Indeed, if $E \subset N$ is positive and $\nu(E) > 0$, then $E \cup P$ is positive and $\nu(E \cup P) = \nu(E) + \nu(P) > m$, which is impossible.

- If $A \subset N$ and $\nu(A) > 0$, there exists $B \subset A$ with $\nu(B) > \nu(A)$. Indeed, since A cannot be positive, there exists $C \subset A$ with $\nu(C) < 0$; thus if $B = A \setminus C$ we have $\nu(B) = \nu(A) - \nu(C) > \nu(A)$.

If N is not negative, then there exists a measurable $F \subset N$ with $\nu(F) > 0$. We specify a sequence of subsets $\{A_j\}$ of N and $\{n_j\} \subset \mathbb{N}$ as follows: n_1 is the smallest integer for which there exists a set $B \subset N$ with $\nu(B) > 1/n_1$, and A_1 is such a set ($A_1 = B$). By the second observation above, since $\nu(A_1) > 0$, there exists B with $\nu(B) > \nu(A_1)$. Let n_2 be the smallest integer for which there exists a set $B \subset A_1$ with $\nu(B) > \nu(A_1) + 1/n_2$, and let A_2 be such B. Proceeding inductively, n_j is the smallest integer for which there exists $A_j \subset A_{j-1}$ with $\nu(A_j) > \nu(A_{j-1}) + 1/n_j$.

Let $A = \bigcap_{j=1}^{\infty} A_j$. Then

$$\infty > \nu(A) = \lim_{j\to\infty} \nu(A_j) = \sum_{j=1}^{\infty} \frac{1}{n_j},$$

so for this convergent series we have $n_j \to \infty$ as $j \to \infty$. But $\nu(A) > 0$ and $A \subset N$, so there exists $B \subset A$ with $\nu(B) > \nu(A) + 1/n$ for some $n \in \mathbb{N}$. For j sufficiently large we have $n < n_j$, and $B \subset A_{j-1}$. This contradicts with the minimality of n_j. Therefore N is negative.

Finally, if P', N' is another pair of sets as in the statement of the theorem, we have $P \setminus P' \subset P$ and $P \setminus P' \subset N'$, so $P \setminus P'$ is both positive and negative, hence null; likewise for $P' \setminus P$. □

If P is a positive set for ν and N is a negative set for ν, and $P \cap N = \varnothing, X = P \cup N$, then the decomposition $X = P \cup N$ is called a **Hahn decomposition** for ν.

Definition 5.4

We say that two signed measures μ and ν on $(X, \mathcal{M})$ are ***mutually singular****, or that ν is* ***singular with respect to*** *μ, or vice verse, if there exist $E, F \in \mathcal{M}$ such that $E \cap F = \varnothing, E \cup F = X$, E is null for μ and F is null for ν. We express this relation ship by the notation*

$$\mu \perp \nu.$$

♣

Example 5.4 Let μ, ν be complex measures on $(\mathbb{R}, \mathcal{B})$ be given by $\mu(E) = m(E \cap (-\infty, 0)), \nu(E) = m(E \cap (2,3))$, where m is the Lebesgue measure. Then for the pair of sets $(-\infty, 0)$, $[0, \infty)$ we have $(-\infty, 0)$ is null for μ and $[0, \infty)$ is null for ν, hence $\mu \perp \nu$. Note that neither μ nor ν is singular with respect to m.

Theorem 5.2 (Jordan decomposition)

If ν is a signed measure, there exist unique positive measures ν^+ and ν such that $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$.

♡

Proof Let $X = P \cup N$ be a Hanh decomposition for ν, and define $\nu^+(E) = \nu(E \cap P), \nu^-(E) = -\nu(E \cap N)$. Then

$$\nu(E) = \nu(E \cap X) = \nu(E \cap (P \cup N)) = \nu(E \cap P) + \nu(E \cap N) = \nu^+(E) - \nu^-(E),$$

and $\nu^+(N) = \nu(N \cap P) = 0, \nu^-(P) = -\nu(P \cap N) = 0$, hence $\nu^+ \perp \nu^-$.

If also $\nu = \mu^+ - \mu^-$ and $\mu^+ \perp \mu^-$, let $E, F \in \mathcal{M}$ be such that $E \cap F = \varnothing$, $E \cup F = X$, and $\mu^+(F) = \mu^-(E) = 0$ (this is from the definition of singularity). Then $X = E \cup F$ another Hahn decomposition for ν, so $P \Delta E$ is ν-null. Therefore, for any $A \in \mathcal{M}$, $\mu^+(A) = \mu^+(A \cap E) = \nu(A \cap E) = \nu(A \cap P) = \nu^+(A)$, and likewise $\nu^- = \mu^-$. $\square$

The measures ν^+ and ν^- are called the positive and negative variations of ν, and $\nu = \nu^+ - \nu^-$ is called the **Jordan decomposition** of ν.

5.3 The Banach Space of Complex Measures

5.3.1 Total Variation

Definition 5.5 (total variation, signed measure)

*If ν is a signed measure on $(X, \mathcal{M})$, we define the **total variation** of ν to be the function $|\nu|$ defined by*

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

♣

We introduce an equivalent definition, which we shall use as the definition of the total variation of a complex measure.

Proposition 5.3

If ν is a signed measure on $(X, \mathcal{M})$ and $|\nu|$ is the total variation, then

$$|\nu|(E) = \sup \left\{ |\nu(E_1)| + \cdots + |\nu(E_n)| : n \in \mathbb{N}, \{E_j\} \textit{ disjoint}, \bigcup_{j=1}^{n} E_j \subset E \right\}.$$

♠

Proof Let $\{E_j\}_{j=1}^n$ be disjoint with $\bigcup_{j=1}^n \subset E$, let $P \cup N$ be a Hahn decomposition for ν so that $\nu^+(N) = 0$ and $\nu^-(P) = 0$. Then

$$\begin{aligned}
|\nu(E_1)| + \cdots + |\nu(E_n)| &= |\nu(P \cap E_1) + \nu(N \cap E_1)| + \cdots + |\nu(P \cap E_n) + \nu(N \cap E_n)| \\
&\leq (|\nu(P \cap E_1)| + \cdots + |\nu(P \cap E_n)|) + (|\nu(\nu(N \cap E_1)| + \cdots + |\nu(N \cap E_n)|) \\
&= (\nu^+(E_1) + \cdots + \nu^+(E_n)) + (\nu^-(E_1) + \cdots + \nu^-(E_n)) \\
&\leq \nu^+(E) + \nu^-(E) = |\nu|(E).
\end{aligned}$$

Taking supremum we get

$$|\nu|(E) \geq \sup \left\{ |\nu(E_1)| + \cdots + |\nu(E_n)| : n \in \mathbb{N}, \{E_j\} \text{ disjoint}, \bigcup_{j=1}^{n} E_j \subset E \right\}.$$

On the other hand, take $E_1 = P \cap E$, $E_2 = N \cap E$, then

$$|\nu(E_1)| + |\nu(E_1)| = |\nu(P \cap E)| + |\nu(N \cap E)| = \nu^+(E) + \nu^-(E) = |\nu|(E),$$

showing the reverse inequality, the proof is complete. $\square$

Definition 5.6 (total variation, complex measure)

*Let ν be a complex measure on $(X, \mathcal{M})$. The **total variation** of ν is the function*

$|\nu| : \mathcal{M} \to [0, \infty]$ *defined by*

$$|\nu|(E) = \sup\left\{ |\nu(E_1)| + \cdots + |\nu(E_n)| : n \in \mathbb{N}, \{E_j\} \text{ disjoint}, \bigcup_{j=1}^{n} E_j \subset E \right\}.$$

♣

Here are some useful properties of the total variation:

Properties

1. $|\nu(E)| \leq |\nu|(E)$.
2. $|\nu|(E) = \nu(E)$ *if* ν *is a finite positive measure.*
3. $|\nu|(E) = 0$ *if and only if* $\nu(A) = 0$ *for every* $A \in \mathcal{M}$ *with* $A \subset E$.

Proof

1. Since $E \subset E$, $|\nu(E)| \leq |\nu|(E)$.
2. Since ν is positive, by (1) we have $\nu(E) \leq |\nu|(E)$. Let $E_1, \cdots, E_n \in \mathcal{M}$ be disjoint and $\bigcup_{j=1}^{n} E_j \subset E$, then
$$\nu(E_1) + \cdots + \nu(E_n) = \nu\left(\bigcup_{j=1}^{n} E_j\right) \leq \nu(E),$$
taking supremum over all such $\{E_j\}_{j=1}^{n}$ leads to $|\nu|(E) \leq \nu(E)$.
3. Exercise. □

Theorem 5.3

Let ν be a complex measure on $(X, \mathcal{M})$, then $|\nu|$ is a measure on $\mathcal{M}$. ♡

Proof $|\nu|(\varnothing) = 0$ is easy to see. Let $\{E_n\}_{n=1}^{\infty} \subset \mathcal{M}$ be disjoint, we want to show

$$\sum_{n=1}^{\infty} |\nu|(E_n) = |\nu|\left(\bigcup_{n=1}^{\infty} E_n\right).$$

This is an equality related to infinite sums, supremums, and infinite unions, so we need to get rid of supremums using "an ε of room" technique. Let $\varepsilon > 0$, by the definition of supremum, we can choose sets $A_{1,n}, \cdots, A_{k(n),n} \in \mathcal{M}$ with $\bigcup_{k=1}^{k(n)} A_{k(n),n} \subset E_n$ such that

$$\sum_{k=1}^{k(n)} |\nu(A_{k,n})| \geq |\nu|(E_n) - \frac{\varepsilon}{2^n}.$$

Note that

$$\bigcup_{n=1}^{\infty} \bigcup_{k=1}^{k(n)} A_{k(n),n} \subset \bigcup_{n=1}^{\infty} E_n,$$

hence

$$\sum_{n=1}^{\infty} \sum_{k=1}^{k(n)} |\nu(A_{k,n})| \leq |\nu|\left(\bigcup_{n=1}^{\infty} E_n\right).$$

On the other hand,

$$\sum_{n=1}^{\infty} \sum_{k=1}^{k(n)} |\nu(A_{k,n})| \geq \sum_{n=1}^{\infty} |\nu|(E_n) - \varepsilon,$$

thus
$$|\nu|\left(\bigcup_{n=1}^{\infty} E_n\right) \geq \sum_{n=1}^{\infty} |\nu|(E_n) - \varepsilon.$$
Since ε is arbitrary,
$$|\nu|\left(\bigcup_{n=1}^{\infty} E_n\right) \geq \sum_{n=1}^{\infty} |\nu|(E_n).$$
For the other direction, Choose disjoint $B_1, \cdots, B_N \in \mathcal{M}$ with $\bigcup_{n=1}^{N} B_n \subset \bigcup_{n=1}^{\infty} E_n$ such that
$$|\nu(B_1)| + \cdots + |\nu(B_N)| \geq |\nu|\left(\bigcup_{n=1}^{\infty} E_n\right) - \varepsilon.$$
Notice that $E_k \cap \bigcup_{n=1}^{N} B_n \subset E_k$ implies
$$\sum_{n=1}^{N} |\nu(B_n \cap E_k)| \leq |\nu|(E_k), \quad k \in \mathbb{N}.$$
Summing over k leads to
$$\begin{aligned}
\sum_{k=1}^{\infty} |\nu|(E_k) &\geq \sum_{k=1}^{\infty}\sum_{n=1}^{N} |\nu(B_n \cap E_k)| \\
&= \sum_{n=1}^{N}\sum_{k=1}^{\infty} |\nu(B_n \cap E_k)| \\
&= \sum_{n=1}^{N} |\nu(B_n)| \\
&\geq |\nu|\left(\bigcup_{n=1}^{\infty} E_n\right) - \varepsilon,
\end{aligned}$$
completing the proof. □

Remark One can definitely use the definition $|\nu| = \nu^+ + \nu^-$ to show immediately that $|\nu|$ is a measure. I choose the more complicated proof here to exhibit a technique of dealing with infinite sums, supremums, and infinite unions.

Example 5.5 Let μ be a measure on a measurable space $(X, \mathcal{M})$ and $f \in L^1(\mu)$. Let $d\nu = f d\mu$, then
$$|\nu|(E) = \int_E |f| \, d\mu.$$
for every $E \in \mathcal{S}$.

Proof Let $P = \{x : f(x) \geq 0\}$ and $N = \{x : f(x) < 0\}$, then $X = P \cup N$ is a Hahn decomposition for ν and we can set $\nu^+(E) = \nu(E \cap P), \nu^-(E) = \nu(N \cap P)$. Then
$$\begin{aligned}
|\nu|(E) &= \nu^+(E) + \nu^-(E) \\
&= \int_{E\cap P} |f| d\mu + \int_{E \cap N} |f| d\mu \\
&= \int_E |f| d\mu.
\end{aligned}$$

In particular, $|\nu|(X) = \int_X |f| \, d\mu = \|f\|_{L^1}$. □

Now we endow $\mathcal{M}_F(\mathcal{S})$ with a norm structure.

Definition 5.7

*Let ν be a complex measure on a measurable space $(X, \mathcal{S})$, the **total variation norm** of ν, denoted $\|\nu\|$, is defined by*

$$\|\nu\| = |\nu|(X).$$

♣

Example 5.6 If μ is a finite measure, then $\|\mu\| = \mu(X)$.

Example 5.7 If μ is a measure, $h \in L^1(\mu)$, and $d\nu = h\, d\mu$, then $\|\nu\| = \|h\|_{L^1}$.

Recall that a norm is a real-valued function, so it cannot attain the value ∞. We now justify that the total variation "norm" is indeed a norm.

Theorem 5.4

Suppose $(X, \mathcal{S})$ is a measurable space and $\nu \in \mathcal{M}_{\mathbb{F}}(\mathcal{S})$, then $\|\nu\| < \infty$.

♡

Proof First suppose ν is a signed measure, then $|\nu|(X) = \nu^+(X) + \nu^-(X)$. Since ν^+ and ν^- are positive signed measures, $\nu(X)$ is finite. If ν is a complex measure, then for any disjoint sets $\{E_j\}_{j=1}^n$ with $\bigcup_{j=1}^n \subset X$,

$$\begin{aligned}
&|\operatorname{Re}\nu(E_1) + i \operatorname{Im}\nu(E_1)| + \cdots + |\operatorname{Re}\nu(E_n) + i \operatorname{Im}\nu(E_n)| \\
&\leq |\operatorname{Re}\nu(E_1)| + \cdots + |\operatorname{Re}\nu(E_n)| + |\operatorname{Im}\nu(E_1)| + \cdots + |\operatorname{Im}\nu(E_n)| \\
&\leq |\operatorname{Re}\nu|(X) + |\operatorname{Im}\nu|(X) < \infty.
\end{aligned}$$

□

5.3.2 Completeness of $\mathcal{M}_{\mathbb{F}}$

Theorem 5.5

Suppose $(X, \mathcal{S})$ is a measurable space. Then $\mathcal{M}_{\mathbb{F}}(\mathcal{S})$ is a Banach space with the total variation norm.

♡

Proof Let $\{\nu_n\}$ be a Cauchy sequence in $\mathcal{M}_{\mathbb{F}}(\mathcal{S})$ and $\varepsilon > 0$, then

$$\|\nu_n - \nu_m\| = |\nu_n - \nu_m|(X) < \varepsilon \text{ for all large } m, n.$$

For any $E \in \mathcal{S}$,

$$|\nu_n(E) - \nu_m(E)| \leq |\nu_n - \nu_m|(E) < \varepsilon,$$

hence $\{\nu_n(E)\}_{n=1}^\infty$ is a Cauchy sequence in $\mathbb{F}$, and thus we can define $\nu(E) = \lim_{n\to\infty} \nu_n(E)$. Let $\{E_k\}_{k=1}^\infty \subset \mathcal{S}$ be disjoint, we have to show

$$\nu(\bigcup_{k=1}^\infty E_k) = \sum_{k=1}^\infty \nu(E_k),$$

which is

$$\lim_{n\to\infty} \nu_n(\bigcup_{k=1}^\infty E_k) = \sum_{k=1}^\infty \lim_{n\to\infty} \nu_n(E_k).$$

This equality involves limit, countable unions and countable sums, so we need to use "an ε of room" technique: truncate the infinite sum to interchange the limit and finite sum.

Fix n, recall that the series $\sum_{k=1}^\infty |\nu_n(E_k)|$ converges absolutely, so there exists $N \in \mathbb{N}$

such that $\sum_{k=N}^{\infty}|\nu_n(E_k)|<\varepsilon$. For all $m\geq n$ we have

$$\begin{aligned}\sum_{k=N}^{\infty}|\nu_m(E_k)| &\leq \sum_{k=N}^{\infty}|(\nu_m-\nu_n)(E_k)|+\sum_{k=N}^{\infty}|\nu_n(E_k)|\\ &\leq \sum_{k=N}^{\infty}|\nu_m-\nu_n|(E_k)+\varepsilon\\ &= |\nu_m-\nu_n|(\bigcup_{k=N}^{\infty}E_k)+\varepsilon<2\varepsilon.\end{aligned}$$

Now

$$\begin{aligned}\left|\nu(\bigcup_{k=1}^{\infty}E_k)-\sum_{k=1}^{N-1}\nu(E_k)\right| &= \left|\lim_{n\to\infty}\nu_n(\bigcup_{k=1}^{\infty}E_k)-\sum_{k=1}^{N-1}\lim_{n\to\infty}\nu_n(E_k)\right|\\ &= \left|\lim_{n\to\infty}\nu_n(\bigcup_{k=1}^{\infty}E_k)-\lim_{n\to\infty}\sum_{k=1}^{N-1}\nu_n(E_k)\right|\\ &= \lim_{n\to\infty}\left|\nu_n(\bigcup_{k=1}^{\infty}E_k)-\nu_n(\bigcup_{k=1}^{N-1}E_k)\right|\\ &= \lim_{n\to\infty}\left|\nu_n(\bigcup_{k=N}^{\infty}E_k)\right|<2\varepsilon.\end{aligned}$$

Finally, we show $\|\nu-\nu_k\|\to 0$ as $k\to\infty$. There exists $M\in\mathbb{N}$ such that $\|\nu_m-\nu_n\|<\varepsilon$ for all $m,n\geq M$. Let $k\geq M$, and $\{E_j\}_{j=1}^{N}\subset\mathcal{S}$ be disjoint, then

$$\sum_{j=1}^{N}|(\nu-\nu_k)(E_j)|=\lim_{n\to\infty}\sum_{j=1}^{N}|(\nu_n-\nu_k)(E_j)|<\varepsilon,$$

completing the proof. □

5.4 Radon-Nikodym Theorem and Conditional Expectations

5.4.1 Lebesgue Decomposition

Definition 5.8 (absolute continuity)

*Suppose ν is a complex measure on a measurable space $(X,\mathcal{M})$ and μ is a positive measure on $(X,\mathcal{M})$. Then ν is called **absolutely continuous** with respect to μ if $\mu(E)=0$ implies that $\nu(E)=0$, where $E\in\mathcal{M}$.* ♣

Example 5.8 $\nu\ll\mu$ if and only if $|\nu|\ll\mu$.

Proof Suppose $\nu\ll\mu$. Let $\mu(E)=0$, then $|\nu|(E)=0$ since $\nu(E)=0$. Conversely, since $|\nu(E)|\leq|\nu|(E)$, we have $\nu(E)=0$ whenever $\mu(E)=0$. □

Example 5.9 $\nu\ll\mu$ if and only if $\nu^+\ll\mu$ and $\nu^-\ll\mu$.

Proof Since $|\nu|=\nu^++\nu^-$, $\nu^+\ll\mu$ and $\nu^-\ll\mu$ implies $|\nu|\ll\mu$. Conversely, suppose $\nu\ll\mu$. If $|\nu|(E)=0$, then $\nu^+(E)=\nu^-(E)=0$, completing the proof. □

Proposition 5.4

Let ν be a complex measure and μ a positive measure on $(X, \mathcal{M})$, then $\nu \ll \mu$ if and only if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|\nu(E)| < \varepsilon$ whenever $\mu(E) < \delta$. ♠

Proof Since $\nu \ll \mu$ iff $|\nu| \ll \mu$ and $|\nu(E)| \leq |\nu|(E)$, it suffices to assume that $\nu = |\nu|$ is positive. Clearly the $\varepsilon - \delta$ condition implies that $\nu \ll \mu$. On the other hand, if the $\varepsilon - \delta$ condition is not satisfied, there exists $\varepsilon > 0$ such that for all $n \in \mathbb{N}$ we can find $E_n \in \mathcal{M}$ with $\mu(E_n) < 2^{-n}$ and $\nu(E_n) \geq \varepsilon$. Let $F_k = \bigcup_k^\infty E_n$ and $F = \bigcap_1^\infty F_k$. Then $\mu(F_k) < \sum_k^\infty 2^{-n} = 2^{1-k}$, so $\mu(F) = 0$; but $\nu(F_k) \geq \varepsilon$ for all k and hence, since ν is finite, $\nu(F) = \lim \nu(F_k) \geq \varepsilon$. Thus it is false that $\nu \ll \mu$. □

Proposition 5.5

Suppose μ is a positive measure and ν is a complex measure on $(X, \mathcal{M})$. If $\nu \ll \mu$ and $\nu \perp \mu$, then $\nu = 0$. ♠

Proof Let E, F be disjoint and $E \cup F = X$ such that E is null for μ and F is null for ν, hence any subset of F is ν-null, so $|\nu|(F) = 0$. Since $|\nu| \ll \mu$, $|\nu|(E) = 0$, thus $|\nu|(X) = 0$, which implies $\nu = 0$. □

Theorem 5.6 (Lebesgue decomposition theorem)

Suppose μ is a positive measure on $(X, \mathcal{M})$ and ν is a complex measure on $(X, \mathcal{M})$, then there exist unique complex measures ν_a and ν_s on $(X, \mathcal{M})$ such that $\nu = \nu_a + \nu_s$ and

$$\nu_a \ll \mu \text{ and } \nu_s \perp \mu.$$ ♡

Proof Let $b = \sup\{|\nu|(B) : B \in \mathcal{M}, \mu(B) = 0\}$. Choose $\{B_n\}_{n=1}^\infty \subset \mathcal{M}$ such that

$$|\nu|(B_n) \geq b - \frac{1}{n} \text{ and } \mu(B_n) = 0.$$

Let $B = \bigcup_{n=1}^\infty B_n$, then $\mu(B) = 0$ and $|\nu|(B) = b$. Let $A = X \setminus B$, define complex measures ν_a, ν_s on $\mathcal{M}$ by

$$\nu_a(E) = \nu(E \cap A), \quad \nu_s(E) = \nu(E \cap B).$$

Clearly $\nu = \nu_a + \nu_s$. If $E \in \mathcal{M}$, then

$$\mu(E) = \mu(E \cap A) + \mu(E \cap B) = \mu(E \cap A)$$

since $\mu(B) = 0$. Now $A \cup B = X, A \cap B = \varnothing, \nu_s(A) = 0$ and $\mu(B) = 0$, hence $\nu_s \perp \mu$.

To show that $\nu_a \ll \mu$, let $E \in \mathcal{M}$ and $\mu(E) = 0$. Then $\mu(B \cup E) = 0$ and hence

$$b \geq |\nu|(B \cup E) = |\nu|(B) + |\nu|(E \setminus B) = b + |\nu|(E \setminus B),$$

which implies that $|\nu|(E \setminus B) = 0$. Thus

$$\nu_a(E) = \nu(E \cap A) = \nu(E \setminus B) = 0,$$

hence $\nu_a \ll \mu$.

Suppose ν_1, ν_2 are complex measures on $\mathcal{M}$ with $\nu_1 \ll \mu, \nu_2 \perp \mu$ and $\nu = \nu_1 + \nu_2$, then $\nu_1 - \nu_a = \nu_s - \nu_2$. Since $\nu_1 - \nu_a \ll \mu$ and $\nu_s - \nu_2 \perp \mu$, it follows that $\nu_1 = \nu_a$ and $\nu_2 = \nu_s$. □

5.4.2 Radon-Nikodym Theorem

Theorem 5.7 (Radon-Nikodym)

Suppose μ is a σ-finite measure on a measurable space $(X, \mathcal{M})$. Suppose ν is a complex measure on $(X, \mathcal{M})$ such that $\nu \ll \mu$. Then there exists $h \in L^1(\mu)$ such that $d\nu = h\, d\mu$, and any two such functions are equal μ-a.e. ♡

Proof **Case (1): finite positive measures.** Suppose μ, ν are finite positive measures. Define $\varphi : L^2(\nu + \mu) \to \mathbb{R}$ by

$$\varphi(f) = \int f d\nu,$$

then by Cauchy-Schwarz inequality,

$$\int |f| d\nu \leq \int |f| d(\nu + \mu) \leq \sqrt{\nu(X) + \mu(X)} \|f\|_{L^2(\nu+\mu)} < \infty,$$

so φ is well-defined. Furthermore, if $f = g$ $(\nu+\mu)$-a.e., then $\varphi(f) - \varphi(g) = \int (f-g) d\nu = 0$ because $f = g$ ν-a.e. Thus φ is a linear functional on $L^2(\nu + \mu)$. Since $|\varphi(f)| \leq \int |f| d\nu$, φ is bounded. By Riesz representation theorem, there exists $g \in L^2(\nu + \mu)$ such that

$$\int f\, d\nu = \int fg\, d(\nu + \mu), \quad f \in L^2(\nu + \mu).$$

Then $\int f d\nu = \int fg\, d\nu + \int fg\, d\mu$, hence

$$\int f(1 - g)\, d\nu = \int fg\, d\mu. \tag{5.1}$$

- If $f = \chi_{\{x: g(x) \geq 1\}}$, then $\int f(1-g)\, d\nu \leq 0$ and $\int fg\, d\mu \geq 0$, hence $\int f(1-g)\, d\nu = \int fg\, d\mu = 0$. Then $\mu(\{x : g(x) \geq 1\}) = 0$.
- If $f = \chi_{\{x: g(x) < 0\}}$, then similarly $\int fg\, d\mu = 0$, which implies that $\mu(\{x : g(x) \geq 1\}) = 0$.

Because $\nu \ll \mu$, the two observations imply that

$$\nu(\{x : g(x) \geq 1\}) = 0, \quad \nu(\{x : g(x) < 0\}) = 0.$$

Thus we may assume that $0 \leq g(x) < 1$ for all $x \in X$. Define $h : X \to [0, \infty)$ by

$$h(x) = \frac{g(x)}{1 - g(x)}.$$

Let $E \in \mathcal{M}$, for each $k \in \mathbb{N}$ let

$$f_k(x) = \begin{cases} \frac{\chi_E(x)}{1-g(x)} & \text{if } \frac{\chi_E(x)}{1-g(x)} \leq k, \\ 0 & \text{otherwise.} \end{cases}$$

Then $f_k \in L^2(\nu + \mu)$. Plug this into (5.1), we have

$$\int f_k(1 - g)\, d\nu = \int f_k g\, d\mu.$$

Letting $k \to \infty$ and using the monotone convergence theorem gives

$$\int_E 1\, d\nu = \int_E h\, d\mu.$$

Thus $d\nu = h d\mu$.

Case (2): σ-finite. Assume μ is a σ-finite measure, then there exists a sequence $X_1 \subset X_2 \subset \cdots$ of sets in $\mathcal{M}$ such that $X = \bigcup_{k=1}^{\infty}$ and each $\mu(X_k) < \infty$. Let ν_k and μ_k be the restrictions of ν and μ to the σ-algebra on X_k consisting of those sets in $\mathcal{M}$ which are subsets of X_k. Then $\nu_k \ll \mu_k$. By case (1) here exists a nonnegative function $h_k \in L^1(\mu_k)$ with $d\nu_k = h_k d\mu_k$. If $j < k$ (so that $X_j \subset X_k$), then

$$\int_E h_j \, d\mu = \nu(E) = \int_E h_k \, d\mu, \quad (E \in \mathcal{M}, E \subset X_j).$$

Thus $\mu(\{x \in X_j : h_j(x) \neq h_k(x)\}) = 0$. Define a function $h : X \to [0, \infty)$ by setting

$$\begin{aligned} h(x) &= h_1(x) \quad \forall x \in X_1, \\ h(x) &= h_2(x) \quad \forall x \in X_2 \setminus X_1, \\ &\cdots \\ h(x) &= h_k(x) \quad \forall x \in X_k \setminus X_{k-1}, \\ &\cdots \end{aligned}$$

Then h is $\mathcal{M}$-measurable and

$$\mu(\{x \in X_k : h(x) \neq h_k(x)\}) = 0$$

for each $k \in \mathbb{N}$. Now let $E \in \mathcal{M}$ be arbitrary, then E is a subset of some X_N. Apply case (1) to X_N, we have

$$\int_E 1 \, d\nu = \int_E h_N \, d\mu = \int_E h \, d\mu,$$

completing the proof of case (2).

Case (3): ν is a signed measure. We have the identity

$$\nu = \frac{1}{2}(|\nu| + \nu) - \frac{1}{2}(|\nu| - \nu),$$

where $|\nu| + \nu \ll \mu$ and $|\nu| - \nu \ll \mu$, hence there are $h^+, h^- \in L^1(\mu)$ such that

$$d(|\nu| + \nu) = h^+ \, d\mu, \quad d(|\nu| - \nu) = h^- \, d\mu.$$

Take $h = \frac{1}{2}(h^+ - h^-) \, d\mu$, then $d\nu = h d\mu$.

Case (4): ν is a complex measure. Applying case (3) to $\operatorname{Re}\nu$ and $\operatorname{Im}\nu$ yields two functions $h_1, h_2 \in L^1(\mu)$ with $d(\operatorname{Re}\nu) = h_1 d\mu$ and $d(\operatorname{Im}\nu) = h_2 d\mu$. Taking $h = h_1 + i h_2$ completes the proof. We leave the proof of uniqueness as an exercise. □

When $d\nu = h d\mu$, the notation $h = \frac{d\nu}{d\mu}$ is called the *Radon-Nikodym derivative* of ν with respect to μ.

We can combine Lebesgue decomposition theorem and Radon-Nikodym theorem:

Theorem 5.8 (Lebesgue-Radon-Nikodym)

Let ν be a complex measure and μ a σ-finite positive measure on $(X, \mathcal{M})$. There exist unique complex measures ν_s, ν_a on $(X, \mathcal{M})$ such that

$$\nu_s \perp \mu, \quad \nu_a \ll \mu, \quad \nu = \nu_s + \nu_a.$$

Moreover, there is a μ-integrable function $f : X \to \mathbb{R}$ such that $d\nu_a = f d\mu$, and any two such functions are equal μ-a.e. ♡

5.4.3 Conditional Expectations

Chapter 6 L^p Spaces

6.1 Basic Properties of L^p Spaces

Let $(X, \mathcal{M}, \mu)$ be a measure space. If f is a measurable function on X and $0 < p < \infty$, we define

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

It is possible that $\|f\|_p = \infty$.

Definition 6.1 (L^p space)

Define

$$L^p(X, \mathcal{M}, \mu) = \{f : X \to \mathbb{C} | f \text{ is measurable and } \|f\|_p < \infty\}.$$

♣

We abbreviate $L^p(X, \mathcal{M}, \mu)$ by $L^p(\mu)$, $L^p(X)$ or L^p when this will cause no confusion.

L^p is a vector space, since if $f, g \in L^p$, then

$$|f + g|^p \leq [2 \max(|f|, |g|)^p] \leq 2^p(|f|^p + |g|^p),$$

so that $f+g \in L^p$. It is obvious that $\|f\|_p = 0$ if and only if $f = 0$ a.e. and $\|cf\|_p = |c|\|f\|_p$, so the only question is the triangle inequality. The case $p \geq 1$ will be proved below (Hölder's inequality), and we will see why the triangle inequality fails for $p < 1$.

Let $a > 0$, $b < 0$ and $0 < p < 1$. For $t > 0$ we have $t^{p-1} > (a+t)^{p-1}$. Integrating from 0 to b we obtain $a^p + b^p > (a+b)^p$. Thus if E, F are disjoint sets with $0 < \mu(E), \mu(F) < \infty$ and we set $a = \mu(E)^{1/p}$, $b = \mu(F)^{1/p}$, then

$$\|\chi_E + \chi_F\|_p = (a^p + b^p)^{1/p} > a + b = \|\chi_E\|_p + \|\chi_F\|_p.$$

6.1.1 Basic Inequalities

Definition 6.2 (conjugate exponents)

Let $1 \leq p, q \leq \infty$. *If*

$$\frac{1}{p} + \frac{1}{q} = 1,$$

we say that p and q are conjugate exponents. Here we use the convention $1/\infty = 0$, *so*

$$\frac{1}{0} + \frac{1}{\infty} = 1, \quad \frac{1}{\infty} + \frac{1}{0} = 1.$$

♣

To prove the triangle inequality for $p \geq 1$, we need Hölder's inequality. Recall the gemetric mean inequality: for $A, B \geq 0$,

$$A^{1/2}B^{1/2} \leq A + B.$$

Here is a generalized version: for $A, B \geq 0$ and $\theta \in [0, 1]$,

$$A^{\theta}B^{1-\theta} \leq \theta A + (1 - \theta)B.$$

If $B \neq 0$, by homogeneity we can divide both sides by B to get

$$\left(\frac{A}{B}\right)^\theta \leq \theta\left(\frac{A}{B}\right) + (1-\theta),$$

which can be proved by differentiating

$$f(x) = x^\theta - \theta x - (1-\theta).$$

The maximum value of f occurs at $t = 1$. Hence the equality holds if and only if $A = B$.

Proposition 6.1 (Hölder's inequality)

Suppose $1 < p < \infty$ and $1 < q < \infty$ are conjugate exponents. If $f \in L^p$ and $g \in L^q$, then $fg \in L^1$ and $\|fg\|_1 \leq \|f\|_p\|g\|_q$ (with equality if and only if $\alpha|f|^p = \beta|g|^q$ for some constants α, β).

Proof We may assume f, g are nonzero, then it suffices to prove

$$\left\|\frac{fg}{\|f\|_p\|g\|_q}\right\|_1 = \left\|\frac{f}{\|f\|_p}\frac{g}{\|g\|_q}\right\|_1 \leq 1,$$

so we may assume that $\|f\|_p = \|g\|_q = 1$. Let $A = |f(x)|^p, B = |g(x)|^q$, then

$$|f(x)g(x)| \leq \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q.$$

Integrating, we have

$$\|fg\|_1 \leq 1.$$

By the generalized geometric mean equality, we see that equality holds if and only if $|f|^p = |g|^q$. □

Proposition 6.2 (Minkowski's inequality)

Suppose $f, g \in L^p$, where $1 \leq p < \infty$. Then $f + g \in L^p$ and

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof The idea is to construct conjugate exponents.

$$\begin{aligned}|f(x)+g(x)|^p &= |f(x)+g(x)||f(x)+g(x)|^{p-1}\\ &\leq |f(x)||f(x)+g(x)|^{p-1} + |g(x)||f(x)+g(x)|^{p-1}\\ &= |f(x)||f(x)+g(x)|^{p/q} + |g(x)||f(x)+g(x)|^{p/q},\end{aligned}$$

where q is the conjugate exponent to $p : p - 1 = p/q$. Integrating and applying Hölder's inequality on RHS,

$$\begin{aligned}\|f+g\|_p^p &\leq (\|f\|_p + \|g\|_p)\left\|(f+g)^{p/q}\right\|_q\\ &= (\|f\|_p + \|g\|_p)\|f+g\|_p^{p/q},\end{aligned}$$

hence

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

□

6.1.2 Completeness and Density

Theorem 6.1

For $1 \leq p < \infty$, L^p is a Banach space. ♡

Proof Let $\{f_n\} \subset L^p$ be a Cauchy sequence, then $\|f_n - f_m\|_p < \varepsilon$ for all large n, m. Extract a subsequence $\{f_{n_k}\}$ such that

$$\left\|f_{n_{k+1}} - f_{n_k}\right\|_p \leq 2^{-k} \quad (k \in \mathbb{N}).$$

Let

$$f(x) = f_{n_1}(x) + \sum_{k=1}^{\infty}(f_{n_{k+1}}(x) - f_{n_k}(x)), \quad S_N f(x) = f_{n_1}(x) + \sum_{k=1}^{N}(f_{n_{k+1}}(x) - f_{n_k}(x))$$

and

$$g(x) = |f_{n_1}(x)| + \sum_{k=1}^{\infty}|f_{n_{k+1}}(x) - f_{n_k}(x)|, \quad S_N g(x) = |f_{n_1}(x)| + \sum_{k=1}^{N}|f_{n_{k+1}}(x) - f_{n_k}(x)|.$$

Then

$$\|S_N g\|_p = \left\| |f_{n_1}| + \sum_{k=1}^{N} |f_{n_{k+1}}(x) - f_{n_k}(x)| \right\|_p,$$

hence by Minkowski's inequality we have

$$\|S_N g\|_p \leq \|f_{n_1}\|_p + \left\| \sum_{k=1}^{N} |f_{n_{k+1}} - f_{n_k}| \right\|_p \leq \|f_{n_1}\|_p + \sum_{k=1}^{N} \left\|f_{n_{k+1}} - f_{n_k}\right\|_p \leq \|f_{n_1}\|_p + 2.$$

Clearly $(S_N g(x))^p$ increases to $g(x)^p$ for every x, so by the monotone convergence theorem, $\|S_N g\|_p \to \|g\|_p$ as $N \to \infty$. Hence $\|g\|_p \leq \|f_{n_1}\|_p + 2 < \infty$. This shows $g \in L^p$, and thus g is finite almost everywhere. By the construction of $S_N f$, we have $S_{N-1} f(x) = f_{n_N}(x) \to f(x)$ a.e.

Now we show $f_{n_N} \to f$ in L^p as $N \to \infty$. Since

$$\|S_N f - f\|_p \leq (|g| + |g|)^p = 2^p |g|^p \in L^1,$$

by the dominated convergence theorem,

$$\int |S_N f - f|^p \to 0 \quad (N \to \infty).$$

Finally, if n is sufficiently large, then there is some M such that $\|f_n - f_{n_M}\|_p < \varepsilon$ and $\|f_{n_M} = f\|_p < \varepsilon$, hence

$$\|f_n - f\|_p \leq \|f_n - f_{n_M}\|_p + \|f_{n_M} - f\|_p < \varepsilon,$$

completing the proof. □

Proposition 6.3

For $1 \leq p < \infty$, the set of simple functions is dense in L^p. ♠

Proof Clearly simple functions are in L^p. If $f \in L^p$, then there is a sequence $\{f_n\}$ of simple functions such that $f_n \to f$ a.e. and $|f_n| \leq |f|$. Then $f_n \in L^p$ and $|f_n - f| \leq 2^p |f|^p \in L^1$, so by the dominated convergence theorem, $\|f_n - f\|_p \to 0$. □

6.1.3 L^∞ Space

If f is a measurable function on X, we define

$$\|f\|_\infty = \inf\{a \geq 0 : \mu(\{x : |f(x)| > a\}) = 0\},$$

and we set $\inf \varnothing = \infty$. In practice, we seldom use this definition. Observe that if $\|f\|_\infty = B$, then

$$\{x : |f(x)| > B\} = \bigcup_{n=1}^{\infty} \{x : |f(x)| > B + 1/n\},$$

hence

$$\mu(\{x : |f(x)| > B\}) = \lim_{n\to\infty} \mu(\{x : |f(x)| > B + 1/n\}) = 0.$$

That is to say, the infimum B can be attained. Moreover, we conclude that $|f(x)| \leq \|f\|_\infty$ almost everywhere.

Proposition 6.4

$f \in L^\infty$ if and only if there is a bounded measurable function g such that $f = g$ a.e.

Proof If f equals a bounded measurable function g a.e., then clearly $\|f\|_\infty < \infty$. Conversely, $f \in L^\infty$ implies $|f(x)| \leq \|f\|_\infty < \infty$ a.e., so we can take $g = f$ on $\{x : |f(x)| \leq \|f\|_\infty\}$ and $g = 0$ elsewhere. □

Example 6.1 Let

$$f(x) = \begin{cases} x^3 & x \in \mathbb{Q}, \\ \arctan x, & x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$$

then $\operatorname{ess\,sup} f = \pi/2$.

Proposition 6.5 (Hölder's inequality, L^∞)

If f, g are measurable functions on X, then $\|fg\|_1 \leq \|f\|_1\|g\|_\infty$. If $f \in L^1$ and $g \in L^\infty$, $\|fg\|_1 = \|f\|_1\|g\|_\infty$ if and only if $|g(x)| = \|g\|_\infty$ a.e. on the set where $f(x) \neq 0$.

Proof If either $f \notin L^1$ or $g \notin L^\infty$, then $\|fg\|_1 \leq \infty$, so the inequality holds. Let $f \in L^1$ and $g \in L^\infty$, then

$$\int |fg| = \int |f||g| \leq \|g\|_\infty \int |f| = \|f\|_1\|g\|_\infty.$$

Suppose $|g(x)| = \|g\|_\infty$ a.e. on $\{x : f(x) \neq 0\}$, then $\|g\|_\infty \int |f| = \int |fg| = \|fg\|_1$. Conversely, if $|g(x)| \neq \|g\|_\infty$ on $\{x : f(x) \neq 0\}$ for all x in a set of positive measure, then there is a constant C such that $\mu(\{x : \|g\|_\infty - g(x)\}) > 0$, then

$$\int |fg| - \int \|g\|_\infty |f| = \int |f(x)|(\|g\|_\infty - g(x))\, d\mu > 0,$$

completing the proof. □

Proposition 6.6

$\|\cdot\|_\infty$ is a norm on L^∞.

Proof (i) It is obvious that $\|f\|_\infty$ for all $f \in L^\infty$. Let $\|f\|_\infty = 0$, then $\mu(\{x : |f(x)| >$

$0\}) = 0$, hence $f = 0$ a.e.

(ii)

$$\begin{aligned}\|\lambda f\|_\infty &= \inf\{a \geq 0 : \mu(|\lambda f|(x) > a) = 0\} \\ &= \inf\{a \geq 0 : \mu(|f|(x) > \frac{a}{|\lambda|}) = 0\} \\ &= \inf\{|\lambda| a : \mu(|f|(x) > a) = 0\} \\ &= |\lambda| \inf\{a : \mu(|f|(x) > a) = 0\} \\ &= |\lambda| \|f\|_\infty.\end{aligned}$$

(iii) Observe that

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty \ ,$$

so $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$, which is to be shown. $\square$

Theorem 6.2

$\|f_n - f\|_\infty \to 0$ iff there exists $E \in \mathcal{M}$ such that $\mu(E^c) = 0$ and $f_n \to f$ uniformly on E. ♡

Proof Let $\|f_n - f\|_\infty \to 0$, let $\varepsilon > 0$ then there is N such that $\|f_n - f\|_\infty < \varepsilon$ for all $n \geq N$, hence $|f_n(x) - f(x)| < \varepsilon$ a.e. For each $n \geq N$ let $F_n = \{x : |f_n(x) - f(x)| \geq \varepsilon\}$, then $\mu(F_n) = 0$. Let $F = \cup_{n=N+1}^\infty F_n$, then $\mu(F) = 0$. Let $E = F^c$, then

$$|f_n(x) - f(x)| < \varepsilon \ \forall x \in E,$$

so $f_n \to f$ uniformly in E and $\mu(E^c) = 0$.

Now suppose $f_n \to f$ uniformly except on a set of measure 0. Let $\varepsilon > 0$ we have

$$\sup_{x \in E} |f_n(x) - f(x)| < \varepsilon$$

for all large n. Since $\mu(E^c) = 0$, it follows that

$$\mu(\{x \in X : |f_n(x) - f(x)| > \sup_{x \in E} |f_n(x) - f(x)|\}) = 0,$$

so $\sup_{x \in E} |f_n(x) - f(x)| \geq \|f_n - f\|_\infty$ for all large n. That is, $\|f_n - f\|_\infty < \varepsilon$ for all large n, thus $\|f_n - f\|_\infty \to 0$. $\square$

Theorem 6.3

L^∞ is a Banach space. ♡

Proof Let $\{f_n\}$ be a Cauchy sequence in L^∞, then $\forall \varepsilon > 0 \, \exists N \, \forall n, m > N : \|f_n - f_m\|_\infty < \varepsilon$. Then for each $n, m > N$, $|f_n(x) - f_m(x)| < \varepsilon$ a.e. Let

$$F_{m,n} = \{x \in X : |f_n(x) - f_m(x)| \geq \varepsilon\},$$

then $\mu(F_{m,n}) = 0$, thus $\bigcup_{m,n>N} F_{m,n}$ is of measure 0. Denote its complement by E, we have $\{f_n\}$ is uniformly Cauchy in E, and since $\mathbb{F}$ is complete, there exists f such that $f_n \to f$ uniformly on E and $\mu(E^c) = 0$, by the previous theorem we have $\|f_n - f\|_\infty$. Therefore, L^∞ is a Banach space. $\square$

6.2 Dual Space of L^p

Throughout this section let $(X, \mathcal{M}, \mu)$ be a fixed measure space.

Definition 6.3 (semifinite)

If for each $E \in \mathcal{M}$ with $\mu(E) = \infty$ there exists $F \in \mathcal{M}$ with $F \subset E$ and $0 < \mu(F) < \infty$, μ is called ***semifinite****.* ♣

Exercise 6.1 Every σ-finite measure is semifinite.

Proof Suppose μ is a σ-finite measure, Then there is an increasing sequence $X_1 \subset X_2 \subset \cdots$ such that $\bigcup_{j=1}^{\infty} X_j = X$ and each $\mu(X_j) < \infty$. Let $\mu(E) = \infty$, then for any $j \in \mathbb{N}$, $E \cap X_j$ is measurable, and $0 < \mu(E \cap X_j) < \infty$, hence μ is semifinite. □

Exercise 6.2 [1] If μ is a semifinite measure and $\mu(E) = \infty$, for any $C > 0$ there exists $F \subset E$ with $C < \mu(F) < \infty$.

Proof

Let p, q be conjugate exponents. For each $g \in L^q$, we define a linear functional ϕ_g on L^p by

$$\phi_g(f) = \int fg, \quad (f \in L^p).$$

By Hölder's inequality,

$$|\phi_g(f)| \leq \int |fg| \leq \|g\|_q \|f\|_p,$$

hence the operator norm $\|\phi_g\| \leq \|g\|_q$. This shows each $g \in L^q$ gives a bounded linear functional $\phi_g \in (L^p)^*$. If we can show every bounded linear functional on L^p has the above form, then we would have $L^q \cong (L^p)^*$.

First, we compute the operator norm $\|\phi_g\|$.

Proposition 6.7

Suppose p, q are conjugate exponents and $1 \leq q < \infty$. If $g \in L^q$, then

$$\|g\|_q = \|\phi_g\| = \sup\left\{ \left| \int fg \right| : \|f\|_p = 1 \right\}.$$

If μ is semifinite, this result holds for $q = \infty$. ♠

Proof We already have $\|\phi_g\| \leq \|g\|_q$, equality is trivial if $g = 0$ a.e. We need to find a function f with $\|f\|_p = 1$ and $\left| \int fg \right| = 1$. Denote $\operatorname{sgn} g = g/|g|$. Let

$$f = \frac{|g|^{q-1} \overline{\operatorname{sgn} g}}{\|g\|_q^{q-1}},$$

then

$$\|f\|_p^p = \frac{\int |g|^{(q-1)p} |\overline{g}|^p / |g|^p}{\|g\|_q^{q-1}} = \frac{\int |g|^q}{\|g\|_q^q} = 1.$$

And

$$\int fg = \int \frac{|g|^{q-1} |g|^2 / |g|}{\|g\|_q^{q-1}} = \|g\|_q,$$

[1] Folland, 1.14

hence $\|\phi_g\| = \sup |\displaystyle\int fg| \geq \|g\|_q$.

If $q = 1$, then $f = \overline{\operatorname{sgn} g} = \dfrac{\overline{g}}{|g|}$, $\|f\|_\infty = 1$, and $\displaystyle\int fg = \int \frac{\overline{g}g}{|g|} = \|g\|_1$.

If $q = \infty$, for $\varepsilon > 0$ let $A = \{x : |g(x)| > \|g\|_\infty - \varepsilon\}$. Then $\mu(A) > 0$. (Otherwise $\|g\|_\infty < \|g\|_\infty$), so if μ is semifinite there is $B \subset A$ with $0 < \mu(B) < \infty$. Let

$$f = \mu(B)^{-1}\chi_B \overline{\operatorname{sgn} g},$$

then

$$\|f\|_1 = \frac{\mu(B)}{\int_B \frac{|\overline{g}|}{|g|}} = \frac{\mu(B)}{\mu(B)} = 1,$$

so

$$\|\phi_g\| \geq \int fg = \frac{1}{\mu(B)} \int_B |g| \geq (\|g\|_\infty - \varepsilon)\mu(B)\mu(B)^{-1} = \|g\|_\infty - \varepsilon.$$

Since ε is arbitrary, $\|\phi_g\| = \|g\|_\infty$. □

Theorem 6.4

Let p, q be conjugate exponents. Suppose that g is a measurable function on X such that $fg \in L^1$ for all f in the space Σ of simple functions that vanish outside a set of finite measure, and the quantity

$$M_q(g) = \sup\left\{\left|\int fg\right| : f \in \Sigma, \|f\|_p = 1\right\}$$

is finite. (We can view this quantity as the operator norm of ϕ_g, but restricted to a smaller space Σ)

Also, suppose either $S_g = \{x : g(x) \neq 0\}$ is $\sigma-$finite or that μ is semifinite. Then $g \in L^q$ and $M_q(g) = \|g\|_q$. ♡

Proof **Case (1).** First we show that if f is a bounded measurable function that vanishes outside a set E of finite measure and $\|f\|_p = 1$, then $\left|\int fg\right| \leq M_q(g)$. Choose a sequence $\{f_n\}$ of simple functions such that $|f_n| \leq |f|$ and $f_n \to f$ a.e. (in particular, f_n vanishes outside E, because there is nothing to approximate outside E). Since $|f_n| \leq \|f\|_\infty \chi_E$ and $\chi_E g \in L^1$, by the dominated convergence theorem we have

$$\left|\int fg\right| = \lim_{n\to\infty} \left|\int f_n g\right| \leq M_q(g).$$

Case (2): $q < \infty$. We may assume that S_g is σ-finite since the condition automatically holds when μ is semifinite (we will separate this fact later). Let $E_1 \subset E_2 \subset \cdots$ be an increasing sequence with $S_g = \bigcup_{n=1}^\infty E_n$ and each $\mu(E_n) < \infty$. Choose a sequence of simple functions $\{\phi_n\}$ such that $\phi_n \to g$ and $|\phi_n| \leq |g|$ and let $g_n = \phi_n \chi_{E_n}$. Then $g_n \to g$ pointwise, $|g_n| \leq |g|$ and g vanishes outside E_n. Define

$$f_n = \frac{|g_n|^{q-1}\overline{\operatorname{sgn} g}}{\|g_n\|_q^{q-1}}.$$

Then $\|f_n\|_p = 1$ and by Fatou's lemma,

$$\begin{aligned}\|g\|_q = \left(\int |g|^q\right)^{1/q} &= \left(\int \liminf |g_n|^q\right)^{1/q} \\ &\le \liminf \|g_n\|_q = \liminf \int |f_n g_n| \\ &\le \liminf \int |f_n g| = \liminf \int f_n g \le M_q(g).\end{aligned}$$

On the other hand, Hölder's inequality gives $M_q(g) \le \|g\|_q$.

Case (3): $q = \infty$. Given $\varepsilon > 0$, let $A = \{x : |g(x)| \ge M_\infty(g) + \varepsilon\}$. If $\mu(A)$ were positive, we could choose $B \subset A$ with $0 < \mu(B) < \infty$. Set $f = \mu(B)^{-1}\chi_B \overline{\operatorname{sgn} g}$, then $\|f\|_1 = 1$, and

$$\int fg = \mu(B)^{-1}\int_B |g| \ge M_\infty(g) + \varepsilon.$$

But this is impossible by Case (1). Hence $\|g\|_\infty \le M_\infty(g)$. Finally, $|\int fg| \le \|g\|_\infty \int |f| = \|g\|_\infty$, completing the proof. □

Exercise 6.3 In Theorem 6.4, if μ is semifinite, then $S_g = \{x : g(x) \neq 0\}$ is σ-finite.

Proof Let $E \in \mathcal{M}$ and $\mu(E) = \infty$, then there exists a measurable $F \subset E$ such that $0 < \mu(F) < \infty$. If $\mu(S_g) < \infty$, then we are done. Suppose $\mu(S_g) = \infty$. Write S_g as a countable union of increasing sequence:

$$S_g = \bigcup_{n=1}^{\infty} \{x : |g(x)| > 1/n\},$$

then there exists a measurable $F \subset S_g$ with $0 < \mu(F) < \infty$, hence there is some $N \in \mathbb{N}$ such that $\{x : |g(x)| > 1/N\}$

Theorem 6.5

Let p, q be conjugate exponents. If $1 < p < \infty$, then for each $\phi \in (L^p)^*$ there exists $g \in L^q$ such that $\phi(f) = \int fg$ for all $f \in L^p$, and hence L^q is isometrically isomorphic to $(L^p)^*$. If μ is σ-finite, the same conclusion holds for $p = 1$. ♡

Proof **Case (1): μ is finite.** If $\phi \in (L^p)^*$ and E is a measurable set, let $\nu(E) = \phi(\chi_E)$. For any disjoint sequence $\{E_j\}$, if $E = \bigcup_{j=1}^\infty E_j$, we have $\chi_E = \sum_{j=1}^\infty \chi_{E_j}$. This series also converges in L^p norm:

$$\begin{aligned}\left\|\chi_E - \sum_{j=1}^n \chi_{E_j}\right\|_p = \left\|\sum_{j=n+1}^\infty \chi_{E_j}\right\|_p &= \left(\int \left|\sum_{j=n+1}^\infty \chi_{E_j}\right|^p\right)^{1/p} \\ &= \left(\int \sum_{j=n+1}^\infty \chi_{E_j}^p\right)^{1/p} \\ &= \mu\left(\bigcup_{j=n+1}^\infty E_j\right)^{1/p} \to 0 \quad (n \to \infty)\end{aligned}$$

(at this point we need the assumption that $p < \infty$). Since ϕ is linear and continuous and

$\sum_{j=1}^{n} \chi_{E_j} \to \sum_{j=1}^{\infty} \chi_{E_j}$ in L^p norm, it follows that

$$\phi\left(\sum_{j=1}^{n} \chi_{E_j}\right) \to \phi\left(\sum_{j=1}^{\infty} \chi_{E_j}\right) \quad (n \to \infty).$$

As a sequence in $\mathbb{C}$, we also have

$$\sum_{j=1}^{n} \phi(\chi_{E_j}) \to \sum_{j=1}^{\infty} \phi(\chi_{E_j}) \quad (n \to \infty),$$

hence

$$\nu(E) = \phi\left(\sum_{j=1}^{\infty} \chi_{E_j}\right) = \sum_{j=1}^{\infty} \phi(\chi_{E_j}) = \sum_{j=1}^{\infty} \nu(E_j).$$

Thus ν is a complex measure. If $\mu(E) = 0$, then $\chi_E = 0$ a.e., so $\nu(E) = 0$; that is, ν is absolutely continuous with respect to μ. By the **Radon-Nikodym theorem** there exists $g \in L^1(\mu)$ such that

$$\phi(\chi_E) = \nu(E) = \int_E g \, d\mu$$

for all E. If f is simple, then $\phi(f) = \int fg \, d\mu$, and $|\int fg| \leq \|\phi\| \|f\|_p$, so $fg \in L^1$. By Theorem 6.4, $g \in L^q$. Since the set of simple functions is dense in L^p, by passing to a limit we have $\phi(f) = \int fg$ for all $f \in L^p$.

Case (2): μ is σ-finite. Let $\{E_n\}$ be increasing, $0 < \mu(E_n) < \infty$, and $X = \bigcup_{n=1}^{\infty} E_n$. Identify $L^p(E_n), L^q(E_n)$ with the subspaces of $L^p(X), L^q(X)$ consisting of functions that vanishes outside E_n. Case (1) shows that for each n there exists $g_n \in L^q(E_n)$ such that $\phi(f) = \int fg_n$ for each $f \in L^p(E_n)$, and $\|g_n\|_q = \left\|\phi|_{L^p(E_n)}\right\| \leq \|\phi\|$. By Radon-Nikodym theorem, g_n is unique modulo alterations on nullsets. Denote ν_n the complex measure on E_n as obtained in Case (1), then $d\nu_n = g_n \, d\mu$ and $d\nu_m = g_m \, d\mu$. Let $n < m$, then $\nu_n = \nu_m$ on E_n, hence $g_n = g_m$ a.e. on E_n for all $n < m$. We define g a.e. on X by setting $g = g_n$ on E_n. By the monotone convergence theorem,

$$\|g\|_q = \lim_{n\to\infty} \|g_n\|_q \leq \|\phi\|,$$

so $g \in L^q$. Moreover, if $f \in L^p$, then by the dominated convergence theorem, $f\chi_{E_n} \to f$ in L^p norm and hence

$$\phi(f) = \lim_{n\to\infty} \phi(f\chi_{E_n}) = \lim_{n\to\infty} \int_{E_n} fg = \int fg.$$

Case (3): μ is arbitrary and $p > 1$. For each σ-finite set $E \subset X$ there is an a.e.-unique $g_E \in L^q(E)$ such that $\phi(f) = \int fg_E$ for all $f \in L^p(E)$ and $\|g_E\|_q \leq \|\phi\|$. If F is σ-finite and $F \supset E$, then $g_F = g_E$ a.e. on E, so $\|g_F\|_q \geq \|g_E\|_q$. Let

$$M = \sup\{\|g_E\|_q : E \text{ is } \sigma\text{-finite}\}.$$

Since $\|g_E\|_q \leq \|\phi\|$ for all σ-finite E, $M \leq \|\phi\|$.

Choose a sequence of σ-finite $\{E_n\}$ so that $\|g_{E_n}\|_q \to M$ and set $F = \bigcup_{n=1}^{\infty} E_n$. Then F is σ-finite and $\|g_F\|_q \geq \|g_{E_n}\|_q$ for all n, hence

$$\|g_F\|_q \geq \lim_{n\to\infty} \|g_{E_n}\|_q = M,$$

so $\|g_F\|_q = M$. Now if A is a σ-finite set containing F, we have

$$\int |g_F|^q + \int |g_{A\setminus F}|^q = \int |g_a|^q \leq M^q = \int |g_F|^q,$$

and thus $g_{A\setminus F} = 0$ and $g_A = g_F$ a.e. (here we use that fact that $q < \infty$). But if $f \in L^p$, then $A = F \cup \{x : f(x) \neq 0\}$ is σ-finite, so

$$\phi(f) = \int f g_A = \int f g_F.$$

Thus we may take $g = g_F$, completing the proof. □

6.3 Selected Exercises

Exercise 6.4 When does equality hold in Minkowski's inequality? ($p = 1, 1 < p < \infty, p = \infty$)

Proof

1. Let $p = 1$ and $f, g \in L^1$, then clearly
$$\int |f+g| \leq \int |f| + |g| = \int |f| + \int |g|.$$
The equality $\int |f+g| = \int |f| + |g|$ holds if and only if $|f+g| = |f| + |g|$ a.e. Taking square on both sides we have $fg = |fg|$, so equality holds if and only if fg is nonnegative a.e.
2. Let $1 < p < \infty$. The first inequality comes from
$$|f+g|^p \leq (|f| + |g|)|f+g|^{p-1},$$
so equality holds if and only if fg is nonnegative. The second inequality comes from
$$\|f(f+g)^{p/q}\| \leq \|f\|_p \|(f+g)^{p/q}\|_q \text{ and } \|g(f+g)^{p/q}\| \leq \|g\|_p \|(f+g)^{p/q}\|_q,$$
where the equality holds iff $|f|^p = \alpha|f+g|^p$ and $|g|^p = \beta|f+g|^p$. Taking pth root we have $b|f| = a|g|$ for some nonnegative constants a, b.
3. Let $p = \infty$. Since $\|\|_\infty$ is a norm, we have
$$\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty.$$
We can view f and g as bounded measurable functions with $\sup|f| = \|f\|_\infty$ and $\sup|g| = \|g\|_\infty$. If equality holds, then there exists $x \in X$ such that $|f(x) + g(x)| = |f(x)| + |g(x)|$, hence $f(x)g(x)$ is nonnegative.

Exercise 6.5 If $f \in L^p \cap L^\infty$ for some $p < \infty$, so that $f \in L^q$ for all $q > p$, then

$$\|f\|_\infty = \lim_{q \to \infty} \|f\|_q.$$

Proof **Case (1):** $\mu(X) < \infty$**.** Since

$$\left(\int |f|^q d\mu\right)^{1/q} = \|f\|_\infty \left(\int \left(\frac{|f|}{\|f\|_\infty}\right)^q d\mu\right)^{1/q} \leq \mu(X)^{1/q}\|f\|_\infty,$$

$\lim_{q\to\infty} \|f\|_q \leq \|f\|_\infty$. On the other hand, let $X_\varepsilon = \{x : |f(x)| \geq (1-\varepsilon)\|f\|_\infty\}$, then $\mu(X_\varepsilon) > 0$ (otherwise $\|f\|_\infty \leq (1-\varepsilon)\|f\|_\infty$, impossible). Now

$$\begin{aligned}\left(\int |f|^q d\mu\right)^{1/q} &\geq \left(\int_{X_\varepsilon} |f|^q d\mu\right)^{1/q} \\ &\geq \mu(X_\varepsilon)^{1/q}\left(\|f\|_\infty^q (1-\varepsilon)^q\right)^{1/q} \\ &= \mu(X_\varepsilon)^{1/q}\|f\|_\infty (1-\varepsilon).\end{aligned}$$

Hence

$$\limsup_{q\to\infty}\left(\int |f|^q d\mu\right)^{1/q} \leq \|f\|_\infty,$$

$$\liminf_{q\to\infty}\left(\int |f|^q d\mu\right)^{1/q} \geq (1-\varepsilon)\|f\|_\infty.$$

Since ε is arbitrary, it follows that

$$\limsup_{q\to\infty}\|f\|_q = \liminf_{q\to\infty}\|f\|_q = \|f\|_\infty,$$

therefore $\lim_{q\to\infty}\|f\|_q = \|f\|_\infty$.

Case (2): X is arbitrary. Since $f \in L^p$, $d\nu = |f|^p\, d\mu$ is a finite measure. Write

$$\left(\int |f|^q d\mu\right)^{1/q} = \left(\int |f|^q \frac{d\nu}{|f|^p}\right)^{1/q} = \left(\int |f|^{q-p} d\nu\right)^{\frac{1}{q-p}\frac{q-p}{q}},$$

then by case (1),

$$\lim_{q\to\infty}\left(\int |f|^{q-p} d\nu\right)^{\frac{1}{q-p}} = \|f\|_\infty.$$

Denote $A(q) = \left(\int |f|^{q-p} d\nu\right)^{\frac{1}{q-p}}$, then

$$\begin{aligned}\lim_{q\to\infty}\left(\int |f|^q\right)^{1/q} &= \lim_{q\to\infty} e^{(1-p/q)\log A(q)} \\ &= e^{\lim_{q\to\infty}(1-p/q)\log A(q)} \\ &= e^{\log\|f\|_\infty} = \|f\|_\infty.\end{aligned}$$

□

Exercise 6.6 Suppose $1 \leq p < \infty$. If $f_n, f \in L^p$ and $f_n \to f$ a.e., then $\|f_n - f\|_p \to 0$ iff $\|f_n\|_p \to \|f\|_p$.

Chapter 7 Radon Measures

Due to the topological structure on $\mathbb{R}^d$, the Lebesgue measure m enjoys some regularity propreties: Let $\varepsilon > 0$ and E be a Lebesgue measurable set, then

1. there exists an open set $U \supset E$ such that $m(U \setminus E) < \varepsilon$;
2. there exists a compact set $K \subset E$ such that $m(E \setminus K) < \varepsilon$.

In this chapter we will move forward to a locally compact Hausdorff space and study a class of measures with certain regular properties. We will also learn some approximation theorems, such as "the continuous functions with compact support is dense in L^p".

7.1 Locally Compact Hausdorff Spaces

Introduction

- *topological spaces*
- *Hausdorff spaces*
- *local compactness*
- *Urysohn's lemma*

7.1.1 Compactness

Let X be a topological space. A set $K \subset X$ is compact if every open cover of K contains a finite subcover. A neighborhood of a point $p \in X$ is any open subset of X which contains p. X is a Hausdorff space if for every $p, q \in X$ with $p \neq q$, there are neighborhoods U of p and V of q such that $U \cap V = \varnothing$.

Definition 7.1 (LCH)

*X is called a **locally compact Hausdorff** space, or LCH space, if X is a Hausdorff space and every point of X has a neighborhood whose closure is compact.* ♣

Theorem 7.1 (finite intersection property)

If $\{K_\alpha\}$ is a collection of compact subsets of a Hausdorff space X and if $\bigcap_\alpha K_\alpha = \varnothing$, then some finite subcollection of $\{K_\alpha\}$ also has empty intersection. ♡

Proof Put $V_\alpha = K_\alpha^c$, fix a member $K_1 \in \{K_\alpha\}$. Clearly $K_1 \subset X = (\bigcap_\alpha K_\alpha)^c = \bigcup_\alpha V_\alpha$, thus $K_1 \subset V_{\alpha_1} \cup \cdots V_{\alpha_n}$ for some finite subcollection. Then

$$K_1 \cap K_{\alpha_1} \cap \cdots \cap K_{\alpha_n} = \varnothing.$$

□

Theorem 7.2 (separation of a compact set and a point)

Suppose X is Hausdorff, K is a compact subset of X, and $p \in K^c$. Then there are open sets U and W such that $p \in U$, $K \subset W$, and $U \cap W = \varnothing$. ♡

Proof Let $x \in K$, then there exist open sets $V_x \ni x, U_x \ni p$ such that $V_x \cap U_x = \varnothing$. Since K is compact, there is a finite subcollection of $\{V_x : x \in K\}$ covering K, say $K \subset \bigcup_{i=1}^n V_{x_i}$, and $\bigcup_{i=1}^n V_{x_i} \cap \bigcap_{i=1}^n U_{x_i} = \varnothing$. Thus we can take $U = \bigcap_{i=1}^n U_{x_i}$ and $V = \bigcup_{i=1}^n V_{x_i}$. □

Theorem 7.3

Suppose U is open in a LCH space X, $K \subset U$, and K is compact. Then there is an open set V with compact closure such that

$$K \subset V \subset \overline{V} \subset U.$$

♡

Proof Every point x of K has a neighborhood $A_x \subset U$ with compact closure, and since K is compact, K is covered by the union of finitely many of these neighborhoods, say $K \subset \bigcup_{i=1}^{n} A_i \subset U$. let $G = \bigcup_{i=1}^{n} A_i$, then $\overline{G}$ is compact. If $U = X$, take $V = G$.

If $U \neq X$, Let $C = U^c$. To each $p \in C$ there corresponds an open set W_p such that $K \subset W_p$ and $p \notin \overline{W_p}$. Consider the collection $\{C \cap \overline{G} \cap \overline{W_p} : p \in C\}$. We claim that $\bigcap_{p \in C}(C \cap \overline{G} \cap \overline{W_p}) = \varnothing$. Suppose the intersection is nonempty, then there exists $x \in C \cap \overline{G} \cap \overline{W_p}$ for all $p \in C$. Then $x \in C$ implies x corresponds to some W_x and $x \notin \overline{W_x}$, but $x \in \overline{W_x}$, a contradiction. Thus $\bigcap_{p \in C}(C \cap \overline{G} \cap \overline{W_p}) = \varnothing$. By the finite intersection property, we can choose $p_1, \cdots, p_n \in C$ such that

$$C \cap \overline{G} \cap \overline{W_{p_1}} \cap \cdots \cap \overline{W_{p_n}} = \varnothing.$$

The set $V = G \cap W_{p_1} \cap \cdots \cap W_{p_n}$ has the required properties because

$$\overline{V} = \overline{G \cap W_{p_1} \cap \cdots \cap W_{p_n}} \subset \overline{G} \cap \overline{W_{p_1}} \cap \cdots \cap \overline{W_{p_n}} \subset C^c = U.$$

□

7.1.2 Urysohn's Lemma

Definition 7.2 (semicontinuous)

Let f be a real-valued function on a topological space.

If $\{x : f(x) > \alpha\}$ is open for every real α, then f is said to be ***lower semicontinuous****.*

If $\{x : f(x) < \alpha\}$ is open for every real α, then f is said to be ***upper semicontinuous****.*

♣

Properties

1. *A real-valued function f is continuous if and only if it is both upper and lower semicontinuous.*
2. *Let $\{f_i\}_{i \in \mathcal{I}}$ be a collection of lower semicontinuous functions, then $\sup_{i \in \mathcal{I}} f_i$ is lower continuous.*
3. *Let $\{g_j\}_{j \in \mathcal{J}}$ be a collection of upper semicontinuous functions, then $\inf_{j \in \mathcal{J}} g_j$ is lower continuous.*

Proof

1. If f is both upper and lower semicontinuous, then $\{x : a < f(x) < b\} = f^{-1}((a, b))$ is open for every $a < b$, hence is continuous. The other direction is obvious.
2. Observe that $\{x : (\sup_{i \in \mathcal{I}} f_i)(x) > \alpha\} = \bigcup_{i \in \mathcal{I}}\{x : f_i(x) > \alpha\}$.
3. Exercise.

□

Example 7.1 Characteristic unctions of open sets are lower semicontinuous, and characteristic functions of closed sets are upper semicontinuous. More precisely, let U be open and $\alpha \in \mathbb{R}$.

- If $\alpha < 0$, then $\{x \in X : \chi_U(x) > \alpha\} = X$.

- If $0 \le \alpha < 1$, then $\{x \in X : \chi_U(x) > \alpha\} = U$.
- If $\alpha \ge 1$, then $\{x \in X : \chi_U(x) > \alpha\} = \varnothing$.

Definition 7.3 (support)

*The **support** of a complex-valued function f on a topological space(denoted by* $\operatorname{supp} f$*) X is the closure of the set*

$$\{x : f(x) \neq 0\}.$$

Denote $C_c(X)$ the collection of all continuous complex-valued functions on X whose support is compact. ♣

Exercise 7.1 Show that $\operatorname{supp}(f+g) \subset (\operatorname{supp} f) \cup (\operatorname{supp} g)$.

We introduce some notations:

Notation	Meaning
$K \prec f$	K is compact, $f \in C_c(X), 0 \le f \le 1, f(x) = 1\ \forall x \in K$
$f \prec V$	V is open, $f \in C_c(X), 0 \le f \le 1, \operatorname{supp} f \subset V$
$K \prec f \prec V$	both $K \prec f$ and $f \prec V$ hold

Theorem 7.4 (Urysohn's lemma)

Suppose X is a LCH space, V is open in X, $K \subset V$ and K is compact. Then there exists an $f \in C_c(X)$ such that

$$K \prec f \prec V.$$ ♡

Proof Let $r_1 = 0, r_2 = 1, r_3, r_4, r_5, \cdots$ be an enumeration of $\mathbb{Q} \cap [0,1]$. By Theorem 7.3, there exist open V_1 with $\overline{V_1}$ being compact such that

$$K \subset V_1 \subset \overline{V_1} \subset V.$$

Apply Theorem 7.3 again to the inclusion $\overline{V_1} \subset V$ we obtain a precompact set V_0 such that

$$\overline{V_1} \subset V_0 \subset \overline{V_0} \subset V.$$

We construct a sequence of sets $\{V_{r_k}\}$ inductively. Suppose $n \ge 2$ and $V_{r_1}, \cdots, V_{r_n}$ has been chosen such that

$$r_i < r_j \text{ implies } \overline{V_{r_j}} \subset V_{r_i}.$$

Then one of $0 = r_1, 1 = r_2, \cdots, r_n$ will be the largest one which is smaller than r_{n+1}, denote it by r_i, and another will be the smallest one larger than r_{n+1}, denote it by r_j. Then $r_i < r_j$, which implies $\overline{V_{r_j}} \subset V_{r_i}$. Then there exists $V_{r_{n+1}}$ such that

$$\overline{V_{r_j}} \subset V_{r_{n+1}} \subset \overline{V_{r_{n+1}}} \subset V_{r_i},$$

and by the choice of r_i and r_j we have $r_i < r_{n+1} < r_j$. Continuing, we obtian a collection $\{V_r : r \in \mathbb{Q} \cap [0,1]\}$ of open sets with $K \subset V_1, \overline{V_0} \subset V$, each $\overline{V_r}$ is compact and $s > r$ implies $\overline{V_s} \subset V_r$.

Define

$$f(x) = \begin{cases} r & x \in V_r, \\ 0, & x \notin V_r, \end{cases}, \quad g(x) = \begin{cases} 1 & x \in \overline{V_s}, \\ s, & x \notin \overline{V_s}, \end{cases}$$

and $f = \sup_r f_r, g = \inf_s g_s$. Then each f_r is lower semicontinuous and each g_s is upper semicontinuous. Hence f is lower semicontinuous and g is upper semicontinuous. Observe that $0 \leq f \leq 1$, and $f(x) = 1$ for all $x \in K$ (because $K \subset V_r$ for all r), and f has support in $\overline{V_0}$. Now we show $f = g$.

- If $f_r(x) > g_s(x)$, then $r > s, x \in V_r, x \notin \overline{V_s}$. But $r > s$ implies $V_r \subset V_s$, a contradiction. Hence $f_r \leq g_s$ for all r, s, so $\sup_r f_r \leq g_s$ for all s, therefore
$$f = \sup_r f_r \leq \inf_s g_s = g.$$
- Suppose $f(x) < g(x)$ for some x, then there are $r, s \in \mathbb{Q} : f(x) < r < s < g(x)$. $f(x) < r$ implies $x \notin V_r$, $g(x) > s$ implies $x \in \overline{V_s}$. But $\overline{V_s} \subset V_r$, a contradiction.

Therefore $f = g$, and thus f is continuous, $K \prec f \prec V$. □

Theorem 7.5 (partitions of unity)

Suppose $V_1, \cdots, V_n$ are open subsets of a LCH space X, K is compact, and
$$K \subset V_1 \cup \cdots \cup V_n.$$
Then there exist functions $h_i \prec V_i (i = 1, \cdots, n)$ such that
$$h_1(x) + \cdots + h_n(x) = 1 \quad (x \in K).$$
The collection $\{h_1, \cdots, h_n\}$ is called a ***partition of unity*** *on K, subordinate to the cover $\{V_1, \cdots, V_n\}$.* ♡

Proof Each $x \in K$ has a neighborhood W_x with compact closure $\overline{W_x} \subset V_i$ for some i (depending on x). Choose $x_1, \cdots, x_m$ such that $K \subset W_{x_1} \cup \cdots \cup W_{x_m}$. If $1 \leq i \leq n$, let H_i be the union of those $\overline{W_{x_j}}$ which lie in V_i. By Urysohn's lemma, there are functions g_i such that $H_i \prec g_i \prec V_i$. Define
$$\begin{aligned} h_1 &= g_1 \\ h_2 &= (1 - g_1)g_2 \qquad \cdots \\ h_n &= (1 - g_1)(1 - g_2) \cdots (1 - g_{n-1})g_n. \end{aligned}$$
Then $h_i \prec V_i$ and
$$h_1 + \cdots + h_n = 1 - (1 - g_1)(1 - g_2) \cdots (1 - g_{n-1})g_n.$$
Since $K \subset H_1 \cup \cdots H_n$, for each point $x \in K$ there exists i such that $g_i(x) = 1$, completing the proof. □

7.2 Positive Linear Functionals on $C_c(X)$

From now on X denotes an LCH space. Let μ be a Borel measure that is finite on compact sets and E a Borel subset of X. The measure μ is called **outer regular** on E if
$$\mu(E) = \inf\{\mu(U) : U \supset E, U \text{ open}\}$$
and **inner regular** on E if
$$\mu(E) = \sup\{\mu(K) : K \subset E, K \text{ compact}\}.$$
If μ is both outer and inner regular on all Borel sets, μ is called **regular**.

Definition 7.4

*A **Radon measure** μ on X is a Borel measure that satisfies all the following conditions:*

- *μ is finite on all compact sets,*
- *μ is outer regular on all Borel sets,*
- *μ is inner regular on all open sets.* ♣

Theorem 7.6 (the Riesz representation theorem)

If I is a positive linear functional on $C_c(X)$, then there is a unique Radon measure μ on X such that $I(f) = \int f\, d\mu$ for all $f \in C_c(X)$. ♡

Throughout the proof of this theorem, we use K to denote a compact subset of X and U will stand for an open set in X. The proof is from Rudin's *Real and Complex Analysis* and Folland's *Real Analysis: Modern Techniques and Their Applications*.

Uniqueness

If μ is outer regular and inner regular, then μ is determined on $\mathcal{M}$ by its values on compact sets. Let μ_1, μ_2 be measures for which the theorem holds, then it suffices to prove that $\mu_1(K) = \mu_2(K)$ for all K. Fix K and $\varepsilon > 0$, then by outer regularity, there exists a $U \supset K$ with $\mu_2(U) < \mu_2(K) + \varepsilon$. By Urysohn's lemma, there exists an f so that $K \prec f \prec U$, hence

$$\begin{aligned}\mu_1(K) = \int \chi_K\, d\mu_1 &\leq \int_X f\, d\mu_1 = I(f) = \int f\, d\mu_2 \\ &\leq \int \chi_U\, d\mu_2 = \mu_2(U) < \mu_2(K) + \varepsilon,\end{aligned}$$

thus $\mu_1(K) \leq \mu_2(K)$. Interchanging the roles of μ_1 and μ_2 gives the opposite inequality.

Construction of μ and the Representation

For every open set V in X, define

$$\mu(U) = \sup\{I(f) : f \prec U\},$$

and we then define $\mu^*(E)$ for an arbitrary $E \subset X$ by

$$\mu^*(E) = \inf\{\mu(U) : E \subset U, U \text{ open}\}.$$

If $U \subset V$, then clearly $\mu(U) \leq \mu(V)$, and hence $\mu^*(U) = \mu(U)$ if U is open.

STEP I $\quad$ *μ^* is an outer measure.*

Proof First we show subadditivity of μ on open sets. Let $\{U_j\}$ be a sequence of open sets and $U = \bigcup_{j=1}^{\infty} U_j$, and let $f \in C_c(X), f \prec U, K = \operatorname{supp}(f)$. Since K is compact, $K \subset \bigcup_{j=1}^{n} U_j$ for some $n \in \mathbb{N}$. By partition of unity there exist $g_1, \cdots, g_n \in C_c(X)$ with $g_j \prec U_j$ and $\sum_{j=1}^{n} g_j = 1$ on K. Then $f = \sum_{j=1}^{n} f g_j$ and $f g_j \prec U_j$, so

$$I(f) = \sum_{j=1}^{n} I(f g_j) \leq \sum_{j=1}^{n} \mu(U_j) \leq \sum_{j=1}^{\infty} \mu(U_j).$$

Taking supremum over all $f \prec U$ yields $\mu(U) \leq \sum_{j=1}^{\infty} \mu(U_j)$. The monotonicity of μ^* is obvious.

Now let $\{E_j\}$ be a sequence of sets in X, and let $\varepsilon > 0$, for each j there is an open set $U_j \supset E$ with $\mu(U_j) \leq \mu^*(E_j) + 2^{-j}\varepsilon$. Clearly $\bigcup_{j=1}^{\infty} E_j \subset \bigcup_{j=1}^{\infty} U_j$, hence

$$\mu^*\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \mu^*\left(\bigcup_{j=1}^{\infty} U_j\right) \leq \sum_{j=1}^{\infty} \mu(U_j) \leq \sum_{j=1}^{\infty} \mu^*(E_j) + \varepsilon,$$

completing the proof. □

STEP II *Every open set is μ^*-measurable.*

Proof We need to show that if U is open and E is any subset of X such that $\mu^*(E) < \infty$, then $\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \setminus U)$. First suppose E is open. Then $E \cap U$ is open, so given $\varepsilon > 0$ we can find $f \in C_c(X)$ such that

$$f \prec E \cap U \text{ and } I(f) > \mu(E \cap U) - \varepsilon.$$

Also, $E \setminus \text{supp}\,(f)$ is open, so there exists $g \in C_c(X)$ such that

$$g \prec E \setminus \text{supp}\,(f) \text{ and } I(g) > \mu(E \setminus \text{supp}\,(f)) - \varepsilon.$$

But then $f + g \prec E$, so

$$\begin{aligned} \mu(E) &\geq I(f+g) = I(f) + I(g) \\ &> \mu(E \cap U) + \mu(E \setminus \text{supp}\,(f)) - 2\varepsilon \\ &\geq \mu^*(E \cap U) + \mu^*(E \setminus U) - 2\varepsilon. \end{aligned}$$

Since ε is arbitrary, $\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \setminus U)$. For the general case, if $\mu^*(E) < \infty$ we can find an open $V \supset E$ with $\mu(V) < \mu^*(E) + \varepsilon$, and hence

$$\begin{aligned} \mu^*(E) + \varepsilon > \mu(V) &\geq \mu^*(V \cap U) + \mu^*(V \setminus U) \\ &\geq \mu^*(E \cap U) + \mu^*(E \setminus U). \end{aligned}$$

Letting $\varepsilon \to 0$ completes the proof. □

At this point by Carathéodory's theorem the collection of μ^*-measurable set is a σ-algebra, and this σ-algebra contains all open sets in X, hence contains the Borel σ-algebra $\mathcal{B}_X$. Thus $\mu = \mu^*|_{\mathcal{B}_X}$ is a Borel measure. The measure μ is outer regular by our definition of μ^*.

STEP III *μ satisfies*

$$\mu(K) = \inf\{I(f) : f \geq \chi_K\} \text{ for all compact } K \subset X. \tag{7.1}$$

Proof Observe that $I(g) = I(f) + I(g - f)$, so $f \leq g$ implies $I(f) \leq I(g)$, thus I is monotone. If K is compact, $f \in C_c(X)$ and $f \geq \chi_K$, let $U_\varepsilon = \{x : f(x) > 1 - \varepsilon\}$. Then U_ε is open, and if $g \prec U_\varepsilon$, then $0 \leq g \leq 1$ on U_ε, and $(1-\varepsilon)^{-1} f > 1$ on U_ε, hence $(1-\varepsilon)^{-1} f - g \geq 0$, and hence

$$I(g) \leq (1-\varepsilon)^{-1} I(f) \quad \text{ for all } g \prec U_\varepsilon.$$

Since $f = 1$ on K, $K \subset U_\varepsilon$. It follows that

$$\mu(K) \leq \mu(U_\varepsilon) \leq (1-\varepsilon)^{-1} I(f).$$

Letting $\varepsilon \to 0$ we see that $\mu(K) \leq I(f)$.

On the other hand, let $\varepsilon > 0$, then by the outer regularity there is an open $U \supset K$ with

$\mu(U) \le \mu(K) + \varepsilon$. By Urysohn's lemma there exists $f \in C_c(X)$ such that $\chi_K \le f \prec U$, hence $I(f) \le \mu(U) \le \mu(K) + \varepsilon$. Letting $\varepsilon \to 0$ we have

$$I(f) \le \mu(K) + \varepsilon \qquad \text{for all } \chi_K \le f \prec U.$$

Since $\inf\{I(f) : f \ge \chi_K\} = \inf\{I(f) : f \ge \chi_K \le f \prec U\}$ by the monotonicity of I, we have shown

$$\inf\{I(f) : f \ge \chi_K\} \le \mu(K),$$

completing the proof. $\square$

STEP IV $\quad$ *$I(f) = \int f \, d\mu$ for all $f \in C_c(X)$.*

Proof It suffices to show $I(f) = \int f \, d\mu$ if $f \in C_c(X, [0,1])$. Given $N \in \mathbb{N}$, for $1 \le j \le N$ let

$$K_j = \{x : f(x) > j/N\}, \quad K_0 = \text{supp } (f).$$

Define $f_1, \cdots, f_N \in C_c(X)$ by

$$f_j(x) = \begin{cases} 0, & x \notin K_{j-1}, \\ f(x) - (j-1)/N & x \in K_{j-1} \setminus K_j, \\ 1/N & x \in K_j. \end{cases}$$

Then $\dfrac{\chi_{K_j}}{N} \le f_j \le \dfrac{\chi_{K_{j-1}}}{N}$, hence

$$\frac{1}{N}\mu(K_j) \le \int f_j \, d\mu \le \frac{1}{N}\mu(K_{j-1}).$$

Now we need connect the integral with the linear functional I. Also, if U is an open set containing K_{j-1}, we have $Nf_j \prec U$ and so $I(f_j) \le \dfrac{\mu(U)}{N}$. Hence by outer regularity, taking infimum over all such $U \supset K_{j-1}$ yields

$$I(f_j) \le \frac{1}{N}\mu(K_{j-1}).$$

Again by 7.1, $\mu(K_j)$ is the infimum of $I(f)$ with $f \ge \chi_K$, hence

$$\frac{1}{N}\mu(K_j) \le I(f_j).$$

Together we have

$$\frac{1}{N}\mu(K_j) \le I(f_j) \le \frac{1}{N}\mu(K_{j-1}).$$

Moreover, $f = \sum_{j=1}^N f_j$, so that

$$\frac{1}{N}\sum_{j=1}^N \mu(K_j) \le \int f \, d\mu \le \frac{1}{N}\sum_{j=1}^N \mu(K_{j-1}),$$

$$\frac{1}{N}\sum_{j=1}^N \mu(K_j) \le I(f) \le \frac{1}{N}\sum_{j=1}^N \mu(K_{j-1}).$$

It follows that

$$\left| I(f) - \int f \, d\mu \right| \le \frac{\mu(K_0) - \mu(K_N)}{N} \le \frac{\mu(\text{supp } (f))}{N}.$$

Since $\mu(\text{supp } (f)) < \infty$, letting $N \to \infty$ leads to $I(f) = \int f \, d\mu$. $\square$

7.3 Approximation Theorems

7.3.1 Density of $C_c(X)$ and Lusin's Theorem

Proposition 7.1

Every Radon measure is inner regular on all of its σ-finite sets.

Proof Suppose μ is a Radon measure and E is σ-finite. If $\mu(E) < \infty$, then for eny $\varepsilon > 0$ we can choose an open $U \supset E$ such that $\mu(U) < \mu(E) + \varepsilon$ and a compact $F \subset E$ such that $\mu(F) > \mu(U) - \varepsilon$. Since $\mu(U \setminus E) < \varepsilon$, we can also choose an open $V \supset U \setminus E$ such that $\mu(V) < \varepsilon$. Let $K = F \setminus V$, then K is compact, $K \subset E$, and

$$\mu(K) = \mu(F) - \mu(F \cap V) > \mu(E) - \varepsilon - \mu(V) > \mu(E) - 2\varepsilon.$$

Thus μ is inner regular on E.

If $\mu(E) = \infty$, we can write E as an increasing union of sets E_j with each $\mu(E_j) < \infty$ and $\mu(E) = \lim_{j\to\infty} \mu(E_j) = \infty$. Then for any $N \in \mathbb{N}$ there exists j such that $N < \mu(E_j) < \infty$. By the preceding argument, there is a compact $K \subset E_j$ with $\mu(K) > N$, thus $\sup\{\mu(K) : K \subset E, K \text{ compact}\} = \infty$, completing the proof. □

Theorem 7.7

If μ is a Radon measure on X, $C_c(X)$ is dense in $L^p(\mu)$ for $1 \leq p < \infty$.

Proof It suffices to show that for any Borel set E with $\mu(E) < \infty$, χ_E can be approximated by functions of compact support in L^p norm. Let $\varepsilon > 0$, we can choose a compact $K \subset E$ and open $U \supset E$ such that $\mu(U \setminus K) < \varepsilon$. By Urysohn's lemma we can choose $f \in C_c(X)$ such that $\chi_K \leq f \leq \chi_U$. Then

$$\|\chi_E - f\|_p \leq \mu(U \setminus K)^{1/p} < \varepsilon^{1/p},$$

completing the proof. □

Theorem 7.8 (Lusin)

Suppose f is a complex measurable function on X and μ is a Radon measure on X, let $\mu(A) < \infty$ and $f(x) = 0$ for all $x \notin A$. Let $\varepsilon > 0$, then there exists a $g \in C_c(X)$ such that

$$\mu(\{x : f(x) \neq g(x)\}) < \varepsilon. \tag{7.2}$$

Furthermore, we may arrange it so that

$$\sup_{x\in X} |g(x)| \leq \sup_{x\in X} |f(x)|. \tag{7.3}$$

Proof **Case (1): $0 \leq f < 1$ and A is compact.** Recall that the measurable function f can be approximated by a sequence of simple functions $\{\phi_n\}$, and put $t_1 = \phi_1, t_n = \phi_n - \phi_{n-1}$ for $n = 2, 3, 4, \cdots$ Then $2^n t_n$ is the characteristic function of a set $T_n \subset A$, and $f(x) = \sum_{n=1}^{\infty} t_n(x)$ for all $x \in X$. Fix an open set V such that $A \subset V$ and $\overline{V}$ is compact. Since T_n is of finite measure, by the regularity of μ, there are compact K_n and open V_n such that

$$K_n \subset T_n \subset V_n \subset V \text{ and } \mu(V_n \setminus K_n) < 2^{-n}\varepsilon.$$

By Urysohn's lemma, there is a function h_n such that $K_n \prec h_n \prec V_n$. Define

$$g(x) = \sum_{n=1}^{\infty} 2^{-n} h_n(x), \quad x \in X.$$

Then $|g(x)| \leq \sum_{n=1}^{\infty} 2^{-n}$, so this series converges uniformly on X, hence g is continuous. Also, supp $(g) \subset \overline{V}$. Since $2^{-n}h_n(x) = t_n(x)$ except in $V_n \setminus K_n$, we have $g(x) = f(x)$ except in $\bigcup_{n=1}^{\infty}(V_n \setminus K_n)$, and $\mu\left(\bigcup_{n=1}^{\infty}(V_n \setminus K_n)\right) < \varepsilon$. Thus $\mu(\{x : f(x) \neq g(x)\}) < \varepsilon$.
Case (2): general case. If A is compact and f is a bounded measurable function, the result holds. Now remove the compactness of A and assume f is a complex measurable function. Since $\mu(A) < \infty$, there is a compact $K \subset A$ with $\mu(A \setminus K) < \varepsilon$. Let $B_n = \{x : |f(x)| > n\}$, then $\cap B_n = \varnothing$ because $f : X \to \mathbb{C}$, so $\lim_{n\to\infty} \mu(B_n) = 0$. Then $f = (1 - \chi_{B_n})f$ on B_n^c, the result follows.

Finally, let $R = \sup_{x\in X} |f(x)|$, and define

$$\varphi(z) = \begin{cases} z & \text{if } |z| \leq R, \\ Rz/|z| & \text{if } |z| > R. \end{cases}$$

Then φ is a continuous surjection from $\mathbb{C}$ to $B(0, R)$. If g satisfies (7.2), we set $g_1 = \varphi \circ g$, then g_1 satisfies (7.2) and (7.3). □

7.4 The Dual of $C_0(X)$

7.4.1 Extension From $C_c(X)$ to $C_0(X)$

We start by introducing a new function space $C_0(X)$. Let X be an LCH space. If $f \in C(X)$, we say that f **vanishes at infinity** if for every $\varepsilon > 0$ the set $\{x : |f(x)| \geq \varepsilon$ is compact, and we define

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

We define the **uniform metric** d on $C_0(X)$ by $d(f, g) = \sup_{x\in X} |f(x) - g(x)|$, and write $\|f\|_u = \sup_{x\in X} |f(x)|$ so that $\|f - g\|_u = \sup_{x\in X} |f(x) - g(x)|$. The uniform metric is defined in the same way on $C_c(X)$.

Proposition 7.2 ($C_c(X)$ is dense in $C_0(X)$)

If X is an LCH space, $C_0(X)$ is the closure of $C_c(X)$ in the uniform metric. In other words, if $f \in C_0(X)$ and $\varepsilon > 0$, then there is $g \in C_c(X)$ such that $\|f - g\|_u < \varepsilon$. ♠

Proof Let $\{f_n\}$ be a sequence in $C_c(X)$ that converges uniformly to f, then f is continuous. For each ε there exists $n \in \mathbb{N}$ such that $\sup_{x\in X} |f_n(x) - f(x)| = \|f_n - f\|_u < \varepsilon$. Then $|f(x)| < \varepsilon$ if $x \neq$ supp (f_n), so $f \in C_0(X)$.
Conversely, for $f \in C_0(X)$, we need to find a sequence $\{f_n\} \subset C_c(X)$ converging to f. The idea is to truncate f to get f_n, and use Urysohn's lemma to modify the continuity of f_n. Let

$$K_n = \left\{x \in X : |f(x)| \geq \frac{1}{n}\right\},$$

then K_n is closed and bounded, hence compact. By Urysohn's lemma there exists $g_n \in C(X)$ such that $g_n = 1$ on K_n and $g_n = 0$ outside another compact set (containing K_n), so

$g_n \in C_c(X)$. Set $f_n = fg_n$, then $fg_n = f$ on K_n and $fg_n \in C_c(X)$. Finally,

$$\|f_n(x) - f(x)\| < \frac{1}{n} \; \forall x \notin K_n,$$

so $\|f_n - f\|_u \to 0$ as $n \to \infty$. □

Lemma 7.1

$C_0(X)$ is complete. ♡

Proof Let $\{f_n\}$ be a Cauchy sequence in $C_0(X)$, then $\{f_n\}$ is uniformly Cauchy. For each x define $f(x) = \lim_{n\to\infty} f_n(x)$, then f is continuous. Let $\varepsilon > 0$, then there exists n such that $|f(x) - f_n(x)| < \varepsilon$ for all $x \in X$. Since $f_n \in C_0(X)$, there exists a compact K_n such that $|f_n(x)| < \varepsilon$ for all $x \in K_n^c$. Thus $|f(x)| < 2\varepsilon$ for all $x \in K_n^c$, hence $f \in C_0(X)$. □

In the proof of the Riesz representation theorem, we have seen that the Radon measure μ also satisfies regularity in terms of the positive linear funcitonal I:

$$\mu(U) = \sup\{I(f) : f \in C_c(X), f \prec U\} \quad \text{for all open } U \subset X,$$

and the linear functional I is precisely given by $I(f) = \int f \, d\mu$, $f \in C_c(X)$. We want to extend I from $C_c(X)$ to $C_0(X)$, and this is guaranteed by the following theorem:

Theorem 7.9 (extension of linear operators)

Let X and Y be a Banach spaces. Let $\mathcal{D} \subseteq X$ be a dense subspace. Suppose $T : \mathcal{D} \to Y$ is a bounded linear operator:

$$\|Tx\|_X \leq C\|x\|_Y, \quad \forall x \in \mathcal{D}.$$

Then T extends uniquely to a bounded linear operator $X \to Y$. ♡

Proof Let $x \in X$, then there exists $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{D}$ with $x_n \to x$ as $n \to \infty$. Define

$$Tx = \lim_{n\to\infty} Tx_n.$$

Since T is bounded, $\{Tx_n\}_{n\in\mathbb{N}}$ is Cauchy, so the limit exists in Y. We also need to justify the limit is independent of choice of approximating sequence, Let $\{y_n\}_{n\in\mathbb{N}} \to x$ as $n \to \infty$, then

$$\begin{aligned} \|Tx_n - Ty_n\|_Y &\leq \|T\|\|x_n - y_n\|_X \\ &\leq \|T\|(\|x_n - x\|_X + \|y - y_n\|_X) \to 0 \quad (n \to \infty), \end{aligned}$$

hence $\lim_{n\to\infty} Tx_n = \lim_{n\to\infty} Ty_n = Tx$.
Finally, by the extension we have $\|Tx_n - Tx\|_Y \to 0$ as $n \to \infty$, so $\lim_{n\to\infty} \|Tx_n\|_Y = \|Tx\|_Y$. Since $\|Tx_n\| \leq C\|x_n\|$, letting $n \to \infty$ leads to

$$\|Tx\| \leq C\|x\|.$$

Let S be another continuous extension. If $x \in X$ and $\{x_n\}_{n\in\mathbb{N}} \subset \mathcal{D}$ converges to x, then $Sx = \lim_{n\to\infty} Sx_n = \lim_{n\to\infty} Tx_n = Tx$, hence $S = T$. □

From the extension theorem we deduce that $\int f \, d\mu$ extends continuously to $C_0(X)$ if and only if it is bounded with respect to the uniform norm. In particular, if we let $U = X$, then

$$\mu(X) = \sup\left\{\int f \, d\mu : f \in C_c(X), f \prec X\right\} = \sup\left\{\int f \, d\mu : f \in C_c(X), 0 \leq f \leq 1\right\}.$$

We observe that if $\mu(X) < \infty$, then I is bounded and $\mu(X)$ is the operator norm of I.

Therefore, the positive bounded linear functionals on $C_0(X)$ are given by integration against finite Radon measures.

Next, we remove the "positive" restriction to give a complete description of $C_0(X)^*$.

7.4.2 Continuous Linear Functionals on $C_0(X)$

We have explored the representation of positive linear functionals on $C_0(X)$, and the following lemma says that any continuous linear functional on $C_0(X)$ decomposes into two positive linear functionals.

Lemma 7.2 (decomposition of linear functionals)

If $I \in C_0(X,\mathbb{R})^*$, there exist positive functionals $I^+, I^- \in C_0(X,\mathbb{R})^*$ such that $I = I^+ - I^-$. ♡

Proof For $f \in C_0(X,[0,\infty))$, we define

$$I^+(f) = \sup\{I(g) : g \in C_0(X,\mathbb{R}), 0 \le g \le f\}.$$

Since I is a bounded linear functional on $C_0(X,\mathbb{R})$, $|I(g)| \le \|I\|\|g\|_u \le \|I\|\|f\|_u$ for all $0 \le g \le f$, and $I(0) = 0$. Hence $0 \le I^+(f) \le \|I\|\|f\|_u$. We claim that I^+ is a linear functional on $C_0(X,[0,\infty))$.

Homogeneity. Let $c \ge 0$, then

$$\begin{aligned} I^+(cf) &= \sup\{I(g) : g \in C_0(X,\mathbb{R}), 0 \le g \le cf\} \\ &= \sup\left\{cI\left(\frac{g}{c}\right) : \frac{g}{c} \in C_0(X,\mathbb{R}), 0 \le \frac{g}{c} \le cf\right\} \\ &= c\sup\left\{I\left(\frac{g}{c}\right) : \frac{g}{c} \in C_0(X,\mathbb{R}), 0 \le \frac{g}{c} \le cf\right\} = cI^+(f). \end{aligned}$$

Linearity. If $0 \le g_1 \le f_1$ and $0 \le g_2 \le f_2$, then $0 \le g_1 + g_2 \le f_1 + f_2$, so that $I^+(f_1+f_2) \ge I(g_1+g_2) = I(g_1) + I(g_2)$. Taking supremum over g_1 and g_2 gives

$$I^+(f_1+f_2) \ge I^+(f_1) + I^+(f_2).$$

On the other hand, if $0 \le g \le f_1 + f_2$, let $g_1 = \min(g, f_1)$ and $g_2 = g - g_1$. Then $0 \le g_1 \le f_1$ and $0 \le g_2 \le f_2$, so

$$I(g) = I(g_1) + I(g_2) \le I^+(f_1) + I^+(f_2),$$

again taking supremum over g, we have

$$I^+(f_1+f_2) \le I^+(f_1) + I^+(f_2).$$

Therefore, $I^+(f_1+f_2) = I^+(f_1) + I^+(f_2)$.

If $f \in C_0(X,\mathbb{R})$, then $f^+, f^- \in C_0(X,[0,\infty))$, and we define

$$I^+(f) = I^+(f^+) - I^+(f^-).$$

If $g, h \ge 0$ and $f = g - h$, then $g + f^- = h + f^+$, hence $I^+(g) = I^+(f^-) = I^+(h) + I^+(f^+)$. Thus $I^+(f) = I^+(g) - I^+(h)$, and thus I^+ is linear on $C_0(X,\mathbb{R})$. Moreover,

$$|I^+(f)| \le \max(I^+(f^+), I^+(f^-)) \le \|I\| \max(\|f^+\|_u, \|f^-\|_u) = \|I\|\|f\|_u,$$

so that $\|I^+\| \le \|I\|$. Finally, let $I^- = I^+ - I$, then $I^- \in C_0(X,\mathbb{R})^*$, and I^+, I^- are positive. □

Definition 7.5

A ***signed Radon measure*** *ν is a signed Borel measure such that ν^+ and ν^- are Radon, and a* ***complex Radon measure*** *is a complex Borel measure whose real and imaginary parts are Radon measures. We denote the space of complex Radon measures on X by $\mathcal{M}(X)$.* ♣

Proposition 7.3

If μ is a complex Borel measure, then μ is Radon iff $|\mu|$ is Radon. Moreover, $\mathcal{M}(X)$ is a vector space and $\|\mu\| = |\mu|(X)$ defines a norm on $\mathcal{M}(X)$. ♠

Proof Suppose ν is a finite positive Borel measure, then ν is Radon if and only if for every Borel set E and every $\varepsilon > 0$ there exist a compact K and an open U such that $K \subset E \subset U$ and $\nu(U \setminus K) < \varepsilon$. Now we prove the proposition. Let μ be a complex Radon measure, then $\mu = \mu_1 + i\mu_2$, where μ_1, μ_2 are signed Radon measures, then $\mu_1 = \mu_1^+ - \mu_1^-, \mu_2 = \mu_2^+ - \mu_2^-$, where $\mu_1^\pm, \mu_2^\pm$ are finite positive Radon measures. Then applying the the observation we just made gives

$$\begin{aligned}
\mu_1^+(U_1^+ \setminus K_1^+) &< \frac{\varepsilon}{4}, \\
\mu_1^-(U_1^- \setminus K_1^-) &< \frac{\varepsilon}{4}, \\
\mu_2^+(U_2^+ \setminus K_2^+) &< \frac{\varepsilon}{4}, \\
\mu_2^-(U_2^- \setminus K_2^-) &< \frac{\varepsilon}{4}.
\end{aligned}$$

Let $U = U_1^+ \cap U_1^- \cap U_2^+ \cap U_2^-$ and $K = K_1^+ \cup K_1^- \cup K_2^+ \cup K_2^-$, then

$$|\mu|(U \setminus K) \le \mu_1^+(U_1^+ \setminus K_1^+) + \mu_1^-(U_1^- \setminus K_1^-) + \mu_2^+(U_2^+ \setminus K_2^+) + \mu_2^-(U_2^- \setminus K_2^-) < \varepsilon.$$

Conversely, if $\mu = \mu_1^+ - \mu_1^- + i(\mu_2^+ - \mu_2^-)$ and $|\mu|(U \setminus K) < \varepsilon$, then

$$\begin{aligned}
\mu_1^+(U \setminus K) &< \frac{\varepsilon}{4}, \\
\mu_1^-(U \setminus K) &< \frac{\varepsilon}{4}, \\
\mu_2^+(U \setminus K) &< \frac{\varepsilon}{4}, \\
\mu_2^-(U \setminus K) &< \frac{\varepsilon}{4}.
\end{aligned}$$

hence $\mu_1^\pm, \mu_2^\pm$ are complex Radon measures. Since complex Radon measures are still complex measures, $\mathcal{M}(X)$ is a normed vector space (in fact it is a Banach space). □

At this point we review the Radon-Nikodym theorem and its consequences. If ν is a complex measure and μ is a σ-finite measure on X, and $\nu \ll \mu$, then there exists $f \in L^1(\mu)$ such that $d\nu = f d\mu$, and we denote this f by $d\nu/d\mu$. By Example 5.5 we have $|\nu|(E) = \int_E |f|\, d\mu$ for every measurable set E, thus we can write $d|\nu| = |f| d\mu$. However, if we only have a complex ν on X, it is possible to find a positive measure μ and $f \in L^1(\mu)$ such that $d\nu = f\, d\mu$? Indeed, write $\nu = \nu_r + i\nu_i$ and take $\mu = |\nu_r| + |\nu_i|$, then $\nu \ll \mu$ and by Radon-Nikodym theorem there exists such an $f \in L^1(\mu)$. Now we use this measure μ to connect ν and $|\nu|$. Clearly $\nu \ll |\nu|$, so there exists $g \in L^1(|\nu|)$ such that $g = d\nu/d|\nu|$, and then we have

$$f d\mu = d\nu = g d|\nu| = g|f|d\mu,$$

so $g|f| = f$ μ-a.e., and hence $|\nu|$-a.e. (since $\mu = |\nu_r| + |\nu_i|$). But clearly $|f| > 0$ $|\nu|$-a.e., (otherwise $|\nu| = 0$ is trivial), then $|g| = 1$ $|\nu|$-a.e. This gives the following lemma, which will be used to prove the Riesz representation theorem.

Lemma 7.3

*If ν is a complex measure on X, then there is a measurable function g such that $|g(x)| = 1$ for all $x \in X$ and $d\nu = g\ d|\nu|$. This is called the **polar decomposition** of ν.* ♡

Proof See the above argument. □

Definition 7.6

Define integration with respect to a complex measure μ by

$$\int f\ d\mu = \int fg\ d|\mu|,$$

where g is a function in the polar decomposition of μ. ♣

Theorem 7.10 (the Riesz representation theorem: $C_0(X)^* \simeq \mathcal{M}(X)$)

Let X be an LCH space, and for $\mu \in \mathcal{M}(X)$ and $f \in C_0(X)$ let $I_\mu(f) = \int f\ d\mu$. Then $\mathcal{M}(X)$ is isometrically isomorphism to $C_0(X)^*$, where the isomorphism is given by $\mu \mapsto I_\mu$. ♡

Proof By the extension of continuous linear functionals from $C_c(X)$ and $C_0(X)$ and Lemma 7.2, $I \in C_0(X)^*$ is of the form I_μ. On the other hand, if $\mu \in \mathcal{M}(X)$, then

$$\left|\int f\ d\mu\right| \le \int |f|\ d|\mu| \le \|f\|_u \|\mu\|,$$

so I_μ is a bounded linear functional on $C_0(X)^*$ and $\|I_\mu\| \le \|\mu\|$. Since $\mu \ll |\mu|$, by Radon-Nikodym theorem there is an $h \in L^1(|\mu|)$ such that $d\mu = h\ d|\mu|$. By the above lemma, then $|h| = 1$ $|\mu|$-a.e. Don't forget that the total variation norm $|\mu|(X) < \infty$! Hence h and $\overline{h}$ vanish outside a set of finite measure, so by Lusin's theorem, for any $\varepsilon > 0$ there exists $f \in C_c(X)$ such that $\|f\|_u = 1$ and $f = \overline{h}$ except on a set E with $|\mu|(E) < \varepsilon/2$. Then

$$\begin{aligned}
\|\mu\| &= \int_X d|\mu| = \int |h|^2\ d\mu = \int h\overline{h}\ d\mu \\
&= \int \overline{h}\ d\mu \le \left|\int f\ d\mu\right| + \left|\int (f - \overline{h})\ d\mu\right| \\
&\le \left|\int f\ d\mu\right| + \int |f - \overline{h}|\ d|\mu| \\
&\le \left|\int f\ d\mu\right| + 2|\mu|(E) \\
&< \left|\int f\ d\mu\right| + \varepsilon \\
&\le \|I_\mu\| + \varepsilon.
\end{aligned}$$

It follows that $\|\mu\| \leq \|I_\mu\|$, completing the proof. $\square$

Chapter 8 Baire Category Theorem

8.1 Baire Category Theorem

reference: Functional Analysis, Stein, Chapter 4

Let X be a metric space with a metric d. Suppose E is a subset of X.

Definition 8.1

We say that E is dense in X if $\overline{E} = X$. Also, E is nowhere dense if $(\overline{E})^\circ = \varnothing$. ♣

Definition 8.2

- *A set $E \subset X$ is of the first category in X if E is a countable union of nowhere dense sets in X. A set of the first category is sometimes said to be "meager".*
- *A set E that is not of the first category in X is referred to as being of the second category in X.*
- *A set $E \subset X$ is defined to be generic if E^c is of the first category.* ♣

Proposition 8.1

E is closed and nowhere dense iff E^c is open and dense. ♠

Proof We show $(E^\circ)^c = \overline{E^c}$. Suppose $x \notin E^\circ$, then for every $r > 0$, $B(x,r)$ is not contained in E, so $B(x,r) \cap E^c \neq \varnothing$, hence $x \in \overline{E^c}$. Conversely, let $x \in \overline{E^c}$, then for every $r > 0$, $B(x,r) \cap E^c \neq \varnothing$, so $B(x,r)$ cannot be contained in E, then $x \notin E^\circ$.

Theorem 8.1 (Baire's Category Theorem)

Every complete metric space X is of the second category in itself: X cannot be written as the countable union of nowhere dense sets. ♡

Proof Suppose to the contrary that $X = \cup_{n=1}^\infty F_n$, where F_n is closed and nowhere dense. F_1 is closed and nowhere dense $\implies$ F_1^c is open and dense, hence F_1^c contains a ball $B_1 = B(r)$. F_2^c is open and dense, so $F_2^c \cap B_1$ is open and nonempty, hence there is a ball $B_2 = B(r/2) \subset F_2^c \cap B_1$. Choose any $x_n \in B_n$, then $\{x_n\}$ is Cauchy, so $x_n \to x \in \cap_{n=1}^\infty \overline{B_n}$, hence $x \notin \cup_{n=1}^\infty F_n = X$.

Corollary 8.1

In a complete metric space, a generic set is dense. ♡

Proof Suppose $E \subset X$ is generic but not dense. Then there is a closed ball $\overline{B} \subset E^c$. Since E is generic, $E^c = \cup_{n=1}^\infty F_n$, where F_n is nowhere dense, hence

$$\overline{B} = \bigcup_{n=1}^{\infty} (F_n \cap \overline{B}).$$

Now the complete metric space $\overline{B}$ is a countable union of nowhere dense sets, contradicting Baire's category theorem.

8.2 Applications of Baire's Category Theorem

8.2.1 Continuity of the limit of a sequence of continuous functions

reference: Functional Analysis, Stein.

Definition 8.3 (oscillation)

Define the oscillation of the function f at a point x by

$$\operatorname{osc}(f)(x) = \lim_{r\to 0} \omega(f)(r, x),$$

where

$$\omega(f)(r, x) = \sup_{y,z\in B_r(x)} |f(y) - f(z)|.$$

♣

Properties

1. $\operatorname{osc}(f)(x) = 0$ *if and only if* f *is continuous at* x.
2. *The set* $E_\varepsilon = \{x \in X : \operatorname{osc}(f)(x) < \varepsilon\}$ *is open. That is,* $\operatorname{osc}(f)$ *is upper semi-continuous.*

Proof For (1), write down the $\varepsilon - \delta$ definition. For (2), If $x \in E_\varepsilon$, then $B_{r/2}(x) \subset E_\varepsilon$.

Lemma 8.1

Suppose $\{f_n\}$ *is a sequence of continuous functions on a complete metric space* X*, and* $f_n(x) \to f(x)$ *for each* x*. Then, given an open ball* $B \subset X$ *and* $\varepsilon > 0$*, there exists an open ball* $B_0 \subset B$ *and an integer* $m \geq 1$ *so that* $|f_m(x) - f(x)| \leq \varepsilon$ *for all* $x \in B_0$.

♡

Proof Let $Y \subset B$ be a closed ball. Define

$$E_l = Y \cap \{x \in X : \sup_{j,k\geq l} |f_j(x) - f_k(x)| \leq \varepsilon\}.$$

Since for each $x \in X$, $\{f_n(x)\}$ is Cauchy, we have

$$X = \bigcup_{l=1}^{\infty} \{x \in X : \sup_{j,k\geq l} |f_j(x) - f_k(x)| \leq \varepsilon\},$$

hence

$$Y = \bigcup_{l=1}^{\infty} E_l.$$

But Y is complete, so it cannot be written as a countable union of nowhere dense sets, hence some E_m is not nowhere dense. From the continuity we see that each E_l is closed, so $E_m^\circ \neq \varnothing$. Take an open ball $B_0 \subset E_m$, then

$$\sup_{j,k\geq m} |f_j(x) - f_k(x)| \leq \varepsilon \ \forall x \in B_0.$$

Let $k \to \infty$, we have $|f_m(x) - f(x)| \leq \varepsilon \ \forall x \in B_0$.

Theorem 8.2

Suppose that $\{f_n\}$ is a sequence of continuous complex-valued functions on a complete metric space X and

$$\lim_{n\to\infty} f_n(x) = f(x)\ \forall x \in X.$$

Then, the set of points of continuity of f is a generic set in X. In other words, the set of points of discontinuity of f is of the first category. ♡

Proof Let $F_n = \{x \in X : \operatorname{osc}(f)(x) \geq 1/n\}$, then the set of discontinuities of f can be written as

$$\mathcal{D} = \bigcup_{n=1}^{\infty} F_n,$$

thus we only need to show that each F_n is nowhere dense. Suppose F_n has a nonempty interior. Choose an open ball $B \subset F_n$. Set $\varepsilon = 1/4n$ in the lemma, $\exists$ an open ball $B_0 \subset B$ and $m \geq 1$ so that

$$|f_m(x) - f(x)| \leq 1/4n\ \forall x \in B_0.$$

Since f_m is contiuous, $\exists B' \subset B_0$ so that

$$|f_m(y) - f_m(z)| \leq 1/4n\ \forall y, z \in B'.$$

Then,

$$|f(y) - f(z)| \leq |f(y) - f_m(y)| + |f_m(y) - f_m(z)| + |f_m(z) - f_m(z)| < \frac{1}{n}\ \forall y, z \in B'.$$

If x' denotes the center of B', we have $\operatorname{osc}(f)(x') < 1/n$, so $x' \notin F_n$, a contradiction.

8.2.2 Nowhere differentiable continuous functions

Theorem 8.3

The set of functions in $C([0,1])$ that are nowhere differentiable is generic. ♡

Let $\mathcal{D}$ denote the set in the theorem, then $\mathcal{D}^c$ is the set of $C([0,1]$-functions that are differentiable at least at one point. If we show $\mathcal{D}^c$ is of the first category, then $\mathcal{D} = (\mathcal{D}^c)^c$ is generic. ·

Let E_N denote the set of all continuous functions so that there exists $0 \leq x^* \leq 1$ with

$$|f(x) - f(x^*)| \leq N|x - x^*|\ \forall x \in [0,1].$$

Now if $f \in \mathcal{D}$, then f is differentiable at some point $a \in [0,1]$. Taylor expanding f at a gives

$$f(x) = f(a) + f'(a)(x-a) + o(x-a)\ \forall x \in [0,1],$$

thus

$$|f(x) - f(a)| \leq |f'(a) + C||x - a|,$$

so f is in some E_N. Now we have the inclusion

$$\mathcal{D} \subset \bigcup_{N=1}^{\infty} E_N.$$

It suffices to show that each E_N is nowhere dense. This will be achieved by showing successively:

1. E_N is a closed set.
2. $(E_N)^\circ = \varnothing$.

Thus $\cup E_N$ is of the first category, so is $\mathcal{D}$.

Proof [property (1)] Suppose that $\{f_n\}$ is a sequence of functions in E_N with $\|f_n - f\| \to 0$. We must show that $f \in E_N$. Since $f_n \in E_N$, there exists $x_n^* \in [0,1]$ with

$$|f_n(x) - f_n(x_n^*)| \le N|x - x_n^*| \; \forall x \in [0,1].$$

Now we have a bounded sequence $\{x_n^* : n \in \mathbb{N}\}$, so we may choose a subsequence $\{x_{n_k}^*\}$ that converges to some $x^* \in [0,1]$. Then,

$$|f(x) - f(x^*)| \le |f(x) - f_{n_k}(x)| + |f_{n_k}(x) - f_{n_k}(x^*)| + |f_{n_k}(x^*) - f(x^*)|.$$

Now

$$|f_{n_k}(x) - f_{n_k}(x^*)| \le |f_{n_k}(x) - f_{n_k}(x_{n_k}^*)| + |f_{n_k}(x_{n_k}^*) - f_{n_k}(x^*)|.$$

Therefore, $f_{n_k} \in E_N$ implies

$$|f_{n_k}(x) - f_{n_k}(x^*)| \le N|x - x_{n_k}^*| + N|x_{n_k}^* - x^*|.$$

Putting all these estimates together, we obtain

$$|f(x) - f(x^*)| \le \varepsilon + N|x - x_{n_k}^*| + N|x_{n_k}^* - x^*|$$

for all large k. Letting $k \to \infty$ we get

$$|f(x) - f(x^*)| \le \varepsilon + N|x - x^*|.$$

Since ε is arbitrary, we conclude that $f \in E_N$.

Definition 8.4 (zig-zag functions)

Let $\mathcal{P}$ denote the subspace of $C([0,1])$ that consists of all continuous piecewise-linear functions. Also, for each $M > 0$, let $\mathcal{P}_M \subset \mathcal{P}$ denote the set of all continuous piecewise-linear functions, each of whose line segments have slopes either $\ge M$ or $\le -M$. ♣

Lemma 8.2

For every $M > 0$, $\mathcal{P}_M$ is dense in $C([0,1])$. ♡

Proof Let $\varepsilon > 0$. f is uniformly continuous on $[0,1] \implies$ there exists $\delta > 0$ so that $|f(x) - f(y)| < \varepsilon$ if $|x - y| < \delta$. Choose n so large that $1/n < \delta$, and define g as a linear function on each $[k/n, (k+1)/n]$ for $k = 0, \cdots, n-1$ with $g(k/n) = f(k/n)$ and $g((k+1)/n) = f((k+1)/n)$. Then $\|f - g\| = \sup_{0 \le x \le 1} |f(x) - g(x)| \le \varepsilon$.
It now suffices to approximate g on $[0,1]$ by zig-zag functions in $\mathcal{P}_M$. If $g(x) = ax + b$ for $0 \le x \le 1/n$, consider two segments

$$\varphi_\varepsilon(x) = g(x) + \varepsilon \quad \text{and} \quad \psi_\varepsilon(x) = g(x) - \varepsilon.$$

Then, see figure. We obtain $h \in \mathcal{P}_M$ so that

$$\psi_\varepsilon(x) \le h(x) \le \varphi_\varepsilon(x) \; \forall 0 \le x \le 1/n,$$

and therefore $|h(x) - g(x)| \le \varepsilon$ in $[0, 1/n]$. Then we begin at $h(1/n)$ and repeat this argument on $[1/n, 2/n]$. Continuing in this fashion, we obtain a function $h \in \mathcal{P}_M$ with

$\|h-g\| \leq \varepsilon$. Hence $\|f-h\| \leq 2\varepsilon$.

Proof [property (2)] Given any $f \in E_N$ and $\varepsilon > 0$, we first choose a fixed $M > N$. Then, there exists $h \in \mathcal{P}_M$ so that $\|f-h\| < \varepsilon$, and $h \notin E_N$ since $M > N$. Therefore, no open ball around f is entirely contained in E_N.

8.3 Fundamental Theorems

8.3.1 The Uniform Boundedness Principle

Theorem 8.4 (Uniform Boundedness Principle)

Let X be a Banach space, Y a normed vector space. Suppose that $\mathcal{F}$ is a collection of continuous linear operators from X to Y. If

$$\sup_{T\in\mathcal{F}} \|T(x)\|_Y < \infty \quad \text{for each } x \in X,$$

then

$$\sup_{T\in\mathcal{F}} \|T\| < \infty.$$

♡

Proof Let $X_n = \{x \in X : \sup_{T\in\mathcal{F}} \|Tx\|_Y \leq n\}$, then

$$X = \bigcup_{n\in\mathbb{N}} X_n$$

and X_n is closed (write X_n as an intersection and use continuity of T). By BCT, some X_m has nonempty interior. That is, $\exists x_0 \in X_m$ and $\varepsilon > 0$ such that

$$\overline{B_\varepsilon(x_0)} = \{x \in X : \|x - x_0\| \leq \varepsilon\} \subset X_m.$$

So if $\|Tx\|_Y \leq m$ for all $T \in \mathcal{F}$ and $\|x - x_0\| \leq \varepsilon$, then $x \in \overline{B_\varepsilon(x_0)}$.
Now let $u \in X$ with $\|u\| \leq 1$ and $T \in \mathcal{F}$, then

$$\begin{aligned}\|Tu\|_Y &= \varepsilon^{-1}\|T(x_0+\varepsilon u) - Tx_0\|_Y \\ &\leq (\|T(x_0+\varepsilon u\|_Y + \|Tx_0\|_Y) \\ &\leq \varepsilon^{-1}(m+m).\end{aligned}$$

Taking supremum over $\|u\| \leq 1$ and T over $\mathcal{F}$ gives

$$\sup_{T\in\mathcal{F}} \|T\| \leq 2\varepsilon^{-1}m < \infty.$$

Theorem 8.5 (Uniform Boundedness Principle, linear functional)

Suppose that X is a Banach space and $\mathcal{L}$ a collection of continuous linear functionals on X. If $\sup_{\ell\in\mathcal{L}} |\ell(f)| < \infty$ for all f in some set of the second category, then

$$\sup_{\ell\in\mathcal{L}} \|\ell\| < \infty.$$

♡

Proof Let E be of the second category. Let

$$E_M = \{f \in X : \sup_{\ell\in\mathcal{L}} |\ell(f)| \leq M\},$$

then $E = \bigcup_{M=1}^{\infty} E_M$, and each E_M is closed. Therefore, some E_{M_0} must have nonempty interior. There exists $f_0 \in X$ and $r > 0$ such that $B_r(f_0) \subset E_{M_0}$. For all $\ell \in \mathcal{L}$ we have

$|\ell(f_0)| \le M_0$ if $\|f - f_0\| < r$. Now for all $\|g\| < r$ and all $\ell \in \mathcal{L}$ we have

$$\|\ell(g)\| \le \|\ell(g + f_0)\| + \|\ell(-f_0)\| \le 2M_0,$$

thus $\sup_{\ell \in \mathcal{L}} \|\ell\| < \infty$.

8.3.1.1 a really simple proof

https://arxiv.org/pdf/1005.1585.pdf

8.3.1.2 divergence of Fourier series

Theorem 8.6

Let $X = C([-\pi, \pi])$ with the sup norm.

1. *Given any point $x_0 \in [-\pi, \pi]$, there is a continuous function whose Fourier series diverges at x_0.*

2. *In fact, the set of continuous functions whose Fourier series diverge on a dense set in $[-\pi, \pi]$ is generic in X.*

♡

WLOG let $x_0 = 0$. For each N define a linear functional $\ell_N : X \to \mathbb{R}$ by

$$\ell_N(f) = S_N(f)(0),$$

we show

1. ℓ_N is continuous;
2. $\|\ell_N\| = L_N = \frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(x)| dx$, where D_N is the nth Dirichlet kernel;
3. $L_N \ge C \log N$.

Lemma 8.3

ℓ_N is continuous and $\|\ell_N\| = L_N$.

♡

Proof

$$\begin{aligned}
|\ell_N(f)| &= |(f * D_N)(0)| \\
&\le \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(y) D_N(0 - y)| dy \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(y) D_N(y)| dy \\
&\le \sup_{-\pi \le y \le \pi} |f(y)| \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |D_N(y)| dy \right) \\
&= L_N \|f\|,
\end{aligned}$$

so ℓ_N is bounded, hence is continuous. We have shown $\|\ell_N\| \le L_N$. Next, we find a sequence $\{f_k\}$ with $\|f_k\| \le 1$ and $\ell(f_k) \to L_N$ as $k \to \infty$.
(idea: take out the absolute value in L_N.) Define

$$g(y) = \begin{cases} 1, & \text{if } D_N(y) > 0 \\ -1, & \text{if } D_N(y) < 0 \end{cases},$$

then g is measurable and $\int |g| \leq 4\pi$.

$$\begin{aligned}\int_{-\pi}^{\pi} |D_N(y)|dy &= \int_{\{D_N(y)>0\}} D_N(y)dy + \int_{\{D_N(y)<0\}} -D_N(y)dy \\ &= \int_{-\pi}^{\pi} g(y)D_N(y)dy.\end{aligned}$$

Now choose a sequence $\{f_k\} \subset C_c([-\pi,\pi])$ with $\|f_k - g\|_{L^1} \to 0$ as $k \to \infty$, then

$$\begin{aligned}|\ell_N(f_k) - L_N| &= \int_{-\pi}^{\pi} (f_k(x) - g(x))D_N(x)dx \\ &\leq \int_{-\pi}^{\pi} |f_k(x) - g(x)||D_N(x)|dx \\ &\leq \sup_{-\pi \leq x \leq \pi} |D_N(x)|\|f_k - g\|_{L^1} \to 0 \quad (k \to \infty).\end{aligned}$$

This shows that $\|\ell_N\| = L_N$ for each N.

Lemma 8.4

$L_N \geq C \log N$.

Proof Recall that $D_N(x) = \dfrac{\sin(N+1/2)x}{\sin(x/2)}$. In the following computation, the constant C varies from steps.

$$\begin{aligned}\int_{0}^{\pi} |D_N(y)|dy &= \int_{0}^{\pi} \frac{|\sin(N+1/2)y|}{|\sin(y/2)|}dy \\ &\geq C\int_{0}^{\pi} \frac{|\sin(N+1/2)y|}{|y|}dy \\ &= C\int_{0}^{(N+1/2)\pi} \frac{|\sin x|}{|x|}dx \\ &\geq C\sum_{k=0}^{N-1}\int_{k}^{k+1} \frac{|\sin x|}{|x|}dx \\ &\geq C\sum_{k=0}^{N-1}\int_{k}^{k+1} \frac{|\sin x|}{k+1}dx \\ &= C\sum_{k=0}^{N-1} \frac{1}{k+1} \\ &\geq C\int_{1}^{N} \frac{1}{x}dx \\ &= C\log N.\end{aligned}$$

Now we prove the theorem. Suppose every $f \in C([-\pi,\pi])$ has its Fourier series converging at 0, then $S_N(f)(0) = \ell_N(f)$ would be finite for each N. By uniform bounded principle, $\|\ell_N\| < \infty$ for all N, a contradiction.

Let $F_n = \{f \in C([-\pi,\pi])$: Fourier series of f diverges at $x_n\}$. Suppose that $\{f : \sup_N |\ell_N(f)| < \infty\}$ were of the second category, then $\sup_N \|\ell_N\| < \infty$, a contradiction. Hence $\{f : \sup_N |\ell_N(f)| < \infty\}$ is of the first category, and as a subset, $\{f \in C([-\pi,\pi])$: Fourier series of f converges at $x_n\}$ is of the first category. That is, F_n^c is of the first category, hence F_n is generic. In particular, take a dense set to be $\mathbb{Q} = \{x_0, x_1, \cdots\}$, then $\bigcap_{n=0}^{\infty} F_n$ is the set of continuous f whose Fourier series diverges at $x_0, x_1, \cdots$. Since

$\bigcup_{n=0}^{\infty} F_n^c$ is of the first category, $\bigcap_{n=0}^{\infty} F_n$ is generic.

8.3.2 The Open Mapping Theorem

Let X and Y be Banach spaces with norms $\|\|_X, \|\|_Y$, and $T : X \to Y$ a mapping.

Definition 8.5

A mapping T that maps open sets to open sets is called an open mapping.
A mapping T is surjective if $T(X) = Y$, and injective if $T(x) = T(y)$ implies $x = y$. ♣

Recall that a mapping T has an inverse if and only if T is bijective.

Proposition 8.2

If T is linear and bijective, then the inverse T^{-1} is also linear. ♠

Proof Let $y_1 = T(x_1), y_2 = T(x_2)$, then $x_1 = T^{-1}(y_1), x_2 = T^{-1}(y_2)$. Also, $y_1 + y_2 = T(x_1 + x_2)$, so $x_1 + x_2 = T^{-1}(y_1 + y_2) = T^{-1}(y_1) + T^{-1}(y_2)$.

Proposition 8.3 (linear operations on sets)

We list some useful expressions. Let $T : X \to Y$ and $B_X(r) := B_X(0, r), B_Y(r) := B_Y(0, r)$.

1. $T(B_X(r)) = rT(B_X(1))$;
2. $T(B_X(a, r)) = T(a) + rT(B_X(1))$. ♠

Proof

1. Note that $Tx = rT(x/r)$.
2. Since $B_X(a, r) = B_X(r) + a$, every element in $B_X(a, r)$ can be written as $x + a$, where $x \in B_X(r)$. Then,
$$T(x + a) = T(a) + rT\left(\frac{x}{r}\right).$$

Theorem 8.7 (Open Mapping Theorem)

Suppose X and Y are Banach spaces, and $T : X \to Y$ is a continuous linear operator. If T is surjective, then T is an open mapping. ♡

Proof It suffices to show that $T(B_X(1))$ contains an open ball centered at the origin.

Corollary 8.2

If X, Y are Banach spaces, and $T : X \to Y$ is a continuous bijective linear transformation, then the inverse $T^{-1} : Y \to X$ is also continuous. Hence there are constants $c, C > 0$ with
$$c\|f\|_X \le \|Tf\|_Y \le C\|f\|_X.$$
♡

8.3.2.1 Decay of Fourier coefficients of L^1-functions

Recall Riemann-Lebesgue lemma: if $f \in L^1([-\pi, \pi])$, then $\lim_{|n|\to\infty} |\widehat{f}(n)| = 0$.
Question: given any sequence of complex numbers $\{a_n : n \in \mathbb{Z}\}$ that vanishes at infinity,

does there exist $f \in L^1([-\pi, \pi])$ with $\widehat{f}(n) = a_n$ for all n? Reformulate the question: let $X_1 = L^1([-\pi, \pi])$ and X_2 be the sequence space with $|a_n| \to 0$ as $|n| \to \infty$. X_2 is equipped with the sup norm $\|\{a_n\}\|_\infty = \sup_{n\in\mathbb{Z}} |a_n|$. Then, we ask wheter the map $T : X_1 \to X_2$ defined by

$$T(f) = \{\widehat{f}(n)\}_{n\in\mathbb{Z}}$$

is surjective.

Theorem 8.8

The map $T : X_1 \to X_2$ given by $T(f) = \{\widehat{f}(n)\}$ is linear, continuous, and injective, but not surjective. ♡

Proof T is clearly linear, and

$$\|Tf\|_\infty = \sup_n |\widehat{f}(n)| = \sup_n \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)e^{inx} dx \right| \le \|f\|_{L^1}.$$

If $T(f) = 0$, then $\widehat{f}(n) = 0$ for all n, by the uniqueness of Fourier coefficient, $f = 0$ in L^1. If T were surjective, then there is a constant $c > 0$ that satisfies

$$c\|f\|_{L^1} \le \|Tf\|_\infty \quad \forall f \in X_1.$$

If we set $f = D_N = \sum_{|n|\le N} e^{inx}$, then $\|D_N\|_{L^1} = L_N \to \infty$ as $N \to \infty$, a contradiction.

8.3.3 The Closed Graph Theorem

Definition 8.6

The graph of T is defined to be as a subset of $X \times Y$ by

$$G_T = \{(x, Tx) : x \in X\}.$$

The linear map T is closed if its graph is a closed subset in $X \times Y$. In other words, T is closed if whenever $\{x_n, Tx_n\}$ converges to x and y, then $y = Tx$. ♣

Theorem 8.9 (closed graph theorem)

Suppose X, Y are Banach spaces. if $T : X \to Y$ is a closed linear map, then T is continuous. ♡

Proof $X \times Y$ is the Banach space with the norm $\|(x, y)\| = \|x\| + \|y\|$. Since the graph of T is a closed subspace, G_T is a Banach space. Consider projections $P_X : G_T \to X$, $P_Y : G_T \to Y$ defined by

$$P_X(x, Tx) = x, \quad P_Y(x, Tx) = Tx.$$

Then P_X and P_Y are continuous and linear. Since P_X is bijective, P_X^{-1} is continuous. Since $T = P_Y \circ P_X^{-1}$, T is continuous.

8.3.3.1 Grothendieck's theorem on closed subspaces of L^p

Theorem 8.10

Let $(X, \mathcal{F}, \mu)$ be a finite measure space, that is, $\mu(X) < \infty$. Suppose that

1. *E is a closed subspace of $L^p(X, \mu)$ for some $1 \le p < \infty$, and*
2. *E is contained in $L^\infty(X, \mu)$.*

Then E is finite dimensional.

Let $I : E \to L^\infty(X, \mu)$ be the identity map $I(f) = f$. Then I is clearly linear. To see I is closed, let $(f_n, If_n) \to (f, g)$, that is, $f_n \to f$ in L^p and $f_n \to g$ in L^∞. Then there is a subsequence $f_{n_k} \to f$ a.e., and $f \to g$ a.e. Hence $g = f = If$, so I is a closed map. Then there is an $M > 0$ so that

$$\|f\|_\infty \le M\|f\|_p \quad \forall f \in E.$$

Lemma 8.5

Under the assumptions of the theorem, there exists $A > 0$ so that

$$\|f\|_\infty \le M\|f\|_2 \quad \forall f \in E.$$

Proof idea: connect L^∞ and L^2 with L^p by constructing conjugate exponents If $1 \le p \le 2$, let $r = 2/p, r^* = 1 - 2/p = 2/(2-p)$. By Hölder's inequality,

$$\|f^p\|_1 \le \|f^p\|_r \|1\|_{r^*},$$

which is

$$\int |f|^p \le \left(\int |f|^2\right)^{p/2} \left(\int 1\right)^{\frac{2-p}{2}},$$

so $\|f\|_p \le B\|f\|_2$ for some $B > 0$.
If $2 < p < \infty$, we note that

$$|f(x)|^p = |f(x)|^{p-2}|f(x)|^2 \le \|f\|_\infty^{p-2}|f(x)|^2.$$

Integrating this inequality gives

$$\|f\|_p^p \le \|f\|_\infty^{p-2}\|f\|_2^2,$$

so

$$\|f\|_\infty^p \le M^p\|f\|_\infty^{p-2}\|f\|_2^2,$$

then

$$\|f\|_\infty^2 \le M^p\|f\|_2^2,$$

which gives $\|f\|_\infty \le A\|f\|_2$. □

We now turn to the proof of the theorem. Suppose $f_1, \cdots, f_n$ is an orthonormal set in L^2 of functions in E, and let $\mathbb{B}$ denote the unit ball in $\mathbb{C}^n$:

$$= \left\{ \zeta = (\zeta_1, \cdots, \zeta_n) \in \mathbb{C}^n : \sum_{j=1}^n |\zeta_j|^2 \le 1 \right\}.$$

Define $T :\to \text{span}\,\{f_1, \cdots, f_n\} \subset E \cap L^2$ by $T\zeta = \sum_{j=1}^n \zeta_j f_j$. Then, for each fixed $\zeta \in$,

$$\|T\zeta\|_{L^2}^2 = \sum_{j=1}^n |\zeta_j|^2 \int |f_j(x)|^2 dx = \sum_{j=1}^n |\zeta_j|^2 \|f_j\|_{L^2}^2 = \sum_{j=1}^n |\zeta_j|^2 \le 1.$$

From the above we also see that T is a continuous linear operator. By the lemma, $\|T\zeta\|_{L^\infty} \leq A\|T\zeta\|_{L^2} \leq A$. That is, $|(T\zeta)(x)| \leq A$ for a.e. $x \in X$. Hence for each ζ, there is a measurable set X_ζ with $\mu(X_\zeta^c) = 0$ such that $|(T\zeta)(x)| \leq A$ for all $x \in X_\zeta$.
Take a countable dense subset, say, $\mathbb{Q}^{2n} \cap = \{\zeta_j : j \in \mathbb{N}\}$, then $|T\zeta_j(x)| \leq A$ for all $x \in X_{\zeta_j}$, where X_{ζ_j} is a full set in X. Let $X' = \cap_{j=1}^\infty X_{\zeta_j}$, then for each $j \in \mathbb{N}$:

$$|T\zeta_j(x)| \leq A \quad \forall x \in X'.$$

Now let $\zeta \in$ be arbitrary. Then, there is a sequence $\{\zeta_{j_k}\} \subset \mathbb{Q}^{2n} \cap$ with $\zeta_{j_k} \to \zeta$ as $k \to \infty$. Since T is continuous, $T\zeta_{j_k} \to T\zeta$ in L^2. By the lemma, $T\zeta_{j_k} \to T\zeta$ in L^∞.

Chapter 8 Exercise

1. (Stein) Let $\{x_j\}$ denote an enumeration of the rational numbers in $\mathbb{R}$, and consider the sets

$$U_n = \bigcup_{j=1}^\infty \left(x_j - \frac{1}{n2^j}, x_j + \frac{1}{n2^j}\right),$$

and

$$U = \bigcap_{n=1}^\infty U_n.$$

Show that U is generic but has Lebesgue measure 0.

Proof $m(U) = 0$ is easy to see.

2. (Stein) Suppose F is a closed subset and $\mathcal{O}$ an open subset of a complete metric space.

(a). F is of the first category if and only if F has empty interior.

Proof Let $F = \cup_{n=1}^\infty F_n$, where F_n is nowhere dense. Suppose that F° were not empty, then some closed ball $\overline{B} \subset F$, so

$$\overline{B} = \bigcup_{n=1}^\infty (F_n \cap \overline{B}).$$

From $\overline{F_n \cap \overline{B}} = \overline{F_n} \cap \overline{B}$ and $(\overline{F_n} \cap \overline{B})^\circ \subset (\overline{F_n})^\circ \cap (\overline{B})^\circ = \varnothing$ we see that $\overline{B}$ is of the first category, a contradiction since $\overline{B}$ is complete. Conversely, $F = \cup_{n=1}^\infty F_n$ with $F_1 = F$ and $F_n = \varnothing$ for all $n > 1$.

(b). $\mathcal{O}$ is of the first category if and only if $\mathcal{O}$ is empty.

Proof Let $\mathcal{O}$ be of the first category and suppose $\mathcal{O}$ is not empty, then $\mathcal{O}$ contains a closed ball $\overline{B}$, thus $\overline{B}$ is of the first category, a contradiction.

(c). F is generic if and only if $F = X$; and $\mathcal{O}$ is generic if and only if $\mathcal{O}^c$ contains no interior.

Proof Suppose F is generic, then F^c is of the first category, so $F^c = \varnothing$, thus $F = X$.

$\mathcal{O}$ is generic $\implies$ $\mathcal{O}^c$ is of the first category $\implies$ $\mathcal{O}^c$ has empty interior.

3. (Stein) Show that the conclusion of the Baire category theorem continues to hold if X_0 is a metric space that arises as an open subset of a complete metric space X. [Hint: Apply the BCT to the closure of X_0 in X.]

Proof Note that $\overline{X_0}$ is complete. Suppose X_0 were of the first category:

$$X_0 = \bigcup_{n=1}^{\infty} A_n,$$

where A_n is nowhere dense. Then,

$$\overline{X_0} = \overline{\bigcup_{n=1}^{\infty} A_n} \supset \bigcup_{n=1}^{\infty} \overline{A_n}.$$

4. (Stein) Prove that every continuous function on $[0, 1]$ can be approximated uniformly by continuous nowhere differentiable functions. Do so by either:
   (a). using Theorem 1.5 (book).
   (b). using only the fact that a continuous nowhere differentiable function exists.

   Proof

   (a). Recall that in a complete metric space, a generic set is dense. Therefore, the set of continuous nowhere differentiable functions is dense in $C([0, 1])$.

   (b).

5. Let X be a complete metric space.
   (a). a dense G_δ is generic.
   (b). Hence a countable dense set is an F_σ, but not a G_δ.
   (c). If E is a generic set, then there exists $E_0 \subset E$ with E_0 a dense G_δ.

   Proof

   (a). Let $\bigcap_{n=1}^{\infty} \mathcal{O}_n$ be dense. If there is some $\mathcal{O}_m$ not dense in X, then $\overline{\bigcap_{n=1}^{\infty} \mathcal{O}_n} \subset \overline{\mathcal{O}_m} \neq X$, a contradiction. Therefore, each $\mathcal{O}_n$ is dense. Then, $\mathcal{O}_n^c$ is closed and nowhere dense. Now $\bigcup_{n=1}^{\infty} \mathcal{O}_n^c$ is of the first category, hence its complement is generic.

   (b). Since a metric space is Hausdorff, each point $\{x\}$ is closed. Then $A = \cup_{n=1}^{\infty} \{x_n\}$ is of course an F_σ. Suppose that A is a G_δ, then $A = \cap_{n=1}^{\infty} G_n$, so each G_n is open and dense, hence G_n^c is closed and nowhere dense. Then

   $$A^c \cap A = \bigcap_{n=1}^{\infty} (G_n \cap X \backslash \{x_n\}) = \varnothing,$$

   so

   $$\bigcup_{n=1}^{\infty} (G_n^c \cup \{x_n\}) = X,$$

   but X is complete, a contradiction.

   (c). Since E is generic, E is dense and $E^c = \cup_{n=1}^{\infty} A_n$, where A_n is nowhere dense, so A_n^c is dense. Hence $E = \cap_{n=1}^{\infty} A_n^c$ is dense. Let $B_n = A_n^c$, then $\overline{B_n^\circ} = \overline{B_n} = X$, so $E_0 = \cap_{n=1}^{\infty} B_n^\circ$ is the desired subset of E.

6. (Stein) The function

$$f(x) = \begin{cases} 0 & \text{if } x \text{ is irrational} \\ 1/q & \text{if } x = p/q \text{ is rational and expressed in lowest form} \end{cases}$$

is continuous precisely at the irrationals. In contrast to this, prove that there is no function on $\mathbb{R}$ that is continuous precisely at the rationals.

Proof Since f is continuous at x if and only if $\operatorname{osc}(f)(x) = 0$, we can write the set of continuities of f as

$$C(f) = \bigcap_{n=1}^{\infty} \{x \in \mathbb{R} : \operatorname{osc}(f)(x) < 1/n\},$$

which is a G_δ. Suppose there exists f_0 continuous precisely at the rationals, then $C(f_0) = \mathbb{Q}$ is a G_δ. But from the last exercise we know that a countable dense set is not a G_δ.

7. (Stein, 4-7) Let $E \subset [0,1]$ and let I be any closed non-trivial interval in $[0,1]$.
   (a). Suppose E is of the fisrt category in $[0,1]$. Show that for every I, the set $E \cap I$ is of the first category in I.
   (b). Suppose E is generic in $[0,1]$. Show that for every I, the set $E \cap I$ is generic in I.
   (c). Construct a set E in $[0,1]$ so that for all I, the set $E \cap I$ is neither of the first category nor generic in I.
8. (Stein) A Hamel basis for a vector space X is a collection $\mathcal{H}$ of vectors in X such that any $x \in X$ can be written as a unique finite linear combination of elements in $\mathcal{H}$. Prove that a Banach space cannot have a countable Hamel basis.

   Proof Suppose that X is a Banach space with a countable Hamel basis $\{e_k : k \in \mathbb{N}\}$. Let $x \in X$, then $x \in \operatorname{span}\{h_1, \cdots, h_n\}$ for some $n \in \mathbb{N}$. Hence

$$X = \bigcup_{n=1}^{\infty} \operatorname{span}\{h_1, \cdots, h_n\}.$$

   Since every finite-dimensional vector space is closed, it suffices to show that each $\operatorname{span}\{h_1, \cdots, h_n\}$ has empty interior. By linearity, we only need to show that each span contains no balls centered at the origin. Suppose that $\operatorname{span}\{h_1, \cdots, h_n\}$ contains $B(0,1)$, then $h_{n+1}/2\|h_{n+1}\| \in B(0,1)$ but is not in that span. Then X is of the first category, a contradiction.
9. (Fall 2016) Show that there is a continuous real-valued function on $[0,1]$ that is not monotone on any open interval $(a,b) \subset [0,1]$.

   Proof Suppose there were no such functions. Let $E_{a,b} = \{f \in C([0,1]) : f \text{ is monotone in } (a,b)\}$. Then

$$C([0,1]) = \bigcup_{a,b\in\mathbb{Q}, a<b} E_{a,b}.$$

   First we show that $E_{a,b}$ is closed. Let $\{f_k\} \subset E_{a,b}$ with $f_k \to f$ uniformly, then each f_k is either increasing or decreasing in (a,b). If all f_k's are increasing (or decreasing), then $f_k(x) \leq f_k(y)$ if $x \leq y (x \geq y$ resp.) for all k, so $f(x) \leq f(y)$ if $x \leq y (x \geq y$ resp.), thus $f \in E_{a,b}$.
   Otherwise, let $I = \{k : f_k \text{ increasing in } (a,b)\}$ and $J = \{k : f_k \text{ decreasing in } (a,b)\}$. Then, $\{f_i : i \in I\}$ and $\{f_j : j \in J\}$ are subsequences of $\{f_k\}$. Note that $I \cup J = \mathbb{N}$.
   (a). WLOG let J be finite, then I is infinite and thus f is increasing in (a,b).
   (b). Suppose that I and J are both infinite. Since $f_k \to f$ pointwisely, for each $x \in (a,b)$, every subsequence of $\{f_k(x)\}$ converges to $f(x)$. Hence

$$\lim_{i\to\infty, i\in I} f_i(x) = f(x), \qquad \lim_{j\to\infty, j\in J} f_j(x) = f(x).$$

   Therefore, f is constant in (a,b), thus belongs to $E_{a,b}$.

   Next, we show that $E_{a,b}$ has no interior. Suppose that $E_{a,b}$ contains an open ball

$B(f_0, r) = \{f \in C([0,1]) : \|f - f_0\| < r\}$, where $f_0 \in E_{a,b}$. Then, we can find a zig-zag function g in (a, b) with $\sup_{a<x<b} |g(x) - f_0(x)| < r$, but g is not monotone in (a, b). Now $C([0,1])$ is of the first category, a contradiction.

10. (Stein) Let $X = C([0,1])$ over $\mathbb{R}$ with the sup norm. Let $\mathcal{M}$ be the collection of functions that are not monotonic (increasing or decreasing) in any interval $[a, b]$, where $0 \le a < b \le 1$. Prove that $\mathcal{M}$ is generic in X.
Hint: Let $\mathcal{M}_{[a,b]}$ denote the subset of X consisting of functions that not monotonic in $[a, b]$. Then $\mathcal{M}_{[a,b]}$ is dense in X, while $\mathcal{M}^c_{[a,b]}$ is closed.
Proof

11. (Stein 4-9) Consider $L^p([0,1])$ with Lebesgue measure. Note that if $f \in L^p$ with $p > 1$, then $f \in L^1$. Show that the set of $f \in L^1$ so that $f \notin L^p$, is generic.
Proof Let q be the conjugate exponent to p. Consider the set
$$E_N = \{f \in L^1 : \int_I |f| \le N m(I)^{1/q} \text{ for all intervals } I\}.$$
If $f \in L^p (p > 1)$, then $f \in L^1$. By the Hölder's inequality,
$$\int |f\chi_I| \le \|f\|_p \|\chi_I\|_q = \left(\int |f|^p\right)^{1/p} m(I)^{1/q} \le N\|\chi_I\|_q$$
for some N since $f \in L^p$. This shows that
$$L^p \subset \bigcup_{N=1}^{\infty} E_N.$$
Now we show E_N is closed. Let $\{f_n\} \subset E_N$ and $f_n \to f$ in L^1. Let I be any interval in $[0,1]$, then
$$\begin{aligned}\int_I |f| &\le \left(\int |f_n - f + f|\right) m(I)^{1/q} \\ &\le \left[\int |f_n - f| + \int |f|\right] m(I)^{1/q} \\ &\le N m(I)^{1/q}\end{aligned}$$
if n is sufficiently large.
Finally, we show E_N is nowhere dense in L^1. Suppose that E_N contains an open ball
$$B(f_0, \varepsilon) = \{f \in L^1 : \|f - f_0\|_{L^1} < \varepsilon\},$$
where $f_0 \in E_N$.

12. (Fall 2020) Suppose that X, Y, Z are Banach spaces, and $T : X \times Y \to Z$ is a mapping such that
(a). For each $x \in X$, the map $y \mapsto T(x, y)$ is a bounded linear map: $Y \to Z$.
(b). For each $y \in Y$, the map $x \mapsto T(x, y)$ is a bounded linear map: $X \to Z$.
Prove that there exists a constant C such that
$$\|T(x, y)\|_Z \le C\|x\|_X \|y\|_Y.$$
Proof

Chapter 9 Hausdorff Measures

The reference of this chapter is *Real Analysis: Measure Theory, Integration, and Hilbert Spaces* by Elias M. Stein and Rami Shakarchi.

9.1 Metric Outer Measure

Let (X, d) be a metric space, recall that the **distance** between two sets A and B is defined by $d(A, B) = \inf\{d(x, y) : x \in A, y \in B\}$.

Definition 9.1

*An outer measure μ^* on X is a **metric outer measure** if it satisfies*

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B) \quad \text{if } d(A, B) > 0.$$

♣

Theorem 9.1

If μ^* is a metric outer measure on a metric space X, then every Borel set in X is measurable. Hence $\mu^*|_{\mathcal{B}_X}$ is a measure.

♡

Proof Since $\mathcal{B}_X$ can be generated by closed sets, it suffices to show that every closed set $F \subset X$ is μ^*-measurable. Given $A \subset X$ with $\mu^*(A) < \infty$, we show that

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

Let $B_n = \{x : A \setminus F : d(x, F) \geq 1/n\}$, then B_n increases and $\bigcup_{n=1}^{\infty} B_n = A \setminus F$. And since $d(x, F) \geq 1/n$ for all $x \in B_n$, we have $d(B_n, F) \geq 1/n$. Thus by the metric outer measure condition,

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

Let $C_n = B_{n+1} \setminus B_n$ and let $x \in C_{n+1}$ (so that $d(x, F) < 1/(n+1)$). If $d(x, y) < \frac{1}{n(n+1)}$, then

$$d(y, F) \leq d(x, y) + d(x, F) \leq \frac{1}{n(n+1)} + \frac{1}{n+1} = \frac{1}{n}.$$

That is to say, if $d(x, y) < \frac{1}{n(n+1)}$, then $y \notin B_n$. This is equivalent as saying "if $y \in B_n$, then $d(x, y) \geq \frac{1}{n(n+1)}$." Taking infimum over all $x \in C_{n+1}$ and $y \in B_n$ yields

$$d(C_{n+1}, B_n) \geq \frac{1}{n(n+1)}.$$

Invoking the metric outer measure condition again, we have

$$\begin{aligned}
\mu^*(B_{2k+1}) &\geq \mu^*(C_{2k} \cup B_{2k-1}) = \mu^*(C_{2k}) + \mu^*(B_{2k-1}) \\
&\geq \mu^*(C_{2k}) + \mu^*(C_{2k-2} \cup B_{2k-3}) \geq \cdots \\
&\geq \sum_{j=1}^{k} \mu^*(C_{2j}).
\end{aligned}$$

Similarly, $\mu^*(B_{2k}) \geq \sum_{j=1}^{k} \mu^*(C_{2j-1})$. Because $\mu^*(B_n) \leq \mu^*(A) < \infty$ for all n, both

$\sum_{j=1}^{k} \mu^*(C_{2j})$ and $\sum_{j=1}^{k} \mu^*(C_{2j-1})$ converges. Now

$$\begin{aligned}\mu^*(A \setminus F) &= \mu^*\left(\bigcup_{n=1}^{\infty} B_n\right) \\ &\leq \mu^*(B_n) + \mu^*\left(\bigcup_{j=1}^{\infty} B_{j+1} \setminus B_j\right) \\ &\leq \mu^*(B_n) + \sum_{j=n}^{\infty} \mu^*(C_j).\end{aligned}$$

Let $n \to \infty$, then $\sum_{j=n}^{\infty} \mu^*(C_j) \to 0$. Thus

$$\mu^*(A \setminus F) \leq \liminf_{n\to\infty} \mu^*(B_n) \leq \limsup_{n\to\infty} \mu^*(B_n) \leq \mu^*(A \setminus F).$$

□

9.2 Hausdorff Measure

Definition 9.2

For any subset E of $\mathbb{R}^d$, we define the ***exterior α-dimensional Hausdorff measure*** *of E by*

$$m_\alpha^*(E) = \lim_{\delta\to 0} \inf\left\{\sum_k d(F_k)^\alpha : E \subset \bigcup_{k=1}^{\infty} F_k, d(F_k) \leq \delta\ \forall k\right\},$$

where d stands for diameter.

♣

For each δ, we have the quantity

$$\mathcal{H}_\alpha^\delta(E) = \inf\left\{\sum_k d(F_k)^\alpha : E \subset \bigcup_{k=1}^{\infty} F_k, d(F_k) \leq \delta\ \forall k\right\}.$$

As δ decreases, there will be fewer choices of covering sets F_k, so the infimum will increase. Therefore, the limit

$$m_\alpha^*(E) = \lim_{\delta\to 0} \mathcal{H}_\alpha^\delta(E)$$

exists (could be infinite). Also notice that $\mathcal{H}_\alpha^\delta(E) \leq m_\alpha^*(E)$ for all $\delta > 0$.

Proposition 9.1 (monotonicity)

If $E_1 \subset E_2$, then $m_\alpha^*(E_1) \leq m_\alpha^*(E_2)$.

♠

Proof Since any cover of E_2 is also a cover of E_1, taking infimum leads to the inequality.

Proposition 9.2 (sub-additivity)

For any countable family $\{E_j\} \subset \mathbb{R}^d$,

$$m_\alpha^*\left(\bigcup_{j=1}^{\infty} E_j\right) \leq \sum_{j=1}^{\infty} m_\alpha^*(E_j).$$

♠

Proof Fix $\delta > 0$. Cover each E_j with $\{F_{j,k}\}_{k\in\mathbb{N}}$ such that

$$\sum_k d(F_{j,k})^\alpha \le \mathcal{H}_\alpha^\delta(E_j) + \frac{\varepsilon}{2^j}.$$

$\{F_{j,k}\}_{j,k\in\mathbb{N}}$ is a cover of $\bigcup_{j=1}^\infty E_j$, hence

$$\mathcal{H}_\alpha^\delta\left(\bigcup_{j=1}^\infty E_j\right) \le \sum_j \sum_k d(F_{j,k})^\alpha \le \sum_{j=1}^\infty \mathcal{H}_\alpha^\delta(E_j) + \varepsilon \le \sum_{j=1}^\infty m_\alpha^*(E_j) + \varepsilon.$$

Since ε is arbitrary, we have $\mathcal{H}_\alpha^\delta\left(\bigcup_{j=1}^\infty E_j\right) \le \sum_{j=1}^\infty m_\alpha^*(E_j)$. Letting δ tend to 0 proves the countable sub-additivity of m_α^*.

Proposition 9.3 (metric outer measure property)

If $d(E_1, E_2) > 0$, then $m_\alpha^*(E_1 \cup E_2) = m_\alpha^*(E_1) + m_\alpha^*(E_2)$.

Proof Choose $0 < \varepsilon < d(E_1, E_2)$. Let $\{F_k\}_{k\in\mathbb{N}}$ be a cover of $E_1 \cup E_2$ with diameter $< \delta < \varepsilon$, let $F_j' = E_1 \cap F_j$ and $F_j'' = E_2 \cap F_j$. Then $\{F_j'\}$ and $\{F_j''\}$ are covers for E_1 and E_2, respectively, and are disjoint. Hence,

$$\sum_j d(F_j')^\alpha + \sum_i d(F_j'')^\alpha \le \sum_k d(F_k)^\alpha.$$

Taking infimum over all coverings and then letting $\delta \to 0$ yields the desired inequality.

Now m_α^* is a metric exterior measure on $\mathbb{R}^d$, so it is a measure on $\mathcal{B}_{\mathbb{R}^d}$.

Definition 9.3 (Hausdorff Measure)

*The restriction of m_α^* to the Borel sets is called the α-**dimensional Hausdorff measure**, denoted m_α.*

Proposition 9.4

If $\{E_j\}$ is a countable family of disjoint Borel sets, then

$$m_\alpha\left(\bigcup_{j=1}^\infty E_j\right) = \sum_{j=1}^\infty m_\alpha(E_j).$$

Proof This follows from the axiom of a measure.

Proposition 9.5

Hausdorff measure is invariant under translations

$$m_\alpha(E+h) = m_\alpha(E) \quad \text{for all } h \in \mathbb{R}^d,$$

and rotations

$$m_\alpha(RE) = m_\alpha(E),$$

where R is a rotation in $\mathbb{R}^d$. Moreover, it scales as follows:

$$m_\alpha(\lambda E) = \lambda^\alpha m_\alpha(E) \quad \text{for all } \lambda > 0.$$

Proof The diameter of a set E is

$$d(E) = \sup_{x,y\in E} |x-y| = \sup_{x,y\in E} \langle x-y, x-y\rangle.$$

It suffices to check that the diameter satisfies the above relations.

1. Clearly $d(E+h) = d(E)$.
2. A rotation is an orthogonal linear map on $\mathbb{R}^d$, so
$$|Rx - Ry|^2 = \langle R(x-y), R(x-y)\rangle = \langle x-y, x-y\rangle = |x-y|^2,$$
hence $d(RE) = d(E)$.
3. $d(\lambda E) = \sup_{x,y\in E} |\lambda x - \lambda y| = \lambda \sup_{x,y\in E} |x-y| = \lambda d(E)$, so $d(\lambda E)^\alpha = \lambda^\alpha d(E)^\alpha$.

□

For some special α, the α-dimensional Hausdorff measure corresponds to our familiar measures.

Properties *[special cases]*

1. *m_0 is the conuting measure.*
2. *m_1 is the Lebesgue measure(restricted to Borel sets) on $\mathbb{R}$.*

Proof

1. Let $x \in \mathbb{R}^d$, we show that $m_0(\{x\}) = 1$. For each $\delta > 0$, the open ball $B(x,\delta)$ covers $\{x\}$ and $d(B(x,\delta))^\alpha = d(B(x,\delta))^0 = 1$, hence $m_0(\{x\}) = 1$. If E is a finite set, then by the countable additivity, $m(E) = \#E$.
2. $\mathcal{H}_1^\delta(E) = \inf\{\sum_k d(F_k) : E \subset \bigcup_{k=1}^\infty F_k, d(F_k) < \delta\}$, and
$m^*(E) = \inf\{\sum_k m(I_k) : E \subset \bigcup_{k=1}^\infty I_k, I_k \text{ are intervals}\}$. Let E be covered by $\{F_k\}$ with $d(F_k) < \delta$ and
$$\sum_k d(F_k) < \mathcal{H}_1^\delta(E) + \varepsilon.$$

Proposition 9.6

If E is a Borel subset of $\mathbb{R}^d$, then $c_d m_d(E) = m(E)$ for some constant c_d that depends only on d.

Proposition 9.7

If $m_\alpha^*(E) < \infty$ and $\beta > \alpha$, then $m_\beta^*(E) = 0$. Also, if $m_\alpha^*(E) > 0$ and $\beta < \alpha$, then $m_\beta^*(E) = \infty$.

Proof Let $d(F) \leq \delta$. If $\beta > \alpha$, then
$$d(F)^\beta = d(F)^{\beta-\alpha} d(F)^\alpha \leq \delta^{\beta-\alpha} d(F)^\alpha.$$
Consequently
$$\mathcal{H}_\beta^\delta(E) \leq \delta^{\beta-\alpha}\mathcal{H}_\alpha^\delta(E) \leq \delta^{\beta-\alpha} m_\alpha^*(E).$$
Since $m_\alpha^*(E) < \infty$ and $\beta - \alpha > 0$, letting $\delta \to 0$ gives $m_\beta^*(E) = 0$. □

Remark The set $\{\beta > 0 : m_\beta^*(E) = 0\}$ is bounded below, hence its infimum exists. Similarly, $\{\beta \leq d : m_\beta^*(E) = \infty\}$ has the supremum.

9.3 Hausdorff Dimension

Let $E \subset \mathbb{R}^d$ be a Borel set, then there exists a unique α such that

$$m_\beta(E) = \begin{cases} \infty & \text{if } \beta < \alpha, \\ 0 & \text{if } \beta > \alpha. \end{cases}$$

α is given by

$$\alpha = \sup\{\beta : m_\beta(E) = \infty\} = \inf\{\beta : m_\beta(E) = 0\}.$$

We say that E has **Hausdorff dimension** α, or that E has dimension α. We shall write $\alpha = \dim E$. If $0 < m_\alpha(E) < \infty$, we say that E has **strict Hausdorff dimension** α. The term **fractal** is applied to sets of fractional dimension.

9.3.1 Examples

The Cantor set

Theorem 9.2

The Cantor set $\mathcal{C}$ has strict Hausdorff dimension $\alpha = \log 2 / \log 3$ ♡

Definition 9.4

Let f be defined on $E \subset \mathbb{R}^d$. We say that f satisfies Hölder condition γ if

$$|f(x) - f(y)| \le M|x - y|^\gamma \quad \forall x, y \in E.$$

♣

Lemma 9.1

Suppose f defined on a compact set E satisfies Hölder condition with exponent γ. Then

1. $m_\beta(f(E)) \le M^\beta m_\alpha(E)$ *if* $\beta = \alpha/\gamma$.
2. $\dim f(E) \le \frac{1}{\gamma} \dim E$.

♡

Proof Let $\{F_k\}_{k\in\mathbb{N}}$ covers E, then $\{f(E \cap F_k)\}_{k\in\mathbb{N}}$ covers $f(E)$, and

$$|f(x) - f(y)| \le M|x - y|^\gamma \quad \forall x, y \in E \cap F_k,$$

so

$$d(f(E \cap F_k)) \le M d(E \cap F_k)^\gamma \le M d(F_k)^\gamma.$$

Hence

$$\sum_k d(f(E \cap F_k))^{\alpha/\gamma} \le M^{\alpha/\gamma} \sum_k d(F_k)^\alpha.$$

Taking infimum and taking limits, we have

$$m_{\alpha/\gamma}(f(E)) \le M^{\alpha/\gamma} m_\alpha(E).$$

- If $0 < m_{\alpha/\gamma}(f(E)) \le M^{\alpha/\gamma} m_\alpha(E) < \infty$, then $\dim f(E) = \alpha/\gamma$ and $\dim E = \alpha$, thus $\dim f(E) \le \frac{1}{\gamma} \dim E$.
- If $m_{\alpha/\gamma}(f(E)) = 0$ and $m_\alpha(E) = 0$, then then $\dim f(E) \le \alpha/\gamma$.

Lemma 9.2

The Cantor-Lebesgue function F on $\mathcal{C}$ satisfies Hölder condition with $\gamma = \log 2/\log 3$. ♡

Proof F is the limit of a sequence $\{F_n\}$ of piecewise linear functions. F_n increases by at most 2^{-n} on each interval of length 3^{-n}. So the slope of F_n is always bounded by $(3/2)^n$, and hence

$$|F_n(x) - F_n(y)| \leq \left(\frac{3}{2}\right)^n |x-y|.$$

The approximating sequence also satisfies $|F(x) - F_n(x)| \leq 1/2^n$. Then

$$\begin{aligned}|F(x) - F(y)| &\leq |F_n(x) - F_n(y)| + |F(x) - F_n(x)| + |F(y) - F_n(y)| \\ &\leq \frac{3^n}{2^n}|x-y| + \frac{2}{2^n}.\end{aligned}$$

We need to choose n so that $3^n|x-y|$ is of the same order as a constant. Take n so that $3^n|x-y| \in [1,3]$. Then

$$|F(x) - F(y)| \leq \frac{c}{2^n} = \frac{c}{3^{(\log_3 2)n}} := c(3^{-n})^\gamma \leq M|x-y|^\gamma,$$

where $\gamma = \log_3 2 = \log 2/\log 3$. □

Now we prove that $\dim \mathcal{C} = \log 2/\log 3$. We only need to show that $0 < m_{\log 2/\log 3}(\mathcal{C}) < \infty$, which looks not so difficult.

Part (I): $m_\gamma(\mathcal{C}) \leq 1$.

Recall the construction of the Cantor set, at nth step we get 2^n intervals of length 3^{-n} and denote the union of these intervals by C_k, then $\mathcal{C} \subset \bigcap_{k=1}^{2^n} C_k$. Fix $\delta > 0$ and choose $3^{-n} < \delta$, then

$$d(C_k)^\gamma \leq 2^n (3^{-n})^\gamma = 2^n 2^{-n} = 1,$$

hence $m_\gamma(\mathcal{C}) \leq 1$.

Part (II): $m_\gamma(\mathcal{C}) > 0$.

Applying Lemma 9.1 with $E = \mathcal{C}$ and $\alpha = \gamma$, we have

$$m_1(f(\mathcal{C})) = m_1([0,1]) \leq M m_\gamma(\mathcal{C}),$$

thus $m_\gamma(\mathcal{C}) > 0$, and we find that $\dim \mathcal{C} = \log 2/\log 3$.

Rectifiable curves

Chapter 10 Topology in Analysis

This chapter is a copy of Chapter 4 of Real Analysis, Folland.

10.1 Topological Spaces

Let X be a nonempty set. A topology on X is a family $\mathbb{T}$ of subsets of X that

- $\varnothing, X \in \mathbb{T}$.
- $\mathbb{T}$ is closed under arbitrary unions.
- $\mathbb{T}$ is closed under finite intersections.

The pair $(X, \mathbb{T})$ is called a topological space.

Definition 10.1 (sets in a TS)

Let $A \subset X$.

1. *The members of $\mathbb{T}$ are called open sets, and the complement of a open set is called a closed set.*
2. *The interior of A is the union of all open sets contained in A (largest open set contained in A).*
3. *The closure of A is the intersection of all closed sets containing A (smallest closed set containing A).*
4. *If $\overline{A} = X$, A is called dense in X.*
5. *If $(\overline{A})^\circ = \varnothing$, A is called nowhere dense.*
6. *x is called a limit point of A if $A \cap (U \backslash \{x\}) \neq \varnothing$ for every neighborhood U of x. The set of all limit points of A is denoted A'.*

♣

Proposition 10.1

1. $(A^\circ)^c = \overline{A^c}$.
2. $(\overline{A})^c = (A^c)^\circ$.
3. $\overline{A} = A \cup A'$.
4. *A is closed iff $A' \subset A$.*

♠

Definition 10.2 (weak and strong)

Let $\mathbb{T}_1, \mathbb{T}_2$ are two topologies on X.

1. *If $\mathbb{T}_1 \subset \mathbb{T}_2$, then we say that $\mathbb{T}_1$ is weaker(coarser) than $\mathbb{T}_2$.*
2. *If $\mathbb{T}_1 \supset \mathbb{T}_2$, then we say that $\mathbb{T}_1$ is stronger(finer) than $\mathbb{T}_2$.*

♣

10.1.1 Base

Definition 10.3 (subbase, base)

If $\mathcal{E} \subset \mathcal{P}(X)$, there is a unique weakest topology $\mathcal{T}(\mathcal{E})$ on X that contains $\mathcal{E}$: the intersection of all topologies on X containing $\mathcal{E}$, which is called the topology generated by $\mathcal{E}$.

$\mathcal{E}$ is called a subbase for $\mathcal{T}(\mathcal{E})$.

A local base for $\mathcal{T}$ at $x \in X$ is a family $\mathcal{N} \subset \mathcal{T}$ such that

- *$x \in V$ for all $V \in \mathcal{N}$;*
- *If U is a neighborhood of x, then $\exists V \in \mathcal{N} : V \subset U$.*

A base for $\mathcal{T}$ is a family $\mathcal{B} \subset \mathcal{T}$ that contains a local base for $\mathcal{T}$ at each $x \in X$. ♣

Proposition 10.2 (characterization of base)

If $\mathcal{T}$ is a topology on X and $\mathcal{E} \subset \mathcal{T}$, then $\mathcal{E}$ is a base for $\mathcal{T}$ iff every nonempty $U \in \mathcal{T}$ is a union of members of $\mathcal{E}$. ♠

Proof Let $\mathcal{E}$ be a base for $\mathcal{T}$, then $\mathcal{E}$ contains a local base at each $x \in X$. Let U be an open set in X, then for each $x \in U$ there is a $V_x \in \mathcal{E}$ such that

$$x \in V_x \subset U.$$

Then $U = \bigcup_{x \in U} V_x$. Conversely, let $x \in X$, then $\{V \in \mathcal{E} : x \in V\}$ is a local base at x.

Proposition 10.3

If $\mathcal{B} \subset \mathcal{P}(X)$, in order for $\mathcal{B}$ to be a base for a topology on X it is necessary and sufficient that:

1. *each $x \in X$ is contained in some $V \in \mathcal{B}$;*
2. *if $U, V \in \mathcal{B}$ and $x \in U \cap V$, there exists $W \in \mathcal{B}$ and $x \in W \subset (U \cap V)$.* ♠

Proof

Topology is also a set-algebraic structure like σ-algebra, so it can definitely be generated by "simple" sets. Moreover, we can describe the topology generated by a family $\mathcal{E}$.

Proposition 10.4 (description of generated topology)

If $\mathcal{E} \subset \mathcal{P}(X)$, the topology $\mathcal{T}(\mathcal{E})$ generated by $\mathcal{E}$ consists of $\varnothing$, X, and all unions of finite intersections of members of $\mathcal{E}$. ♠

We have three ways to show a family of sets $\mathcal{B}$ is a basis:

1. Passing to a neighborhood basis;
2. the most intuitive way: show that every nonempty open set is a union of memebers of $\mathcal{B}$;
3. a convenient way: show that the intersection of two base elements contains another base element.

10.2 Continuous Maps

10.2.1 Weak and Product Topologies

Definition 10.4 (weak topology)

If X is any set and $\{f_\alpha : X \to Y_\alpha\}_{\alpha\in A}$ is a family of maps from X into some topological spaces Y_α, there is a unique weakest topology $\mathcal{T}$ on X makes all the f_α continuous; it is called the ***weak topology*** *generated by $\{f_\alpha\}_{\alpha\in A}$. Namely, $\mathcal{T}$ is the topology generated by sets of the form $f_\alpha^{-1}(U_\alpha)$ where $\alpha \in A$ and U_α is open in Y_α.* ♣

The product topology is an example of weak topology.

Definition 10.5

If $\{X_\alpha\}_{\alpha\in A}$ is any family of topological spaces, the product topology on $X = \prod_{\alpha\in A} X_\alpha$ is the weak topology generated by the coordinate maps $\pi_\alpha : X \to X_\alpha$. ♣

Proposition 10.5 (base for the product topology)

A base for the product topology is given by the sets of the form

$$\bigcap_{j=1}^{n} \pi_{\alpha_j}^{-1}(U_{\alpha_j}), \quad n \in \mathbb{N}, U_{\alpha_j} \in \mathcal{T}_{\alpha_j} \text{ for } 1 \le j \le n.$$

♠

Proof We shall prove the case when $X = \prod_{j=1}^n X_j$, where each X_j is endowed with the topology $\mathcal{T}_j$. Let $\mathcal{B} = \left\{\bigcap_{j=1}^n \pi_j^{-1}(U_j) : U_j \in \mathcal{T}_j, 1 \le j \le n\right\}$.

- Each $x \in X$ is contained in some member of $\mathcal{B}$. Write $x = (x_1, \cdots, x_n)$, then there exists $U_j \in \mathcal{T}_j$ such that $x_j \in U_j$, hence

$$(x_1, \cdots, x_n) \in U_1 \times \cdots \times U_n = \bigcap_{j=1}^{n} \pi_j^{-1}(U_j).$$

  Here we use the fact that

$$\pi_j^{-1}(U_j) = X_1 \times \cdots \times X_{j-1} \times U_j \times X_{j+1} \times \cdots \times X_n.$$

- By definition, the product topology is the topology generated by sets of the form $\pi_i^{-1}(U_i)$, where $1 \le i \le n$ and $U_i \in \mathcal{T}_i$. Since the product topology contains all unions of finite intersections of members of $\mathcal{B}$, it follows that $\mathcal{B}$ is a base.

□

Proposition 10.6

If X_j is Hausdorff, then $X = \prod_{j=1}^n X_j$ is Hausdorff. ♠

Proof Let $x \ne y$ in X, then $\pi_i(x) \ne \pi_i(y)$ for some i. Then choose disjoint neighborhoods U and V of $\pi_i(x)$ and $\pi_i(y)$ in X_i. We have $\pi_i^{-1}(U) \cap \pi_i^{-1}(V) = \varnothing$ in X. □

Proposition 10.7

If X_j and Y are topological spaces and $X = \prod_{j=1}^n X_j$, then $f : Y \to X$ is continuous iff each $\pi_j \circ f$ is continuous. ♠

Proof If $\pi_j \circ f$ is continuous for each i, then

$$(\pi_j \circ f)^{-1}(U_j) = f^{-1}(\pi_j^{-1}(U_j))$$

is open in Y for each open U_j in X_j. This shows that $f^{-1}(E)$ is open for every E in the generating set of the product topology, hence f is continuous. □

If $X_\alpha = X$ for all $\alpha \in A$, then $\prod_{\alpha\in A}$ is the set X^A of maps from A to X. Think of $A = \mathbb{N}$ and $\prod_{n\in\mathbb{N}} \mathbb{R} = \{(x_1, x_2, \cdots) : x_n \in \mathbb{R}\}$, the space of real number sequences.

Proposition 10.8

If X is a topological space, A is a nonempty set, and $\{f_n\}$ is a sequence in X^A, then $f_n \to f$ in the product topology iff $f_n \to f$ pointwise.

10.2.2 Topologies on Spaces of Continuous Functions

Let X be any set, we introduce some notations:

- $B(X, \mathbb{R})$ is the space of all bounded real-valued functions on X.
- If X is a topological space, denote $C(X, \mathbb{R})$ the space of continuous functions on X.
- If X is a topological space, we define

$$BC(X, \mathbb{F}) = B(X, \mathbb{F}) \cap C(X, \mathbb{F}) \quad (\mathbb{F} = \mathbb{R} \text{ or } \mathbb{C}).$$

- If $f \in B(X)$, we define the uniform norm of f to be

$$\|f\|_u = \sup_{x\in X} |f(x)|.$$

Proposition 10.9

If X is a topological space, $BC(X)$ is a closed subspace of $B(X)$ in the uniform metric; in particular, $BC(X)$ is complete.

Proof Suppose $\{f_n\} \subset BC(X)$ and $\|f_n - f\|_u \to 0$. Given $\varepsilon > 0$, choose N so large that $\|f_n - f\|_u < \varepsilon/3$ for $n < N$. Since f_n is continuous at x, there is a neighborhood U of x such that $|f_n(y) - f_n(x)| < \varepsilon/3$ for $y \in U$. Then

$$|f(y) - f(x)| \leq |f(y) - f_n(y)| + |f_n(y) - f_n(x)| + |f_n(x) - f(x)| < \varepsilon.$$

□

Example 10.1 If X has the trivial topology, then $C(X)$ consists only of constant functions.

Proof Let $f \in C(X)$, then $f^{-1}(U) = X$ for all open sets in $\mathbb{R}$, hence f is constant. □

Theorem 10.1 (Urysohn's lemma)

Let X be a normal space. If A and B are disjoint closed sets in X, there exists $f \in C(X, [0,1])$ such that $f = 0$ on A and $f = 1$ on B.

Theorem 10.2 (Tietze extension theorem)

Let X be a normal space. If A is a closed subset of X and $f \in C(A, [a,b])$, there exists $F \in C(X, [a,b])$ such that $F|A = f$.

10.3 Nets

Definition 10.6 (directed set)

A directed set is a set A equipped with a binary relation $\lesssim$ such that

- $\alpha \lesssim \alpha$ *for all* $\alpha \in A$;
- *if* $\alpha \lesssim \beta$ *and* $\beta \lesssim \gamma$, *then* $\alpha \lesssim \gamma$;
- *for any* $\alpha, \beta \in A$ *there exists* $\gamma \in A$ *such that* $\alpha \lesssim \gamma$ *and* $\beta \lesssim \gamma$.

♣

Definition 10.7 (net)

A net in a set X is a mapping $\alpha \mapsto x_\alpha$ from a directed set A into X. We denote such a mapping by

$$\langle x_\alpha \rangle_{\alpha \in A}.$$

♣

Example 10.2

1. $\mathcal{N}$ is a net with $j \lesssim k$ iff $j \leq k$.
2. $\mathbb{R} \setminus \{a\}$ with $x \lesssim y$ iff $|x - a| \geq |y - a|$.
3. The set of all partitions $\{x_j\}_0^n$ of $[a, b]$ (i.e. $a = x_0 < \cdots < x_n = b$) with
$$\{x_j\}_0^n \lesssim \{y_k\}_0^m \iff \max(x_j - x_{j-1}) \geq \max(y_k - y_{k-1}).$$
4. The set $\mathcal{N}$ of all neighborhoods of a point x in a topological space X, with $U \lesssim V$ iff $U \supset V$. (We say that $\mathcal{N}$ is directed by reverse inclusion.)
5. The Cartesian product $A \times B$ of two directed sets, with $(\alpha, \beta) \lesssim (\alpha', \beta')$ iff $\alpha \lesssim \alpha'$ and $\beta \lesssim \beta'$. (This is *always* the way we make $A \times B$ into a directed set.)

Definition 10.8

Let X be a topological space and $E \subset X$.

1. *A net $\langle x_\alpha \rangle_{\alpha \in A}$ is **eventually** in E if there exists $\alpha_0 \in A$ such that $x_\alpha \in A$ for all $\alpha \gtrsim \alpha_0$;*
2. *$\langle x_\alpha \rangle$ is **frequently** in E if for every $\alpha \in A$ there exists $\beta \gtrsim \alpha$ such that $x_\beta \in E$.*
3. *A point $x \in X$ is a **limit** of $\langle x_\alpha \rangle$(or $\langle x_\alpha \rangle$ **converges** to x, or $x_\alpha \to x$) if for every neighborhood U of x, $\langle x_\alpha \rangle$ is eventually in U.*
4. *x is a **cluster point** of $\langle x_\alpha \rangle$ if for every neighborhood U of x, $\langle x_\alpha \rangle$ is frequently in U.*

♣

Proposition 10.10

If X is a topological space, $E \subset X$, and $x \in X$, then x is an accumulation point of E iff there is a net in $E \setminus \{x\}$ that converges to x, and $x \in \overline{E}$ iff there is a net in E that converges to x.

♠

Proof Let x be an accumulation point of E, let $\mathcal{N}$ be the set of neighborhoods of x, directed by reverse inclusion. For each $U \in \mathcal{N}$, pick $x_U \in (U \setminus \{x\}) \cap E$. Now let V be an arbitrary neighborhood of x, then $x_U \in V$ for all $U \subset V$ (i.e., for all $U \gtrsim V$), hence $\langle x_U \rangle_{U \in \mathcal{N}}$ is eventually in V. Conversely, if $x_\alpha \in E \setminus \{x\}$ and $x_\alpha \to x$, then every punctured neighborhood of x contains some x_α, so x is an accumulation point of E.

Proposition 10.11 (net continuity)

If X and Y are topological spaces and $f : X \to Y$, then f is continuous at x iff for every net $\langle x_\alpha \rangle$ converging to x, $\langle f(x_\alpha) \rangle$ converges to $f(x)$. $\square$

Proof Let f be continuous at x and let V be a neighborhood of $f(x)$, then $f^{-1}(V)$ is a neighborhood of x. Hence, if $x_\alpha \to x$ then $\langle x_\alpha \rangle$ is eventually in $f^{-1}(V)$, so $\langle f(x_\alpha) \rangle$ is eventually in V, and thus $f(x_\alpha) \to f(x)$. On the other hand, if f is not continuous at x, there is a neighborhood V of $f(x)$ such that $f^{-1}(V)$ is not a neighborhood of x, that is, $x \notin (f^{-1}(V))^\circ$ (x is not an interior point), or equivalently $x \in \overline{f^{-1}(V^c)}$. Then there is a net $\langle x_\alpha \rangle$ in $f^{-1}(V^c)$ that converges to x. But then $f(x_\alpha) \notin V$, so $f(x_\alpha)$ does not converge to $f(x)$. $\square$

10.4 Compactness and Locally Compact Hausdorff Spaces

Definition 10.9

A topological space X is said to be ***compact*** *if whenever $\{U_\alpha\}_{\alpha \in A}$ is an open cover of X, there is a finite subset $B \subset A$ such that $X = \bigcup_{\alpha \in B} U_\alpha$. To be brief, we say: X is compact if every open cover of X has a finite subcover.*

A subset Y of a topological space X is called ***compact*** *if for any open cover $\bigcup_{\alpha \in A} U_\alpha \supset Y$, there is a finite $B \subset A$ with $Y \subset \bigcup_{\alpha \in B} U_\alpha$.*

Y is called ***precompact*** *if $\overline{Y}$ is compact.*

A family $\{F_\alpha\}_{\alpha \in A}$ of subsets of X is said to have the **finite intersection property** if $\bigcap_{\alpha \in B} F_\alpha \neq \varnothing$ for all fintie $B \subset A$.

Proposition 10.12 (finite intersection property)

A topological space X is compact iff for every family $\{F_\alpha\}_{\alpha \in A}$ of closed sets with the finite intersection property, $\bigcap_{\alpha \in A} F_\alpha \neq \varnothing$.

Proof Let X be compact and suppose there were a family $\{F_\alpha\}_{\alpha \in A}$ with FIP but $\bigcap_{\alpha \in A} F_\alpha = \varnothing$, then $\bigcup_{\alpha \in A} F_\alpha^c = X$, so there is a finite $B \subset A : \bigcup_{\alpha \in B} F_\alpha^c = X$, hence $\bigcap_{\alpha \in B} F_\alpha = \varnothing$, contradicting the FIP condition.

Suppose X is not compact, then there is an open cover $\bigcup_{\alpha \in A} U_\alpha = X$ has no finite subcover. That is, $\bigcup_{\alpha \in B} U_\alpha \neq X$ for all finite $B \subset A$. Taking complements gives

$$\bigcap_{\alpha \in B} U_\alpha^c \neq \varnothing,$$

hence the family $\{U_\alpha^c\}_{\alpha \in A}$ has the FIP, and $\bigcup_{\alpha \in A} U_\alpha^c = \varnothing$, completing the proof. $\square$

10.4.1 Basic Facts About Compact Spaces

Proposition 10.13

A closed subset of a compact space is compact.

Proposition 10.14 (separation property)

If F is a compact subset of a Hausdorff space X and $x \notin F$, then there are disjoint open sets U, V such that $x \in U$ and $F \subset V$.

Proposition 10.15

Every compact subset of a Hausdorff space is closed.

Proposition 10.16

Every compact Hausdorff space is normal.

Proposition 10.17

If X is compact and Y is Hausdorff, then any continuous bijection $f : X \to Y$ is a homeomorphism.

Theorem 10.3

If X is a topological space, the following are equivalent:

1. *X is compact.*
2. *Every net in X has a cluster point.*
3. *Every net in X has a convergent subnet.*

Definition 10.10 (LCH)

A topological space is called locally compact if every point has a compact neighborhood. We call locally compact Hausdorff spaces ***LCH*** *spaces for short.*

10.4.2 Urysohn's Lemma

In Real Analysis I, we introduced the Urysohn's lemma and prove that $C_c(\mathbb{R}^d)$ is dense in $L^1(\mathbb{R}^d)$. The idea is to show the case of a characteristic function χ_E on a measurable set E using the regularity property, and using Urysohn's lemma to modify χ_E to be a continuous function f so that $\|\chi_E - f\|_{L^1}$ is small. We first present some topological properties of a LCH space.

Proposition 10.18

If X is an LCH space, $U \subset X$ is open, and $x \in U$, there is a compact neighborhood N of x such that $N \subset U$.

Proposition 10.19

If X is an LCH space and $K \subset U \subset X$ where K is compact and U is open, there exists a precompact open V such that $K \subset V \subset \overline{V} \subset U$.

10.4.3 Functions of Compact Support

Proposition 10.20

If X is an LCH space, $C_0(X)$ is the closure of $C_c(X)$ in the uniform metric.

Appendix A Set Theory

A.1 Cartesian Products

Definition A.1

Let $\{X_\alpha\}_{\alpha\in A}$ be an indexed family of sets, their ***Cartesian product*** $\prod_{\alpha\in A} X_\alpha$ *is the set of all maps* $f : A \to \bigcup_{\alpha\in A} X_\alpha$ *such that* $f(\alpha) \in X_\alpha$ *for all* $\alpha \in A$. ♣

Definition A.2

If $X = \prod_{\alpha\in A} X_\alpha$ *and* $\alpha \in A$, *we define the* α*th* ***projection*** *or* ***coordinate map*** $\pi_\alpha : X \to X_\alpha$ *by* $\pi_\alpha(f) = f(\alpha)$. ♣