Real Analysis III

Abstract Measure

Author: Kumiko

Institute: Kitauji IT

Date: 2022 summer

Contents

Chapter 1 Abstract Measures

1.1 Measures and Set-Algebraic Structures

Introduction

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

In Real Analysis I we have seen the Lebesgue measure and I also gave a preview on a particular example of an abstract measure. One of the properties of Lebesgue integral states: If $\{E_n : n \in \mathbb{N}\}$ is a disjoint sequence of subsets of $\mathbb{R}^d$ and $f \in L^1(\mathbb{R}^d)$, then

$$\int_{\bigcup_{n=1}^{\infty} E_n} f(x)dx = \sum_{n=1}^{\infty} \int_{E_n} f(x)dx.$$

Fix this f, define a set function μ on the Lebesgue σ-algebraby $\mu(E) = \int_E f(x)dx$, then

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

which has exactly the same form of the **countable additivity** of the Lebesgue measure. Moreover, clearly we have $\mu(\varnothing) = 0$. The more general definition of a measure should look like:

Definition 1.1 (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 only care about what it does rather what it is.

Example 1.1 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

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

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

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.
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 1.2 (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 $E^c \in \mathcal{M}$.*

$\clubsuit$

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

Definition 1.3 (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 $\{E_n : n \in \mathbb{N}\} \subset \mathcal{M}$, $\mu\left(\bigcup_{n=1}^{\infty} E_n\right)) = \sum_{n=1}^{\infty} \mu(E_n)$.*

$\clubsuit$

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

1.1.1 σ-Algebras

We begin by looking some examples of σ-algebras.

Example 1.2 If X is any set, then $\mathcal{P}(X)$ and $\{\varnothing, X\}$ are σ-algebras.

Example 1.3 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 1.4

*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 1.4 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 1.1

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 1.5

*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 1.2

$\mathcal{B}_{\mathbb{R}}$ 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 exericises.

1.1.2 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 1.5 (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 1.6 (Lebesgue measure) You are very familiar with this!

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

Theorem 1.1 (monotonicity and subadditivity)

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

♡

Proof

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

Theorem 1.2 (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 1.6

*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 1.3

Suppose that $(X, \mathcal{M}, \mu)$ is a measure space. Let $\mathcal{N} = \{N \in \mathcal{M} : \mu(N) = 0\}$ and $\overline{\mathcal{M}} = \{E \cup F : E \in \mathcal{M}, F \subset N$ 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 1.7

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

1.1.3 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 1.7 A vector space V is a set V with addition and scalar multiplication which satisfy 8 properties.

Example 1.8 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 1.9 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].

Definition 1.8 (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 $A^c := X \setminus A$ 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.

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

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 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 1.4

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.

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

(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 $A \in \mathcal{D}(\mathcal{E})$.

1.2 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:

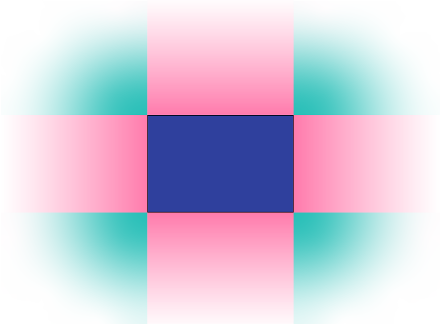
$\mathbb{R}^d$ case	abstraction
cube	elementary figures
volume	a volume function

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 $-\infty \le a_j, b_j \le \infty$ 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 1.9

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)

1.2.1 (volume, semialgebra)

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

Example 1.10 In $\mathbb{R} \cup \{+\infty\}$, $\{(a, b] : -\infty \le a, b \le \infty\}$ is a semialgebra and the length of an interval is a volume.

1.2.2 (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 1.10 (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 1.3

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 1.11

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* $\mu_0(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu_0(A_j)$.

♣

1.2.3 (outer measure, power set)

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

Definition 1.12 (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 1.4

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

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

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

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 1.13

*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 1.1 Prove this. Compare the definition between a measure and a premeasure.

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

Proposition 1.5

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?

1.2.4 (measure, σ-algebra)

Theorem 1.5 (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:

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

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 1.6

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 1.2.3 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 1.2.3 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 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.

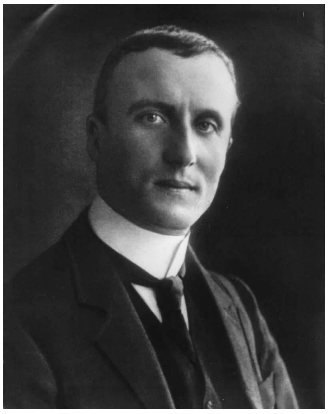


Figure 1.1: Constantin Carathéodory

1.3 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 1.6

$\mathcal{B}_{\mathbb{R}}$ 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 $\mathcal{E} = \{(a, b] : -\infty \le a < b < \infty\}$, which is called **h-intervals**.

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

Exercise 1.2 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.

1.3.1 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 1.7

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 1.8

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 1.9

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.

[7] Here we actually use Theorem 1.6, but we still refer to it as Carathéodory's theorem. Notice that $\mu^*|_{\mathcal{B}_{\mathbb{R}}}$ is not necessarily complete.

Theorem 1.7

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 we define

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

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

♡

Proof Start from $\mu_F((a, b]) = F(b) - F(a)$, using proposition 1.9 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 1.9.

Proof [proof of proposition 1.7] 编号有问题

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:

$$\begin{aligned}
I_1 &= I_1 \cap \bigcup_{j=1}^n J_j = \bigcup_{j=1}^n I_1 \cap J_j, \\
&\cdots \\
I_m &= I_m \cap \bigcup_{j=1}^n J_j = \bigcup_{j=1}^n I_m \cap J_j, \\
J_1 &= J_1 \cap \bigcup_{i=1}^m I_i = \bigcup_{i=1}^m J_1 \cap I_i \\
&\cdots \\
J_n &= J_n \cap \bigcup_{i=1}^m I_i = \bigcup_{i=1}^m J_n \cap I_i.
\end{aligned}$$

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 $F(M) - F(a) \leq \sum_{j=1}^{\infty} \mu_0(I_j) + 2\varepsilon$. □

1.3.2 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 1.14 (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 1.10 (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 1.3 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 $\mu^*(A) \leq \mu^*(E) + \epsilon$.
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

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

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 $\mu^*(A_n) \leq \mu^*(E) + 1/n$, then

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

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 1.4 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 1.15
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.

1.3.3 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

for every $\varepsilon > 0$ there exists an open set $\mathcal{O} \supset E$ 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 1.1

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 $F(b_j + \delta_j) - F(b_j) < \epsilon 2^{-j}$. Then $E \subset \bigcup_1^\infty (a_j, b_j + \delta_j)$ 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 1.8

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 (a_j, b_j) such that $E \subset \bigcup_1^\infty (a_j, b_j)$ and $\mu(E) \leq \sum_1^\infty \mu\left((a_j, b_j)\right) + \epsilon$. If $U = \bigcup_1^\infty (a_j, b_j)$ 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.

1.3.4 Cantor Set and Cantor Function

1.4 Measurable Functions

1.4.1 Definitions

Definition 1.16

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 1.11

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 1.17

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 1.12

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 1.11. □

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

Definition 1.18

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 1.5 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. $\square$

The following criterion is more applicable.

Exercise 1.6 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 $f^{-1}(\{-\infty\}) \in \mathcal{M}$. $\square$

1.4.2 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 1.11 (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 1.19

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

1.4.3 Distributions

Definition 1.20

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 1.7 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 1.21

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 1.9

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 2 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.

2.1 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 $[0, \infty]$.

2.1.1 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 2.1 χ_E is measurable if and only if E is measurable.

Definition 2.1

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 2.1 (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 $\text{range}(f) = \{c_1, \cdots, c_N\}$, 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 2.1 (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 2.2 (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 2.2

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 2.1 (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

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

2.1.2 Stage 2: Nonnegative Functions

Definition 2.3

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 2.2

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

1. *$\int f \leq \int g$ whenever $f \leq g$,*
2. *$\int cf = c \int f$ for all $c \geq 0$,*

3. $\int(f+g)=\int f+\int g$. (use MCT to prove)

Theorem 2.3 (MCT)

If $\{f_n\}$ is a sequence in L^+ with $f_j \leq f_{j+1}$ 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.

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

The partial sums of nonnegative functions form an increasing sequence.

Theorem 2.4

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 2.3

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

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

Theorem 2.5 (Fatou's Lemma)

If $\{f_n\}$ is any sequence in L^+, then
$$\int \liminf f_n \leq \liminf \int f_n.$$

Proof
$$\int \liminf f_n = \lim_{k\to\infty} \int \inf_{n\geq k} f_n \leq \liminf \int f_n.$$
The last inequality follows from $\inf_{n\geq k} f_n \leq f_j \ \forall j \geq k$, then
$$\int \inf_{n\geq k} f_n \leq \int f_j \ \forall j \geq k,$$
hence
$$\int \inf_{n\geq k} f_n \leq \inf_{j\geq k} \int f_j.$$
□

Proposition 2.4

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.

2.1.3 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 2.4

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 2.5

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

2.1.4 Connections Between Measure and Integration

Theorem 2.6

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.

2.2 L^1 Space

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

Proposition 2.6

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

Proposition 2.7

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 2.7 (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 2.8

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 2.9 (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 2.10 (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 2.8 (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, $|f\chi_{A_N}| \leq |f| \in L^1$, so by the dominated convergence theorem

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

Now

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

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

Proposition 2.9 (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

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

and clearly $\mu(E_N) < \infty$. □

Example 2.2 [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

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

The remaining part is $\int_E |f| d\mu$. From basic analysis we know if a sequence of numbers $|c_n| \le M$ 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 **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

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

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

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

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

□

2.3 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 2.5 (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 2.3 If $f \in L^1(\mathbb{R}^d)$, then $\widehat{f}$ is continuous on $\mathbb{R}^d$.

Proof We need to estimate

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

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

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

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

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

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
- $\lim\limits_{h\to 0}\dfrac{F(x+h)-F(x)}{h} = \displaystyle\int \lim_{h\to 0}\frac{f(x+h,y)-f(x,y)}{h}dy$, that is,
$$\frac{dF}{dx}(x) = \int \frac{\partial}{\partial x}f(x,y)dy?$$

We have the following theorem:

Theorem 2.11

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 $|f(x, t)| \le g(x)$ 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 $|(\partial f/\partial t)(x, t)| \le g(x)$ 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,
$$|h_n(x)| \le \sup_{t\in[a,b]}\left|\frac{\partial}{\partial t}f(x,t)\right|\frac{|t_n-t_0|}{|t_n-t_0|} \le g(x),$$
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 2.4 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,

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

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

For the second one,

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

□

2.4 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 2.6

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 2.10

$\bigotimes_{n\in\mathbb{N}} \mathcal{M}_n$ 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 2.11

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 2.1

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 2.12

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. □

2.5 Product Measures

2.5.1 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)

2.5.1.1 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 $\pi(E) = \mu(A)\nu(B)$.

2.5.1.2 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,

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

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

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

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

2.5.1.3 Step 3: (outer measure, power set)

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

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

2.5.1.4 Step 4: (measure, σ-algebra)

Let's copy the proposition 2.10: $\mathcal{M} \otimes \mathcal{N}$ 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 $A \times B$.

2.5.2 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 2.7

*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

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

♣

Proposition 2.13 (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 2.8 (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 2.1 (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 1.4.

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 2.12

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. $\int \mu((E_1)^y) d\nu(y) \le \mu(X)\nu(Y) < \infty$, that is, $\mu((E_1)^y) \in L^1(\nu)$, and $|\mu((E_n)^y)| \le \mu((E_1)^y)$.

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 2.13 (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 → nonnegative simple functions → nonnegative measurable functions.

By Theorem 2.12, 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.

2.5.3 Examples of Fubini's Theorem

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

Example 2.5 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.

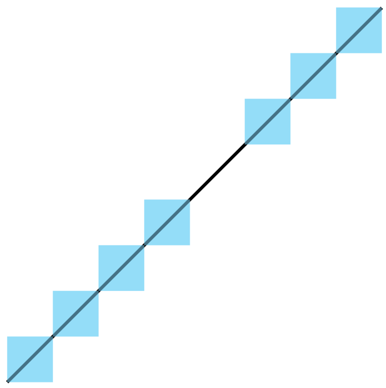


Figure 2.1: diagonal of $[0,1]\times[0,1]$

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,

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

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 2.6 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 2.7 (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 2.8 (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,

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

2.6 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 2.9 (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 $\int f dm$.* ♣

2.6.1 Translation-Invariance

Theorem 2.14 (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 1.6). 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 1.3), then they induces the same measure on $\mathcal{B}_{\mathbb{R}}$ by Theorem 1.6. 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 1.6. 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 (1.6), 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 $N_0 \subset N \in \mathcal{B}_{\mathbb{R}^d}$

---

[2]If you cannot understand this sentence, see section of product measures (2.5) or watch my video.

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

with $m(N)=0$. Now Suppose $E\in\mathcal{L}^d$, then $E=G\cup N$, where G is a Borel set and $m(N)=0$, then N is a subset of some Borel null set M, then $m(N+a)\le m(M+a)=0$. Therefore

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

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 $\int(f\circ\tau_a)dm=\int fdm$ 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 $f\in L^1(m)$. □

2.6.2 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 2.15 (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 $f\ge 0$ or $f\in L^1(m)$, then

$$\int f(x)dx=|\det T|\int f\circ T(x)dx. \tag{2.1}$$

♡

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 (2.1) 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,
$$\begin{aligned}\int(f\circ T_1)(x)dx&=\int\cdots\left(\int f(x_1,\cdots,cx_j,\cdots,x_d)dx_j\right)dx_1\cdots dx_d\\ &=\frac{1}{|c|}\int\cdots\int f(x_1,\cdots,x_j,\cdots,x_d)dx_jdx_1\cdots dx_d\\ &=\frac{1}{|c|}\int f(x)dx\\ &=\frac{1}{|\det T_1|}\int f(x)dx.\end{aligned}$$
- 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 2.2

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 2.3

Lebesgue measure is invariant under rotations. ♡

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

2.6.3 Differentiable Change of Variable

Theorem 2.16

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*

$$\int_{G(\Omega)} f(x)dx = \int_{\Omega} f \circ G(x)|detD_x G|dx.$$

2. *If $E \subset \Omega$ and $E \in \mathcal{L}^d$, then $G(E) \in \mathcal{L}^d$ and*

$$m(G(E)) = \int_E |\det D_x G| dx.$$

♡

2.7 Integration in Polar Coordinates

In $\mathbb{R}^2$, we have the usual polar coordinates

$$x = r\cos\theta, y = r\sin\theta,$$

and calculus tells us

$$dxdy = rdrd\theta.$$

Let the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$. We can represent any point $x \in \mathbb{R}^n \setminus \{0\}$ by its direction and length.

Definition 2.10

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

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

Intuitively, Φ decomposes x into its length and direction. Moreover, Φ is a continuous bijection, so its continuous inverse is $\Phi^{-1}(r, x') = rx'$. This maps gives a homeomorphism of $\mathbb{R}^n \setminus \{0\}$ and the cylinder.

Example 2.9 If $n = 2$, then $\Phi((\sqrt{3}, 1)) = \left(2, (\sqrt{3}/2, 1/2)\right)$, $\Phi^{-1}(2, (\sqrt{3}/2, 1/2)) = 2(\sqrt{3}/2, 1/2) = (\sqrt{3}, 1)$.

The range of Φ^{-1} is contained in $\mathbb{R}^n$ (on which we have defined Lebesgue measure), while the domain of Φ^{-1} is $(0,\infty) \times S^{n-1}$. 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. Hence, we have a Borel measure m_* on $(0, \infty) \times S^{n-1}$ induced by Φ:

$$m_*(E) = m(\Phi^{-1}(E)).$$

Next, we decompose m_* to get a surface measure on S^{n-1}. Let us first define the measure $\rho = \rho_n$ on $(0, \infty)$ by $\rho(E) = \int_E r^{n-1} dr$, which can also be written as $d\rho_n = r^{n-1} dr$.

Example 2.10 Let $n = 2$, then $d\rho_2 = r dr$, thus $d\theta$ will be our surface measure.

Theorem 2.17

There is a unique Borel measure $\sigma = \sigma_{n-1}$ on S^{n-1} 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

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

♡

Proof

Example 2.11 (Gaussian) If $a > 0$, then

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

Solution

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

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 2.11 (Gamma Function)

Define

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

♣

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

Proposition 2.14

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

♠

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 surface measure σ^{n-1} on the sphere S^{n-1}. One checks easily that σ^{n-1} is the weak limit of the measures $\delta^{-1} \mathcal{L}^n L(B(0, 1+\delta) \setminus B(0,1))$ as $\delta \to 0$

Chapter 3 Complex Measures

3.1 Total Variation

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. (take f a negative function).

Definition 3.1

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 $\nu : \mathcal{M} \to \mathbb{C}$.*
- *We still refer to a countably additive function $\nu : \mathcal{M} \to \mathbb{R}$ as a complex measure since every real number is complex.*

♣

Example 3.1 Define ν on the Borel subsets of $[-1, 1]$ by

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

Chapter 4 Hausdorff Measures

4.1 Metric Outer Measure

4.2 Hausdorff Measure

Definition 4.1

*For any subset E of $\mathbb{R}^d$, we define the **exterior α-dimensional Hausdorff measure** of E by*

$$m_\alpha^*(E) = \lim_{\delta \to 0} \inf \left\{ \sum_k d(F_k)^\alpha : E \subset \bigcup_{k=1}^\infty F_k, d(F_k) \le \delta \ \forall k \right\},$$

where d stands for diameter. ♣

For each δ, we have the quantity

$$\mathcal{H}_\alpha^\delta(E) = \inf \left\{ \sum_k d(F_k)^\alpha : E \subset \bigcup_{k=1}^\infty F_k, d(F_k) \le \delta \ \forall k \right\}.$$

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 $\mathcal{H}_\alpha^\delta(E) \le m_\alpha^*(E)$ for all $\delta > 0$.

Proposition 4.1 (monotonicity)

If $E_1 \subset E_2$, then $m_\alpha^*(E_1) \le m_\alpha^*(E_2)$. ♠

Proof Since any cover of E_2 is also a cover of E_1, taking infimum leads to the inequality.

Proposition 4.2 (sub-additivity)

For any countable family $\{E_j\} \subset \mathbb{R}^d$,

$$m_\alpha^* \left(\bigcup_{j=1}^\infty E_j \right) \le \sum_{j=1}^\infty m_\alpha^*(E_j).$$

♠

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 4.3 (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 4.2 (Hausdorff Measure)

*The restriction of m_α^* to the Borel sets is called the **α-dimensional Hausdorff measure**, denoted m_α.*

Proposition 4.4

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 4.5

Hausdorff measure is invariant under translations

$$m_\alpha(E + h) = m_\alpha(E) \quad \textit{for all } h \in \mathbb{R}^d,$$

and rotations

$$m_\alpha(RE) = m_\alpha(E),$$

where R is a rotation in $\mathbb{R}^d$. Moreover, it scales as follows:

$$m_\alpha(\lambda E) = \lambda^\alpha m_\alpha(E) \quad \textit{for all } \lambda > 0.$$

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 4.6

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 4.7

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 $d(F) \le \delta$. If $\beta > \alpha$, then

$$d(F)^\beta = d(F)^{\beta-\alpha} d(F)^\alpha \le \delta^{\beta-\alpha} d(F)^\alpha.$$

Consequently

$$\mathcal{H}_\beta^\delta(E) \le \delta^{\beta-\alpha}\mathcal{H}_\alpha^\delta(E) \le \delta^{\beta-\alpha} m_\alpha^*(E).$$

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, $\{\beta \le d : m_\beta^*(E) = \infty\}$ has the supremum.

4.3 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.

4.3.1 Examples

The Cantor set

Theorem 4.1

The Cantor set $\mathcal{C}$ has strict Hausdorff dimension $\alpha = \log 2/\log 3$

Definition 4.3

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 4.1

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

$$d(f(E\cap F_k)) \le Md(E \cap F_k)^\gamma \le Md(F_k)^\gamma.$$

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 4.2

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

$$|F_n(x) - F_n(y)| \le \left(\frac{3}{2}\right)^n |x-y|.$$

The approximating sequence also satisfies $|F(x) - F_n(x)| \le 1/2^n$. Then

$$\begin{aligned} |F(x) - F(y)| &\le |F_n(x) - F_n(y)| + |F(x) - F_n(x)| + |F(y) - F_n(y)| \\ &\le \frac{3^n}{2^n}|x-y| + \frac{2}{2^n}. \end{aligned}$$

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

$$|F(x) - F(y)| \le \frac{c}{2^n} = \frac{c}{3^{(\log_3 2)n}} := c(3^{-n})^\gamma \le M|x-y|^\gamma,$$

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): $m_\gamma(\mathcal{C}) \le 1$.

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

$$d(C_k)^\gamma \le 2^n(3^{-n})^\gamma = 2^n 2^{-n} = 1,$$

hence $m_\gamma(\mathcal{C}) \leq 1$.
Part (II): $m_\gamma(\mathcal{C}) > 0$.
Applying Lemma 4.1 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 5 Topology in Analysis

This chapter is a copy of Chapter 4 of Real Analysis, Folland.

5.1 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 5.1 (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 5.1

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 5.2 (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 $\mathbb{T}_2$.*

♣

5.1.1 Base

Definition 5.3 (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 5.2 (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 5.3

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 5.4 (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.

5.2 Continuous Maps

5.2.1 Weak and Product Topologies

Definition 5.4 (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 5.5

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 5.5 (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 5.6

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 5.7

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 5.8

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. ♠

5.2.2 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 5.9

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

$$|f(y) - f(x)| \le |f(y) - f_n(y)| + |f_n(y) - f_n(x)| + |f_n(x) - f(x)| < \varepsilon.$$

□

Example 5.1 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 5.1 (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 5.2 (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 $F|A = f$. ♡

5.3 Nets

Definition 5.6 (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 5.7 (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 5.2

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 5.8

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 5.10

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 5.11 (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 $f(x)$. □

♠

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 $f(x)$. □

5.4 Locally Compact Hausdorff Spaces

Definition 5.9 (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.*

♣

5.4.1 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 5.12

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 5.13

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 $K \subset V \subset \overline{V} \subset U$.

5.4.2 Functions of Compact Support

Proposition 5.14

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 αth ***projection*** *or* ***coordinate map*** $\pi_\alpha : X \to X_\alpha$ *by $\pi_\alpha(f) = f(\alpha)$.* ♣