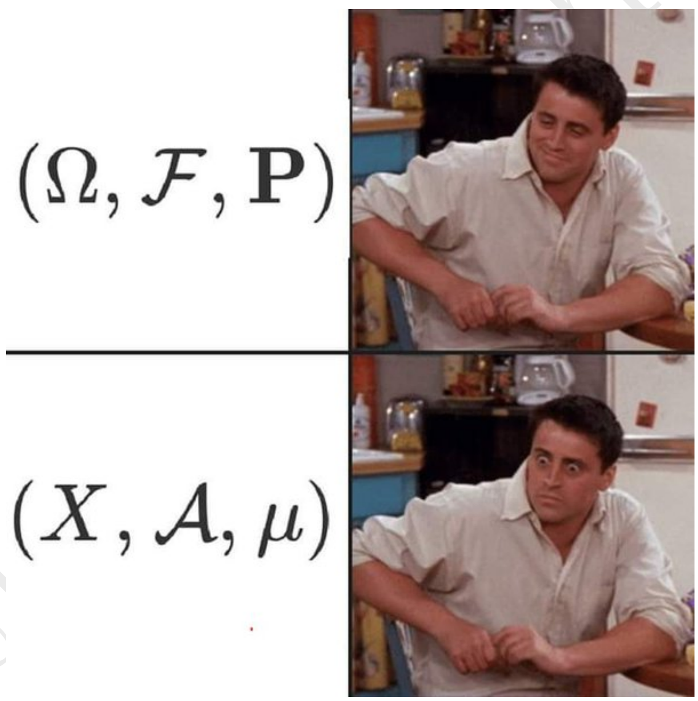
Fantuan's Academia

Fantuan's Analysis Notes series Vol.3

Notes on Measure Theory and Probability

Author: Jingxuan Xu



August 27, 2025

Contents

All the Sections with * are hard sections and can be skipped without losing coherence.

This note is referenced on lectures:

- **IMPA lecture on Measure Theory and Probability Theory** by Prof. Claudio Landim.
- **MATH365 Measure Theory course** taught by Dr. Alexei B. Piunovskiy at University of Liverpool.
- **S&DS600 Advanced Probability course** taught by Prof. Sekhar Tatikonda at Yale University.
- **MATH520 Measure Theory and Integration course** taught by Prof. Charles Smart at Yale University.

and referenced on books:

- **Measure Theory 2nd ed.** by Donald L. Cohn.
- **Measure, Integration and Real Analysis** by Sheldon Axler.
- **Real Analysis: Modern Techniques and Their Applications 2nd ed.** by Gerald B. Folland.
- **Probability: Theory and Examples 5th ed.** by Rick Durrett.

PART I: Measure Theory

Chapter 1

Introduction and Setup

1.1 Why Measure Theory?

The main reason for the development of measure theory is that, **Riemann Integral** has many deficiencies so that we need a more delicate approach to the integration theory to solve these problems.

Example 1.1.1: Some Functions are not Riemann Integrable

The **Dirichlet Function** $f : [0,1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1, & \text{if } x \text{ is rational} \\ 0, & \text{if } x \text{ is irrational} \end{cases}$$

is not Riemann Integrable, since for any partition interval $[a,b] \subseteq [0,1]$ with $a < b$, we will have

$$\inf_{[a,b]} f = 0 \text{ and } \sup_{[a,b]} f = 1$$

The supremum of lower Riemann Sum and infimum of Upper Riemann Sum would never be equal to each other. Therefore, the integral does not exist. However, consider that $\mathbb{Q}$ is countable and $\mathbb{R}\backslash\mathbb{Q}$ is uncountable, we will reasonably guess that the 'area' under this curve should be 0 in some sense. So we need to fix this.

Example 1.1.2: Riemann Integral dos not preserve pointwise limit

Let $r_1, r_2, \cdots$ be a sequence that includes all rational numbers in $[0,1]$ exactly once. For $k \in \mathbb{N}$, definef_k : $[0,1] \to \mathbb{R}$ by

$$f_k(x) = \begin{cases} 1, & \text{if } x \in \{r_1, \cdots, r_k\} \\ 0, & \text{Otherwise} \end{cases}$$

Each f_k is Riemann Integrable, but the pointwise limit is the Dirichlet Function, which is not integrable!

You can see that Riemann Integral does not behave very well on pathological functions. In this note, we will fix these problems using a new integral method called the **Lebesgue Integral**, and derive new theories based on this.

1.2 Motivation: A Non-measurable Set

Suppose that we want to define a *set function* $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$, where $\mathcal{P}(\mathbb{R})$ denotes the power set of $\mathbb{R}$, so that for any set $A \in \mathcal{P}(\mathbb{R})$, it outputs an intuitive 'length', or 'measure' of that set. To make it consistent with our common sense, it should satisfy some properties:

1. *(Interval Measure)*: For any kinds of intervals $(a, b), (a, b], [a, b)$ or $[a, b]$, where $a, b \in (-\infty, \infty)$, we have
$$\mu((a,b)) = \mu((a,b]) = \mu([a,b)) = \mu([a,b]) = b - a$$
2. *(Monotonicity)*: If $A \subseteq B \subseteq \mathbb{R}$, then $\mu(A) \leqslant \mu(B)$.
3. *(Translation Invariant)*: For any $A \subseteq \mathbb{R}$ and $t \in \mathbb{R}$, we have $\mu(t + A) = \mu(A)$
4. *(Finite Additivity)*: For finitely many disjoint sets $A_1, \cdots, A_n$, we have
$$\mu\left(\bigsqcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k)$$

These properties are set to make the function μ be intuitive enough so that there is no pathological behaviors happening. For example, Property 1 set the intuitive 'length' of an interval; Property 2 provides that 'bigger' sets have 'longer' length; Property 3 provides that after a translation, the 'length' is not changed; Property 4 provides that disjoint sets' length can be added together as the length of the union.

For the purpose of analysis and good properties of limiting operations, we want more than finite additivity. Actually, we want the *countable additivity*:

4b. For countably many disjoint sets $A_1, A_2, \cdots$, we have
$$\mu\left(\bigsqcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$$

With this property, limiting operations are easy in hand, so it guarantees some good analytical properties. Note that *countable additivity implies finite additivity*: take $A_{n+1} = A_{n+2} = \cdots = \emptyset$, we can get finite additivity from the countable one. So Property 4 is redundant after we introduced 4b.

Unfortunately, this kind of function does not exist. One of the most famous construction to show this is called the **Vitali Set**, published by Giuseppe Vitali in 1905. The construction of this set is dependent on so-called **Axiom of choice** in set theory. We stated here the form we are going to use.

Axiom 1.2.1: Axiom of Choice

If $\mathcal{E}$ is a set whose elements are disjoint nonempty sets, then there exists a set V that contains exactly one element in each set that is an element of $\mathcal{E}$.

For example, suppose $\mathcal{E} = \{\{1,2\}, \{3,4\}, \{6,7,8\}\}$. Then, it is a set whose elements are disjoint nonempty sets. Then, we can find a set V, for instance, $V = \{1, 3, 8\}$ such that it contains exactly one element in each set that is an element of $\mathcal{E}$. This is intuitive, but when the cardinality of $\mathcal{E}$ is countable, or even uncountable, the statement quickly becomes obscure to validate.

Theorem 1.2.2: Counterexample: Vitali Set

There does not exist a function μ such that it simultaneously satisfies Properties 1,2,3, and 4b stated above.

Proof. We will prove this by contradiction. We will do this by constructing the Vitali set V, so that there is a contradiction on the value of $\mu(V)$.

STEP I: Countable Subadditivity

We first prove that this function satisfies *countable subadditivity*. That is, for countably many sets $A_1, A_2, \cdots$, we have

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) \leqslant \sum_{k=1}^{\infty} \mu(A_k)$$

Note that we do not guarantee that A_k's are disjoint here. To prove this, we construct a list of disjoint set by

$$B_1 = A_1, \quad B_2 = A_2 \backslash A_1, \quad B_3 = A_3 \backslash (A_1 \cup A_2), \quad B_4 = A_4 \backslash (A_1 \cup A_2 \cup A_3) \cdots$$

Then, $\{B_k\}_{k=1}^{\infty}$ is a countable collection of disjoint sets. Moreover, $B_k \subseteq A_k$ for any $k = 1, 2, \cdots$, and

$$\bigcup_{k=1}^{\infty} A_k = \bigsqcup_{k=1}^{\infty} B_k \tag{1.1}$$

Therefore, we can apply countable additivity and monotonicity to get

$$\begin{aligned}
\mu\left(\bigcup_{k=1}^{\infty} A_k\right) &= \mu\left(\bigsqcup_{k=1}^{\infty} B_k\right) && \text{(Equation 4.3)} \\
&= \sum_{k=1}^{\infty} \mu(B_k) && \text{(Countable additivity)} \\
&\leqslant \sum_{k=1}^{\infty} \mu(A_k) && \text{(Monotonicity)}
\end{aligned}$$

STEP II: Equivalence Class

Define the *equivalence relation* on $[-1, 1]$ such that for any $a, b \in [-1, 1]$, $a \sim b$ iff $a - b \in \mathbb{Q}$. It is easy to check the reflexivity, symmetry, and transitivity property of this binary relation, so it is indeed an equivalence relation. The *equivalence class* generated by this equivalence relation is

$$[a] = \{b \in [-1, 1] : a - b \in \mathbb{Q}\}$$

Recall the property of the equivalence class that, for $a, b \in [-1, 1]$, either $[a] \cap [b] = \emptyset$ or $[a] = [b]$. Therefore, this defines a number of disjoint sets on $[-1, 1]$. Clearly $a \in [a]$ for each $a \in [-1, 1]$, so $[-1, 1] = \bigcup_{a \in [-1,1]} [a]$.

STEP III: Axiom of Choice

By the Axiom of Choice, we can construct the set V such that it contains exactly one element in each of the distinct sets in $\{[a] : a \in [-1,1]\}$. Then, $V \cap [a]$ has only one element for each $a \in [-1,1]$. This set V is called the **Vitali Set**.

STEP IV: Lower Bound of μ

Denote $r_1, r_2, \cdots$ the enumeration of all distinct rational numbers on $[-2,2]$, i.e.,

$$[-2,2] \cap \mathbb{Q} = \{r_1, r_2, \cdots\}$$

Then, we can deduce that

$$[-1,1] \subseteq \bigcup_{k=1}^{\infty} (r_k + V) \tag{1.2}$$

This is because, by definition of V, for any $a \in [-1,1]$, there exists a unique $v \in V$ such that $v \in V \cap [a]$. Thus, $a - v \in \mathbb{Q}$. Since $a, v \in [-1,1]$, we have $a - v \in [-2,2]$, so it belongs to one of the $r_k, k = 1, 2, \cdots$. Suppose $a - v = r_q \in \mathbb{Q}$. Then, $a = r_q + v \in r_q + V$. This shows that $[-1,1] \subseteq r_q + V$ for some q, thus it is also contained in the union.

With this Equation 4.4, by Property 1, 2, 3, and countable subadditivity, we have

$$\begin{aligned}
\mu([-1,1]) &= 2 && \text{(Property 1)} \\
&\leqslant \mu\left(\bigcup_{k=1}^{\infty} (r_k + V)\right) && \text{(Monotonicity)} \\
&\leqslant \sum_{k=1}^{\infty} \mu(r_k + V) && \text{(Countable subadditivity)} \\
&= \sum_{k=1}^{\infty} \mu(V) && \text{(Translation invariant)}
\end{aligned}$$

Therefore, $\mu(V) > 0$.

STEP V: Upper Bound of μ

Note that the sets $\{r_k + V\}_{k=1}^{\infty}$ are disjoint. To prove this, suppose there exists an element t such that $t \in (r_j + V) \cap (r_k + V)$ where $j \neq k$, then $t = r_j + v_1 = r_k + v_2$ for some $v_1, v_2 \in V$. This implies $v_1 - v_2 = r_k - r_j \in \mathbb{Q}$. By the definition of V, this could only be the case that $v_1 = v_2$, which implies that $r_j = r_k$, which implies that $j = k$, which is a contradiction.

Since $r_k \in [-2,2]$, $V \subseteq [-1,1]$, we have

$$\bigcup_{k=1}^{\infty} (r_k + V) \subseteq [-3,3]$$

Then, by Property 1,2,3 and 4b, we have

$$\begin{aligned}
\mu\left(\bigcup_{k=1}^{\infty} (r_k + V)\right) &= \sum_{k=1}^{\infty} \mu(r_k + V) && \text{(Countable additivity)} \\
&= \sum_{k=1}^{\infty} \mu(V) && \text{(Translation invariant)} \\
&\leqslant \mu([-3,3]) = 6 && \text{(Property 1 and monotonicity)}
\end{aligned}$$

Since we have already seen $\mu(V) > 0$, the only value $\sum_{k=1}^{\infty} \mu(V)$ can take is ∞, which is a contradiction to this upper bound.

□

This informs us that, if we want to have a 'measure' function on $\mathbb{R}$ that satisfies all these properties, the only thing we can do is to restrict the domain of the function $\mathcal{P}(\mathbb{R})$, such that extremely pathological sets like Vitali set is excluded, and these sets are defined to be 'non-measurable'. This invokes the next section about σ-algebra.

1.3 Classes of Subsets

Which subset should we restrict to? Here is some definitions.

1.3.1 Semi-algebra, Algebra, and σ-algebra

Definition 1.3.1: Semialgebra

Let X be an arbitrary set. A collection $\mathcal{S}$ of subsets of X is a **semi-algebra** on X if

- *Nonempty:* $X \in \mathcal{S}$,
- *Closed under intersection:* $E, F \in \mathcal{S}, \implies E \cap F \in \mathcal{S}$,
- *Complement is a finite disjoint union:* $E \in \mathcal{S}, \implies E^c = \bigsqcup_{j=1}^{n} E_j$, where $E_j \in \mathcal{S}$ and are disjoint.

Example 1.3.2: Examples of Semi-algebra

The only example of semi-algebra that we are interested in is **the collection of all left-open, right-closed intervals**. We will call it *half-open interval semi-algebra* later on for simplicity. On the real line $\mathbb{R}$, we have

$$\mathcal{S} = \{(a,b] : a < b,\ a, b, \in \mathbb{R}\} \cup \{(-\infty, b] : b \in \mathbb{R}\} \cup \{(a, \infty) : a \in \mathbb{R}\} \cup \emptyset \cup \mathbb{R}$$

To prove this is a semi-algebra, first we have $\emptyset \in \mathcal{S}$. Second, If $E, F \in \mathcal{S}$, you can easily see that the intersection of two half-open interval is still half-open interval. Finally, the complement of a half-open interval is either a ray or a union of two rays, which is a finite disjoint union of elements in $\mathcal{S}$.

Definition 1.3.3: Algebra/Field

Let X be an arbitrary set. A collection $\mathcal{R}$ of subsets of X is an **algebra** (or a **field**) on X if

1. **Nonempty:** $\mathcal{R} \neq \emptyset$
2. **Closed under complement:** If $E \in \mathcal{R}$, then $E^c \in \mathcal{R}$
3. **Closed under finite union:** If $E_1, E_2, \cdots, E_n \in \mathcal{R}$, then $\bigcup_{i=1}^{n} E_i \in \mathcal{R}$

Note that an algebra has a few other properties:

(a) **Closed under finite intersection:** If $E_1, E_2, \cdots, E_n \in \mathcal{R}$, then $\bigcap_{i=1}^{n} E_i \in \mathcal{R}$

(b) $\emptyset \in \mathcal{R}$ and $X \in \mathcal{R}$

Proof. (a) Note that

$$\bigcap_{i=1}^{n} E_i = \left(\bigcup_{i=1}^{n} E_i^c\right)^c$$

By De Morgan's Law. Since $\mathcal{R}$ is closed under finite union and complement, it follows that $\bigcap_{i=1}^{n} E_i \in \mathcal{R}$.
(b) Since $\mathcal{R}$ is nonempty, there exists a set $E \in \mathcal{R}$. Since $\mathcal{R}$ is closed under finite union, finite intersection and complement, we have

$$\emptyset = E \cap E^c \in \mathcal{R} \quad \text{and} \quad X = E \cup E^c \in \mathcal{R}$$

$\square$

Definition 1.3.4: σ-Algebra/σ-Field

Let X be an arbitrary set. A collection $\mathcal{A}$ of subsets of X is a **σ-algebra** (or a **σ-field**) on X if

1. **Nonempty:** $\mathcal{A} \neq \emptyset$
2. **Closed under complement:** If $E \in \mathcal{A}$, then $E^c \in \mathcal{A}$
3. **Closed under countable union:** If $E_1, E_2, E_3, \cdots \in \mathcal{A}$, then $\bigcup_{i=1}^{\infty} E_i \in \mathcal{A}$

Note that a σ-algebra has a few other properties:

(a) **Closed under countable intersection:** If $E_1, E_2, E_3, \cdots \in \mathcal{A}$, then $\bigcap_{i=1}^{\infty} E_i \in \mathcal{A}$

(b) $\emptyset \in \mathcal{A}$ and $X \in \mathcal{A}$

(c) **Closed under finite union and intersection:** If $E_1, E_2, \cdots, E_n \in \mathcal{A}$, then $\bigcup_{i=1}^{n} E_i \in \mathcal{A}$ and $\bigcap_{i=1}^{n} E_i \in \mathcal{A}$

Proof. (a) Note that

$$\bigcap_{i=1}^{\infty} E_i = \left(\bigcup_{i=1}^{\infty} E_i^c\right)^c$$

By De Morgan's Law. Since $\mathcal{A}$ is closed under countable union and complement, it follows that $\bigcap_{i=1}^{n} E_i \in \mathcal{A}$.
(b) Since $\mathcal{A}$ is nonempty, there exists a set $E \in \mathcal{A}$. Since $\mathcal{A}$ is closed under countable union, countable intersection and complement, we have

$$\emptyset = E \cap E^c \cap E \cap E \cap \cdots \in \mathcal{A} \quad \text{and} \quad X = E \cup E^c \cup E \cup E \cup \cdots \in \mathcal{A}$$

(c) The closed under finite union and intersection can be naturally deduced by closed under countable union and intersection, by letting $E_i = \emptyset$ for $i > n$ to prove closed under finite union, and letting $E_i = X$ for $i > n$ to prove closeness under finite intersection. $\square$

Finally, note that all σ-algebra are algebra, and all algebra are semi-algebra. Here we see some examples of algebra and σ-algebra. We omit the verification here.

Example 1.3.5: Examples of Algebra and σ-Algebra

σ-algebra:

- Let X be a set, and $\mathcal{A}$ be the power set $\mathcal{P}(X)$, i.e., the collection of all subsets of X. Then, $\mathcal{A}$ is a σ-algebra on X.
- Let X be a set, and $\mathcal{A} = \{\emptyset, X\}$. Then $\mathcal{A}$ is a σ-algebra.
- Let X be a set, whatever countable or uncountable, and let $\mathcal{A}$ be the collection of all subsets E of X such that either E or E^c is countable. Then $\mathcal{A}$ is a σ-algebra.

Algebra but not a σ-algebra:

- Let X be an infinite set. Let $\mathcal{R}$ be the collection of all subsets E of X such that either E or E^c is finite. Then $\mathcal{R}$ is an algebra, but not a σ-algebra.
- Let $\mathcal{R}$ be the collection of all subsets of $\mathbb{R}$ that are unions of finitely many intervals of the form $(a, b]$, $(a, +\infty)$ or $(-\infty, b]$. Then, $\mathcal{R}$ is an algebra, but not a σ-algebra (for example, the open interval (a, b) is a countable union of these sets, but itself is not in $\mathcal{R}$).

Neither algebra nor σ-algebra:

- Let X be an infinite set. Let $\mathcal{C}$ be the collection of all finite subsets of X. Then it is not an algebra since X is not contained in $\mathcal{C}$.
- Let X be an uncountable set. Let $\mathcal{C}$ be the collection of all countable subsets of X. Then, it is not an algebra since X is not contained in $\mathcal{C}$.

1.3.2 Generated Classes of Subsets

We can construct new σ-algebras from the existing ones. The most naive method is using intersection. The proof of the next proposition is extremely trivial.

Proposition 1.3.6: Intersection of σ-algebra

Let X be a set. Then the intersection of an arbitrary nonempty collection of σ-algebras on X is again a σ-algebra on X.

Proof. Let $\{\mathcal{A}_i\}_{i\in I}$ be a nonempty collection of σ-algebras on X, and let $\mathcal{A}$ be the intersection of the σ-algebras $\mathcal{A}_i$, $\mathcal{A} = \bigcap_{i\in I} \mathcal{A}_i$

- Since each $\mathcal{A}_i$ is a σ-algebra, $X \in \mathcal{A}_i$ for each $i \in I$. We have $X \in \mathcal{A}$, $\mathcal{A}$ is not empty.
- Suppose $E \in \mathcal{A}$. Then $E \in \mathcal{A}_i$ for each $i \in I$. Since each $\mathcal{A}_i$ is a σ-algebra, $E^c \in \mathcal{A}_i$ for each $i \in I$. Therefore, $E^c \in \mathcal{A}$. The collection $\mathcal{A}$ is thus closed under complement.
- Suppose $E_1, E_2, E_3, \cdots \in \mathcal{A}$. Then, $E_1, E_2, E_3, \cdots \in \mathcal{A}_i$ for each $i \in I$. Since each $\mathcal{A}_i$ is a σ-algebra, $\bigcup_{i=1}^\infty E_i \in \mathcal{A}_i$ for each $i \in I$. Therefore, $\bigcup_{i=1}^\infty E_i \in \mathcal{A}$. The collection $\mathcal{A}$ is thus closed under countable union.

In conclusion, we have $\mathcal{A}$ is a σ-algebra. □

Note: The union of σ-algebras can fail to be a σ-algebra. For example, consider the set $X = \{1, 2, 3, 4\}$. The collection

$$\mathcal{A} = \{\emptyset, \{1\}, \{2, 3, 4\}, X\}, \quad \mathcal{B} = \{\emptyset, \{2\}, \{1, 3, 4\}, X\}$$

are then σ-algebra on X. The union

$$\mathcal{A} \cup \mathcal{B} = \{\emptyset, \{1\}, \{2\}, \{2, 3, 4\}, \{1, 3, 4\}, X\}$$

fail to be a σ-algebra since the union $\{1\} \cup \{2\} = \{1, 2\}$ does not belong to $\mathcal{A} \cup \mathcal{B}$.

The preceding proposition shows that, there exists a smallest σ-algebra that contains some subset of $\mathcal{P}(X)$.

Corollary 1.3.7: Smallest containing σ-algebra

Let X be a set. Let $\mathcal{E}$ be a collection of subsets of X. Then, there exists a unique smallest σ-algebra on X that contains $\mathcal{E}$. This smallest σ-algebra is exactly the intersection of all σ-algebra on X that contains $\mathcal{E}$.

Proof. Let $\{\mathcal{A}_i\}_{i \in I}$ be the collection of all σ-algebra on X that contains $\mathcal{E}$. Let $\mathcal{A} = \bigcap_{i \in I} \mathcal{A}_i$. By Proposition **??**, $\mathcal{A}$ is a σ-algebra. It is naturally smaller than all σ-algebra that contains $\mathcal{E}$. □

This smallest σ-algebra is so important that we give it a name and a notation.

Definition 1.3.8: Generated σ-algebra

The smallest σ-algebra that contains $\mathcal{E}$ is called the σ-algebra **generated** by $\mathcal{E}$, and is denoted by $\sigma(\mathcal{E})$.

The next trivial lemma would be useful later.

Lemma 1.3.9: Monotone Class Property

Let X be a set. Let $\mathcal{E}$ and $\mathcal{F}$ be two arbitrary subset of $\mathcal{P}(X)$. If $\mathcal{E} \subseteq \sigma(\mathcal{F})$, then $\sigma(\mathcal{E}) \subseteq \sigma(\mathcal{F})$.

Proof. Since $\sigma(\mathcal{F})$ is a σ-algebra that contains $\mathcal{E}$, it also contains the smallest σ-algebra that contains $\mathcal{E}$, i.e., $\sigma(\mathcal{E})$. □

We can extend this notation to more general cases. For example, we can define $\mathcal{R}(\mathcal{E})$ by the **smallest algebra that contains** $\mathcal{E}$, i.e., the intersection of all algebras that contains $\mathcal{E}$. Follow the same proof from Proposition 1.3.6, we can see that this is indeed an algebra.
Note that for generated σ-algebra, we typically do not have closed form representation, which is a main difficulty when dealing with these objects. However, there is a very clear representation for the algebra generated by semi-algebra, which is one of the basis towards definition of Lebesgue measure.

Theorem 1.3.10: Finite Disjoint Union of Semi-algebra is Algebra

Let X be an arbitrary set. Let $\mathcal{S} \subseteq \mathcal{P}(X)$ be a semi-algebra. Let $\mathcal{R}$ be the algebra generated by $\mathcal{S}$. Then,

$$A \in \mathcal{R} \quad \Longleftrightarrow \quad A \text{ can be written as finite disjoint union of memebers in } \mathcal{S}.$$

Proof.
($\Longleftarrow$) This direction is trivial. Suppose $A = \bigsqcup_{j=1}^n E_j$ where $\{E_j\}_{j=1}^n \subseteq \mathcal{S}$ and E_j's are disjoint. Since $E_j \in \mathcal{S}$ for each j, and $\mathcal{R}$ contains $\mathcal{S}$, we have $E_j \in \mathcal{R}$ for each j as well. Since algebra is closed under finite union, we have $A = \bigcup_{j=1}^n E_j \in \mathcal{R}$.
($\Longrightarrow$) This is the main part of the proof. Suppose $A \in \mathcal{R}$. We first define

$$\mathcal{B} = \left\{ \bigsqcup_{j=1}^n F_j : F_j \in \mathcal{S}, F_j \text{ disjoint} \right\}$$

If we prove that $\mathcal{B}$ is an algebra, and $\mathcal{B} \supseteq \mathcal{S}$, since $\mathcal{R}$ is the smallest algebra that contains $\mathcal{S}$, we will have $\mathcal{B} \supseteq \mathcal{R}$. This then indicates that all elements in $\mathcal{R}$ is a finite disjoint union of members in $\mathcal{S}$. Note that $\mathcal{B} \supseteq \mathcal{S}$ is trivial by the definition of $\mathcal{B}$. The only thing we need to prove is that $\mathcal{B}$ is an algebra.

First, $X \in \mathcal{B}$ because $X \in \mathcal{S}$ by the fact that $\mathcal{S}$ is a semi-algebra.

Second, let $A, B \in \mathcal{B}$. We want to show that $A \cap B \in \mathcal{B}$. We can express A and B as

$$A = \bigsqcup_{j=1}^n E_j, E_j \in \mathcal{S}, E_j \text{ disjoint} \quad \text{and} \quad B = \bigsqcup_{k=1}^m F_k, F_k \in \mathcal{S}, F_k \text{ disjoint}$$

Then,

$$A \cap B = \left(\bigsqcup_{j=1}^n E_j \right) \cap \left(\bigsqcup_{k=1}^m F_k \right) = \bigsqcup_{j=1}^n \bigsqcup_{k=1}^m (E_j \cap F_k) \in \mathcal{B}$$

since $E_j \cap F_k \in \mathcal{S}$ and are disjoint.

Finally, let $A \in \mathcal{B}$, we want to show that $A^c \in \mathcal{B}$. Using the expression from above, we have

$$A^c = \left(\bigsqcup_{j=1}^n E_j \right)^c = \bigcap_{j=1}^n E_j^c$$

Since $E_j \in \mathcal{S}$, we can write E_j^c as a finite disjoint union of elements in $\mathcal{S}$ such that

$$E_j^c = \bigsqcup_{k_j=1}^{l_j} F_{j,k_j}, \text{ where } F_{j,k_j} \in \mathcal{S} \text{ and disjoint for each } j$$

Then, we can write A^c as

$$\begin{aligned} A^c = \bigcap_{j=1}^n E_j^c &= \left(\bigsqcup_{k_1=1}^{l_1} F_{1,k_1} \right) \cap \left(\bigsqcup_{k_2=1}^{l_2} F_{2,k_2} \right) \cap \cdots \cap \left(\bigsqcup_{k_n=1}^{l_n} F_{n,k_n} \right) \\ &= \bigsqcup_{k_1=1}^{l_1} \bigsqcup_{k_2=1}^{l_2} \cdots \bigsqcup_{k_n=1}^{l_n} (F_{1,k_1} \cap F_{2,k_2} \cap \cdots \cap F_{n,k_n}) \in \mathcal{B} \end{aligned}$$

since $F_{1,k_1} \cap F_{2,k_2} \cap \cdots \cap F_{n,k_n} \in \mathcal{S}$ and are disjoint. $\square$

1.4 Borel σ-Algebra

We use the preceding theory about generated σ-algebra to define an important family of σ-algebra.

Definition 1.4.1: Borel σ-Algebra/Borel Set

Let X be any topological space. The σ-algebra generated by the collection of all open sets (or equivalently, all closed sets) in X is called the **Borel σ-algebra** on X, denoted by $\mathcal{B}(X)$. Its elements are called **Borel sets**.

Note that the equivalence of the two definition using open sets and closed sets is because, a σ-algebra containing all open sets must contain all closed sets since it is closed under complement. The Borel σ-algebra on $\mathbb{R}$ would be in special importance. The next proposition shows other ways of defining $\mathcal{B}(\mathbb{R})$.

Proposition 1.4.2: Alternative Definitions of $\mathcal{B}(\mathbb{R})$

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

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

Proof. We will prove this iteratively. This proof is also trivial.

- Since open intervals are open sets, $\mathcal{E}_1$ is contained in the collection of all open sets. Therefore $\mathcal{B}(\mathbb{R}) \supseteq \mathcal{E}_1$. By Lemma 1.3.9, we have $\mathcal{B}(\mathbb{R}) \supseteq \sigma(\mathcal{E}_1)$.
- Since each closed interval can be written as countable intersection of open intervals, for example, by

$$[a,b] = \bigcap_{k=1}^{\infty} \left(a - \frac{1}{k}, b + \frac{1}{k}\right)$$

  closed intervals are contained in the σ-algebra generated by $\mathcal{E}_1$. Therefore, $\sigma(\mathcal{E}_1) \supseteq \mathcal{E}_2$. By Lemma 1.3.9, we have $\sigma(\mathcal{E}_1) \supseteq \sigma(\mathcal{E}_2)$.
- Similarly, we can write half-open intervals as

$$(a,b] = \bigcap_{k=1}^{\infty} \left[a + \frac{1}{k}, b\right], \quad [a,b) = \bigcap_{k=1}^{\infty} \left(a - \frac{1}{k}, b - \frac{1}{k}\right]$$

  Therefore, $\sigma(\mathcal{E}_2) \supseteq \mathcal{E}_3$ and $\sigma(\mathcal{E}_3) \supseteq \mathcal{E}_4$. By Lemma 1.3.9, we have $\sigma(\mathcal{E}_2) \supseteq \sigma(\mathcal{E}_3) \supseteq \sigma(\mathcal{E}_4)$.
- We can write sets in $\mathcal{E}_5$ and $\mathcal{E}_8$ as

$$(a,+\infty) = \bigcap_{k=1}^{\infty} \left[a + \frac{1}{k}, k\right), \quad (-\infty,a] = (a,+\infty)^c$$

Therefore, $\sigma(\mathcal{E}_4) \supseteq \mathcal{E}_5$ and $\sigma(\mathcal{E}_5) \supseteq \mathcal{E}_8$. By Lemma 1.3.9, we have $\sigma(\mathcal{E}_4) \supseteq \sigma(\mathcal{E}_5) \supseteq \sigma(\mathcal{E}_8)$.

- We can write sets in $\mathcal{E}_6$ and $\mathcal{E}_7$ as

$$(-\infty, a) = \bigcap_{k=1}^{\infty} \left(-\infty, a - \frac{1}{k}\right], \quad [a, \infty) = (-\infty, a)^c$$

Therefore, $\sigma(\mathcal{E}_8) \supseteq \mathcal{E}_6$ and $\sigma(\mathcal{E}_6) \supseteq \mathcal{E}_7$. By Lemma 1.3.9, we have $\sigma(\mathcal{E}_8) \supseteq \sigma(\mathcal{E}_6) \supseteq \sigma(\mathcal{E}_7)$.

- Finally, since each open interval can be written as

$$(a, b) = \left(\bigcap_{k=1}^{\infty} \left[a + \frac{1}{k}, \infty\right)\right) \cap [b, \infty)^c$$

and each open set in $\mathbb{R}$ can be written as countable union of disjoint open intervals. Therefore, all open sets are contained in the σ-algebra generated by $\mathcal{E}_7$. This means that $\sigma(\mathcal{E}_7) \supseteq \mathcal{B}(\mathbb{R})$.

In conclusion, we have, iteratively,

$$\mathcal{B}(\mathbb{R}) \supseteq \sigma(\mathcal{E}_1) \supseteq \sigma(\mathcal{E}_2) \supseteq \sigma(\mathcal{E}_3) \supseteq \sigma(\mathcal{E}_4) \supseteq \sigma(\mathcal{E}_5) \supseteq \sigma(\mathcal{E}_8) \supseteq \sigma(\mathcal{E}_6) \supseteq \sigma(\mathcal{E}_7) \supseteq \mathcal{B}(\mathbb{R})$$

All these σ-algebras are equivalent. $\square$

There is a standard terminology for the level of hierarchy. A countable intersection of open sets is called a **$\mathbf{G}_\delta$-set**. A countable union of closed sets is called an **$\mathbf{F}_\sigma$-set**. Similarly, we can define a countable union of G_δ-set as $G_{\delta\sigma}$-set, a countable intersection of F_σ-set as $F_{\sigma\delta}$-set, $\cdots$.

Chapter 2

Measure and Extension of Measures

2.1 Measures

2.1.1 Definition and Examples

Definition 2.1.1: Measure

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

1. $\mu(\emptyset) = 0$.
2. **Countable Additivity:** If $\{E_i\}_{i\in\mathbb{N}}$ is a sequence of disjoint sets in $\mathcal{A}$, then $\mu(\bigcup_i E_i) = \sum_i \mu(E_i)$

Relative to this terminology, there is another one with looser constraint.

Definition 2.1.2: Finite Additive Measure

Let X be a set. Let $\mathcal{A}$ be an algebra (not necessarily a σ-algebra) on X. A **finite additive measure** on $\mathcal{A}$ is a function $\mu : \mathcal{A} \to [0, +\infty]$ such that

1. $\mu(\emptyset) = 0$.
2. **Finite Additivity:** If $\{E_i\}_{i=1}^n$ are finitely many disjoint sets in $\mathcal{A}$, then $\mu(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n \mu(E_i)$

Note that all measures are finitely additive (just let $E_i = \emptyset$ for $i > n$, and use the property that $\mu(\emptyset) = 0$). However, a finite additive measure may not be a measure.

Definition 2.1.3: Measurable Space/Measurable Set/Measure Space

Let X be a set. Let $\mathcal{A}$ be a σ-algebra on X. Let μ be a measure on $\mathcal{A}$.

- $(X, \mathcal{A})$ is called a **measurable space**. The sets in $\mathcal{A}$ are called **measurable sets**.
- $(X, \mathcal{A}, \mu)$ is called a **measure space**.

Now let's see some examples of measures.

Example 2.1.4: Examples of Measures

Measures:

- Let X be an arbitrary set. Let $\mathcal{A}$ be a σ-algebra on X. Define function $\mu : \mathcal{A} \to [0, +\infty]$ such that

$$\mu(E) = \begin{cases} n, & \text{if } |E| = n \\ +\infty, & \text{if } E \text{ is an infinite set} \end{cases}$$

  Then μ is a measure. It is called the **counting measure** on $(X, \mathcal{A})$.

- Let X be a nonempty set. Let $\mathcal{A}$ be a σ-algebra on X. Let $x \in X$. Define function $\delta_x : \mathcal{A} \to [0, +\infty]$ such that

$$\delta_x(E) = \begin{cases} 1, & \text{if } x \in E \\ 0, & \text{if } x \notin E \end{cases}$$

  Then δ_x is a measure. It is called the **Dirac measure** or **point mass** at x.

- Let X be an arbitrary set. Let $\mathcal{A}$ be a σ-algebra on X. Define function $\mu : \mathcal{A} \to [0, +\infty]$ such that

$$\mu(E) = \begin{cases} +\infty, & \text{if } E \neq \emptyset \\ 0, & \text{if } E = \emptyset \end{cases}$$

  Then, μ is a measure.

- Let X be an uncountable set, let $\mathcal{A}$ be a σ-algebra that contains subsets E of X such that either E or E^c is countable. Define function $\mu : \mathcal{A} \to [0, +\infty]$ such that

$$\mu(E) = \begin{cases} 0, & \text{if } E \text{ is countable} \\ 1, & \text{if } E^c \text{ is countable} \end{cases}$$

  Then μ is a measure. This is a little bit harder to see. First of all, $\mu(\emptyset) = 0$ since $\emptyset$ is countable. Further, for some set $E \in \mathcal{A}$, if E is uncountable, then E^c is countable. Thus all sets that is disjoint with E must be countable. Therefore, for disjoint sets $E_1, E_2, E_3, \cdots$ in this σ-algebra, there could be only one uncountable set. If there is, the union is uncountable and the measure is 1. If not, the measure is 0, and it is equal to the result of countable sum.

Finite Additive Measures:

- Let $X = \mathbb{Z}^+$, the set of all positive integers. Let $\mathcal{A}$ be the collection of subsets E of X such that either E or E^c is finite. Then $\mathcal{A}$ is an algebra, but not a σ-algebra. Define function $\mu : \mathcal{A} \to [0, +\infty]$ such that

$$\mu(E) = \begin{cases} 1, & \text{if } E \text{ is infinite} \\ 0, & \text{if } E \text{ is finite} \end{cases}$$

  Then it is a finitely additive measure. However, it is not a measure.

Not finite additive measure nor measure:

- Let X be a set that has at least two elements. Let $\mathcal{A}$ be the collection of all subsets of X. Define function $\mu : \mathcal{A} \to [0, +\infty]$ such that
$$\mu(E) = \begin{cases} 1, & \text{if } E \neq \emptyset \\ 0, & \text{if } E = \emptyset \end{cases}$$
Then it is not a finite additive measure since for two nonempty disjoint sets E and F, $\mu(E \cup F) = 1 \neq \mu(E) + \mu(F) = 2$.

2.1.2 Properties of Measure

The basic properties of measures are summarized below.

Theorem 2.1.5: Properties of Measures

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

1. **Monotonicity:** If $E, F \in \mathcal{A}$, and $E \subseteq F$, then $\mu(E) \leqslant \mu(F)$.
2. **Countable Subadditivity:** For arbitrary $\{E_i\}_{i=1}^{\infty} \in \mathcal{A}$, we have
$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) \leqslant \sum_{i=1}^{\infty} \mu(E_i)$$
3. **Continuous from below:** For $\{E_n\}_{n=1}^{\infty} \in \mathcal{A}$ with $E_1 \subseteq E_2 \subseteq E_3 \subseteq \cdots$, we have
$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \lim_{i \to \infty} \mu(E_i)$$
4. **Continuous from above:** For $\{E_n\}_{n=1}^{\infty} \in \mathcal{A}$ with $E_1 \supseteq E_2 \supseteq E_3 \supseteq \cdots$ and $\mu(E_1) < +\infty$, we have
$$\mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{i \to \infty} \mu(E_i)$$

Proof. 1. If $E \subseteq F$, then $F = E \cup (F \backslash E)$. Since E and $F \backslash E$ are disjoint, by countable additivity, we have
$$\mu(F) = \mu(E) + \mu(F \backslash E) \geqslant \mu(E)$$

2. Let $F_1 = E_1$, and
$$F_k = E_k \backslash \left(\bigcup_{i=1}^{k-1} E_i\right), k > 1$$
Then, $\{F_k\}$ are disjoint, with $\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} F_i$ and $F_k \subseteq E_k$ for all k. Therefore, by countable additivity and

monotonicity,

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \mu\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mu(F_i) \leqslant \sum_{i=1}^{\infty} \mu(E_i)$$

3. Set $E_0 = \emptyset$ for convenience. Since the set $\{E_i\}$ is increasing, $\{E_i \backslash E_{i-1}\}$ are disjoint and $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \mu\left(\bigcup_{i=1}^{\infty}(E_i \backslash E_{i-1})\right)$, by countable additivity,

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \mu\left(\bigcup_{i=1}^{\infty}(E_i \backslash E_{i-1})\right) = \sum_{i=1}^{\infty} \mu(E_i \backslash E_{i-1}) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(E_i \backslash E_{i-1}) = \lim_{n\to\infty} \sum_{i=1}^{n} \mu(E_n)$$

4. Let $F_i = E_1 \backslash E_i$. Then, $F_1 \subseteq F_2 \subseteq F_3 \subseteq \cdots$, and $\mu(E_1) = \mu(F_i) + \mu(E_i)$ by finite additivity. Also, $\bigcup_{i}^{\infty} F_i = E_1 \backslash (\bigcap_{i=1}^{\infty} E_i)$. By continuity from below, we have

$$\mu\left(E_1 \backslash \left(\bigcap_{i=1}^{\infty} E_i\right)\right) = \mu\left(\bigcup_{i=1}^{\infty} F_i\right) = \lim_{n\to\infty} \mu(F_n) = \lim_{n\to\infty} (\mu(E_1) - \mu(E_n)) = \mu(E_1) - \lim_{n\to\infty} \mu(E_n) \tag{2.1}$$

Note that, by countable additivity, we have

$$\mu\left(E_1 \backslash \left(\bigcap_{i=1}^{\infty} E_i\right)\right) + \mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \mu(E_1) \tag{2.2}$$

where we can subtract $\mu(\bigcap_{i=1}^{\infty} E_i)$ on both sides since $\mu(E_1) < \infty$ and the sequence is decreasing. Then, combining the equation 4.3 and 4.4, we have

$$\mu(E_1) - \mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \mu(E_1) - \lim_{n\to\infty} \mu(E_n)$$

Since $\mu(E_1) < +\infty$, we can subtract this term from both sides to get our expected result:

$$\mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} \mu(E_n)$$

□

Note that the condition $\mu(E_1) < +\infty$ can be relaxed as $\mu(E_n) < +\infty$ for some $n \in \mathbb{N}$, since we can always omit the first n terms of the sequence and retain the same result. This assumption is necessary. Consider $X = \mathbb{R}$, and $\mathcal{A}$ be all Lebesgue measurable sets (this would be introduced later, or you can see my *real analysis* note). Let μ be the Lebesgue measure (also, this would be introduced later). The only thing we need to know is that this measure assign $\mu([a, b]) = b - a$, $\mu([a, +\infty]) = +\infty$ and $\mu(\mathbb{R}) = +\infty$. Consider a collection of decreasing set

$$E_n = [-n, n]^c, n \in \mathbb{N}$$

Then, each E_n has an infinite measure. However, the intersection is $\emptyset$, which has measure 0. There is a standard terminology that represents the 'sizes' of measures.

Definition 2.1.6: Finite/σ-Finite/Semifinite

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

- If $\mu(X) < +\infty$, μ is called **finite**.
- If $X = \bigcup_{j=1}^{\infty} E_j$ where $E_j \in \mathcal{A}$ for all j, and $\mu(E_j) < +\infty$ for all j, μ is called **σ-finite**.
- If for each $E \in \mathcal{A}$ with $\mu(E) = +\infty$, there exists $F \in \mathcal{A}$ with $F \subseteq E$ such that $0 < \mu(F) < +\infty$, μ is called **semifinite**.

Note that there is a hierarchy here. **Every finite measure is σ-finite**. This is trivial. Also, **every σ-finite measure is semifinite**. This needs a little bit more work. Suppose μ is σ-finite. Then we can write $X = \bigcup_{j=1}^{\infty} E_j$. Now suppose $\mu(E) = +\infty$. Then E can be written as $E = (\bigcup_{j=1}^{\infty} E_j) \cap E = \bigcup_{j=1}^{\infty}(E_j \cap E)$. Since $\mu(E_j) < +\infty$ and $E_j \cap E \subseteq E_j$, by monotonicity, we have $\mu(E_j \cap E) < +\infty$ for all j. If we set

$$A_j = E_j \cap E, \quad B_j = A_j \backslash \left(\bigcup_{k=1}^{j-1} A_j \right)$$

We have $\bigcup_i A_i = \bigcup_i B_i$, and B_i is disjoint. Also, by monotonicity, $\mu(B_j) < +\infty$ for all j. Therefore, by countable additivity,

$$+\infty = \mu(E) = \mu \left(\bigcup_{i=1}^{\infty} A_i \right) = \mu \left(\bigcup_{i=1}^{\infty} B_i \right) = \sum_{i=1}^{\infty} \mu(B_i)$$

With each $\mu(B_i) < +\infty$, to make $\sum_{i=1}^{\infty} \mu(B_i) = +\infty$, there must exist some i such that $0 < \mu(B_i) < +\infty$. Since this $B_i \subseteq E$, we have proved our result.

2.1.3 Completion of measure

Definition 2.1.7: Null set

Let $(X, \mathcal{A}, \mu)$ be a measure space. A set $E \in \mathcal{A}$ such that $\mu(E) = 0$ is called a **null set**.

By subadditivity, we have any countable union of null sets is a null set, since

$$0 \leqslant \mu \left(\bigcup_{i=1}^{\infty} E_i \right) \leqslant \sum_{i=1}^{\infty} \mu(E_i) = 0$$

Definition 2.1.8: Almost Everywhere

If a statement about points $x \in X$ is true except for x in some null set, we say that it is true **almost everywhere** (abbreviation **a.e.**). (Sometimes we also use **almost surely**, abbreviation **a.s.**).

If $\mu(E) = 0$ and $F \subseteq E$ with $F \in \mathcal{A}$, then $\mu(F) = 0$ by monotonicity. However, it needs not to be true that $F \in \mathcal{A}$.

Definition 2.1.9: Complete Measure

A measure whose domain includes all subsets of null sets is called **complete**.

Luckily, we can always enlarge domains of μ to make it complete.

Theorem 2.1.10: Completion of Measure

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\mathcal{N} = \{N \in \mathcal{A} : \mu(N) = 0\}$ be the collection of null sets. Let

$$\bar{\mathcal{A}} = \{E \cup F : E \in \mathcal{A} \text{ and } F \subseteq N \text{ for some } N \in \mathcal{N}\}$$

Then, $\bar{\mathcal{A}}$ is a σ-algebra, and there exists a **unique** extension $\bar{\mu}$ of μ to a complete measure on $\bar{\mathcal{A}}$.

Proof.

- We first prove that $\bar{\mathcal{A}}$ is a σ-algebra.
  1. First, clearly $\bar{\mathcal{A}}$ is not empty.
  2. Second, suppose $E \cup F \in \bar{\mathcal{A}}$, where $E \in \mathcal{A}$ and $F \subseteq N \in \mathcal{N}$. If we replace F and N by $F \backslash E$ and $N \backslash E$ (this can be done since $N \backslash E = N \cap E^c$ is still in $\mathcal{A}$, thus in $\mathcal{N}$ by monotonicity), $E \cup F$ is the same set, so we can assume that $E \cap N = \emptyset$. In this case,
  $$E \cup F = (E \cup N) \cap (N^c \cup F)$$
  Therefore,
  $$(E \cup F)^c = (E \cup N)^c \cup (N \backslash F)$$
  Since on the right hand side, $(E \cup N)^c \in \mathcal{A}$, and $N \backslash F \subseteq N$, we have $(E \cup F)^c \in \bar{\mathcal{A}}$. Thus $\bar{\mathcal{A}}$ is closed under complement.
  3. Finally, since both $\mathcal{A}$ and $\mathcal{N}$ are closed under countable union, and sets in $\bar{\mathcal{A}}$ is just the union of sets in two, $\bar{\mathcal{A}}$ is also closed under countable union.

- Now we define the expected measure $\bar{\mu}$. For $E \cup F \in \bar{\mathcal{A}}$, we define
  $$\bar{\mu}(E \cup F) = \mu(E)$$
  We first need to show that this is well-defined. Suppose $E_1 \cup F_1 = E_2 \cup F_2$, where $F_j \subseteq N_j \in \mathcal{N}$. Then $E_1 \subseteq E_2 \cup N_2$, therefore
  $$\bar{\mu}(E_1 \cup F_1) = \mu(E_1) \leqslant \mu(E_2) + \mu(N_2) = \mu(E_2)$$
  Similarly, we have $E_2 \subseteq E_1 \cup N_1$, therefore,
  $$\bar{\mu}(E_2 \cup F_2) = \mu(E_2) \leqslant \mu(E_1) + \mu(N_1) = \mu(E_1)$$
  Therefore, $\bar{\mu}(E_1 \cup F_1) = \bar{\mu}(E_2 \cup F_2)$, the measure is well-defined.

- Now we need to show that $\bar{\mu}$ is indeed a measure.
  1. First of all, since $\bar{\mu}(E \cup F) = \mu(E)$, it is a function from $\bar{\mathcal{A}}$ to $[0, +\infty]$.
  2. Second, $\bar{\mu}(\emptyset) = \mu(\emptyset) = 0$.

3. Third, let $\{\bar{E}_j\}_{i=1}^{\infty}$ be a disjoint collection of sets in $\bar{\mathcal{A}}$. Then,

$$\bar{\mu}\left(\bigcup_{j=1}^{\infty}\bar{E}_j\right)=\bar{\mu}\left(\bigcup_{j=1}^{\infty}(E_j\cup F_j)\right)=\bar{\mu}\left(\bigcup_{j=1}^{\infty}E_j\cup\bigcup_{j=1}^{\infty}F_j\right)$$

Since $F_j\subseteq N_j\in\mathcal{N}$, and $N=\bigcup_{j=1}^{\infty}N_j$ is a null set. Since $\bigcup_{j=1}^{\infty}F_j\subseteq N$. Hence,

$$\bar{\mu}\left(\bigcup_{j=1}^{\infty}E_j\cup\bigcup_{j=1}^{\infty}F_j\right)=\mu\left(\bigcup_{j=1}^{\infty}E_j\right)=\sum_{j=1}^{\infty}\mu(E_j)=\sum_{j=1}^{\infty}\bar{\mu}(E_j\cup F_j)=\sum_{j=1}^{\infty}\bar{\mu}(\bar{E}_j)$$

Combining those two equations, $\bar{\mu}$ is countable additive.

- Let $\bar{E}$ be a null set in $\bar{\mathcal{A}}$. That is, if we write $\bar{E}=E\cup F$, then $\bar{\mu}(\bar{E})=\mu(E)=0$. Therefore, E is also a null set. Since F is a subset of a null set, $\bar{E}$ is a subset of a null set N. Let $\bar{F}\subseteq\bar{E}$. Then, immediately $\bar{F}\subseteq N$ and $\bar{F}\in\bar{\mathcal{A}}$ since we can write $\bar{F}=\emptyset\cup\bar{F}$. Therefore, the domain of $\bar{\mu}$ contains each subsets of null sets. $\bar{\mu}$ is complete.

- Now we see $\bar{\mu}$ is actually an extension. Since for all $E\in\mathcal{A}$, we can write $E=E\cup\emptyset$. Then $\bar{\mu}(E)=\mu(E)$ Therefore, μ and $\bar{\mu}$ take the same value on the domain $\mathcal{A}$. Therefore, $\bar{\mu}$ is an extension.

- Finally, we will show the uniqueness of $\bar{\mu}$. Suppose λ is another complete measure on $\bar{\mathcal{A}}$ such that it is an extension of μ. Suppose $\bar{E}=E\cup F\in\bar{\mathcal{A}}$, where $E\in\mathcal{A}$ and $F\subseteq N\in\mathcal{N}$. Similarly as before, we can take $E\cap N=\emptyset$. Then,

$$\mu(E)=\lambda(E)\leqslant\lambda(\bar{E})\leqslant\lambda(E)+\lambda(N)=\mu(E)+\mu(N)=\mu(E)$$

Hence, $\lambda(\bar{E})=\mu(E)=\bar{\mu}(\bar{E})$. Since $\bar{E}$ is arbitrary, $\lambda=\mu$.

□

2.2 Outer Measures

As stated in my *real analysis* note, we can always approximate the measure of a set from outside. In that note, we use this thought to produce Lebesgue Outer measure. Here we generalize that thought to the abstract measure theory.

Definition 2.2.1: Outer Measure

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

- $\mu^*(\emptyset)=0$
- **Monotonicity:** For $A\subseteq B\subseteq X$, we have $\mu^*(A)\leqslant\mu^*(B)$
- **Countable Subadditivity:** For a sequence of set $\{E_i\}$ in X, we have $\mu^*(\bigcup_{i=1}^{\infty}E_i)\leqslant\sum_{i=1}^{\infty}\mu^*(E_i)$

Note that the domain of outer measure is $\mathcal{P}(X)$, which is the collection of all subsets of X, not an arbitrary σ-algebra. With this definition, we can generalize our idea of approximating a set by a box in *real analysis*.

Proposition 2.2.2: Generating Outer Measure

Let $\mathcal{E} \subseteq \mathcal{P}(X)$ and $\rho : \mathcal{E} \to [0, \infty]$ be such that $\emptyset \in \mathcal{E}$, $X \in \mathcal{E}$ and $\rho(\emptyset) = 0$. For any $A \subseteq X$, define

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

Then μ^* is an outer measure.

Proof. Since we can always take $E_j = X$ for all j so that $A \subseteq \bigcup_{j=1}^{\infty} E_j$ for all A, the definition makes sense. Now we need to verify the three properties of outer measure.

- Clearly $\mu^*(\emptyset) = 0$ by taking all $E_j = \emptyset$.
- Let $A \subseteq B \subseteq X$. For all indexed sets $\{E_j\}$ such that $B \subseteq \bigcup_{j=1}^{\infty} E_j$, we also have $A \subseteq \bigcup_{j=1}^{\infty} E_j$. This induces that $\mu^*(A) \leqslant \mu^*(B)$.
- To prove the countable subadditivity, let $\{A_j\}_{j=1}^{\infty} \subseteq \mathcal{P}(X)$ and $\epsilon > 0$. For each j there exists $\{E_j^k\}_{k=1}^{\infty} \subseteq \mathcal{E}$ such that $A_j \subseteq \bigcup_{k=1}^{\infty} E_j^k$ and
$$\sum_{k=1}^{\infty} \rho(E_j^k) \leqslant \mu^*(A_j) + \frac{\epsilon}{2^j} \tag{2.3}$$
which sums up (by j) to
$$\sum_{j,k} \rho(E_j^k) \leqslant \sum_{j=1}^{\infty} \mu^*(A_j) + \epsilon \tag{2.4}$$
Since for each j, $A_j \subseteq \bigcup_{k=1}^{\infty} E_j^k$, we have $A = \bigcup_{j=1}^{\infty} A_j \subseteq \bigcup_{j,k} E_j^k$, which means that
$$\mu^*(A) \leqslant \sum_{j,k} \rho(E_j^k) \tag{2.5}$$
Combine Equation 2.4 and 2.5, we have
$$\mu^*(A) \leqslant \sum_{j=1}^{\infty} \mu^*(A_j) + \epsilon$$
Since ϵ is arbitrary, we have the desired countable subadditivity.

□

2.2.1 Carathéodory's Criterion

The fundamental step from outer measures to measures is stated below.

Definition 2.2.3: μ^*-Measurable

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

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

Note that the inequality $\mu^*(E) \leqslant \mu^*(E \cap A) + \mu^*(E \cap A^c)$ holds for any A and E. Moreover, if $\mu^*(E) = \infty$, the reverse inequality is then trivial. So, to prove that A is μ^*-measurable, we only need to show that

$$\mu^*(E) \geqslant \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \subseteq X \text{ such that } \mu^*(E) < \infty$$

Theorem 2.2.4: Carathéodory's Criterion

Let μ^* be an outer measure on X. Then, the collection $\mathcal{E}$ of μ^*-measurable sets is a *σ-algebra*. Moreover, the restriction of μ^* to $\mathcal{E}$ is a *complete measure*.

Proof. 1. We first prove that $\mathcal{E}$ is an algebra.

- Note that $\mu^*(E) = \mu^*(E \cap \emptyset) + \mu^*(E \cap \emptyset^c)$ for all $E \subseteq X$. Thus $\emptyset$ is μ^*-measurable, $\emptyset \in \mathcal{E}$.
- $\mathcal{E}$ is closed under complements since the definition of μ^*-measurable is symmetric between A and A^c.
- To prove that it is closed under finite union, we only need to show that for $A, B \in \mathcal{E}$, we have $A \cup B \in \mathcal{E}$. Let $E \subseteq X$ be arbitrary. Then,

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*(E \cap A \cap B) + \mu^*(E \cap A \cap B^c) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A^c \cap B^c) \quad (2.6)$$

Since $A \cup B = (A \cap B) \cup (A^c \cap B) \cup (A \cap B^c)$, we have

$$\begin{aligned} E \cap (A \cup B) &= E \cap ((A \cap B) \cup (A^c \cap B) \cup (A \cap B^c)) \\ &= (E \cap A \cap B) \cup (E \cap A^c \cap B) \cup (E \cap A \cap B^c) \end{aligned}$$

Therefore, by countable subadditivity, we have

$$\mu^*(E \cap (A \cup B)) \leqslant \mu^*(E \cap A \cap B) + \mu^*(E \cap A^c \cap B) + \mu^*(E \cap A \cap B^c) \quad (2.7)$$

Combine Equation 2.6 and 2.7, we have

$$\mu^*(E) \geqslant \mu^*(E \cap (A \cup B)) + \mu^*(E \cap A^c \cap B^c) = \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$$

which shows that $A \cup B$ is μ^*-measurable. Therefore, $\mathcal{E}$ is closed under finite union.

2. Now we show that $\mathcal{E}$ is a σ-algebra. To do this, it suffices to show that $\mathcal{E}$ is closed under countable disjoint union. Let $\{A_j\}_{j=1}^{\infty}$ be a sequence of disjoint sets in $\mathcal{E}$. Let $B_n = \bigcup_{j=1}^{n} A_j$ and $B = \bigcup_{j=1}^{\infty} A_j$. We know that $B_n \in \mathcal{E}$ since $\mathcal{E}$ is an algebra. Then, for any $E \subseteq X$,

$$\mu^*(E \cap B_n) = \mu^*(E \cap B_n \cap A_n) + \mu^*(E \cap B_n \cap A_n^c) = \mu^*(E \cap A_n) + \mu^*(E \cap B_{n-1})$$

Therefore, induction shows that

$$\mu^*(E \cap B_n) = \mu^*(E \cap A_n) + \mu^*(E \cap A_{n-1}) + \cdots + \mu^*(E \cap B_1) = \sum_{j=1}^{n} \mu^*(E \cap A_j)$$

Therefore,

$$\begin{aligned}\mu^*(E) =& \mu^*(E \cap B_n) + \mu^*(E \cap B_n^c) = \sum_{j=1}^{n} \mu^*(E \cap A_j) + \mu^*(E \cap B_n^c) \\ &\geqslant \sum_{j=1}^{n} \mu^*(E \cap A_j) + \mu^*(E \cap B^c)\end{aligned}$$

where the last inequality holds because of monotonicity of outer measure, where $(E \cap B^c) \subseteq (E \cap B_n^c)$. Let $n \to \infty$, we have

$$\begin{aligned}\mu^*(E) &\geqslant \sum_{j=1}^{\infty} \mu^*(E \cap A_j) + \mu^*(E \cap B^c) \geqslant \mu^* \left(\bigcup_{j=1}^{\infty} (E \cap A_j) \right) + \mu^*(E \cap B^c) \\ &= \mu^*(E \cap B) + \mu^*(E \cap B^c) \geqslant \mu^*(E)\end{aligned}$$

where the second inequality comes from countable subadditivity. All inequalities thus must be equalities in this calculation. This shows that $B = \bigcup_{j=1}^{\infty} A_j \in \mathcal{E}$.

3. Then, we show that $\mu^*|_{\mathcal{E}}$ is actually a measure. From above, take $E = B$, we have

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

Therefore, it is countable additive.

4. We finally prove that $\mu^*|_{\mathcal{E}}$ is complete. If $\mu^*(A) = 0$, for any $E \subseteq X$ we have

$$\mu^*(E) \leqslant \mu^*(E \cap A) + \mu^*(E \cap A^c) = \mu^*(E \cap A^c) \leqslant \mu^*(E)$$

All inequalities thus become equalities. Therefore, $A \in \mathcal{E}$, and $\mu^*|_{\mathcal{E}}$ is a complete measure. □

2.3 Extension of Measure

The goal of this section is to construct the so-called Borel measure on the real line. We will start from the half-open interval semi-algebra, and define a intuitive measure on it. Then, we are going to extend gradually the domain to the σ-algebra generated by the semi-algebra, which is just the Borel σ-algebra, and also extend the measure onto it. Note that it is intuitive that we can define a function $\mu : \mathcal{S} \to [0, \infty]$ such that

$$\mu((a, b]) = b - a$$

where $\mathcal{S}$ is the half-open interval semi-algebra, and a, b can be infinity. Then, we will go from here.

2.3.1 Premeasure

Definition 2.3.1: Pre-measure

Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an *algebra*. A function $\mu_0 : \mathcal{A} \to [0, \infty]$ is a **pre-measure** if

- $\mu_0(\emptyset) = 0$.
- If $\{A_j\}_{j=1}^{\infty}$ is a sequence of disjoint sets in $\mathcal{A}$ such that $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$, then $\mu_0(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu_0(A_j)$.

Therefore, a pre-measure is just a measure as long as the referenced countable union is in the algebra. Now we devote some intuitive but lengthy effort to construct a pre-measure on the algebra generated by finite disjoint union of half-open intervals.

Proposition 2.3.2: Pre-measure on half-open interval Algebra

Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and right-continuous. If $(a_j, b_j]$ $(j = 1, 2, \cdots, n)$ are disjoint half-open intervals, let

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

and $\mu_0(\emptyset) = 0$. Then, μ_0 is a premeasure on the algebra $\mathcal{A}$ generated by finite disjoint union of half-open intervals.

Proof.

- μ_0 *is well-defined and finitely additive:* We need to check μ_0 is well-defined since elements of $\mathcal{A}$ can be represented in more than one way as disjoint union of half-open intervals. We first suppose $\{(a_j, b_j]\}_{j=1}^{n}$ are disjoint and $\bigcup_{j=1}^{n}(a_j, b_j] = (a, b]$. Then, after reindexing, we must have

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

  Therefore, $\sum_{j=1}^{n}[F(b_j) - F(a_j)] = F(b) - F(a)$ in this case. More generally, if $\{I_i\}_{i=1}^{n}$ and $\{J_j\}_{j=1}^{m}$ are finite sequences of disjoint half-open intervals such that $\bigcup_i I_i = \bigcup_j J_j$. Then,

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

  Thus μ_0 is well-defined, and it is finitely additive by construction.

- μ_0 *is countable additive:* Suppose $\{I_j\}_{j=1}^{\infty}$ is a sequence of disjoint half-open intervals with $\bigcup_{j=1}^{\infty} I_j \in \mathcal{A}$. We need to show that $\mu_0(\bigcup_{j=1}^{\infty} I_j) = \sum_{j=1}^{\infty}\mu_0(I_j)$.

  Since $\bigcup_{j=1}^{\infty} I_j \in \mathcal{A}$, it is a finite union of half-open intervals. Therefore, the sequence $\{I_j\}_{j=1}^{\infty}$ can be partitioned into finitely many subsequences such that the union of the intervals in each subsequences is a single half-open interval. Then, we can consider the subsequences separately and using finite additivity of μ_0. Hence, we may assume that $\bigcup_{j=1}^{\infty} I_j = I = (a, b]$ is a single half-open interval.

("$\geqslant$") We have

$$\mu_0(I) = \mu_0\left(\bigcup_{j=1}^{n} I_j\right) + \mu_0\left(I \backslash \bigcup_{j=1}^{n} I_j\right) \geqslant \mu_0\left(\bigcup_{j=1}^{n} I_j\right) = \sum_{j=1}^{n} \mu_0(I_j)$$

where the first equality comes from the fact that both $\bigcup_{j=1}^{n} I_j$ and $I \backslash \bigcup_{j=1}^{n} I_j$ are finite union of half-open intervals, and the last equality follows from the definition of μ_0. Let $n \to \infty$, we have $\mu_0(I) \geqslant \sum_{j=1}^{\infty} \mu(I_j)$.

("$\leqslant$") We first suppose that a, b are finite. Fix $\epsilon > 0$. Since F is right-continuous, there exists $\delta > 0$ such that $F(a+\delta) - F(a) < \epsilon$. If $I_j = (a_j, b_j]$, for each j there exists $\delta_j > 0$ such that $F(b_j + \delta_j) - F(B_J) < \epsilon 2^{-j}$. The open intervals $(a_j, b_j + \delta_j)$ cover the compact set $[a+\delta, b]$, so there is a finite subcover. By discarding any $(a_j, b_j + \delta_j)$ that is contained in a larger one and reindexing j, we may assume that

- The intervals $(a_1, b_1 + \delta_1), \cdots, (a_N, b_N + \delta_N)$ cover $[a+\delta, b]$,
- $b_j + \delta_j \in (a_{j+1}, b_{j+1} + \delta_{j+1})$ for $j = 1, 2, \cdots, N-1$

Then, we have

$$\mu_0(I) = F(b) - F(a) < F(b) - F(a+\delta) + \epsilon \leqslant F(b_N + \delta_N) - F(a_1) + \epsilon$$

where the last inequality holds since F is increasing. If we write the formula telescopingly, we have

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

Since ϵ is arbitrary, take $\epsilon \to 0$ and we are done. If $a = -\infty$, for any $M < \infty$ the intervals $(a_j, b_j + \delta_j)$ cover $[-M, b]$, so the same reasoning gives $F(b) - F(-M) \leqslant \sum_{j=1}^{\infty} \mu_0(I_j) + 2\epsilon$. Similarly, when $b = \infty$, we can obtain $F(M) - F(a) \leqslant \sum_{j=1}^{\infty} \mu_0(I_j) + 2\epsilon$. The desired result is attained again by letting $\epsilon \to 0$ and $M \to \infty$.

□

2.3.2 Carathéodory's Extension Theorem

Carathéodory's Theorem can be applied to extend measures from algebras to σ-algebras. This is achieved by applying Carathéodory Theorem on some outer measure. Note that an outer measure can be induced from pre-measure.

Proposition 2.3.3: Outer Measure induced from Pre-measure

If μ_0 is a pre-measure on $\mathcal{A} \subseteq \mathcal{P}(X)$, it induces an outer measure on X just like in Proposition 2.2.2

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

Then,

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

Proof.

1. Let $E \in \mathcal{A}$. If $E \subseteq \bigcup_{j=1}^{\infty} A_j$ with $A_j \in \mathcal{A}$, let $B_n = E \cap (A_n \backslash \bigcup_{j=1}^{n-1} A_j)$. Then, $B_n \in \mathcal{A}$ and they are disjoint, with $\bigcup_{n=1}^{\infty} B_n = E$. Therefore,
$$\mu_0(E) = \sum_{j=1}^{\infty} \mu_0(B_j) \leqslant \sum_{j=1}^{\infty} \mu_0(A_j)$$
where the first equality comes from countable additivity of pre-measure, and the second inequality comes from monotonicity. Take infimum on both sides, we have
$$\mu_0(E) \leqslant \mu^*(E)$$
To prove the reverse inequality, let $A_1 = E$ and $A_j = \emptyset$ for all $j > 1$. Then, $\mu^*(E) = \inf\{\sum_{j=1}^{\infty} \mu_0(E_j) : E_j \in \mathcal{A}, E \subseteq \bigcup_{j=1}^{\infty} E_j\} \leqslant \sum_{j=1}^{\infty} \mu_0(A_j) = \mu_0(E)$.

2. Let $A \in \mathcal{A}$ and $E \subseteq X$. Let $\epsilon > 0$ be arbitrary. Then, there exists a sequence of sets $\{B_j\}_{j=1}^{\infty} \subseteq \mathcal{A}$ such that $E \subseteq \bigcup_{j=1}^{\infty} B_j$, and
$$\sum_{j=1}^{\infty} \mu_0(B_j) \leqslant \mu^*(E) + \epsilon$$
Since μ_0 is countably additive on $\mathcal{A}$, we have
$$\mu^*(E) + \epsilon \geqslant \sum_{j=1}^{\infty} \mu_0(B_j \cap A) + \sum_{j=1}^{\infty} \mu_0(B_j \cap A^c) \geqslant \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

where the second inequality comes from monotonicity. Since ϵ is arbitrary, A is μ^*-measurable. □

A σ-finite pre-measure can be uniquely extended to the σ-algebra generated by the algebra where the pre-measure is defined on.

Theorem 2.3.4: Carathéodory's Extension Theorem

Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, and μ_0 be a pre-measure on $\mathcal{A}$. Let $\mathcal{E}$ be the σ-algebra generated by $\mathcal{A}$. Then,

- There exists a measure $\mu = \mu^*|_{\mathcal{E}}$, where μ^* is defined in 2.3.3, which is the extension of μ_0.
- If ν is another measure on $\mathcal{E}$ that extends μ_0, the $\nu(E) \leqslant \mu(E)$ for all $E \subseteq \mathcal{E}$, with equality when $\mu(E) < \infty$.
- If μ_0 is σ-finite, then μ is the unique extension of μ_0 to a measure on $\mathcal{E}$.

Proof.

- By Proposition 2.3.3, μ^* is indeed an extension of μ_0, and $\mathcal{A}$ is a subset of σ-algebra $\mathcal{M}$ of μ^*-measurable sets. since $\mathcal{E}$ is generated by $\mathcal{A}$, it is also included in $\mathcal{M}$ by Lemma 1.3.9. By Carathéodory Theorem 2.2.4, $\mu^*|_{\mathcal{M}}$ is a measure. Hence $\mu^*|_{\mathcal{E}}$ is also a measure.

- If $E \in \mathcal{E}$, and $E \subseteq \bigcup_{j=1}^{\infty} A_j$, where $A_j \in \mathcal{A}$, then

$$\nu(E) \leqslant \sum_{j=1}^{\infty} \nu(A_j) = \sum_{j=1}^{\infty} \mu_0(A_j)$$

Take infimum on both sides, we have $\nu(E) \leqslant \mu(E)$.

Now suppose that $\mu(E) < \infty$. Set $A = \bigcup_{j=1}^{\infty} A_j$. Note that A may not belong to $\mathcal{A}$, since $\mathcal{A}$ is just an algebra. Therefore, we use the thought of the limit, and we have

$$\nu(A) = \lim_{n\to\infty} \nu\left(\bigcup_{j=1}^{n} A_j\right) = \lim_{n\to\infty} \mu\left(\bigcup_{j=1}^{n} A_j\right) = \mu(A)$$

where the first and the last equality comes from continuity from below (if we define $B_j = \bigcup_{i=1}^{j} A_i \in \mathcal{A}$, then it is a union of increasing sequence of set). Since $\mu(E) < \infty$, we can choose A_j's so that $\mu(A) \leqslant \mu(E) + \epsilon$. This will lead to

$$\mu(A) = \mu(E) + \mu(A \backslash E) \leqslant \mu(E) + \epsilon \quad \Longrightarrow \quad \mu(A \backslash E) \leqslant \epsilon$$

Then,

$$\mu(E) \leqslant \mu(A) = \nu(A) = \nu(E) + \nu(A \backslash E) \leqslant \nu(E) + \mu(A \backslash E) \leqslant \nu(E) + \epsilon$$

Since ϵ is arbitrary, we have $\mu(E) \leqslant \nu(E)$. Combining the result from statement 1, we have $\nu(E) = \mu(E)$.

- Suppose $X = \bigcup_{j=1}^{\infty} B_j$ with $B_j \in \mathcal{A}$ and $\mu_0(B_j) < \infty$, we can assume that B_j's are disjoint. Then, for any $E \in \mathcal{E}$,

$$\mu(E) = \sum_{j=1}^{\infty} \mu(E \cap B_j) = \sum_{j=1}^{\infty} \nu(E \cap B_j) = \nu(E)$$

where the middle equality comes from the result of statement 2. This indicates that $\nu = \mu$. □

An application of this theorem is just to extend the pre-measure on the algebra of finite disjoint union of h-intervals, to a Borel measure.

Proposition 2.3.5: Borel measure on the Real Line

Let $F : \mathbb{R} \to \mathbb{R}$ be any increasing, right-continuous function. Then,

- There is a *unique* Borel measure μ_F on $\mathbb{R}$ such that $\mu_F((a, b]) = F(b) - F(a)$ for all a, b.
- If G is another such function, we have $\mu_F = \mu_G$ if and only if $F - G$ is 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, with $\mu=\mu_F$.

Proof. First, Each F induces a premeasure on $\mathcal{A}$ by Proposition 2.3.2. This premeasure, by Theorem 2.3.4, can be uniquely extended on $\mathcal{B}(\mathbb{R})$ since the premeasure is σ-finite (since $\mathbb{R}=\bigcup_{j=-\infty}^{\infty}(j,j+1])$. $\mu_F((a,b])=F(b)-F(a)$ by the definition of μ_0 and since that μ_F is an extension of μ_0.
Second, it is clear that F and G induced the same premeasure if and only if $F-G$ is constant.
Finally, the monotonicity of μ implies the monotonicity of F, and the continuity of μ from above and below implies the right continuity of F for $x\geqslant 0$ and $x<0$. $\mu=\mu_F$ on $\mathcal{A}$ clearly, and hence $\mu=\mu_F$ on $\mathcal{B}(\mathbb{R})$ by the uniqueness of extension. □

2.4 Approximation Theorem

The measure extension theorem yields an abstract statement of existence and uniqueness for measures on $\sigma(\mathcal{A})$ that were first defined on a algebra $\mathcal{A}$ only. The following theorem, however, shows that the measure of a set from $\sigma(\mathcal{A})$ can be well approximated by finite and countable operations with sets from $\mathcal{A}$.

Proposition 2.4.1: Approximation Theorems for Measure

Let $\mathcal{A}$ be an algebra on X and let μ be a measure on $\sigma(\mathcal{A})$ that is σ-finite.

(i) For any $A\in\sigma(A)$ with $\mu(A)<\infty$ and any $\epsilon>0$, there exists $E\in\mathcal{A}$ such that $\mu(A\Delta E)\leqslant\epsilon$.

(ii) For any $A\in\sigma(A)$ and $\epsilon>0$, there exists disjoint sets $E_1,E_2,\cdots\in\mathcal{A}$ such that $A\subseteq\bigcup_{n=1}^{\infty}E_n$ and $\mu(\bigcup_{n=1}^{\infty}E_n\backslash A)\leqslant\epsilon$.

Proof. (i) Let $\mu^*(A)$ be the outer measure induced by μ on $\mathcal{A}$, and it coincides on $\sigma(\mathcal{A})$, that is,

$$\mu^*(A)=\inf_{A_i\in\mathcal{A}}\left\{\sum_{i=1}^{\infty}\mu(A_i):A\subseteq\bigcup_{i=1}^{\infty}A_i,A_i\in\mathcal{A}\right\}=\mu(A),\quad\forall\,A\in\sigma(\mathcal{A})$$

By this definition, we can find $\epsilon>0$ such that there exists $A_i\in\mathcal{A}$ such that

$$\mu^*(A)\leqslant\sum_{i=1}^{\infty}\mu(A_i)\leqslant\mu^*(A)+\frac{\epsilon}{2}$$

Since $\mu(A)<\infty$ by assumption, the series in the middle must converge. Therefore, there exists $n_0\in\mathbb{N}$ such that

$\sum_{i=n_0}^{\infty} \mu(A_i) \leqslant \epsilon/2$. We choose $E = \bigcup_{i=1}^{n_0} A_i \in \mathcal{A}$. Then,

$$\mu^*(E \backslash A) = \mu^*\left(\bigcup_{i=1}^{n_0} A_i \backslash A\right) \leqslant \mu^*\left(\bigcup_{i=1}^{\infty} A_i \backslash A\right) = \mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) - \mu^*(A) \leqslant \sum_{i=1}^{\infty} \mu^*(A_i) - \mu^*(A) \leqslant \frac{\epsilon}{2}$$

where the second equality holds since $\mu^*(A) < \infty$. Moreover,

$$\mu^*(A \backslash E) = \mu^*\left(A \backslash \bigcup_{i=1}^{n_0} A_i\right) \leqslant \mu^*\left(\bigcup_{i=1}^{\infty} A_i \backslash \bigcup_{j=1}^{n_0} A_j\right) = \mu^*\left(\bigcup_{j=n_0+1}^{\infty} A_j\right) \leqslant \sum_{j=n_0+1}^{\infty} \mu^*(A_j) \leqslant \frac{\epsilon}{2}$$

Therefore, $\mu(A \Delta E) = \mu(E \backslash A) + \mu(A \backslash E) = \mu^*(E \backslash A) + \mu^*(A \backslash E) \leqslant \epsilon$, since μ and μ^* coincides on $\sigma(\mathcal{A})$.

(ii) Let $A \in \sigma(\mathcal{A})$. Since μ is σ-finite, we can let $E_n \nearrow X$, $E_n \in \sigma(\mathcal{A})$ with $\mu(E_n) < \infty$ for any $n \in \mathbb{N}$. For every $n \in \mathbb{N}$, we choose a covering $\{B_{n,m}\}_{m \in \mathbb{N}}$ such that

$$\mu(A \cap E_n) \geqslant \sum_{m=1}^{\infty} \mu(B_{n,m}) - \frac{\epsilon}{2^n}$$

which can be done by the definition of outer measure. Note that any countable union of sets $\bigcup B_i$ in $\mathcal{A}$ can be written as a countable *disjoint* union of sets in $\mathcal{A}$, which can be easily done by letting $A_i = B_i \backslash \bigcup_{j=1}^{i-1} B_j$, where $A_i \in \mathcal{A}$ by definition of algebra. Therefore, we can find $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A}$ disjoint such that

$$\bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} B_{n,m} = \bigcup_{n=1}^{\infty} A_n$$

Then,

$$\mu\left(\bigcup_{n=1}^{\infty} A_n \backslash A\right) = \mu\left(\bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} B_{n,m} \backslash A\right) \leqslant \mu\left(\bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} (B_{n,m} \backslash (A \cap E_n))\right) \leqslant \sum_{n=1}^{\infty}\left(\left(\sum_{m=1}^{\infty} \mu(B_{n,m})\right) - \mu(A \cap E_n)\right) \leqslant \epsilon$$

□

For a complete measure that is constructed using Carathéodory's criterion, we can also achieve 'exact approximation'. The following approximation theorem is also set as the definition of Lebesgue measure in some context, for example, in Yale University MATH320/520 course Measure Theory and Integration.

Proposition 2.4.2: Exact Approximation of Complete measure

Let $\mathcal{A}$ be an algebra on X and let μ be a measure on $\sigma(\mathcal{A})$ that is σ-finite. For any μ^*-measurable set A, there are sets $A_-, A_+ \in \sigma(\mathcal{A})$ with $A_- \subseteq A \subseteq A_+$ such that $\mu(A_+ \backslash A_-) = 0$.

Proof. Let A be a μ^*-measurable set. Since μ is σ-finite, we can let $E_n \nearrow X$, $E_n \in \sigma(\mathcal{A})$ with $\mu(E_n) < \infty$ for any $n \in \mathbb{N}$. For any $m, n \in \mathbb{N}$, choose $A_{n,m} \in \sigma(\mathcal{A})$ such that $A_{n,m} \supseteq A \cap E_n$ and $\mu^*(A_{n,m}) \leqslant \mu^*(A \cap E_n) + 2^{-n}/m$, which can be done by the definition of outer measure. Define

$$A_m = \bigcup_{n=1}^{\infty} A_{n,m} \in \sigma(\mathcal{A})$$

Then, $A_m \supseteq A$ and $\mu^*(A_m \backslash A) \leqslant 1/m$. Define

$$A_+ = \bigcap_{m=1}^{\infty} A_m \in \sigma(\mathcal{A})$$

Then, $A \subseteq A_+$ and $\mu^*(A_+ \backslash A) = 0$. Similarly, choose $(A_-)^c \in \sigma(\mathcal{A})$ with $(A_-)^c \supseteq A^c$ and $\mu^*((A_-)^c \backslash A^c) = 0$, by replacing all A above by A^c. Then, $A_- \subseteq A \subseteq A_+$, and

$$\mu(A_+ \backslash A_-) = \mu^*(A_+ \backslash A_-) = \mu^*(A_+ \backslash A) + \mu^*(A \backslash A_-) = 0$$

$\square$

2.5 Lebesgue-Stieltjes Measure on the Real Line

2.5.1 Definition and its Equivalant Definition

In the previous section we construct the Borel measure μ_F, from semialgebra, to algebra, and extend this to σ-algebra by Carathéodory's extension theorem. The completion $\hat{\mu}_F$ of the Borel measure μ_F is called the *Lebesgue-Stieltjes measure*.

Definition 2.5.1: Lebesgue-Stieltjes Measure

The completion of μ_F defined above is called the **Lebesgue-Stieltjes measure**. It is denoted by $\bar{\mu}_F$.

This is where the Carathéodory's criterion comes into effect. We will prove that this completion is indeed just all the μ^*-measurable sets. We will prove this generally, not only for the Lebesgue-Stieltjes measure. Before stating and proving the final theorem, we introduce a lemma.

Lemma 2.5.2: Approximation of Outer Measure

Let $\mathcal{A} \subseteq \mathcal{P}(X)$ 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. Then,

(a) For any $E \subseteq X$ and $\epsilon > 0$, there exists $A \in \mathcal{A}_\sigma$ such that $E \subseteq A$ and $\mu^*(A) \leqslant \mu^*(E) + \epsilon$.

(b) If $\mu^*(E) < \infty$, then E is μ^*-measurable *if and only if* there exist $B \in \mathcal{A}_{\sigma\delta}$ such that $E \subseteq B$ and $\mu^*(B \backslash E) = 0$.

(c) If μ_0 is σ-finite, the restriction $\mu^*(E) < \infty$ in (b) is redundant.

Proof. (a) By definition of the outer measure

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

Hence, for each $\epsilon > 0$, there is a sequence $(A_i)_{i=1}^{\infty} \subseteq \mathcal{A}$ with

$$E \subseteq \bigcup_{i=1}^{\infty} A_i, \quad \sum_{i=1}^{\infty} \mu_0(A_i) \leqslant \mu^*(E) + \epsilon$$

Set $A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}_\sigma$. By monotonicity and subadditivity of μ^*, we have

$$\mu^*(A) \leqslant \sum_{i=1}^{\infty} \mu^*(A_i) = \sum_{i=1}^{\infty} \mu_0(A_i) \leqslant \mu^*(E) + \epsilon$$

Since ϵ was arbitrary, this proves (a).

(b) ($\Longrightarrow$) Suppose $\mu^*(E) < \infty$ and E is μ^*-measurable. By part (a), for each $n \in \mathbb{N}$, we can choose $A_n \in \mathcal{A}_\sigma$ such that

$$E \subseteq A_n, \quad \mu^*(A_n) \leqslant \mu^*(E) + 2^{-n}$$

Since E is μ^*-measurable, we have that $\mu^*(A_n) = \mu^*(A_n \cap E) + \mu^*(A_n \cap E^c) = \mu^*(E) + \mu^*(A_n \backslash E)$. Since $\mu^*(E) < \infty$, it leads to $\mu^*(A_n \backslash E) = \mu^*(A_n) - \mu^*(E)$. Therefore, we have

$$\mu^*(A_n \backslash E) = \mu^*(A_n) - \mu^*(E) \leqslant 2^{-n}$$

Set $B = \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}_{\sigma\delta}$. Since $E \subseteq A_n$ for any n, clearly $E \subseteq B$. Moreover, by monotonicity of μ^*, we have

$$\mu^*(B \backslash E) \leqslant \mu^*(A_n \backslash E) \leqslant 2^{-n} \quad \forall n \in \mathbb{N}$$

Take $n \to \infty$, we have $\mu^*(B \backslash E) = 0$.
($\Longleftarrow$) Conversely, suppose there exists $B \in \mathcal{A}_{\sigma\delta}$ such that $E \subseteq B$ and $\mu^*(E \backslash B) = 0$. We need to prove that E is μ^*-measurable. By Proposition 2.3.3, every set in $\mathcal{A}$ is μ^*-measurable. If we denote all the collection of all μ^*-measurable sets by $\mathcal{M}$, then $\mathcal{M}$ is a σ-algebra by Carathéodory's criterion. Then, $\mathcal{A} \subseteq \mathcal{M}$ and thus $\sigma(\mathcal{A}) \subseteq \mathcal{M}$ by Lemma 1.3.9. Since $B \in \mathcal{A}_{\sigma\delta} \subseteq \sigma(\mathcal{A})$, we have that B is μ^*-measurable. Since $\mu^*|_{\mathcal{M}}$ is a complete measure, and $\mu^*(B \backslash E) = 0$, we have that $B \backslash E$ is also μ^*-measurable. Note that

$$B \backslash (B \backslash E) = B \cap (B \cap E^c)^c = B \cap (B^c \cup E) = (B \cap B^c) \cup (B \cap E) = E$$

Therefore, E is is also μ^*-measurable since $\mathcal{M}$ is a σ-algebra. Note that in this direction we don't need the assumption that $\mu^*(E) < \infty$.

(c) Now we need to re-prove the forward direction of (b) without using the assumption that $\mu^*(E) < \infty$. Suppose μ_0 is σ-finite. Then, $X = \bigcup_{n=1}^{\infty} E_n$ where $E_n \in \mathcal{A}$ and $\mu^*(E_n) = \mu_0(E) < \infty$ for all $n \in \mathbb{N}$. Likewise in part (b), we can choose $\{A_{mn}\}_{m,n=1}^{\infty} \subseteq \mathcal{A}_\sigma$ such that

$$E_n \cap E \subseteq A_{mn}, \quad \mu^*(A_{mn}) \leqslant \mu^*(E_n \cap E) + \frac{1}{m2^n}$$

Since E is μ^*-measurable, $E_n \in \mathcal{A} \subseteq \mathcal{M}$, we have E_n is μ^*-measurable, thus $E_n \cap E$ is μ^*-measurable. Therefore, $\mu^*(A_{mn}) = \mu^*(A_{mn} \cap (E_n \cap E)) + \mu^*(A_{mn} \cap (E_n \cap E)^c) = \mu^*(E_n \cap E) + \mu^*(A_{mn} \backslash (E_n \cap E))$. Since $\mu^*(E_n \cap E) \leqslant$

$\mu^*(E_n) < \infty$, we have

$$\mu^*(A_{mn} \backslash (E_n \cap E)) = \mu^*(A_{mn}) - \mu^*(E_n \cap E) \leqslant \frac{1}{m2^n}$$

Let $A_m = \bigcup_{n=1}^{\infty} A_{mn} \in \mathcal{A}_\sigma$. Then, $E \subseteq A_m$. Now, note that

$$A_m \backslash E = \bigcup_{n=1}^{\infty} A_{mn} \backslash \bigcup_{n=1}^{\infty} (E \cap E_n) \subseteq \bigcup_{n=1}^{\infty} (A_{mn} \backslash (E \cap E_n))$$

We can get

$$\mu^*(A_m \backslash E) \leqslant \mu^* \left(\bigcup_{n=1}^{\infty} (A_{mn} \backslash (E \cap E_n)) \right) \leqslant \sum_{n=1}^{\infty} \mu^*(A_{mn} \backslash (E_n \cap E)) \leqslant \sum_{n=1}^{\infty} \frac{1}{m2^n} = \frac{1}{m}$$

Let $B = \bigcap_{m=1}^{\infty} A_m \in \mathcal{A}_{\sigma\delta}$. Then, $E \subseteq B$, and $\mu^*(B \backslash E) \leqslant \mu^*(A_m \backslash E) \leqslant 1/m$ for all $m \in \mathbb{N}$. Send $m \to \infty$, we get the result. □

Theorem 2.5.3: Completion and Carathéodory Criterion

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

Proof. By Theorem 2.1.10, the completion of μ is the unique extention of μ on the set $\bar{\mathcal{A}} = \{E \cup F : E \in \mathcal{A} \text{ and } F \subseteq N \text{ for some } N \in \mathcal{N}\}$ where $\mathcal{N} = \{N \in \mathcal{A} : \mu(N) = 0\}$. Let $\bar{\mu}$ be this extension. To show that $\hat{\mu} = \bar{\mu}$, we must first show that $\hat{\mathcal{A}} = \bar{\mathcal{A}}$.

($\subseteq$) Let $A \in \hat{\mathcal{A}}$. By Lemma 2.5.2 (c), there exists $B \in \mathcal{A}_{\sigma\delta} = \mathcal{A}$ such that $A \subseteq B$ and $\mu^*(B \backslash A) = 0$. This shows that $B \backslash A \in \hat{\mathcal{A}}$ since $\hat{\mu}$ is complete. Again, by Lemma 2.5.2, there exists $E \in \mathcal{A}$ such that $B \backslash A \subseteq E$ and $\mu^*(E \backslash (B \backslash A)) = 0$. Note that since $B, E \in \mathcal{A}$, we have $B \backslash E \in \mathcal{A}$. Moreover, $A \cap E \in \hat{\mathcal{A}}$. Since $A = (B \backslash E) \cup (A \cap E)$, for the goal of proving that A is in the completion, we only need to show that $A \cap E$ is a null set. However, note that

$$\mu^*(A \cap E) \leqslant \mu^*(E \backslash (B \backslash A)) = 0$$

Since $A \cap E \in \hat{\mathcal{A}}$, by Lemma 2.5.2, there exists $F \in \mathcal{A}$ such that $A \cap E \subseteq F$ and $\mu^*(F \backslash (A \cap E)) = 0$. This leads to the fact that $\mu(F) = \mu^*(F) \leqslant \mu^*(F \backslash (A \cap E)) + \mu^*(A \cap E) = 0$. Thus, $F \in \mathcal{N}$, and $A \cap E \subseteq F$. This shows that $A \cap E$ is a null set, and thus $A \in \bar{\mathcal{A}}$.
($\supseteq$) Let $A \in \bar{\mathcal{A}}$. Then, $A = E \cup F$ where $E \in \mathcal{A}$ and $F \subseteq N$ for some $N \in \mathcal{N}$ ($\mu(N) = 0$). Clearly, $A \subseteq E \cup N \in \mathcal{A} = \mathcal{A}_{\sigma\delta}$, and $\mu^*((E \cup N) \backslash A) = \mu^*(N \backslash F) \leqslant \mu^*(N) = 0$. By Lemma 2.5.2, A is μ^*-measurable, so $A \in \hat{\mathcal{A}}$.

After proving $\hat{\mathcal{A}} = \bar{\mathcal{A}}$, note that $\hat{\mu}|_{\mathcal{A}} = \mu$, so $\hat{\mu}$ is also an extension of μ. However, by Theorem 2.1.10, the extension to the complete measure is unique, so it is trivial that $\hat{\mu} = \bar{\mu}$. □

Why this indicates that the Lebesgue-Stieltjes measure is exactly on those μ^*-measurable sets? Suppose we have the premeasure μ_0 on the finite disjoint union of half-open intervals, and we want to extend it to an outer measure. It turns out that this extended outer measure is exactly the same with the outer measure extended by the Borel measure. Here is a standard proof: Let $\mathcal{A}$ be the algebra mentioned above, and μ be the Borel measure. The definitions of the outer measures are

$$\mu^*(E) = \inf\left\{\sum_{i=1}^{\infty}\mu_0(A_i) : E \subseteq \bigcup_{i=1}^{\infty} A_i,\ A_i \in \mathcal{A}\right\}, \quad \mu^*_{\text{Borel}}(E) = \inf\{\mu(B), E \subseteq B,\ B \text{ Borel}\}$$

First, if $E \subseteq \bigcup_{i=1}^{\infty} A_i$ with $A_i \in \mathcal{A}$ for each i, then each A_i is Borel, and $\mu(A_i) = \mu_0(A_i)$, thus

$$\mu^*_{\text{Borel}}(E) \leqslant \sum_{i=1}^{\infty}\mu(A_i) = \sum_{i=1}^{\infty}\mu_0(A_i)$$

Take infimum on both sides, we have $\mu^*_{\text{Borel}}(E) \leqslant \mu^*(E)$.

Second, because μ is the Carathéodory extension of μ_0, every Borel set B is μ^*-measurable, and satisfies $\mu(B) = \mu^*(B)$. Hence, for any Borel set $B \supseteq E$, we have

$$\mu^*(E) \leqslant \mu^*(B) = \mu(B)$$

Take infimum on both sides, we have $\mu^*(E) \leqslant \mu^*_{\text{Borel}}(E)$. These two together indicate that these two extended outer measures are equivalent.

Therefore, combining this result from Theorem 2.5.3, since Borel measure is indeed σ-finite, the restriction of the outer measure induced by Borel measure on the μ^*-measurable sets is just the completion, and this restriction is equivalent to the restriction of the outer measure induced by the premeasure μ_0.

2.5.2 Regularity Property

Lebesgue-Stieltjes measures enjoy some useful regularity properties that we now investigate. In this section we denote the Lebesgue-Stieltjes measure associated to the increasing, right-continuous function F by μ, and the domain of μ by $\mathcal{A}_\mu$. Thus, for any $E \in \mathcal{A}_\mu$,

$$\mu(E) = \inf\left\{\sum_{j=1}^{\infty}[F(b_j) - F(a_j)] : E \subseteq \bigcup_{j=1}^{\infty}(a_j, b_j]\right\} = \inf\left\{\sum_{j=1}^{\infty}\mu((a_j, b_j]) : E \subseteq \bigcup_{j=1}^{\infty}(a_j, b_j]\right\}$$

A lemma about representation of this meausre is needed for further proofs.

Lemma 2.5.4: Representation of Lebesgue-Stieltjes Measure using Open Intervals

For any $E \in \mathcal{A}_\mu$,

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

Proof. ($\leqslant$) Call the RHS $\nu(E)$. Each (a_j, b_j) is a countable disjoint union of half open intervals $I_j^k = (c_j^k, c_j^{k+1}]$ where $\{c_j\}$ is any sequence such that $c_j^1 = a_j$ and c_j^k increases to b_j when $k \to \infty$. Thus, $E \subseteq \bigcup_{j,k=1}^{\infty} I_j^k$, and

$$\sum_{j=1}^{\infty}\mu((a_j, b_j)) = \sum_{j,k=1}^{\infty}\mu(I_j^k) \geqslant \mu(E)$$

Take infimum on both sides, we have $\mu(E) \leqslant \nu(E)$.
($\geqslant$) By definition, given $\epsilon > 0$, there exists $\{(a_j, b_j]\}_{j\in\mathbb{N}}$ with $E \subseteq \bigcup_{j=1}^{\infty}(a_j, b_j]$ and $\sum_{j=1}^{\infty}\mu((a_j, b_j]) \leqslant \mu(E) + \epsilon$. Since F is right-continuous, for each j there exists $\delta_j > 0$ such that $F(b_j + \delta_j) - F(b_j) < \epsilon 2^{-j}$. Then, $E \subseteq \bigcup_{j=1}^{\infty}(a_j, b_j + \delta_j)$ and

$$\nu(E) \leqslant \sum_{j=1}^{\infty}\mu((a_j, b_j + \delta_j)) = \sum_{j=1}^{\infty}(F(b_j + \delta_j) - F(b_j) + F(b_j) - F(a_j)) < \sum_{j=1}^{\infty}(F(b_j) - F(a_j)) + \epsilon \leqslant \mu(E) + 2\epsilon$$

which completes the proof. □

Theorem 2.5.5: Outer/Inner Regularity

If $E \in \mathcal{A}_\mu$, then

$$\mu(E) = \inf\{\mu(U) : U \supseteq E,\ U \text{ open}\} = \sup\{\mu(K) : K \subseteq E,\ K \text{ compact}\}$$

Proof. By Lemma 2.5.4, for any $\epsilon > 0$, there exists intervals (a_j, b_j) such that $E \subseteq \bigcup_{j=1}^{\infty}(a_j, b_j)$ and $\mu(E) \leqslant \sum_{j=1}^{\infty}\mu((a_j, b_j)) + \epsilon$. If $U = \bigcup_{j=1}^{\infty}(a_j, b_j)$, then U is open, $U \supseteq E$ and $\mu(U) \leqslant \mu(E) + \epsilon$. On the other hand, $\mu(U) \geqslant \mu(E)$ whenever $U \supseteq E$. These two show the first equality.

For the second equality, first suppose that E is bounded. If E is also closed, then E is compact, and the equality is obvious. Otherwise, there is $n \in \mathbb{N}$ such that $E \subseteq [-n, n]$. Then, $[-n, n]\backslash E \in \mathcal{A}_\mu$. By the proved first equality, for any $\epsilon > 0$, there exists open set U such that

$$\mu([-n, n]\backslash E) \leqslant \mu(U) \leqslant \mu([-n, n]\backslash E) + \epsilon$$

Everything in this inequality is finite, since E is bounded. Therefore,

$$\mu(E) = \mu([-n, n]) - \mu([-n, n]\backslash E) \geqslant \mu([-n, n]) - \mu(U) \geqslant \mu([-n, n]) - \mu([-n, n]\backslash E) - \epsilon = \mu(E) - \epsilon$$

Note that $[-n, n]\backslash U$ is closed and bounded, thus U is compact. Since $\mu([-n, n]\backslash U) = \mu([-n, n]) - \mu(U)$, combining with the last inequality, we have seen that for any $\epsilon > 0$, there is a compact set $K = [-n, n]\backslash U$ such that $K \subseteq E$ and $\mu(E) \geqslant \mu(K) \geqslant \mu(E) - \epsilon$, which proves the second equality in the theorem (with assumption that E is bounded).

If E is not bounded, then define, for $j \in \mathbb{Z}$, $E_j = E \cap (j, j+1]$. Each $E_j \in \mathcal{A}_\mu$ and each E_j is bounded. Moreover, they are disjoint. By the part proved with the assumption of boundedness, there exists $K_j \subseteq E_j$ such that K_j is compact and

$$\mu(E_j) \geqslant \mu(K_j) \geqslant \mu(E_j) - \frac{\epsilon}{2^{|j|+2}}$$

Let $H_n = \bigcup_{-n}^{n-1} K_j$. We have that H_n is compact, and $H_n \subseteq E$. We also have

$$
\begin{aligned}
\mu(E) \geqslant \mu(H_n) &= \sum_{j=-n}^{n-1} \mu(K_j) = \sum_{j=-n}^{n-1} \mu(E_j) - \sum_{j=-n}^{n-1} \frac{\epsilon}{2^{|j|+2}} \\
&= \mu(E \cap [-n, n]) - \sum_{j=-n}^{n-1} \frac{\epsilon}{2^{|j|+2}} \geqslant \mu(E \cap [-n, n]) - \sum_{j=-\infty}^{\infty} \frac{\epsilon}{2^{|j|+2}} \\
&= \mu(E \cap [-n, n]) - \frac{\epsilon}{2^{0+2}} - \sum_{j=1}^{\infty} \frac{\epsilon}{2^{j+2}} - \sum_{j=-\infty}^{-1} \frac{\epsilon}{2^{-j+2}} = \mu(E \cap [-n, n]) - \frac{1}{4}\epsilon - 2 \times \frac{1}{4}\epsilon \\
&= \mu(E \cap [-n, n]) - \frac{3}{4}\epsilon
\end{aligned}
$$

Since $\{E \cap [-n, n]\}_{n \in \mathbb{N}}$ is a non-decreasing sequence of sets, and $E = \bigcup_{n=1}^{\infty}(E \cap [-n, n])$. By continuity from below, we have that $\lim_{n\to\infty} \mu(E \cap [-n, n]) = \mu(E)$. Taking limit on both sides, we have

$$
\mu(E) - \epsilon \leqslant \lim_{n\to\infty} \mu(H_n) \leqslant \mu(E)
$$

which completes the proof. □

Theorem 2.5.6: Structure of Measurable Sets

If $E \subseteq \mathbb{R}$, then TFAE:

(a) $E \in \mathcal{A}_\mu$.

(b) $E = V \backslash N_1$ where V is a G_δ set and $\mu(N_1) = 0$.

(c) $E = H \cup N_2$ where H is an F_σ set and $\mu(N_2) = 0$.

Proof. Obviously (b) and (c) both imply (a) since μ is complete. Conversely, we first prove the bounded case. Suppose $E \in \mathcal{A}_\mu$ and $\mu(E) < \infty$. By Theorem 2.5.5, for $j \in \mathbb{N}$ we can choose an open $U_j \supseteq E$ and a compact $K_j \subseteq E$ such that

$$
\mu(U_j) - 2^{-j} \leqslant \mu(E) \leqslant \mu(K_j) + 2^{-j}
$$

Let $V = \bigcap_{j=1}^{\infty} U_j \in G_\delta$ and $H = \bigcup_{j=1}^{\infty} K_j \in F_\sigma$. Then, $H \subseteq E \subseteq V$ and $\mu(V) = \mu(H) = \mu(E) < \infty$. Therefore, $\mu(V \backslash E) = \mu(V) - \mu(E) = 0$ and $\mu(E \backslash H) = \mu(E) - \mu(H) = 0$. Let $N_1 = V \backslash E$ and $N_2 = E \backslash H$, this proves the case for $\mu(E) < \infty$.

For the general case, for each $j \in \mathbb{N}$, set $E_j = E \cap [-j, j]$. Then, each E_j is measurable and has $\mu(E_j) < \infty$. By the previous proof, we have that there exists $H_j \subseteq E_j$, $H_j \in F_\sigma$, $\mu(E_j \backslash H_j) = 0$. Now let $H = \bigcup_{j=1}^{\infty} H_j$. Then, H is a countable union of F_σ sets, hence itself F_σ. Moreover,

$$
E = \bigcup_{j=1}^{\infty} E_j = \bigcup_{j=1}^{\infty} (H_j \cup (E_j \backslash H_j)) = H \cup \bigcup_{j=1}^{\infty} (E_j \backslash H_j)
$$

But each $E_j \backslash H_j$ has measure zero, the countable union of them also have measure zero. Therefore, $E = H \cup N_2$ where H is an F_σ set and $\mu(N_2) = 0$.

Now we apply previous case to E^c, which is also measurable. By previous proof, we have that there exists $H' \subseteq E^c$, $H' \in F_\sigma$, $\mu(E^c \backslash H') = 0$. Taking complement, we have

$$E = (E^c)^c = (H' \cup (E^c \backslash H'))^c = H'^c \cap (E^c \backslash H')^c = H'^c \backslash (E^c \backslash H')$$

By De-Morgan's law, if $H' \in F_\sigma$, then $H'^c \in G_\delta$. Therefore, let $V = H'^c$ and $N_1 = E^c \backslash H'$, we see that we can write $E = V \backslash N_1$ such that $V \in G_\delta$ and $\mu(N_2) = 0$. □

We end this section with the re-discover of Lebesgue measure (which should be taught in real analysis before), and introduce the extended real line for further uses.

Definition 2.5.7: Lebesgue Measure

The complete measure μ_F associated with the function $F(x) = x$ is called the **Lebesgue measure**.

It is sometimes convenient to consider the extended real system $\bar{\mathbb{R}} = [-\infty, \infty]$. In tradition, we define $0 \cdot \infty = 0$. We define Borel sets $\mathcal{B}(\bar{\mathbb{R}}) = \{E \subseteq \mathbb{R} : \bar{E} \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})\}$. The Borel measure will not change its value with adding one point $-\infty$ or $+\infty$. The corresponding Lebesuge-Stieltjes measures are just completion of this. It can be similarly verified that $\mathcal{B}(\bar{\mathbb{R}})$ is generated by the rays $(a, \infty]$ or $[-\infty, a)$ with $a \in \mathbb{R}$.

Chapter 3

Integration

In this section, we will define abstract integration, and explore its properties.

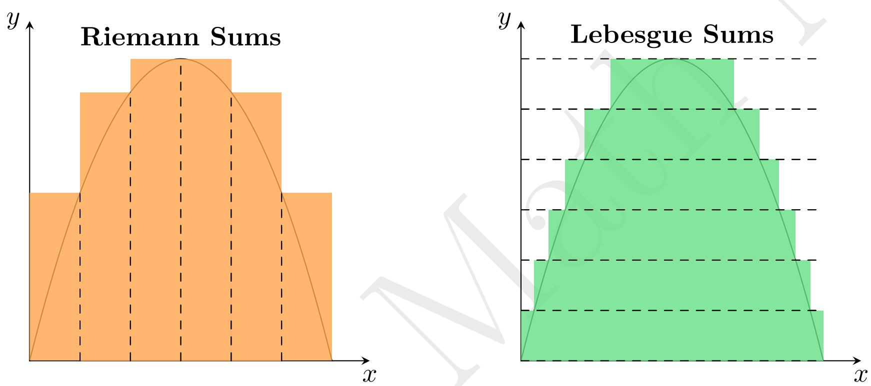


Figure 3.1: Riemann's Integration and Lebesgue's Integration

In introductory analysis course, the Riemann integration, as shown in figure 3.1 above, is defined using a 'vertical cut' on the function, where the height of each rectangle is defined using sup/inf of the function value over that region. When the sup/inf converges to a same value, we say that this function is 'Riemann integrable'.
Instead, Lebesgue gives a new thought of using the 'horizontal cut', to solve the problems such as the Riemann non-integrability of Dirichlet's function. However, this definition also brings a new challenge, that for each cut of the function value, the corresponding 'height' may not be an interval, or even may not be measurable. This gives us the motivation to define the *measurable function*.

3.1 Measurable Functions

Definition 3.1.1: Measurable Function

Let $(X, \mathcal{A}, \mu)$ be a measure space. The function $f : X \to \bar{\mathbb{R}}$ is **$\mathcal{A}$-measurable** (or just **measurable**) if for any $E \in \mathcal{B}(\bar{\mathbb{R}})$, we have $f^{-1}(E) \in \mathcal{A}$.

Normally, it is difficult to show that one function is measurable by this definition, since we need to verify for every Borel set, the corresponding inverse image is measurable. Fortunately, there exists easier ways using the following lemma. Moreover, the proof technique for this lemma is important, and will occur many times later.

- **Proof Technique: Generating Class Argument**
  To prove that all $A \in \mathcal{A}$ has some property P, we only need to prove that, there exists $\mathcal{E} \subseteq \mathcal{A}$, such that

  1. All sets in $\mathcal{E}$ has the property P.
  2. $\mathcal{A} \subseteq \sigma(\mathcal{E})$.
  3. $\mathcal{C} = \{A \in \mathcal{A} : A \text{ has the property } P\}$ is a σ-algebra.

  Then, $\mathcal{E} \subseteq \mathcal{C} \subseteq \mathcal{A} \subseteq \sigma(\mathcal{E})$, and $\sigma(\mathcal{E}) \subseteq \sigma(\mathcal{C}) = \mathcal{C}$, so $\mathcal{C} = \mathcal{A}$.

Lemma 3.1.2: Criterion of Measurability

Let $(X, \mathcal{A}, \mu)$ be a measure space. A function $f : X \to \bar{\mathbb{R}}$ is measurable if and only if

1. $f^{-1}((-\infty, x]) \in \mathcal{A}$ for all $x \in \mathbb{R}$, or
2. $f^{-1}((-\infty, x)) \in \mathcal{A}$ for all $x \in \mathbb{R}$, or
3. $f^{-1}([x, \infty)) \in \mathcal{A}$ for all $x \in \mathbb{R}$, or
4. $f^{-1}((x, \infty)) \in \mathcal{A}$ for all $x \in \mathbb{R}$.

Proof. ($\Longrightarrow$) Suppose f is measurable, then 1,2,3 and 4 are trivial.
($\Longleftarrow$) We only deal with 1, the other are similar. Suppose 1 holds. Then, since $\mathcal{B}(\bar{\mathbb{R}}) = \sigma((-\infty, x])$, by the generating class argument, we only need to prove that $\mathcal{C} = \{A \in \mathcal{B}(\bar{\mathbb{R}}) : f^{-1}(A) \in \mathcal{A}\}$ is a σ-algebra.

- First, $\emptyset \in \mathcal{C}$ since $\emptyset \in \mathcal{B}(\bar{\mathbb{R}})$, and $f^{-1}(\emptyset) = \emptyset \in \mathcal{A}$.
- Second, suppose $A \in \mathcal{C}$, then $A^c \in \mathcal{B}(\bar{\mathbb{R}})$, and $f^{-1}(A^c) = (f^{-1}(A))^c \in \mathcal{A}$.
- Finally, if $A_1, A_2, \cdots \in \mathcal{C}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{B}(\bar{\mathbb{R}})$, and $f^{-1}(\bigcup_{n=1}^{\infty} A_n) = \bigcup_{n=1}^{\infty} f^{-1}(A_n) \in \mathcal{A}$.

which completes the proof. □

3.2 Properties of Measurable Functions

In this section, we will state some important and useful properties of measurable functions. First of all, the measurability is closed under some simple algebra.

Proposition 3.2.1: Preservation of Measurability under Algebra

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, g : X \to \bar{\mathbb{R}}$ be measurable functions, and $\alpha \in \mathbb{R}$ be a constant. Then, (1) αf (2) $f + g$ (3) f^2 (4) $1/f$ (5) fg are all measurable.

Proof. (1): Suppose f is measurable, we have

$$\{\omega : \alpha f(\omega) \leqslant x\} = \begin{cases} \{x \geqslant 0\} \in \mathcal{A}, & \alpha = 0 \\ \{\omega : f(\omega) \leqslant x/\alpha\} \in \mathcal{A}, & \alpha > 0 \\ \{\omega : f(\omega) \geqslant x/\alpha\} \in \mathcal{A}, & \alpha < 0 \end{cases}$$

These are all measurable sets.

(2) For $f + g$, we have

$$\{\omega \in X : f(\omega) + g(\omega) < x\} = \bigcup_{r \in \mathbb{Q}} (\{\omega : f(\omega) < r\} \cap \{\omega : g(\omega) < x - r\}) \in \mathcal{A}$$

The RHS is measurable since both $\{\omega : f(\omega) < r\}$ and $\{\omega : g(\omega) < x - r\}$ are measurable, and measurability is closed under intersection and countable union. However, the implication above is not immediately clear. We will rigorously prove this.
($\subseteq$) First of all, take a point $\omega \in \{f + g < x\}$. Then, it indicates that $f(\omega) < x - g(\omega)$. Since rational number is dense, we can take some $r \in \mathbb{Q}$ such that $f(\omega) < r < x - g(\omega)$. This shows that $f(\omega) < r$ and $g(\omega) < x - r$. Since this is true for some r, we have $\omega \in \bigcup_{r \in \mathbb{Q}}(\{f < r\} \cap \{g < x - r\})$.
($\supseteq$) Then, suppose $\omega \in \bigcup_{r \in \mathbb{Q}}(\{f < r\} \cap \{g < x - r\})$. Immediately we have $f + g < r + x - r = x$, so $\omega \in \{f + g < x\}$. This completes the proof of the implication.

(3) $\{\omega \in X : f(\omega)^2 < x\} = \begin{cases} \emptyset, & x \leqslant 0 \\ \{\omega : -\sqrt{x} < f(\omega) < \sqrt{x}\}, & x > 0 \end{cases}$, all measurable.

(4) $\{\omega \in X : 1/f(\omega) < x\} = \begin{cases} \{\omega : f(\omega) < 0\} \cup \{\omega : f(\omega) > 1/x\} \in \mathcal{A}, & x > 0 \\ \{\omega : f(\omega) < 0\} \in \mathcal{A}, & x = 0 \\ \{\omega : 1/x < f(\omega) < 0\} \in \mathcal{A}, & x < 0 \end{cases}$, all measurable.

(5) $fg = \frac{1}{2}((f + g)^2 - f^2 - g^2)$, all components are measurable. □

Second, it is worth to note that measurability is also closed on limit operations.

Proposition 3.2.2: Preservation of Measurability under Limit

Let $(X, \mathcal{A}, \mu)$ be a measure space. Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of measurable functions in the space. Then, (1) $\sup_n f_n$ (2) $\inf_n f_n$ (3) $\limsup_n f_n$ (4) $\liminf_n f_n$ are all measurable. Moreover, (5) if $f_n \to f$, then f is measurable.

Proof. (1) and (2): $\{\omega : \sup_n f_n(\omega) > x\} = \bigcup_{n=1}^{\infty}\{\omega : f_n(\omega) > x\} \in \mathcal{A}$. $\inf_n f_n = -\sup_n\{-f_n\} \in \mathcal{A}$.
(3) and (4): $\limsup_n f_n = \inf_k \sup_{n \geqslant k} f_n \in \mathcal{A}$, $\liminf_n f_n = \sup_k \inf_{n \geqslant k} f_n \in \mathcal{A}$.
(5): If $f_n \to f$, then $f = \sup_n f_n = \inf_n f_n \in \mathcal{A}$. □

Corollary 3.2.3: Preservation of Measurability under max and min

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, g : X \to \bar{\mathbb{R}}$ be measurable. Then, $\max\{f, g\}$, $\min\{f, g\}$ are measurable.

Proof. We could take the sequence $\{f_n\}$ such that $f_1 = f$, $f_2 = f_3 = \cdots = g$. Then $\sup_n f_n = \max\{f, g\}$. Similar for the min function. $\square$

Remark: The preservation of measurability only works in countable operations. It does not necessarily hold for uncountable operations. For example, let $(\mathbb{R}, \mathcal{L}, \lambda)$ be the Lebesgue measure space. Let $A \notin \mathcal{L}$. Since all countable sets are Lebesgue measurable, A is uncountable. We define a function f_t such that

$$t \in \mathbb{R}, \quad f_t : \mathbb{R} \to \mathbb{R}, \quad f_t(x) = \begin{cases} 1, & x = t \\ 0, & x \neq t \end{cases}, \quad \sup_{t \in A} f_t = \mathbb{1}_A$$

Then, $\sup_t f_t$ is not measurable, but each f_t is measurable.

For future reference, we present a useful decomposition of functions.

Definition 3.2.4: Positive/Negative Part

Let $f : X \to \bar{\mathbb{R}}$. The **positive** and **negative part** of f is defined to be

$$f^+(x) = \max\{f(x), 0\}, \quad f^-(x) = \max\{-f(x), 0\}$$

Corollary 3.2.5: Preservation of Measurability under Positive/Negative Part and Absolute Value

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to \bar{\mathbb{R}}$ be a measurable function. Then, $f^+, f^-, |f|$ are all measurable.

Proof. f^+, f^- are measurable by Corollary 3.2.3, $|f|$ is measurable since $f = f^+ + f^-$, and by Proposition 3.2.1. $\square$

Finally, in topological spaces and Lebesgue measure space, there are other important properties.

Proposition 3.2.6: Continuity Implies Measurability

Let X be a topological space and $(X, \mathcal{A}, \mu)$ be a measure space. Suppose $\mathcal{A} \supseteq \mathcal{B}$, i.e., the σ-algebra contains the one that generated by all open sets. Let $f : X \to \bar{\mathbb{R}}$. Then, if f is continuous, then f is measurable.

Proof. $f^{-1}((-\infty, a))$ is always an open set for any a by the definition of continuity. Open sets is contained in $\mathcal{A}$. $\square$

Proposition 3.2.7: Almost Everywhere Equivalent Implies Measurability

Let $(\mathbb{R}, \mathcal{L}, \lambda)$ be the Lebesgue measure space. Let $f : \mathbb{R} \to \bar{\mathbb{R}}$ be $\mathcal{L}$-measurable. Let $g : \mathbb{R} \to \bar{\mathbb{R}}$ with $g = f$ a.e. Then, g is $\mathcal{L}$-measurable.

Proof. Let $A \in \mathcal{B}(\bar{\mathbb{R}})$. Since $g = f$ a.e., there exists $E \in \mathcal{L}$ such that $\lambda(E^c) = 0$ and $f(\omega) = g(\omega)$ for all $\omega \in E$. Then,

$$\begin{aligned} g^{-1}(A) = \{\omega : g(\omega) \in A\} &= (\{\omega : g(\omega) \in A\} \cap E) \cup (\{\omega : g(\omega) \in A\} \cap E^c) \\ &= \underbrace{(\{\omega : f(\omega) \in A\} \cap E)}_{\in \mathcal{A}} \cup \underbrace{(\{\omega : g(\omega) \in A\} \cap E^c)}_{\in \mathcal{A} \text{ since } \mathcal{L} \text{ is complete}} \end{aligned}$$

which completes the proof. $\square$

3.3 Simple Functions and Approximation

To formally define the integral, we need to first introduce an elementary family of functions.

Definition 3.3.1: Simple Functions

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $\{E_j\}_{j=1}^n$ be measurable and forms a finite partition of X, i.e., E_j's are disjoint, and the union of these sets is X. A **simple function** is of the form

$$f = \sum_{j=1}^{n} c_j \mathbb{1}_{E_j}, \quad c_j \in \mathbb{R}$$

Note that **simple functions are measurable**, since

$$f^{-1}(A) = \bigcup_{k: c_k \in A} E_k \in \mathcal{A}$$

Some Remark: The definition above is the **canonical form** of a simple functions. More generally, it can be written as $f = \sum_{j=1}^n a_j \mathbb{1}_{F_j}$, where F_j may not be disjoint. It is easy to show that every simple function can be written in the canonical form. Moreover, a function can have more than one representation in canonical form, which brings some trouble of defining integrals. This would be dealt later.

Here comes the core theorem of this section. It also forms the basis of defining the integral. The theorem below says that, every nonnegative measurable function can be written as the limit of an increasing sequence of simple functions.

Theorem 3.3.2: Approximation by Simple Function

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let f be a nonnegative measurable function in this space. Then, there exists a sequence $\{f_n\}_{n\in\mathbb{N}}$, where each f_n is a nonnegative simple function, such that $f_n \nearrow f$ pointwise.

Proof. We prove this in a constructive way. Define the function

$$f_n(x) = \begin{cases} n, & f(x) \geqslant n \\ \dfrac{k}{2^n}, & \dfrac{k}{2^n} \leqslant f(x) < \dfrac{k+1}{2^n}, \quad 0 \leqslant k \leqslant n2^n - 1 \end{cases}$$

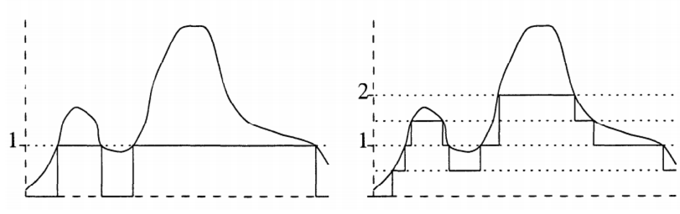


Figure 3.2: Function f and its approximation f_1 and f_2

STEP I: f_n can be written as

$$f_n(x) = \sum_{k=0}^{n2^n-1} \frac{k}{2^n}\mathbb{1}\left\{\frac{k}{2^n} \leqslant f(x) < \frac{k+1}{2^n}\right\} + n\mathbb{1}\{f(x) \geqslant n\}$$

Since $\mathbb{1}\left\{\frac{k}{2^n} \leqslant f(x) < \frac{k+1}{2^n}\right\}$ and $\mathbb{1}\{f(x) \geqslant n\}$ are all measurable sets (by the measurability of f and Lemma 3.1.2), we can see that each f_n is indeed nonnegative and simple.

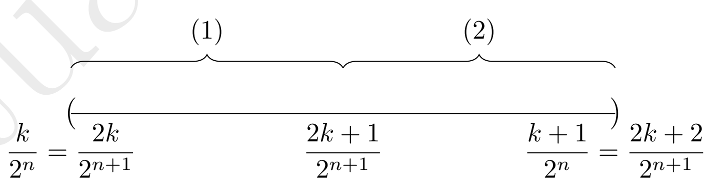
STEP II: Now we will show that f_n converges to f.

- If $f(x) = +\infty$, then $f_n(x) = n$ for all $n \geqslant 1$. Therefore, $f_n(x) \to f(x)$ in this case.
- If $f(x) < +\infty$, then we can always choose n_0 large enough such that $f(x) < n_0$. Let $n \geqslant n_0$. Then, there exists $0 \leqslant k \leqslant n2^n - 1$ such that $k \leqslant 2^n f(x) < k+1$, which means that $k = \lfloor 2^n f(x) \rfloor$. By definition, $f_n(x) = \frac{k}{2^n} = \frac{\lfloor 2^n f(x)\rfloor}{2^n}$. Using the fact that $\lfloor 2^n f(x)\rfloor \leqslant 2^n f(x) < \lfloor 2^n f(x)\rfloor + 1$, we have:
  - $f_n(x) = \frac{\lfloor 2^n f(x)\rfloor}{2^n} \leqslant \frac{2^n f(x)}{2^n} = f(x)$.
  - $f_n(x) > \frac{2^n f(x) - 1}{2^n} = f(x) - \frac{1}{2^n}$

  Therefore, $f_n(x) \leqslant f(x) < f_n(x) + \frac{1}{2^n}$, which shows that $f_n(x) \to f(x)$.

STEP III: Finally, we prove that $\{f_n\}$ is an increasing sequence of function. Similar as above, we separate into two cases.

- If $f(x) = +\infty$, then $f_n(x) = n$, and $f_{n+1}(x) = n+1 > f_n(x)$ for all n, it is increasing at x.
- If $f(x) < +\infty$, we separate into three cases:
  - *CASE I:* $n < n+1 \leqslant f(x)$. Then, $f_n(x) = n$, $f_{n+1}(x) = n+1 > f_n(x)$.
  - *CASE II:* $n < f(x) < n+1$. Then, $f_n(x) = n$, $f_{n+1}(x) = \frac{\lfloor 2^{n+1} f(x)\rfloor}{2^{n+1}} \geqslant \frac{\lfloor n2^{n+1}\rfloor}{2^{n+1}} = \frac{n2^{n+1}}{2^{n+1}} = n = f_n(x)$.
  - *CASE III:* $f(x) \leqslant n < n+1$. This is the most difficult case. We will further separate into two cases.



The proof idea is, no matter $f(x)$ is in area (1) or (2), $f_n(x) = \frac{k}{2^n}$. However, when n increases by 1, the interval is divided by two parts, i.e., (1) and (2), and if $f(x)$ is in area (1), $f_{n+1}(x) = f_n(x)$ since the f_n chooses the smallest value of that interval; if $f(x)$ is in area (2), $f_{n+1}(x) > f_n(x)$ since now, the smallest point changed to the middle point $\frac{2k+1}{2^{n+1}}$. Below we put this idea into rigour.

We have

$$f_n(x) = \frac{\lfloor 2^n f(x)\rfloor}{2^n}, \quad f_{n+1}(x) = \frac{\lfloor 2^{n+1} f(x)\rfloor}{2^{n+1}}$$

On one hand, suppose $\frac{k}{2^n} \leqslant f(x) < \frac{2k+1}{2^{n+1}}$, then $k \leqslant 2^n f(x) < k+1$, so $f_n(x) = \frac{\lfloor 2^n f(x) \rfloor}{2^n} = \frac{k}{2^n}$. And, $2k \leqslant 2^{n+1} f(x) < 2k+1$, thus $f_{n+1}(x) = \frac{\lfloor 2^{n+1} f(x) \rfloor}{2^{n+1}} = \frac{2k}{2^{n+1}} = \frac{k}{2^n} = f_n(x)$.
On the other hand, suppose $\frac{2k+1}{2^{n+1}} \leqslant f(x) < \frac{2k+2}{2^{n+1}}$, then $k \leqslant 2^n f(x) < k+1$, and $f_n(x) = \frac{\lfloor 2^n f(x) \rfloor}{2^n} = \frac{k}{2^n}$. And, $2k+1 \leqslant 2^{n+1} f(x) \leqslant 2k+2$. Thus, $f_{n+1}(x) = \frac{\lfloor 2^{n+1} f(x) \rfloor}{2^{n+1}} = \frac{2k+1}{2^{n+1}} > f_n(x)$.

In conclusion, f_n is nonnegative, simple, increasing, and converging to f, which completes the proof. □

3.4 Three Steps Towards the Definition of Integral

3.4.1 Integral of Nonnegative Simple Functions

Definition 3.4.1: Integral of Nonnegative Simple Functions

Let $(X, \mathcal{A}, \mu)$ be a measure space. The integral of a simple function $f : X \to \bar{\mathbb{R}}$, where $f(x) = \sum_{j=1}^{n} c_j \mathbb{1}_{E_j}(x)$, $c_j \geqslant 0$, is defined as

$$\int f \, \mathrm{d}\mu = \sum_{j=1}^{n} c_j \mu(E_j)$$

where we use the convention $0 \cdot \infty = 0$.

The convention is 'weird' but reasonable: if $f = 0$ on a set with infinite measure, we should identify the integral over this set as 0.

Moreover, this integral is **well-defined**: Suppose the function can be also written in the form $f = \sum_{k=1}^{m} d_k \mathbb{1}_{F_k}$ in canonical form. Then,

$$\mu(E_j) = \mu\left(E_j \cap \left(\bigcup_{k=1}^{m} F_k\right)\right) = \mu\left(\bigcup_{k=1}^{m} (E_j \cap F_k)\right) = \sum_{k=1}^{m} \mu(E_j \cap F_k)$$

where the last equality holds since $E_j \cap F_k$ are disjoint. Therefore, we can rewrite the integral as

$$\int f \, \mathrm{d}\mu = \sum_{j=1}^{n} c_j \mu(E_j) = \sum_{j=1}^{n} \sum_{k=1}^{m} c_j \mu(E_j \cap F_k) = \sum_{j=1}^{n} \sum_{k=1}^{m} d_k \mu(E_j \cap F_k) = \sum_{k=1}^{m} \mu(F_k)$$

where the third equality holds otherwise the two forms of f would not define the same function.

We will show that the integral defined here is linear.

Proposition 3.4.2: Linearity of Integral on Simple Functions

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, g : X \to \bar{\mathbb{R}}$ be two nonnegative simple functions. Then,

$$\int (f+g) \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu + \int g \, \mathrm{d}\mu, \quad \int kf \, \mathrm{d}\mu = k \int f \, \mathrm{d}\mu, \forall\, k \geqslant 0$$

Proof. Suppose $f = \sum_{j=1}^{n} c_j \mathbb{1}_{E_j}$ and $g = \sum_{k=1}^{m} d_k \mathbb{1}_{F_k}$, where $\{E_j\}$ and $\{F_k\}$ are partitions of X. Then, $E_j =$

$\sum_{k=1}^{m} E_j \cap F_k$, and $F_k = \sum_{j=1}^{m} E_j \cap F_k$. Thus, we can re-write these functions as

$$f = \sum_{j=1}^{n} \sum_{k=1}^{m} c_j \mathbb{1}_{E_j \cap F_k}, \quad g = \sum_{j=1}^{n} \sum_{k=1}^{m} d_k \mathbb{1}_{E_j \cap F_k}$$

Thus,

$$f + g = \sum_{j=1}^{n} \sum_{k=1}^{m} (c_j + d_k) \mathbb{1}_{E_j \cap F_k}$$

and we have

$$\begin{aligned} \int (f+g)\,\mathrm{d}\mu &= \sum_{j=1}^{n} \sum_{k=1}^{m} (c_j + d_k)\mu(E_j \cap F_k) = \sum_{j=1}^{n} c_j \sum_{k=1}^{m} \mu(E_j \cap F_k) + \sum_{k=1}^{m} d_k \sum_{j=1}^{n} \mu(E_j \cap F_k) \\ &= \sum_{j=1}^{n} c_j \mu(E_j) + \sum_{k=1}^{m} d_k \mu(F_k) = \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu \qquad \text{(Countable Additivity)} \end{aligned}$$

For the second statement, we have

$$kf = \sum_{j=1}^{n} kc_j \mathbb{1}_{E_j}, \quad \Longrightarrow \quad \int kf\,\mathrm{d}\mu = \sum_{j=1}^{n} kc_j \mu(E_j) = k \sum_{j=1}^{n} c_j \mu(E_j) = k \int f\,\mathrm{d}\mu$$

which completes the proof. □

3.4.2 Integral of Nonnegative Measurable Functions

Motivated by the approximation theorem 3.3.2, every nonnegative measurable function can be approximated by a sequence of increasing simple functions, we may define

$$\int f\,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n\,\mathrm{d}\mu$$

for any nonnegative measurable function f, and its simple approximation f_n. However, there are two major problems with this definition:

- The limit may not exist (even in the sense of extended real number).
- The limit may depend on the chosen approximating sequence.

However, the following lemma shows that these two worries are redundant.

Lemma 3.4.3: Well-Defined Integral Limit

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

- Suppose $f, g : X \to \bar{\mathbb{R}}$ are two nonnegative simple functions and $f \leqslant g$. Then, $\int f\,\mathrm{d}\mu \leqslant \int g\,\mathrm{d}\mu$.
- Suppose $\{f_n\}_{n=1}^{\infty}, \{g_n\}_{n=1}^{\infty}$ are two sequences of nonnegative increasing simple functions such that $f_n, g_n \nearrow f$. Then, $\lim_n \int f_n\,\mathrm{d}\mu = \lim_n \int g_n\,\mathrm{d}\mu$.

Proof. • For the first statement, suppose $f = \sum_{j=1}^n c_j \mathbb{1}_{E_j}, E_j \in \mathcal{A}$, $\{E_j\}$ being a partition of X. Let $g = \sum_{k=1}^m d_k \mathbb{1}_{F_k}$, $F_k \in \mathcal{A}$, $\{F_k\}$ being a partition of X. When $E_j \cap F_k \neq \emptyset$, we have $c_j \leqslant d_k$ since $f \leqslant g$. Thus, we can write

$$f = \sum_{j=1}^n \sum_{k=1}^m c_j \mathbb{1}_{E_j \cap F_k}, \quad g = \sum_{j=1}^n \sum_{k=1}^m d_k \mathbb{1}_{E_j \cap F_k}$$

and conclude that

$$\int g \, \mathrm{d}\mu = \sum_{j=1}^n \sum_{k=1}^m d_k \mu(E_j \cap F_k) \geqslant \sum_{j=1}^n \sum_{k=1}^m c_j \mu(E_j \cap F_k) = \int f \, \mathrm{d}\mu$$

• For the second statement, we need to first prove a lemma:

Let $\{f_n\}$ be a sequence of increasing non-negative simple functions, and let g be a non-negative simple function such that $g \leqslant \lim_n f_n$. Then, we have $\int g \, \mathrm{d}\mu \leqslant \lim_n \int f_n \, \mathrm{d}\mu$.

To prove this, we start with a simple case that $g = c\mathbb{1}_E$, where $E \in \mathcal{A}$ and $c \geqslant 0$. If $c = 0$, then $\int g \, \mathrm{d}\mu = 0$. Then $\int g \, \mathrm{d}\mu \leqslant \lim_n \int f_n \, \mathrm{d}\mu$ is trivially satisfied. So we focus on the case where $c > 0$. If $c > 0$, we always have $g \leqslant \lim_n \mathbb{1}_E f_n$ (If $x \in E$, $g \leqslant \lim_n f_n = \mathbb{1}_E f_n$. If $x \notin E$, $g = 0 = \lim_n \mathbb{1}_E f_n$). We can write $\mathbb{1}_E f_n$ as

$$\mathbb{1}_E f_n = \mathbb{1}_E \sum_{k=1}^{m_n} c_{n,k} \mathbb{1}_{F_{n,k}} = \sum_{k=1}^{m_n} c_{n,k} \mathbb{1}_{F_{n,k} \cap E}$$

which is also a simple function. Fix $0 < \epsilon < c$, let $A_n = \{x \in E : f_n(x) \geqslant c - \epsilon\}$. We note that

- $A_n \subseteq A_{n+1}$, it is an increasing sequence obviously, since f_n is increasing.
- $\bigcup_n A_n = E$. First, $\bigcup_n A_n \subseteq E$ by definition. Second $\bigcup_n A_n \supseteq E$: take $x \in E$, we have $g(x) = c \leqslant \lim_n f_n$. Thus, there exists n_0 such that when $N > n_0$, $f_N \geqslant c - \epsilon$ by the definition of limit, so $x \in A_N$.

Now, by monotonicity proved in statement 1, we have

$$\int f_n \, \mathrm{d}\mu \geqslant \int f_n \mathbb{1}_{A_n} \, \mathrm{d}\mu \geqslant \int (c-\epsilon) \mathbb{1}_{A_n} \, \mathrm{d}\mu = (c-\epsilon)\mu(A_n) \quad \Longrightarrow \quad (c-\epsilon) \lim_n \mu(A_n) = (c-\epsilon)\mu(E) \leqslant \lim_n \int f_n \, \mathrm{d}\mu$$

where the first and second inequality follows from monotonicity of integral, the implication is acieved by taking limit on both sides, and $\lim_n \mu(A_n) = \mu(E)$ by continuity from below.

- If $\mu(E) = \infty$, then $\lim_n \int f_n \, \mathrm{d}\mu = \infty$, then $\int g \, \mathrm{d}\mu \leqslant \lim_n \int f_n \, \mathrm{d}\mu$ is trivially satisfied.
- If $\mu(E) < \infty$, then $c\mu(E) - \epsilon\mu(E) \leqslant \lim_n \int f_n \, \mathrm{d}\mu$. Take $\epsilon \to 0$, we have $\int g \, \mathrm{d}\mu = c\mu(E) \leqslant \int f_n \, \mathrm{d}\mu$.

Now, we generalize to the case where $g = \sum_{k=1}^n c_k \mathbb{1}_{E_k}$, a general nonnegative simple function. We then have

$$\int g \, \mathrm{d}\mu = \sum_{k=1}^n c_k \mu(E_k) = \sum_{k=1}^n \int c_k \mathbb{1}_{E_k} \, \mathrm{d}\mu$$

Observe that $\lim_n f_n \mathbb{1}_{E_k} \geqslant c_k \mathbb{1}_{E_k}$ (For $x \in E_k$, $c_k \mathbb{1}_k(x) = g(x) \leqslant \lim_n f_n(x) = \lim_n f_n(x) \mathbb{1}_{E_k}(x)$. For $x \notin E_k$, both are 0). Thus, we can directly apply the previous result on $c_k \mathbb{1}_{E_k}$ and $\lim_n f_n \mathbb{1}_{E_k}$,

$$\int g \, \mathrm{d}\mu = \sum_{k=1}^n \int c_k \mathbb{1}_{E_k} \leqslant \sum_{k=1}^n \lim_n \int f_n \mathbb{1}_{E_k} \, \mathrm{d}\mu \qquad \text{(Proved when } g = c\mathbb{1}_E\text{)}$$

$$= \lim_n \sum_{k=1}^n \int f_n \mathbb{1}_{E_k} \, \mathrm{d}\mu \qquad \text{(Sum is finite, change order of sum and limit)}$$

$$= \lim_n \int \sum_{k=1}^n f_n \mathbb{1}_{E_k} \, \mathrm{d}\mu = \lim_n \int f_n \, \mathrm{d}\mu \qquad \text{(Linearity of integral on simple functions)}$$

which completes the proof. □

The first statement shows that the integral is monotonically increasing, so the limit exists in a extended real number sense. Second, the limit also does not depend on specific sequence. Therefore, we can just write the definition as below:

Definition 3.4.4: Integral of Nonnegative Measurable Functions

Let $(X, \mathcal{A}, \mu)$ be a measure space. The integral of a nonnegative measurable function $f : X \to \bar{\mathbb{R}}$ is defined as

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

This definition agrees the case when f is simple, as the family of simple functions over which the supremum is taken includes f itself.

3.4.3 Integral of Arbitrary Measurable Functions

The final step is easy. We just treat positive and negative part separately.

Definition 3.4.5: Integral of Arbitrary Measurable Functions

Let $(X, \mathcal{A}, \mu)$ be a measure space. The integral of a measurable function $f : X \to \bar{\mathbb{R}}$ is defined as

$$\int f \, \mathrm{d}\mu = \int f^+ \, \mathrm{d}\mu - \int f^- \, \mathrm{d}\mu$$

where f^+, f^- are positive and negative part, respectively. This is defined as long as one of the $\int f^+ \, \mathrm{d}\mu$ or $\int f^- \, \mathrm{d}\mu$ is finite. f is called **integrable** if $\int f^+ \, \mathrm{d}\mu, \int f^- \, \mathrm{d}\mu < \infty$.

A final definition is about an integral on a specific region.

Definition 3.4.6: Integral on Region

Let $(X, \mathcal{A}, \mu)$ be a measure space. The integral of a measurable function $f : X \to \bar{\mathbb{R}}$ on the set $E \in \mathcal{A}$ is defined as

$$\int_E f \, \mathrm{d}\mu = \int \mathbb{1}_E f \, \mathrm{d}\mu$$

3.5 Properties of Integral

We have proved linearity and monotonicity of integral on simple nonnegative functions before (Proposition 3.4.2 and Lemma 3.4.3). We will then start from nonnegative measurable functions.

3.5.1 Properties for Integral on Nonnegative Measurable Functions

We first recover the linearity and monotonicity from previous sections.

Property 3.5.1: Linearity of Integral for Nonnegative Measurable Functions

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, g : X \to \bar{\mathbb{R}}$ be two nonnegative measurable functions. Let $c \geqslant 0$ be an arbitrary constant. Then,

$$\int (f+g)\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu, \quad \int cf\,\mathrm{d}\mu = c\int f\,\mathrm{d}\mu$$

Proof. Let $f_n \nearrow f$ and $g_n \nearrow g$ be two sequences of nonnegative increasing simple functions that approximate f and g, respectively. Then, $f_n + g_n$ is a sequence of simple, nonnegative and increasing functions either, and $f_n + g_n \nearrow f + g$. Simiarly, $cf_n \nearrow cf$ is also a sequence of nonnegative increasing simple functions. Therefore,

$$\int (f+g)\,\mathrm{d}\mu = \lim_n \int (f_n + g_n)\,\mathrm{d}\mu = \lim_n \left(\int f_n\,\mathrm{d}\mu + \int g_n\,\mathrm{d}\mu \right) = \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu$$

where the first equality follows from definition, the second equality follows from Lemma 3.4.3 statement 1 (linearity of integral on nonnegative simple functions), and the third equality follows again by definition. Similarly,

$$\int cf\,\mathrm{d}\mu = \lim_n \int cf_n\,\mathrm{d}\mu = c\lim_n \int f_n\,\mathrm{d}\mu = c\int f\,\mathrm{d}\mu$$

which completes the proof. □

Lemma 3.5.2: Nonnegativity Implies Nonnegative Integral

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f \geqslant 0$ be a measurable function. Then, $\int f\,\mathrm{d}\mu \geqslant 0$.

Proof. Let $f_n \nearrow f$ be a sequence of nonnegative increasing simple functions. Since $\int f_n\,\mathrm{d}\mu \geqslant 0$ by definition, we have $\int f\,\mathrm{d}\mu = \lim_n \int f_n\,\mathrm{d}\mu \geqslant 0$. □

Property 3.5.3: Monotonicity of Integral for Nonnegative Measurable Functions

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, g : X \to \bar{\mathbb{R}}$ be two nonnegative measurable functions. Suppose $0 \leqslant g \leqslant f$, then

$$\int g\,\mathrm{d}\mu \leqslant \int f\,\mathrm{d}\mu$$

Proof. We have

$$f = g + (f-g) \quad \Longrightarrow \quad \int f\,\mathrm{d}\mu = \int (g + f - g)\,\mathrm{d}\mu = \int g\,\mathrm{d}\mu + \int (f-g)\,\mathrm{d}\mu \geqslant \int g\,\mathrm{d}\mu$$

where the second inequality follows from linearity 3.5.1, and the final inequality follows from that $f - g \geqslant 0$ and Lemma 3.5.2. □

To end this section, we show that the integral on a zero-measure set is always 0.

Property 3.5.4: Zero Integral on Zero Measure Set

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to \bar{\mathbb{R}}$ be a nonnegative measurable function. Suppose $\mu(E) = 0$. Then, $\int_E f \, \mathrm{d}\mu = 0$.

Proof. $\int_E f \, \mathrm{d}\mu = \int \mathbb{1}_E f \, \mathrm{d}\mu$. Let $g_n \nearrow \mathbb{1}_E f$ be a sequence of simple functions. Note that if $x \notin E$, $g_n(x) \leqslant \mathbb{1}_E f(x) = 0$. Thus, $g_n = g_n \mathbb{1}_E$. We have

$$\int g_n \, \mathrm{d}\mu := \sum_{k=1}^{m_n} c_{n,k} \mu(F_{n,k}) = \int \mathbb{1}_E g_n \, \mathrm{d}\mu = \sum_{k=1}^{m_n} c_{n,k} \underbrace{\mu(F_{n,k} \cap E)}_{=0} = 0$$

with the convention that $0 \cdot \infty = 0$. Therefore, by definition, $\int_E f \, \mathrm{d}\mu = \lim_n \int \mathbb{1}_E g_n \, \mathrm{d}\mu = 0$. □

3.5.2 Properties for Integral on Arbitrary Measurable Functions

Other than the case where f is nonnegative, when the integrand is arbitrary, the integral can be not well-defined. We can absolutely work on the case where the integral does exist. However, for simplicity, we will work on the case when f is integrable. The case where the integral exists is easy to extend.

Property 3.5.5: Linearity of Integral for Measurable Functions

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, g : X \to \bar{\mathbb{R}}$ be two measurable functions. Let $c \in \mathbb{R}$ be an arbitrary constant. Suppose f, g are integrable. Then, $f + g$, cf are also integrable, and

$$\int (f+g) \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu + \int g \, \mathrm{d}\mu, \quad \int cf \, \mathrm{d}\mu = c \int f \, \mathrm{d}\mu$$

Proof. We first show that $f + g$ and cf are integrable.
If f, g are integrable, then $\int f^+ \, \mathrm{d}\mu, \int f^- \, \mathrm{d}\mu, \int g^+ \, \mathrm{d}\mu, \int g^- \, \mathrm{d}\mu < \infty$. Then, $(f+g)^+ = f^+ + g^+$, so $\int (f+g)^+ \, \mathrm{d}\mu = \int f^+ \, \mathrm{d}\mu + \int g^+ \, \mathrm{d}\mu < \infty$ by Property 3.5.1. Similarly, we also have $(f+g)^- = f^- + g^-$, and $\int (f+g)^- \, \mathrm{d}\mu < \infty$.
Now consider cf. If $c = 0$, it is trivial that cf is integrable. If $c > 0$, then $(cf)^+ = cf^+$ and $(cf)^- = cf^-$, so $\int (cf)^+ \, \mathrm{d}\mu = c \int f^+ \, \mathrm{d}\mu < \infty$ and $\int (cf)^- \, \mathrm{d}\mu = c \int f^- \, \mathrm{d}\mu < \infty$ by Property 3.5.1. If $c < 0$, then $(cf)^+ = -cf^-$, and $(cf)^- = -cf^+$, we can get the result similarly.

Now we prove the two equations. We will first prove that:

$$f, g \geqslant 0, \text{ we have } \int (f-g) \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu - \int g \, \mathrm{d}\mu, \text{ when LHS and RHS are all well-defined}$$

Let $h = f - g = h^+ - h^-$. Then, $f + h^- = g + h^+$. This is well-defined for all cases: (1) when f, g are finite, it is trivial. (2) when $f(x) = \infty$ for some x, g finite, then $h(x) = \infty$, so $h^+(x) = \infty$, $h^-(x) = 0$, $f + h^-$ and $g + h^+$ are all well-defined. (3) When $g(x) = \infty$ for some x, f finite, then $h(x) = -\infty$, $h^+(x) = 0$, $h^-(x) = \infty$. Similarly, $f + h^-$ and $g + h^+$ are all well-defined. These can be also extended to cases where $f = -\infty$, etc.
Thus,

$$\int (f + h^-) \, \mathrm{d}\mu = \int (g + h^+) \, \mathrm{d}\mu \implies \int f \, \mathrm{d}\mu + \int h^- \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu + \int h^+ \, \mathrm{d}\mu \qquad \text{(Property 3.5.1)}$$

$$\implies \int f\,\mathrm{d}\mu - \int g\,\mathrm{d}\mu = \int h^+\,\mathrm{d}\mu - \int h^-\,\mathrm{d}\mu = \int h\,\mathrm{d}\mu = \int (f-g)\,\mathrm{d}\mu$$

With this conclusion, if we write $f+g = f^+ + g^+ - f^- - g^-$ (also, this make sense for all infinite cases), then

$$\begin{aligned}\int (f+g)\,\mathrm{d}\mu &= \int \left[(f^+ + g^+) - (f^- + g^-)\right]\,\mathrm{d}\mu = \int (f^+ + g^+)\,\mathrm{d}\mu - \int (f^- + g^-)\,\mathrm{d}\mu \\ &= \int f^+\,\mathrm{d}\mu + \int g^+\,\mathrm{d}\mu - \int f^-\,\mathrm{d}\mu - \int g^-\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu\end{aligned}$$

where the second equality follows the previous result, and the third equation follows linearity of integral 3.5.1.

Finally, to deal with cf, we separate into three cases.

- When $c = 0$, it is trivial.
- When $c > 0$, we have

$$\int cf\,\mathrm{d}\mu = \int (cf)^+\,\mathrm{d}\mu - \int (cf)^-\,\mathrm{d}\mu = \int cf^+\,\mathrm{d}\mu - \int cf^-\,\mathrm{d}\mu = c\left(\int f^+\,\mathrm{d}\mu - \int f^-\,\mathrm{d}\mu\right) = c\int f\,\mathrm{d}\mu$$

- When $c < 0$, we have

$$\int cf\,\mathrm{d}\mu = \int (cf)^+\,\mathrm{d}\mu - \int (cf)^-\,\mathrm{d}\mu = \int -cf^-\,\mathrm{d}\mu - \int -cf^+\,\mathrm{d}\mu = c\int f\,\mathrm{d}\mu$$

which completes the proof. $\square$

Property 3.5.6: Integrability Implies Finite Almost Everywhere

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to \bar{\mathbb{R}}$ be integrable. Then, $|f| < \infty$ a.e..

Proof. Since $|f| = f^+ + f^-$, and f is integrable, we have $|f|$ is integrable. Then,

$$n\mu\{|f| \geqslant n\} = \int_{\{|f|\geqslant n\}} n\,\mathrm{d}\mu \leqslant \int_{\{|f|\geqslant n\}} |f|\,\mathrm{d}\mu \leqslant \int |f|\,\mathrm{d}\mu \quad \implies \quad \mu\{|f| \geqslant n\} \leqslant \frac{1}{n}\int |f|\,\mathrm{d}\mu$$

Let $A_n = \{|f| \geqslant n\}$. Then, $A_n \supseteq A_{n+1}$ is decreasing. Moreover, $\bigcap_{n=1}^{\infty} A_n = \{|f| = \infty\}$. Thus,

$$\mu\{|f| = \infty\} = \lim_{n\to\infty} \mu\{|f| \geqslant n\} \leqslant \lim_{n\to\infty} \frac{1}{n}\int |f|\,\mathrm{d}\mu = 0$$

where the last term equals 0 since $|f|$ is integrable. $\square$

We can also recover the monotonicity.

Property 3.5.7: Monotonicity of Integral for Measurable Functions

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, g : X \to \bar{\mathbb{R}}$ be two measurable functions. Suppose $g \leqslant f$, then

$$\int g\,\mathrm{d}\mu \leqslant \int f\,\mathrm{d}\mu$$

Proof. By Lemma 3.5.2, since $f-g \geqslant 0$, we have $\int(f-g)\,\mathrm{d}\mu \geqslant 0$. Then, by linearity, $\int(f-g)\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu - \int g\,\mathrm{d}\mu \geqslant 0$. We have $\int f\,\mathrm{d}\mu \geqslant \int g\,\mathrm{d}\mu$. $\square$

Property 3.5.8: Additivity of Integration

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to \bar{\mathbb{R}}$ be an integrable function. Let $A, B \in \mathcal{A}$ be two disjoint sets, $A \cap B = \emptyset$. Then, $\mathbb{1}_{\mathbb{A}} f$ is integrable, and

$$\int_{A \cup B} f\,\mathrm{d}\mu = \int_A f\,\mathrm{d}\mu + \int_B f\,\mathrm{d}\mu$$

Proof. We first prove that $\mathbb{1}_A f$ is integrable. This is trivial since $\mathbb{1}_A f \leqslant f$, we have $\int_A f\,\mathrm{d}\mu \leqslant \int f\,\mathrm{d}\mu < \infty$. Then, we note that

$$\int_{A \cup B} f\,\mathrm{d}\mu = \int \mathbb{1}_{A \cup B} f\,\mathrm{d}\mu$$

and

$$(\mathbb{1}_{A \cup B} f)^+ = \mathbb{1}_{A \cup B} f^+ = \mathbb{1}_A f^+ + \mathbb{1}_B f^+$$

Similarly,

$$(\mathbb{1}_{A \cup B} f)^- = \mathbb{1}_{A \cup B} f^- = \mathbb{1}_A f^- + \mathbb{1}_B f^-$$

Therefore,

$$\begin{aligned}
\int \mathbb{1}_{A \cup B} f\,\mathrm{d}\mu &= \int (\mathbb{1}_{A \cup B} f)^+\,\mathrm{d}\mu - \int (\mathbb{1}_{A \cup B} f)^-\,\mathrm{d}\mu \\
&= \int \mathbb{1}_A f^+\,\mathrm{d}\mu + \int \mathbb{1}_B f^+\,\mathrm{d}\mu - \int \mathbb{1}_A f^-\,\mathrm{d}\mu - \int \mathbb{1}_B f^-\,\mathrm{d}\mu = \int_A f\,\mathrm{d}\mu + \int_B f\,\mathrm{d}\mu
\end{aligned}$$

which completes the proof. $\square$

Property 3.5.9: Triangle Inequality

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to \bar{\mathbb{R}}$ be a measurable function. Then,

$$\left|\int f\,\mathrm{d}\mu\right| \leqslant \int |f|\,\mathrm{d}\mu$$

Proof. We have

$$\left|\int f\,\mathrm{d}\mu\right| = \left|\int f^+\,\mathrm{d}\mu - \int f^-\,\mathrm{d}\mu\right| \leqslant \int f^+\,\mathrm{d}\mu + \int f^-\,\mathrm{d}\mu = \int (f^+ + f^-)\,\mathrm{d}\mu = \int |f|\,\mathrm{d}\mu$$

$\square$

Property 3.5.10: Zero Integral implies Zero a.e.

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to \bar{\mathbb{R}}$ be a nonnegative measurable function, with $\int f\,\mathrm{d}\mu = 0$. Then, $f = 0$ a.e..

Proof. Let $E_n = \{x \in X : f(x) \geqslant 1/n\}$. Then, by monotonicity, $\int \mathbb{1}_{E_n} f \, \mathrm{d}\mu \leqslant \int f \, \mathrm{d}\mu$. On the other hand, since $\mathbb{1}_{E_n} f \geqslant \frac{1}{n}\mathbb{1}_{E_n}$, we have $\int \mathbb{1}_{E_n} f \, \mathrm{d}\mu \geqslant \frac{1}{n}\int \mathbb{1}_{E_n} \, \mathrm{d}\mu$. Thus,

$$\frac{1}{n}\mu(E_n) = \int \frac{1}{n}\mathbb{1}_{E_n} \, \mathrm{d}\mu \leqslant \int \mathbb{1}_{E_n} f \, \mathrm{d}\mu \leqslant \int f \, \mathrm{d}\mu = 0 \quad \Longrightarrow \quad \mu(E_n) = 0$$

Let

$$E = \bigcup_{n=1}^{\infty} E_n = \{x \in X : f(x) > 0\}$$

Then,

$$\mu(E) \leqslant \sum_{n=1}^{\infty} \mu(E_n) = 0$$

which shows that $f = 0$ almost everywhere. □

We extend one of the previous result to arbitrary measurable function for further use.

Corollary 3.5.11: Zero Integral on Zero Measure Set

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to \bar{\mathbb{R}}$ be a measurable function. Suppose $\mu(E) = 0$. Then, $\int_E f \, \mathrm{d}\mu = 0$.

Proof. This has been proved in Property 3.5.4 for nonnegative measurable functions. Now if f is arbitrary, then $\int_E f \, \mathrm{d}\mu = \int_E f^+ \, \mathrm{d}\mu - \int_E f^- \, \mathrm{d}\mu = 0$ since both f^+ and f^- are nonnegative. □

Property 3.5.12: Almost everywhere Equal implies Same Integral

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, g : X \to \bar{\mathbb{R}}$ be two measurable functions, and $f = g$ a.e.. Then, $\int f \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu$.

Proof. Let $E = \{f = g\}$, since $f = g$ a.e., we have $\mu(E^c) = 0$. Then,

$$\begin{aligned} \int f \, \mathrm{d}\mu &= \int (\mathbb{1}_E + \mathbb{1}_{E^c}) f \, \mathrm{d}\mu = \int \mathbb{1}_E f \, \mathrm{d}\mu + \int \mathbb{1}_{E^c} f \, \mathrm{d}\mu \\ &= \int \mathbb{1}_E f \, \mathrm{d}\mu = \int \mathbb{1}_E g \, \mathrm{d}\mu = \int \mathbb{1}_E g \, \mathrm{d}\mu + \int \mathbb{1}_{E^c} g \, \mathrm{d}\mu = \int g \, \mathrm{d}\mu \end{aligned}$$

where we use Corollary 3.5.11 to deduce that $\int_{E^c} f \, \mathrm{d}\mu = \int_{E^c} g \, \mathrm{d}\mu = 0$, and use the fact that $f = g$ on the set E. □

Property 3.5.13: Integrability under Domination

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f, h : X \to \bar{\mathbb{R}}$ be measurable functions, and $|h| \leqslant f$. Then, if f is integrable, h is also integrable.

Proof. We have $h^+ \leqslant f$, thus $\int h^+ \, \mathrm{d}\mu \leqslant \int f \, \mathrm{d}\mu < \infty$. Similarly, $h^- \leqslant f$, so $\int h^- \, \mathrm{d}\mu \leqslant \int f \, \mathrm{d}\mu < \infty$. □

Corollary 3.5.14: Integrability of Bounded Functions on Finite-Measure Sets

Let $(X, \mathcal{A}, \mu)$ be a measure space. Let $f : X \to \bar{\mathbb{R}}$ be a measurable function. If $|f| \leqslant c$ on E, where $\mu(E) < \infty$, and $|f| = 0$ on E^c, then f is integrable.

Proof. Since $f^+, f^- \leqslant |f|$, we only need to show that $\int |f| < \infty$. However, note that

$$\int |f| \, \mathrm{d}\mu = \int |f|(\mathbb{1}_E + \mathbb{1}_{E^c}) \, \mathrm{d}\mu = \int \mathbb{1}_E |f| \, \mathrm{d}\mu + \int \mathbb{1}_{E^c} |f| \, \mathrm{d}\mu \leqslant \int c\mathbb{1}_E \, \mathrm{d}\mu + 0 = c\mu(E) < \infty$$

□

3.6 Convergence Theorems

Sometimes we cannot easily change the order of limit and integration. In this section, we have three powerful theorems that deal with this problem. For simplicity, we will assume that we work in a measure space $(X, \mathcal{A}, \mu)$ later on, unless otherwise stated.

3.6.1 Monotone Convergence Theorem (MCT)

Theorem 3.6.1: Monotone Convergence Theorem

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of nonnegative measurable functions. Suppose that $f_n \leqslant f_{n+1}$ for all n, and $f_n \nearrow f$. Then,

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu$$

Proof. There are two approaches that are very elegant, and I want to include both in my notes.

PROOF I: Use Purely the Convergence of Simple Functions

We first note that f is nonnegative and measurable, by Proposition 3.2.2. Now, we construct sequences of increasing nonnegative simple functions $\{g_{n,k}\}_{k\in\mathbb{N}}$ for $n \in \mathbb{N}$, such that $g_{n,k} \nearrow f_n$ in k for every n. It is shown as in below.

simple functions

$$\begin{array}{ccccc}
g_{11} & g_{12} & g_{13} & \cdots & \to f_1 \\
g_{21} & g_{22} & g_{23} & \cdots & \to f_2 \\
g_{31} & g_{32} & g_{33} & \cdots & \to f_3 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
 & & & & f
\end{array}$$

We have

$$\lim_{k\to\infty} \int g_{1k} \, \mathrm{d}\mu = \int f_1 \, \mathrm{d}\mu, \quad \cdots, \quad \lim_{k\to\infty} \int g_{nk} \, \mathrm{d}\mu = \int f_n \, \mathrm{d}\mu, \quad \cdots$$

We construct $g_k = \max_{1\leqslant n\leqslant k} g_{nk}$ (the maxima in the purple boxes shown in the figure above). Note that

- g_k are simple functions. It is also increasing, since

$$g_{k+1} = \max\left\{\max_{1\leqslant n\leqslant k} g_{n,k+1}, g_{k+1,k+1}\right\} \geqslant \max\left\{\max_{1\leqslant n\leqslant k} g_{n,k}, g_{k+1,k+1}\right\} \geqslant \max_{1\leqslant n\leqslant k} g_{n,k} = g_k$$

  where the first inequality holds since $g_{n,k+1} \geqslant g_{n,k}$ by definition.

- $g_k \nearrow f$. To prove this, we consider a row n in the 'simple function convergence matrix' above, and consider $n \leqslant k$, i.e., the tail of that row. Then, $g_{n,k} \leqslant f_n \leqslant f_k$. Thus, $g_k = \max_{1\leqslant n\leqslant k} g_{n,k} \leqslant f_n \leqslant f_k$. This shows that $g_k \leqslant f_k$ for $k \geqslant n$. Moreover, if we fix k, and consider $m \leqslant k$, we have $g_{m,k} \leqslant g_k$ by definition, so

$$g_{m,k} \leqslant g_k \leqslant f_k \quad \overset{\text{take } k\to\infty}{\Longrightarrow} \quad f_m \leqslant \lim_{k\to\infty} g_k \leqslant f \overset{\text{take } m\to\infty}{\Longrightarrow} \lim_{k\to\infty} g_k = f$$

Therefore,

$$\int f \,\mathrm{d}\mu = \lim_{k\to\infty} \int g_k \,\mathrm{d}\mu \leqslant \lim_{k\to\infty} \int f_k \,\mathrm{d}\mu$$

On the other hand, by monotonicity of integral (Property 3.5.7), it is easy to see that $\lim_{k\to\infty} \int f_k \,\mathrm{d}\mu \leqslant \int f \,\mathrm{d}\mu$ since $f_k \leqslant f$ for all $k \in \mathbb{N}$. Combining the two, we have

$$\int f \,\mathrm{d}\mu = \lim_{k\to\infty} \int f_k \,\mathrm{d}\mu$$

PROOF II: Use a 'Small epsilon Room' in Convergence

To construct this proof, we first need two lemmas:

Lemma 1: Let ϕ be a nonnegative simple function. If $A_1, A_2, \cdots$ are disjoint measurable sets and $A = \bigcup_{n=1}^{\infty} A_n$, then

$$\int_A \phi \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_{A_n} \phi \,\mathrm{d}\mu$$

To prove this lemma, we denote the canonical form of the simple function

$$\phi \mathbb{1}_A = \sum_{k=1}^{M} a_k \mathbb{1}_{E_k \cap A}, \quad \text{and} \quad \phi \mathbb{1}_{A_n} = \sum_{k=1}^{M} a_k \mathbb{1}_{E_k \cap A_n}$$

Therefore,

$$\begin{aligned}\int_A \phi \,\mathrm{d}\mu = \int \phi \mathbb{1}_A \,\mathrm{d}\mu &= \sum_{k=1}^{M} a_k \mu(E_k \cap A) = \sum_{k=1}^{M} a_k \sum_{n=1}^{\infty} \mu(E_k \cap A_n) \\ &= \sum_{n=1}^{\infty} \sum_{k=1}^{M} a_k \mu(E_k \cap A_n) = \sum_{n=1}^{\infty} \int \phi \mathbb{1}_{A_n} \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_{A_n} \phi \,\mathrm{d}\mu\end{aligned}$$

where the last equality in the first row comes from countable additivity of measure, and the first equality in the second row comes from nonnegativity of the terms (so we can change the order of summation).

Lemma 2: Let ϕ be a nonnegative simple function. If $A_1 \subseteq A_2 \subseteq \cdots$ are nested measurable sets and $A = \bigcup_{n=1}^{\infty} A_n$, *then*

$$\int_A \phi \, \mathrm{d}\mu = \lim_{n\to\infty} \int_{A_n} \phi \, \mathrm{d}\mu$$

To prove this lemma, note that if $\int_{A_k} \phi \, \mathrm{d}\mu = \infty$ for some k, then there is nothing to prove. So, we may assume that $\int_{A_k} \phi \, \mathrm{d}\mu < \infty$ for every k. If we set $A_0 = \emptyset$, then

$$\bigcup_{k=1}^{\infty} A_k = \bigcup_{j=1}^{\infty} (A_j \backslash A_{j-1})$$

and the sets on the RHS above are disjoint. Further,

$$A_{j+1} = A_j \cup (A_{j+1} \backslash A_j)$$

and the sets on RHS are disjoint. Therefore, by Property 3.5.8, we have

$$\int_{A_{j+1}} \phi \, \mathrm{d}\mu = \int_{A_j} \phi \, \mathrm{d}\mu + \int_{A_{j+1}\backslash A_j} \phi \, \mathrm{d}\mu$$

where all of these integrals are finite. Moreover, by Lemma 1, we see that

$$\int_A \phi \, \mathrm{d}\mu = \sum_{j=1}^{\infty} \int_{A_j \backslash A_{j-1}} \phi \, \mathrm{d}\mu = \lim_{N\to\infty} \sum_{j=1}^{N} \left(\int_{A_j} \phi \, \mathrm{d}\mu - \int_{A_{j-1}} \phi \, \mathrm{d}\mu \right) = \lim_{N\to\infty} \int_{A_N} \phi \, \mathrm{d}\mu - \int_{A_0} \phi \, \mathrm{d}\mu = \lim_{N\to\infty} \int_{A_N} \phi \, \mathrm{d}\mu$$

where in the last equality we use Corollary 3.5.11 since $\mu(A_0) = 0$.

Now we start to prove the Monotone Convergence Theorem. Similarly, we have $\lim_{k\to\infty} \int f_k \, \mathrm{d}\mu \leqslant \int f \, \mathrm{d}\mu$ by monotonicity. We only need to prove that $\int f \, \mathrm{d}\mu \leqslant \lim_{k\to\infty} \int f_k \, \mathrm{d}\mu$. Let ϕ be any simple function such that $0 \leqslant \phi \leqslant f$. We want to show that the supremum of integral of all such ϕ (which is the definition of integral of f) is less than the limit of integral of f_k. However, here is a difficulty: *f may be itself simple or 'locally simple', so ϕ may overlap with f somewhere, and in this case $f_k \leqslant \phi$ on that region for all k, so we cannot construct a 'less than or equal to inequality' based on this.* To solve this problem, we introduce a shrink factor $0 < \alpha < 1$ to shrink ϕ. Set $E_n = \{f_n \geqslant \alpha\phi\}$, and observe that

$$E_1 \subseteq E_2 \subseteq \cdots$$

Further, $\bigcup_n E_n = X$ (this is where we use the assumption $\alpha < 1$). The Lemma 2 thus implies that $\int_{E_n} \phi \, \mathrm{d}\mu \to \int_E \phi \, \mathrm{d}\mu$. Consequently,

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu = \limsup_{n\to\infty} \int f_n \, \mathrm{d}\mu \geqslant \limsup_{n\to\infty} \int_{E_n} f_n \, \mathrm{d}\mu \geqslant \limsup_{n\to\infty} \int_{E_n} \alpha\phi \, \mathrm{d}\mu = \alpha \int_E \phi \, \mathrm{d}\mu$$

where lim can be changed to lim sup since the limit exists. Let $\alpha \to 1$, we see that $\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu \geqslant \int \phi \, \mathrm{d}\mu$. Finally, taking the supremum over all such simple function ϕ, we obtain the inequality $\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu \geqslant \int f \, \mathrm{d}\mu$ □

3.6.2 Fatou's Lemma

If we do not have monotone sequence of functions, we cannot change the order of integration and limit. However, we at least have an inequality.

Theorem 3.6.2: Fatou's Lemma

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of nonnegative measurable functions. Then,

$$\int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu$$

Proof. Define

$$f(x) = \liminf_{n\to\infty} f_n(x) = \lim_{k\to\infty} \inf_{n\geqslant k} f_n(x) = \lim_{k\to\infty} g_k(x)$$

where

$$g_k = \inf_{n\geqslant k} f_n$$

The function g_k is nonnegative, monotonically increasing to f. By Monotone convergence theorem,

$$\int f \, \mathrm{d}\mu = \lim_{k\to\infty} \int g_k \, \mathrm{d}\mu$$

However, $g_k \leqslant f_k$ by definition, and therefore $\int g_k \, \mathrm{d}\mu \leqslant \int f_k \, \mathrm{d}\mu$ for every k. Consequently,

$$\int f \, \mathrm{d}\mu = \lim_{k\to\infty} \int g_k \, \mathrm{d}\mu = \liminf_{k\to\infty} \int g_k \, \mathrm{d}\mu \leqslant \liminf_{k\to\infty} \int f_k \, \mathrm{d}\mu$$

which completes the proof. □

Example 1: Let

$$f_n = \begin{cases} 1, & x > n \\ 0, & x \leqslant n \end{cases}$$

Then, $\int f_n \, \mathrm{d}\mu = \infty$ for any n. However, $\liminf f_n = 0$, so

$$\int \liminf f_n \, \mathrm{d}\mu = 0 < \infty = \liminf \int f_n \, \mathrm{d}\mu$$

Note that we can replace $f_n \geqslant 0$ by $f_n \geqslant g$ where g is some integrable function, and the argument still works.

Corollary 3.6.3: Fatou's Lemma with Relaxed Condition

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions such that $f_n \geqslant g$ for all n, where g is an integrable function. Then,

$$\int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu$$

Proof. Define $h_n(x) = f_n(x) - g$. Then, $h_n(x) \geqslant 0$ for all n. Apply Fatou's Lemma, we have

$$\begin{aligned} &\int \liminf_{n\to\infty} h_n \, \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int h_n \, \mathrm{d}\mu \quad \Longrightarrow \quad \int \liminf_{n\to\infty} (f_n - g) \, \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int (f_n - g) \, \mathrm{d}\mu \\ \Longrightarrow \quad &\int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu - \int g \, \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu - \int g \, \mathrm{d}\mu \end{aligned}$$

which results in the same inequality after adding $\int g \, \mathrm{d}\mu$ on both sides. □

Corollary 3.6.4: Reversed Fatou's Lemma

Let $\{h_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions such that $h_n \leqslant g$ for all n, where g is an integrable function. Then,

$$\int \limsup_{n\to\infty} h_n \, \mathrm{d}\mu \geqslant \limsup_{n\to\infty} \int h_n \, \mathrm{d}\mu$$

Proof. $f_n = -h_n \geqslant -g$, where $-g$ is also integrable. Therefore, by the Fatou's lemma with relaxed condition, we have

$$\int \liminf_{n\to\infty}(-h_n) \, \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int (-h_n) \, \mathrm{d}\mu \quad \Longrightarrow \quad \int -\limsup_{n\to\infty} h_n \, \mathrm{d}\mu \leqslant -\limsup_{n\to\infty} \int h_n \, \mathrm{d}\mu$$

which leads to the desired inequality after multiplying both sides by -1. □

3.6.3 Dominated Convergence Theorem (DCT)

The assumption of Monotone convergence theorem is too strong and may not be applicable sometimes. Here is a weaker one.

Theorem 3.6.5: Dominated Convergence Theorem (DCT)

Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions. If $f_n \to f$ and $|f_n| \leqslant g$ for all n, where g is an integrable function. Then, f is integrable, and

$$\lim_{n\to\infty} \int f_n \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu$$

Proof. We have $f_n \leqslant g$ and $f_n \geqslant -g$. Therefore, by Fatou's lemma with relaxed condition,

$$\int \limsup_{n\to\infty} f_n \, \mathrm{d}\mu \geqslant \limsup_{n\to\infty} \int f_n \, \mathrm{d}\mu, \quad \text{and} \quad \int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu$$

Note that when limit exists, $\limsup = \lim = \liminf$. Therefore,

$$\limsup_{n\to\infty} \int f_n \, \mathrm{d}\mu \leqslant \int \limsup_{n\to\infty} f_n \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu = \int \liminf_{n\to\infty} f_n \, \mathrm{d}\mu \leqslant \liminf_{n\to\infty} \int f_n \, \mathrm{d}\mu$$

However, $\limsup \geqslant \liminf$, so everything in this chain becomes equality. □

Remark: For MCT, Fatou's Lemma and DCT, all asumptions can be replaced by 'almost everywhere' version, since zero measure set would not influence the value of integration. For example, in MCT, we can set $f_n \geqslant 0$ a.e., and $f_n \nearrow f$ a.e.; In Fatou's lemma, we can have $f_n \geqslant g$ a.e.; In DCT, we can have $f_n \to f$ a.e. and $|f_n| \leqslant g$ a.e..

Chapter 4

Related Spaces and Measures

4.1 A Step Backwards to Carathéodory's Extension Theorem

Before we formally go into this chapter, it is convenient to go back to Carathéodory's extension theorem, and to have some more elegant idea of how we should extend measures on semi-algebra to the σ-algebra generated by that.

Theorem 4.1.1: Extension from Semialgebra to σ-algebra

Let $\mathcal{S}$ be a semi-algebra and let μ defined on $\mathcal{S}$ have $\mu(\emptyset) = 0$. Suppose

(i) If $S \in \mathcal{S}$ is a finite disjoint union of sets $\{S_i\}_{i=1}^n \subseteq \mathcal{S}$, then $\mu(S) = \sum_{i=1}^n \mu(S_i)$.

(ii) If $\{S_i\}_{i=1}^\infty \subseteq S$ are disjoint and $S \in \mathcal{S}$, with $S = \bigsqcup_{i=1}^\infty S_i$, then $\mu(S) \leqslant \sum_{i=1}^\infty \mu(S_i)$.

Then, μ has a unique extension $\bar{\mu}$ that is a pre-measure on $\mathcal{R}(\mathcal{S})$. If the extension is σ-finite, then there exists a unique extension ν that is a measure on $\sigma(\mathcal{S})$.

Proof. Theorem 1.3.10 told us that $\mathcal{R}(\mathcal{S})$ is the collection of all finite disjoint unions of sets in $\mathcal{S}$. Therefore, we define $\bar{\mu}$ on $\mathcal{R}(\mathcal{S})$ by

$$\bar{\mu}(A) = \sum_{i=1}^n \mu(S_i), \quad S_i \in \mathcal{S},\ A = \bigsqcup_{i=1}^n S_i,\ S_i \text{ disjoint}$$

STEP I: Well-definedness First, we need to show that $\bar{\mu}$ is well-defined. Suppose that, also, $\bar{\mu}(A) = \sum_{i=1}^n \mu(T_i)$, where $T_i \in \mathcal{S}$, $A = \bigsqcup_{i=1}^n T_i$. Then, observe that,

$$S_i = \bigsqcup_{j=1}^n (S_i \cap T_j), \quad T_j = \bigsqcup_{i=1}^n (S_i \cap T_j)$$

Therefore,

$$\sum_{i=1}^n \mu(S_i) = \sum_{i,j=1}^n \mu(S_i \cap T_j) = \sum_{j=1}^n \mu(T_j)$$

STEP II: Finite Additivity Then, we prove the following property:

If $A \in \mathcal{R}(\mathcal{S})$ and $\{B_i\}_{i=1}^n \subseteq \mathcal{R}(\mathcal{S})$ disjoint, with $A = \bigsqcup_{i=1}^n B_i$. Then, $\bar{\mu}(A) = \sum_{i=1}^n \bar{\mu}(B_i)$

Suppose that $B_i = \bigsqcup_{j=1}^n S_{i,j}$, where $S_{i,j} \in \mathcal{S}$ are disjoint. Then, $A = \bigsqcup_{i=1}^n B_i = \bigsqcup_{i,j=1}^n S_{i,j}$. Then, by the definition of $\bar{\mu}$, we have

$$\bar{\mu}(A) = \sum_{i,j=1}^n \mu(S_{i,j}) = \sum_{i=1}^n \left(\sum_{j=1}^n \mu(S_{i,j}) \right) = \sum_{i=1}^n \bar{\mu}(B_i)$$

STEP III: Finite Subadditivity Then, we prove the following property:

If $A \in \mathcal{R}(\mathcal{S})$ and $\{B_i\}_{i=1}^n \subseteq \mathcal{R}(\mathcal{S})$, not necessarily disjoint, with $A \subseteq \bigcup_{i=1}^n B_i$. Then, $\bar{\mu}(A) \leqslant \sum_{i=1}^n \bar{\mu}(B_i)$

We begin with $n = 1$ case. Let $B_1 = B$, $B = A + (B \cap A^c)$. Note that $B \cap A^c = \mathcal{R}(\mathcal{S})$. So,

$$\bar{\mu}(A) \leqslant \bar{\mu}(A) + \bar{\mu}(B \cap A^c) = \bar{\mu}(B)$$

where the last equality is by STEP II. To handle $n > 1$, let $F_k = B_1^c \cap B_2^c \cap \cdots \cap B_{k-1}^c \cap B_k$. Then, note that F_i's are disjoint, and $\bigcup_{i=1}^n B_i = F_1 \sqcup F_2 \sqcup \cdots \sqcup F_n$, and $A = A \cap (\cup_{i=1}^n B_i) = (A \cap F_1) \sqcup (A \cap F_2) \sqcup \cdots \sqcup (A \cap F_n)$. Then, we have

$$\begin{aligned} \bar{\mu}(A) &= \sum_{k=1}^n \bar{\mu}(A \cap F_k) && \text{(By (a))} \\ &\leqslant \sum_{k=1}^n \bar{\mu}(F_k) && \text{((By (b) first part } n = 1\text{))} \\ &= \bar{\mu}\left(\bigcup_{i=1}^n B_i \right) && \text{(By (a))} \end{aligned}$$

STEP IV: Countable Additivity Now we extend the additivity to countably many disjoint unions. Suppose $A \in \mathcal{R}(\mathcal{S})$ and $A = \bigsqcup_{i=1}^\infty B_i$, where $B_i \in \mathcal{R}(\mathcal{S})$ are disjoint. Let $B_i = \bigsqcup_{j=1}^n S_{i,j}$ where $S_{i,j} \in \mathcal{S}$ are disjoint. Then, $A = \bigsqcup_{i=1}^\infty \bigsqcup_{j=1}^n S_{i,j}$. So, replacing B_i's by $S_{i,j}$'s, we may assume that $B_i \in \mathcal{S}$. Since $A \in \mathcal{R}(\mathcal{S})$, we can write $A = \bigsqcup_{j=1}^n T_j$, where $T_j \in \mathcal{S}$ are disjoint. Moreover, $T_j = \bigsqcup_{i=1}^\infty (T_j \cap B_i)$. By the property (ii), we have

$$\mu(T_j) \leqslant \sum_{i=1}^\infty \mu(T_j \cap B_i)$$

Summing over j and exchanging the order of summing, we have

$$\bar{\mu}(A) = \sum_{j=1}^n \mu(T_j) \leqslant \sum_{j=1}^n \sum_{i=1}^\infty \mu(T_j \cap B_i) = \sum_{i=1}^\infty \sum_{j=1}^n \mu(T_j \cap B_i) = \sum_{i=1}^\infty \mu(B_i)$$

where the last equality follows from property (i). To prove the opposite inequality, let $A_n = B_1 \sqcup B_2 \sqcup \cdots \sqcup B_n$, and $C_n = A \cap A_n^c$. Note that $C_n \in \mathcal{R}(\mathcal{S})$. By finitely additivity of $\bar{\mu}$, we then have

$$\bar{\mu}(A) = \bar{\mu}(C_n) + \sum_{i=1}^n \bar{\mu}(B_i) \geqslant \sum_{i=1}^n \bar{\mu}(B_i)$$

Taking $n \to \infty$, we have the desired reversed inequality. After these four steps, we have extended μ to a pre-measure on an algebra. By Carathéodory extension theorem, if it is σ-finite, we can further extend it to σ-algebra. □

The takeaway is that, if we construct a function μ on a semi-algebra that is countably additive and $\mu(\emptyset) = 0$, then we can then extend it to an algebra, and then σ-algebra if it is σ-finite.

4.2 Product Measures

Suppose we have two measure spaces $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$. We want to consider the product space

$$X = X_1 \times X_2 = \{(x_1, x_2) : x_1 \in X_1, x_2 \in X_2\}$$

and define the corresponding *product σ-algebra* $\mathcal{A}_1 \otimes \mathcal{A}_2$ and *product measure* μ on it. We will first construct the product σ-algebra by considering the *rectangles*.

Definition 4.2.1: Rectangle

Let $E_1 \in \mathcal{A}_1$ and $E_2 \in \mathcal{A}_2$. Then, the sets of the form

$$E_1 \times E_2 = \{(x_1, x_2) : x_1 \in E_1, x_2 \in E_2\}$$

are called **rectangles**.

Now it's natural to first consider the 'measure' on rectangles defined as

$$\mathcal{A}_1 \times \mathcal{A}_2 = \{E_1 \times E_2 : E_1 \in \mathcal{A}_1, E_2 \in \mathcal{A}_2\}, \quad \mu(E_1 \times E_2) = \mu_1(E_1)\mu_2(E_2)$$

Now, if we can show that $\mathcal{A}_1 \times \mathcal{A}_2$ is a semi-algebra, and construct a pre-measure on the algebra generated by $\mathcal{A}_1 \times \mathcal{A}_2$, we can then use Carathéodory's extension theorem to extend it to a measure.

Proposition 4.2.2: $\mathcal{A}_1 \times \mathcal{A}_2$ is a semi-algebra

The set of all rectangles, $\mathcal{A}_1 \times \mathcal{A}_2$, is a semi-algebra on the space $X_1 \times X_2$.

Proof. First, $\emptyset = \emptyset \times \emptyset \in \mathcal{A}_1 \times \mathcal{A}_2$ since $\mathcal{A}_1$ and $\mathcal{A}_2$ are both σ-algebras, hence $\emptyset \in \mathcal{A}_1$ and $\emptyset \in \mathcal{A}_2$.
Second, for $A, B \in \mathcal{A}_1 \times \mathcal{A}_2$, we can write $A = E_1 \times E_2$, $B = F_1 \times F_2$, where $E_1, F_1 \in \mathcal{A}_1$ and $E_2, F_2 \in \mathcal{A}_2$. Then,

$$A \cap B = (E_1 \times E_2) \cap (F_1 \times F_2) = (E_1 \cap F_1) \times (E_2 \cap F_2)$$

since $\mathcal{A}_1$ and $\mathcal{A}_2$ are σ-algebra, we have $E_1 \cap F_1 \in \mathcal{A}_1$ and $E_2 \cap F_2 \in \mathcal{A}_2$. Hence, $A \cap B \in \mathcal{A}_1 \times \mathcal{A}_2$.
Finally, for $A \in \mathcal{A}_1 \times \mathcal{A}_2$, we can write $A = E_1 \times E_2$ where $E_1 \in \mathcal{A}_1$ and $E_2 \in \mathcal{A}_2$. Then,

$$A^c = (E_1 \times E_2^c) \cup (E_1^c \times E_2) \cup (E_1^c \times E_2^c)$$

which is shown in the figure below. These three sets are disjoint, and each of them belongs to $\mathcal{A}_1 \times \mathcal{A}_2$ by the property of σ-algebra. These three arguments show that $\mathcal{A}_1 \times \mathcal{A}_2$ is a semi-algebra. $\square$

Remark: The figure is just for display purpose. E_1 and E_2 can be any measurable sets, and $E_1 \times E_2$ can be not a 'rectangle shape' even if they are called rectangle.

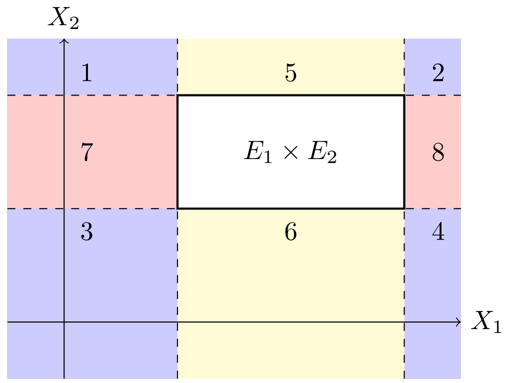


Figure 4.1: Blue: $E_1^c \times E_2^c$. Yellow: $E_1 \times E_2^c$. Red: $E_1^c \times E_2$

Definition 4.2.3: Product σ-algebra

Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be two measure spaces. The **product σ-algebra** on the space $X_1 \times X_2$ is defined as

$$\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{A}_1 \times \mathcal{A}_2)$$

Recall we have defined the measure on semi-algebra $\mathcal{A}_1 \times \mathcal{A}_2$ such that $\mu(E_1 \times E_2) = \mu_1(E_1) \times \mu_2(E_2)$. Now the goal is to extend this measure to a pre-measure on the algebra generated by $\mathcal{A}_1 \times \mathcal{A}_2$, which is just the set of finite disjoint union of rectangles.

Lemma 4.2.4: Section Lemma

Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be two measure spaces. Suppose $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$. Define the **$x$-section** of A and **y-section** of A by

$$A_x = \{y \in X_2 : (x, y) \in A\} \subseteq X_2, \quad A^y = \{x \in X_1 : (x, y) \in A\} \subseteq X_1$$

respectively. Then, for all $x \in X_1$, we have $A_x \in \mathcal{A}_2$. Similarly, for all $y \in X_2$, we have $A^y \in \mathcal{A}_1$.

Proof. Since we are going to prove a property for all the sets in a σ-algebra, we are going to use the generating class argument. Let

$$\mathcal{C} = \{A \in \mathcal{A}_1 \otimes \mathcal{A}_2 : A \text{ satisfies these desired properties}\}$$

(a) We first show that $\mathcal{C} \supseteq \mathcal{A}_1 \times \mathcal{A}_2$. Take $A \in \mathcal{A}_1 \times \mathcal{A}_2$. Then, $A = E_1 \times E_2$ with $E_1 \in \mathcal{A}_1$ and $E_2 \in \mathcal{A}_2$. Then,

$$A_x = \begin{cases} E_2, & x \in E_1 \\ \emptyset, & x \in E_1^c \end{cases} \in \mathcal{A}_2, \quad A^y = \begin{cases} E_1, & y \in E_2 \\ \emptyset, & y \in E_2^c \end{cases} \in \mathcal{A}_1$$

Therefore, $A \in \mathcal{C}$, which means that $\mathcal{C} \supseteq \mathcal{A}_1 \times \mathcal{A}_2$.

(b) Next, we will show that $\mathcal{C}$ is a σ-algebra.

- $X = X_1 \times X_2 \in \mathcal{C}$. This is because, fix $x \in X_1$, we have $X_x = (X_1 \times X_2)_x = X_2 \in \mathcal{A}_2$, and similar for y.

- Let $A \in \mathcal{C}$. We need to show that $A^c \in \mathcal{C}$. Note that

$$(A^c)_x = \{y \in X_2 : (x,y) \in A^c\} = \{y \in X_2 : (x,y) \notin A\} = \{y \in X_2 : (x,y) \in A\}^c = (A_x)^c \in \mathcal{A}_2$$

  and similarly for y.

- Finally, we need to show that for $\{A_j\}_{j\in\mathbb{N}} \subseteq \mathcal{C}$, we have $\bigcup_{j=1}^{\infty} A_j \in \mathcal{C}$. Fix $x \in X_1$, we have

$$\left(\bigcup_{j=1}^{\infty} A_j\right)_x = \left\{y \in X_2 : (x,y) \in \bigcup_{j=1}^{\infty} A_j\right\} = \bigcup_{j=1}^{\infty}\{y \in X_2 : (x,y) \in A_j\} = \bigcup_{j=1}^{\infty}(A_j)_x \in \mathcal{A}_2$$

  and similarly for y.

Therefore, $C \supseteq \sigma(\mathcal{A}_1 \times \mathcal{A}_2) = \mathcal{A}_1 \otimes \mathcal{A}_2$. ∎

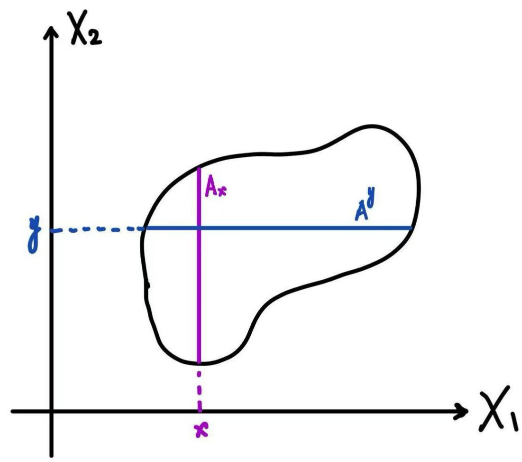


Figure 4.2: x-section and y-section of A

Now we show that μ is σ-additive on the semi-algebra.

Proposition 4.2.5: Finitely Additive on Semi-algebra

Let $A, \{A_j\}_{j=1}^n \subseteq \mathcal{A}_1 \times \mathcal{A}_2$ and $A = \bigcup_{j=1}^n A_j$, where A_j's are disjoint. Then,

$$\mu(A) = \sum_{j=1}^n \mu(A_j)$$

Proof. Since $A \in \mathcal{A}_1 \times \mathcal{A}_2$, we can write $A = E \times F$ where $E \in \mathcal{A}_1$ and $F \in \mathcal{A}_2$. Then,

$$A_x = \begin{cases} \emptyset, & x \notin E \\ F, & x \in E \end{cases}$$

From the proof of Lemma 4.2.4, we have seen that $A_x = \left(\bigcup_{j=1}^n A_j\right)_x = \bigcup_{j=1}^n (A_j)_x$. Notice that the union in $\bigcup_{j=1}^\infty (A_j)_x$ is disjoint, since for $j \neq k$,

$$(A_j)_x \cap (A_k)_x = \{y \in X_2 : (x,y) \in A_j\} \cap \{y \in X_2 : (x,y) \in A_k\} = \emptyset$$

since A_j and A_k are disjoint. Now, write each A_j as $A_j = E_j \times F_j$, where $E_j \in \mathcal{A}_1$ and $F_j \in \mathcal{A}_2$. Then,

$$(A_j)_x = \begin{cases} \emptyset, & x \notin E_j \\ F_j, & x \in E_j \end{cases}$$

Thus,

$$A_x = \bigcup_{j=1}^n (A_j)_x = \bigcup_{j=1}^n F_j \mathbb{1}_{E_j}(x)$$

Now suppose $x \in E$, then $A_x = F$, since μ_2 is additive,

$$\mu_2(F) = \sum_{j=1}^n \mu_2(F_j)\mathbb{1}_{E_j}(x) = \mathbb{1}_E(x)\mu_2(F)$$

where all $\mathbb{1}_{E_j}(x) = 1$ since $x \in E$. Taking integral both sides, we have

$$\int \mathbb{1}_E(x)\mu_2(F)\,\mathrm{d}\mu_1 = \int \sum_{j=1}^n \mu_2(F_j)\mathbb{1}_{E_j}(x)\,\mathrm{d}\mu_1$$

On the LHS, we have $\int \mathbb{1}_E(x)\mu_2(F)\,\mathrm{d}\mu_1 = \mu_1(E)\mu_2(F) = \mu(A)$. On the RHS, we have $\int \sum_{j=1}^n \mu_2(F_j)\mathbb{1}_{E_j}(x)\,\mathrm{d}\mu_1 = \sum_{j=1}^n \int \mu_2(F_j)\mathbb{1}_{E_j}(x)\,\mathrm{d}\mu_1 = \sum_{j=1}^n \mu(A_j)$ by linearity. This shows that $\mu(A) = \sum_{j=1}^n \mu(A_j)$. □

Proposition 4.2.6: Countable additive on Semi-algebra

Let $A, \{A_j\}_{j=1}^\infty \subseteq \mathcal{A}_1 \times \mathcal{A}_2$ and $A = \bigcup_{j=1}^\infty A_j$, where A_j's are disjoint. Then,

$$\mu(A) = \sum_{j=1}^\infty \mu(A_j)$$

Proof. We borrow all the notation from the previous proof. Similar with the previous proof, by the proof of Lemma 4.2.4, we have seen that

$$A_x = \bigcup_{j=1}^\infty (A_j)_x = \bigcup_{j=1}^\infty F_j \mathbb{1}_{E_j}(x)$$

The integral then becomes

$$\mu(A) = \int \mathbb{1}_E(x)\mu_2(F)\,\mathrm{d}\mu_1 = \int \sum_{j=1}^\infty \mu_2(F_j)\mathbb{1}_{E_j}(x)\,\mathrm{d}\mu_1$$

Instead of using linearity, we use the monotone convergence theorem. Since $\sum_{j=1}^n \mu_2(F_j)\mathbb{1}_{E_j}\,\mathrm{d}\mu_1 \nearrow \sum_{j=1}^\infty \mu_2(F_j)\mathbb{1}_{E_j}\,\mathrm{d}\mu_1$, we can interchange the sum and the integral, and get $\mu(A) = \sum_{j=1}^\infty \mu(A_j)$. □

Therefore, μ **satisfies the properties in Theorem 4.1.1, and we can use Theorem 4.1.1 to extend it onto**

$\mathcal{R}(\mathcal{A}_1 \times \mathcal{A}_2)$. Moreover, if μ_1 and μ_2 are σ-finite, we have $X_1 = \bigcup_{j=1}^{\infty} E_j$ where $\mu_1(E_j) < \infty$ for all j, and $X_2 = \bigcup_{k=1}^{\infty} F_k$, $\mu_2(F_k) < \infty$ for all k. Then,

$$X = X_1 \times X_2 = \bigcup_{j,k=1}^{\infty} E_j \times F_k, \quad \mu(E_j \times F_k) = \mu_1(E_j)\mu_2(F_k) < \infty$$

Thus, X is σ-finite for measure μ, and we can use **Carathéodory's extension theorem** to uniquely extend to $\mathcal{A}_1 \otimes \mathcal{A}_2$. We write this extension as $\mu_1 \otimes \mu_2$.

This can be then easily generalized to any finite product spaces. Let $(X_i, \mathcal{A}_i, \mu_i)$, where $i = 1, 2, \cdots, n$ be measure spaces. Then, we define the product σ-algebra on the space $\prod_{i=1}^{n} X_i$ by

$$\mathcal{A} = \bigotimes_{i=1}^{n} \mathcal{A}_i = \sigma\left(\prod_{i=1}^{n} \mathcal{A}_i\right)$$

Now if there is a measure μ defined on $\prod_{i=1}^{n} X_i$ such that

$$\mu\left(\prod_{i=1}^{n} E_i\right) = \prod_{i=1}^{n} \mu_i(E_i), \quad E_i \in \mathcal{A}_i$$

Then, if σ-finite, this can be uniquely extended to the product σ-algebra, and we denote this measure by $\bigotimes_{i=1}^{n} \mu_i$.

4.3 Countable Product Measures

Now we can extend this to infinite case. We first deal with the *countable* one. In this case, we need to restrict to a *probability measure*.

Definition 4.3.1: Probability Measure

Let $(X, \mathcal{A}, \mu)$ be a measure space. μ is called a **probability measure** if $\mu(X) = 1$.

Let $(X_j, \mathcal{A}_j, \mu_j)$ for $j = 1, 2, \cdots$ be measure spaces. Denote $X = X_1 \times X_2 \times \cdots = \prod_{j=1}^{\infty} X_j$. Now we want to, similarly, construct a semi-algebra in this set.

Definition 4.3.2: Cylinder Set

A **cylinder set** is a set $E \subseteq X$ such that it is the form of

$$E = E_1 \times E_2 \times \cdots \times E_n \times X_{n+1} \times X_{n+2} \times \cdots, \quad E_j \in \mathcal{A}_j$$

That is, cylinder set is a product of countably many sets, with only finite number of them different from the whole set. Later on, we will denote

$$X^{(n)} = X_n \times X_{n+1} \times \cdots = \prod_{j=n}^{\infty} X_j$$

for simplicity.

Proposition 4.3.3: Cylinder Sets Form a Semi-algebra

The collection of all cylinder sets

$$\mathcal{C} = \{E : E = E_1 \times E_2 \times \cdots \times E_n \times X_{n+1} \times X_{n+2} \times \cdots,\ n \geqslant 1,\ E_j \in \mathcal{A}_j,\ 1 \leqslant j \leqslant n\}$$

is a semi-algebra.

Proof. First, $X \in \mathcal{C}$ since $X = X_1 \times X_2 \times \cdots$ where $X_j \in \mathcal{A}_j$.
Second, we need to show that if $A, B \in \mathcal{C}$, then $A \cap B \in \mathcal{C}$. Suppose

$$A = A_1 \times \cdots \times A_n \times X^{(n+1)},\ A_j \in \mathcal{A}_j,\ 1 \leqslant j \leqslant n, \quad B = B_1 \times \cdots \times B_m \times X^{(m+1)},\ B_k \in \mathcal{A}_k,\ 1 \leqslant k \leqslant m$$

Without loss of generality, suppose $n \leqslant m$. Then, $A = A_1 \times \cdots \times A_n \times X_{n+1} \times \cdots \times X_m \times X^{(m+1)}$. Then, we can write the intersection as

$$\begin{aligned} A \cap B &= (A_1 \times \cdots \times A_n \times X_{n+1} \times \cdots \times X_m \times X^{(m+1)}) \cap (B_1 \times \cdots \times B_m \times X^{(m+1)}) \\ &= (A_1 \cap B_1) \times \cdots \times (A_n \cap B_n) \times B_{n+1} \times \cdots \times B_m \times X^{(m+1)} \end{aligned}$$

Since $A_i, B_i \in \mathcal{A}_i$ and $\mathcal{A}_i$ are algebra, we have $A_i \cap B_i \in \mathcal{A}_i$, which means that $A \cap B \in \mathcal{C}$.
Finally, we want to prove that for $A \in \mathcal{C}$, we have A^c is a finite disjoint union of cylinder sets. To prove this, suppose $A = A_1 \times \cdots \times A_n \times X^{(n+1)}$, and then

$$A^c = \bigsqcup_{\substack{\sigma_1 = +1, -1 \\ \vdots \\ \sigma_n = +1, -1 \\ \sum_{j=1}^n \mathbb{1}\{\sigma_j = -1\} \geqslant 1}} B^{1,\sigma_1} \times \cdots \times B^{n,\sigma_n} \times X^{(n+1)}, \quad B^{j,1} = A_j,\ B^{j,-1} = A_j^c$$

This is just a generalization from the last section. Since $A_j, A_j^c \in \mathcal{A}_j$, A is a disjoint union of cylinder sets. $\square$

Now, we want to define a set function $\mu : \mathcal{C} \to [0, 1]$ similarly as before. Let $A \in \mathcal{C}$ where $A = A_1 \times A_2 \times \cdots \times A_n \times X^{(n+1)}$, $A_j \in \mathcal{A}_j$. Then, we define

$$\mu(A) = \prod_{j=1}^n \mu_j(A_j) \prod_{j=n+1}^\infty \mu_j(X_j) = \prod_{j=1}^n \mu_j(A_j)$$

Proposition 4.3.4: Additivity on Semi-algebra

μ is finitely additive on $\mathcal{C}$.

Proof. Let $A \in \mathcal{C}$ and $A^{(1)}, \cdots A^{(k)} \in \mathcal{C}$ with $A = \bigsqcup_{j=1}^k A^{(j)}$, where $A^{(j)}$'s are disjoint. Suppose

$$A = A_1 \times \cdots \times A_n \times X^{(n+1)}, \quad A^{(j)} = A_1^{(j)} \times \cdots \times A_{n_j}^{(j)} \times X^{(n_j+1)},\ 1 \leqslant j \leqslant k$$

Let $m = \max\{n, n_j, 1 \leqslant j \leqslant k\}$. Then, we can write

$$A = A_1 \times \cdots \times A_m \times X^{(m+1)}, \quad A^{(j)} = A_1^{(j)} \times \cdots \times A_m^{(j)} \times X^{(m+1)},\ 1 \leqslant j \leqslant k$$

where $A_k = X_k$ for $k \geqslant n$ and $A_k^{(j)} = X_k$ for $k \geqslant n_j$. By definition, we have

$$\mu(A) = \prod_{\ell=1}^{m} \mu_\ell(A_\ell), \quad \mu(A^{(j)}) = \prod_{\ell=1}^{m} \mu_\ell(A_\ell^{(j)})$$

Notice that if we write $E = A_1 \times \cdots \times A_m$, and $E^{(j)} = A_1^{(j)} \times \cdots A_m^{(j)}$, then $E = \bigsqcup_{j=1}^{k} E^{(j)}$, and the product measure $\tilde{\mu}$ defined on $X_1 \times \cdots X_m$ is just

$$\tilde{\mu}(E) = \mu(A) = \prod_{\ell=1}^{m} \mu_\ell(A_\ell), \quad \tilde{\mu}(E^{(j)}) = \mu(A^{(j)}) = \prod_{\ell=1}^{m} \mu_\ell(A_\ell^{(j)})$$

Since we have shown that this product measure $\tilde{\mu}$ is finitely additive, we have

$$\mu(A) = \tilde{\mu}(E) = \sum_{j=1}^{k} \tilde{\mu}(E^{(j)}) = \sum_{j=1}^{k} \mu(A^{(j)})$$

which shows the additivity. $\square$

Note that on the proof above, we just use the finite product result. However, if we want to prove σ-additivity, there would be a problem: If $A = \bigsqcup_{j=1}^{\infty} A^{(j)}$, the subscript n_j in $A^{(j)}$ might become larger and larger when $j \to \infty$, resulting in a non-cylinder form. Therefore, we need alternative approach for proving its σ-additivity.

First, we extend this measure onto the algebra $\mathcal{R}(\mathcal{C})$ generated by $\mathcal{C}$. Let $\mu : \mathcal{R}(\mathcal{C}) \to [0, 1]$. If $A \in \mathcal{R}(\mathcal{C})$ and $A = \bigsqcup_{j=1}^{n} A^{(j)}$, where $A^{(j)} \in \mathcal{C}$ and they are disjoint, then we define

$$\mu(A) = \sum_{j=1}^{n} \mu(A^{(j)})$$

Then, the finite additivity just trivially follows from Proposition 4.3.4 and the definition. The only thing remains is the σ-additivity. Before proving that, we need some lemma.

Definition 4.3.5: Continuous at E

Let $\mathcal{C} \subseteq \mathcal{P}(X)$. Let $\mu : \mathcal{C} \to [0, \infty]$ be a set function. Let $E \in \mathcal{C}$. Then,

- μ is **continuous from below at** E if

$$\forall \{E_n\}_{n=1}^{\infty} \subseteq \mathcal{C},\ E_n \nearrow E,\ \text{we have } \lim_{n\to\infty} \mu(E_n) = \mu(E)$$

- μ is **continuous from above at** E if

$$\forall \{E_n\}_{n=1}^{\infty} \subseteq \mathcal{C},\ E_n \searrow E,\ \text{ and } \mu(E_{n_0}) < \infty \text{ for some } n_0,\ \text{we have } \lim_{n\to\infty} \mu(E_n) = \mu(E)$$

If μ is both continuous from below and continuous from above at E, we say that μ is **continuous** at E.

The lemma we need is for the criterion from finite additivity to σ-additivity. We will just use the condition (c), but we state other two for further use.

Lemma 4.3.6: Criterion of σ-additivity

Let $\mathcal{R} \subseteq \mathcal{P}(X)$ be an algebra. Let $\mu : \mathcal{R} \to [0, \infty]$ be a finitely additive function on $\mathbb{R}$. Then,

(a) If μ is σ-additive, then μ is continuous at E for all $E \in \mathcal{R}$.

(b) If μ is continuous from below at all $E \in \mathcal{R}$, then μ is σ-additive.

(c) If μ is continuous from above at $\emptyset$ and $\mu(X) < \infty$, then μ is σ-additive.

Proof. (a) Suppose μ is σ-additive. We first prove that it is continuous from below. Let $E \in \mathcal{R}$, $\{E_n\}_{n=1}^{\infty} \subseteq \mathcal{R}$, and $E_n \nearrow E$. Then, we define

$$F_1 = E_1,\ F_2 = E_2 \backslash E_1,\ \cdots,\ F_k = E_k \backslash E_{k-1}, \cdots$$

Then, F_k's are disjoint, and $E = \bigcup_{n=1}^{\infty} E_n = \bigsqcup_{k=1}^{\infty} F_k$. By σ-additivity, we have

$$\begin{aligned}
\mu(E) = \sum_{k=1}^{\infty} \mu(F_k) &= \lim_{n\to\infty} \sum_{k=1}^{n} \mu(F_k) \\
&= \lim_{n\to\infty} \mu\left(\bigsqcup_{k=1}^{n} F_k\right) && (\sigma\text{-additivity implies finite additivity}) \\
&= \lim_{n\to\infty} \mu(E_n) && \left(\textstyle\bigsqcup_{k=1}^{n} F_k = E_n \text{ by definition}\right)
\end{aligned}$$

Now we prove that it is continuous from above. Let $E \in \mathcal{R}$, $\{E_n\}_{n=1}^{\infty} \subseteq \mathcal{R}$, and $E_n \searrow E$ with $\mu(E_{n_0}) < \infty$ for some n_0. Without loss of generality, we assume that $n_0 = 1$. Then, we construct

$$G_1 = E_1 \backslash E_2,\ G_2 = E_1 \backslash E_3, \cdots,\ G_k = E_1 \backslash E_{k+1}, \cdots$$

Since $\{E_n\}$ is a decreasing sequence, $\{G_k\}$ is an increasing sequence, where $G_k \nearrow E_1 \backslash E$. Since $\mathcal{R}$ is an algebra, we have $G_k \in \mathcal{R}$ for all k. Therefore, by the continuity from below which we have already proved, we have

$$\lim_{k\to\infty} \mu(G_k) = \mu(E_1 \backslash E) \tag{4.1}$$

Since μ is finitely additive, we have $\mu(G_k) + \mu(E_{k+1}) = \mu(E_1)$. Since $\mu(E_{k+1}) < \infty$ by assumption and decreasing monotonicity of the sequence $\{E_n\}$, we can subtract $\mu(E_{k+1})$ on both sides, and get

$$\mu(G_k) = \mu(E_1) - \mu(E_{k+1}) \tag{4.2}$$

Similarly, we can have $\mu(E_1 \backslash E) = \mu(E_1) - \mu(E)$. Combining all these equations, we have

$$\lim_{k\to\infty} (\mu(E_1) - \mu(E_{k+1})) = \mu(E_1) - \lim_{k\to\infty} \mu(E_k) = \mu(E_1) - \mu(E) \quad \Longrightarrow \quad \lim_{k\to\infty} \mu(E_k) = \mu(E)$$

where the implication is true since $\mu(E_1) < \infty$.

(b) Suppose μ is continuous from below at all $E \in \mathcal{R}$. Let $E \in \mathcal{R}$, $\{E_k\}_{k=1}^{\infty} \subseteq \mathcal{R}$, and $E = \bigsqcup_{k=1}^{\infty} E_k$. Denote

$F_n = \bigsqcup_{k=1}^n E_k \in \mathcal{R}$. Then, $F_n \nearrow E$ by definition. Since μ is continuous from below, we have

$$\lim_{n\to\infty} \mu(F_n) = \lim_{n\to\infty} \mu\left(\bigsqcup_{k=1}^n E_k\right) = \lim_{n\to\infty} \sum_{k=1}^n \mu(E_k) = \sum_{k=1}^\infty \mu(E_k) = \mu(E)$$

where the second equality uses the finite additivity of μ.

(c) Suppose μ is continuous from above at $\emptyset$ and $\mu(X) < \infty$. Let $E \in \mathcal{R}$, $\{E_k\}_{k=1}^\infty \subseteq \mathcal{R}$, and $E = \bigsqcup_{k=1}^\infty E_k$. Let $F_n = \bigsqcup_{k=n}^\infty E_k$. Notice that $F_n \in \mathcal{R}$ because $F_n = E \setminus \bigsqcup_{j=1}^{n-1} E_j$. Then, $F_n \searrow \emptyset$, and $\mu(F_1) < \infty$. By continuity from above at $\emptyset$, we have $\lim_{n\to\infty} \mu(F_n) = 0$. We then have,

$$\mu(E) = \mu\left(\bigsqcup_{k=1}^n E_k \cup \bigsqcup_{k=n+1}^\infty E_k\right) = \mu\left(\bigsqcup_{k=1}^n E_k\right) + \mu\left(\bigsqcup_{k=n+1}^\infty E_k\right) = \sum_{k=1}^n \mu(E_k) + \mu(F_{n+1})$$

where all the equality comes from finite additivity of μ. This holds for all n, so we can let $n \to \infty$. Then, we have

$$\lim_{n\to\infty} \sum_{k=1}^n \mu(E_k) = \sum_{k=1}^\infty \mu(E_k), \quad \lim_{n\to\infty} \mu(F_{n+1}) = 0$$

which implies the σ-additivity. $\square$

We will use (c) to prove that μ is countably additive.

Lemma 4.3.7: Continuous from above at empty set

$\mu : \mathcal{R}(\mathcal{C}) \to [0,1]$ defined above is continuous from above at $\emptyset$.

Proof. **STEP I: Notation** We define μ on the space $X = X_1 \times X_2 \times \cdots$. Then, we can also have things defined on the space $X_2 \times X_3 \times \cdots = X^{(2)}$. Then, we can let $\mathcal{C}^{(2)}$ be the collection of cylinder sets on this space, and $\mathcal{R}(\mathcal{C}^{(2)})$ be the algebra generated by $\mathcal{C}^{(2)}$. Finally, we can define measure $\mu^{(2)}$ on this space. Let $E \in \mathcal{C}^{(2)}$, where $E = E_2 \times E_3 \times \cdots \times E_m \times X^{(m+1)}$. Then, we define

$$\mu^{(2)}(E) = \prod_{j=2}^m \mu_j(E_j)$$

Similarly, we can define $X^{(k)}$, $\mathcal{C}^{(k)}$, $\mathcal{R}(\mathcal{C}^{(k)})$ and $\mu^{(k)}$, for $k \geqslant 2$.
STEP II: Section Take $A \in \mathcal{C}$, where $A = E_1 \times E_2 \times \cdots \times E_n \times X^{(n+1)}$, $E_j \in \mathcal{A}_j$. Then, define $A(x_1) \subseteq X^{(2)}$ as the section of A at point $x_1 \in X_1$. That is,

$$A(x_1) = \begin{cases} E_2 \times E_3 \times \cdots \times E_n \times X^{(n+1)}, & x_1 \in E_1 \\ \emptyset, & x_1 \notin E_1 \end{cases}$$

Note that for all $x_1 \in X_1$, we have $A(x_1) \in \mathcal{C}^{(2)}$.
STEP III: Section Claim We have the following claim:

$$\text{If } A \in \mathcal{C}, \text{ then } x \mapsto \mu^{(2)}(A(x)) \text{ is } \mathcal{A}_1\text{-measurable, and } \mu(A) = \int \mu^{(2)}(A(x))\, \mathrm{d}\mu_1(x)$$

To prove this claim, note that

$$\mu^{(2)}(A(x)) = \begin{cases} \prod_{j=2}^{\infty} \mu_j(E_j), & x \in E_1 \\ 0, & x \notin E_1 \end{cases} = \mathbb{1}_{E_1}(x) \prod_{j=2}^{n} \mu_j(E_j)$$

is a simple function. So, it is measurable, and its integral is

$$\int \mu^{(2)}(A(x))\,\mathrm{d}\mu_1(x) = \mu_1(E_1) \prod_{j=2}^{n} \mu_j(E_j) = \prod_{j=1}^{n} \mu_j(E_j) = \mu(A)$$

STEP IV: Section Claim Generalization on Algebra Now, if $A \in \mathcal{R}(\mathcal{C})$, the claim still holds. Let $A = \bigsqcup_{j=1}^{n} A^{(j)}$, where $A^{(j)} \in \mathcal{C}$ are disjoint. Then,

$$A(x) = \left(\bigsqcup_{j=1}^{n} A^{(j)}\right)(x) = \bigsqcup_{j=1}^{n} (A^{(j)}(x)), \quad \text{where } A^{(j)}(x) \text{ disjoint}$$

which we have proved before. Then,

$$\begin{aligned} \int \mu^{(2)}(A(x))\,\mathrm{d}\mu_1(x) &= \int \mu^{(2)}\left(\left(\bigsqcup_{j=1}^{n} A^{(j)}\right)(x)\right)\mathrm{d}\mu_1(x) = \int \mu^{(2)}\left(\bigsqcup_{j=1}^{n} (A^{(j)}(x))\right)\mathrm{d}\mu_1(x) \\ &= \int \sum_{j=1}^{n} \mu^{(2)}(A^{(j)}(x))\,\mathrm{d}\mu_1(x) && \text{(additivity of } \mu^{(2)}\text{)} \\ &= \sum_{j=1}^{n} \int \mu^{(2)}(A^{(j)}(x))\,\mathrm{d}\mu_1(x) && \text{(linearity of integral)} \\ &= \sum_{j=1}^{n} \mu(A^{(j)}(x)) = \mu(A) && \text{(additivity of } \mu\text{)} \end{aligned}$$

STEP V: State the Goal Consider a sequence of sets $\{A^{(n)}\}_{n=1}^{\infty} \subseteq \mathcal{R}(\mathcal{C})$, and $A^{(n)} \searrow \emptyset$. Then, we need to prove that $\mu(A^{(n)}) \to 0$. However, we choose to prove its contrapositive:

$$\{A^{(n)}\}_{n=1}^{\infty} \subseteq \mathcal{R}(\mathcal{C}),\ A^{(n)} \text{ is decreasing. If there exists } \epsilon > 0 \text{ such that } \mu(A^{(n)}) \geqslant \epsilon, \text{ then } \bigcap_{n=1}^{\infty} A^{(n)} \neq \emptyset$$

STEP VI: Prove the Goal Note that for all $x \in X_1$, we have $A^{(n)}(x) \in \mathcal{R}(\mathcal{C}^{(2)})$. We define

$$B^{(n)} = \{x \in X_1 : \mu^{(2)}(A^{(n)}(x)) \geqslant \epsilon/2\}$$

Then, $B^{(n)} \in \mathcal{A}_1$ since we have proved that $\mu^{(2)}(A^{(n)}(x))$ is measurable. Moreover, $B^{(n)} \supseteq B^{(n+1)}$ since $A^{(n)}$ is decreasing. By assumption, $\mu(A^{(n)}) \geqslant \epsilon$, so we have

$$\epsilon \leqslant \mu(A^{(n)}) = \int \mu^{(2)}(A^{(n)}(x))\,\mathrm{d}\mu_1(x) \qquad \text{(STEP IV)}$$

$$\begin{aligned}
&= \int_{B^{(n)}} \mu^{(2)}(A^{(n)}(x))\,\mathrm{d}\mu_1(x) + \int_{(B^{(n)})^c} \mu^{(2)}(A^{(n)}(x))\,\mathrm{d}\mu_1(x) \\
&\leqslant \int_{B^{(n)}} 1\,\mathrm{d}\mu_1(x) + \int_{(B^{(n)})^c} \frac{\epsilon}{2}\,\mathrm{d}\mu_1(x) && (\mu^{(2)}(A^{(n)}(x)) \leqslant 1 \text{ and definition of } B^{(n)}) \\
&= \mu_1(B^{(n)}) + \frac{\epsilon}{2}\mu_1((B^{(n)})^c) \\
&= \mu_1(B^{(n)}) + \frac{\epsilon}{2}(1 - \mu_1(B^{(n)}))
\end{aligned}$$

Rearranging, we have

$$\frac{\epsilon}{2} \leqslant \left(1 - \frac{\epsilon}{2}\right)\mu_1(B^{(n)}) \leqslant \mu_1(B^{(n)}) \quad \Longrightarrow \quad \mu_1(B^{(n)}) \geqslant \frac{\epsilon}{2}$$

Since $B^{(n)} \in \mathcal{A}_1$, and it is decreasing, and its measure is bounded away by some positive factor, we have

$$\bigcap_{n=1}^{\infty} B^{(n)} \neq \emptyset$$

- ***ITERATION I:*** There exists $x_1 \in \bigcap_{n=1}^{\infty} B^{(n)}$. By the definition of $B^{(n)}$, we have

  $$\mu^{(2)}\left(A^{(n)}(x_1)\right) \geqslant \frac{\epsilon}{2} \text{ for all } n \in \mathbb{N}$$

  Therefore, from sequence $A^{(n)} \in \mathcal{R}(\mathcal{C})$, we obtained the sequence $A^{(n)}(x_1) \in \mathcal{R}(\mathcal{C}^{(2)})$, which is also decreasing, and bounded away by some positive factor.

- ***ITERATION k:*** We can do the steps above iteratively, and we will have: there exists $(x_1, x_2, \cdots, x_k) \in X_1 \times \cdots X_k$ such that the sequence $A^{(n)}(x_1, \cdots, x_k) \in \mathcal{R}(\mathcal{C}^{(k+1)})$ is decreasing, and $\mu^{(k+1)}(A^{(n)}(x_1, \cdots, x_k)) \geqslant \epsilon/2^k$ for all n.

STEP VII: The final claim We finally claim that:

$$(x_1, x_2, \cdots) \in \bigcap_{n=1}^{\infty} A^{(n)}$$

To prove this claim, we fix n. Since $A^{(n)} \in \mathcal{R}(\mathcal{C})$, it is a union of cylinders. Thus, there must be some point m that $A^{(n)} = A_1^{(n)} \times A_2^{(n)} \times \cdots \times A_m^{(n)} \times X^{(m+1)}$. Therefore, we have the following observation:

> there exists $m \geqslant 1$, such that if $(y_1, y_2, \cdots) \in A^{(n)}$, and another point $(z_1, z_2, \cdots)$ with $z_j = y_j$ for all $1 \leqslant j \leqslant m$, then $(z_1, z_2, \cdots) \in A^{(n)}$

We can apply this observation on $A^{(n)}$. Since we have shown that $\mu^{(k+1)}(A^{(n)}(x_1, \cdots, x_k)) \geqslant \epsilon/2^k$ for all n, we have $A^{(n)}(x_1, \cdots, x_k) \neq \emptyset$ for all n. Then, by the observation, we can choose iteration m so that $A^{(n)}(x_1, \cdots, x_m) \neq \emptyset$, which means that there exists $(y_{m+1}, y_{m+2}, \cdots) \in A^{(n)}(x_1, \cdots, x_m)$. By defintion, this indicates that $(x_1, \cdots, x_m, y_{m+1}, y_{m+2}, \cdots) \in A^{(n)}$. By the observation, we can replace all the y's by x's, and finally have $(x_1, x_2, \cdots) \in A^{(n)}$. Since this n is arbitrary, we have proved our final claim. $\square$

Theorem 4.3.8: σ-additivity of the measure

μ is σ-additive on $\mathcal{R}(\mathcal{C})$.

Proof. By Lemma 4.3.6 (c) and 4.3.7. □

Then, if μ is σ-additive, we can apply Carathéodory extension theorem to uniquely extend it to $\sigma(\mathcal{C})$. We usually denote this measure space by

$$\left(\prod_{i=1}^{\infty} X_i, \bigotimes_{i=1}^{\infty} \mathcal{A}_i, \bigotimes_{i=1}^{\infty} \mu_i\right)$$

4.4 Monotone Class Theorem

Before going into the next section, we introduce a tool that is very useful for the future proof works.

Definition 4.4.1: Monotone Class

Let $\mathcal{M} \subseteq \mathcal{P}(X)$. Then, $\mathcal{M}$ is a **monotone class** if it satisfies the following properties:

- If $\{A_j\}_{j=1}^{\infty} \subseteq \mathcal{M}$ and $A_j \nearrow A$, then $A \in \mathcal{M}$.
- If $\{B_j\}_{j=1}^{\infty} \subseteq \mathcal{M}$ and $B_j \searrow B$, then $B \in \mathcal{M}$.

Remark: It is easy to see that, if we have a collection of monotone classes $\{\mathcal{M}_\alpha\}_{\alpha\in I}$, then $\bigcap_{\alpha\in I} \mathcal{M}_\alpha$ is also a monotone class. Therefore, as algebra and σ-algebra, we can have a similar terminology: for a set $\mathcal{C} \subseteq \mathcal{P}(X)$, we have the monotone class generated by $\mathcal{C}$, which is denoted by $\mathcal{M}(\mathcal{C})$.

Theorem 4.4.2: Monotone Class Theorem

Let $\mathcal{R} \subseteq \mathcal{P}(X)$ be an algebra. Let $\mathcal{M}(\mathcal{R})$ be the monotone class generated by $\mathcal{R}$. Let $\sigma(\mathcal{R})$ be the σ-algebra generated by $\mathcal{R}$. Then,

$$\mathcal{M}(\mathcal{R}) = \sigma(\mathcal{R})$$

Proof. ($\supseteq$): This direction is trivial. Since $\sigma(\mathcal{R})$ is a monotone class, and $\sigma(\mathcal{R}) \supseteq \mathcal{R}$, by definition we have $\sigma(\mathcal{R}) \supseteq \mathcal{M}(\mathcal{R})$.
($\subseteq$): To show this direction, we use the *generating class argument*. We have already known that $\mathcal{M}(\mathcal{R}) \supseteq \mathcal{R}$. If we can show that $\mathcal{M}(\mathcal{R})$ is a σ-algebra, then this can lead to $\sigma(\mathcal{R}) \subseteq \mathcal{M}(\mathcal{R})$ since $\sigma(\mathcal{R})$ is the smallest σ-algebra.
STEP I: Set Definition Take $E \in \mathcal{M}(\mathcal{R})$. We define

$$\mathcal{G}(E) = \{F \in \mathcal{M}(\mathcal{R}) : E\backslash F, E\cap F, F\backslash E \in \mathcal{M}(\mathcal{R})\}$$

STEP II: Claim I Our first claim is

$$\text{If } E \in \mathcal{R}, \text{ then } \mathcal{G}(E) \supseteq \mathcal{M}(\mathcal{R})$$

To prove this, we use the *generating class argument*. Thus, we need to show that

- $\mathcal{G}(E) \supseteq \mathcal{R}$.
- $\mathcal{G}(E)$ is a monotone class.

To prove this first one, take $H \in \mathcal{R}$. Then, both $H, E \in \mathcal{R}$. Since $\mathcal{R}$ is an algebra, we have $E\backslash H$, $E \cap H$, $H\backslash E \in \mathcal{M}(\mathcal{R})$. To prove the second one. Take $H_k \nearrow H$, where $H_k \in \mathcal{G}(E)$. Then,

$$E\backslash H_k \in \mathcal{M}(\mathcal{R}), \text{ and } E\backslash H_k \searrow E\backslash H \in \mathcal{M}(\mathcal{R}) \text{ since } \mathcal{M}(\mathcal{R}) \text{ is a monotone class}$$

$$E \cap H_k \in \mathcal{M}(\mathcal{R}), \text{ and } E \cap H_k \nearrow E \cap H \in \mathcal{M}(\mathcal{R}) \text{ since } \mathcal{M}(\mathcal{R}) \text{ is a monotone class}$$

$$H_k\backslash E \in \mathcal{M}(\mathcal{R}), \text{ and } H_k\backslash E \nearrow H\backslash E \in \mathcal{M}(\mathcal{R}) \text{ since } \mathcal{M}(\mathcal{R}) \text{ is a monotone class}$$

These three shows that H satisfies the property of $\mathcal{G}(E)$, which means that $H \in \mathcal{G}(E)$. Similarly, we can take $G_k \searrow G$, $G_k \in \mathcal{G}(E)$ and show that $G \in \mathcal{G}(E)$. These show that $\mathcal{G}(E)$ is a monotone class.
STEP III: Claim II The second claim is

$$\text{If } E \in \mathcal{M}(\mathcal{R}), \text{ then } \mathcal{G}(E) \supseteq \mathcal{M}(\mathcal{R})$$

Again, we use the *generating class argument*. To prove the first one, take $H \in \mathcal{R}$. We need to show that $H \in \mathcal{G}(E)$. That is, we need to show that $E\backslash H, E \cap H, H\backslash E \in \mathcal{M}(\mathcal{R})$. However, we have proved in STEP II that $\mathcal{G}(H) \supseteq \mathcal{M}(\mathcal{R})$ since $H \in \mathcal{R}$, we have that $E\backslash H, E \cap H, H\backslash E \in \mathcal{M}(\mathcal{R})$ are all true for $E \in \mathcal{M}(\mathcal{R})$. To prove the second one, it is exactly the same as in STEP II.
STEP IV: Show that $\mathcal{M}(\mathcal{R})$ is an algebra To show that $\mathcal{M}(\mathcal{R})$ is an algebra, we need to show:

- $X \in \mathcal{M}(\mathcal{R})$: This one is trivial since $X \in \mathcal{R}$, and $\mathcal{M}(\mathcal{R})$ is generated by $\mathcal{R}$.

- Take $E \in \mathcal{M}(\mathcal{R})$. We need to show that $E^c \in \mathcal{M}(\mathcal{R})$. Because we have proved that $\mathcal{G}(E) \supseteq \mathcal{M}(\mathcal{R})$ for all $E \in \mathcal{M}(\mathcal{R})$ in Claim II, we know that $E \in \mathcal{G}(X)$ since $X \in \mathcal{M}(\mathcal{R})$. However,

$$\mathcal{G}(X) = \{F \in \mathcal{M}(\mathcal{R}) : F^c, F, \emptyset \in \mathcal{M}(\mathcal{R})\}$$

  This indicates that $E^c \in \mathcal{M}(\mathcal{R})$.

- Finally, If $E, F \in \mathcal{M}(\mathcal{R})$, we need to show that $E \cap F \in \mathcal{M}(\mathcal{R})$. To prove this, observe that by Claim II, we have $\mathcal{G}(E) \supseteq \mathcal{M}(\mathcal{R})$ for all $E \in \mathcal{M}(\mathcal{R})$. Moreover, since $F \in \mathcal{M}(\mathcal{R})$, we have $\mathcal{G}(E) \supseteq \mathcal{M}(\mathcal{R})$, thus $F \in \mathcal{G}(E)$. By definition of $\mathcal{G}(E)$, $E \cap F \in \mathcal{M}(\mathcal{R})$.

STEP V: Show that $\mathcal{M}(\mathcal{R})$ is a σ-algebra Finally, we need to show its countable additivity: If $\{A_j\}_{j=1}^{\infty} \subseteq \mathcal{M}(\mathcal{R})$, then we have $\bigcup_{j=1}^{\infty} A_j \in \mathcal{M}(\mathcal{R})$. To prove this, we have known that $\bigcup_{j=1}^{n} A_j \in \mathcal{M}(\mathcal{R})$ since $\mathcal{R}$ is an algebra. Since $\mathcal{M}(\mathcal{R})$ is a monotone class, we have $\bigcup_{j=1}^{n} A_j \nearrow \bigcup_{j=1}^{\infty} A_j \in \mathcal{M}(\mathcal{R})$. □

4.5 Tonelli's and Fubini's Theorem

In this section, we will talk about interchanging the order of integration. Let $(X_1, \mathcal{A}_1, \mu_1), (X_2, \mathcal{A}_2, \mu_2)$ be two measure spaces with μ_1, μ_2 σ-finite. Then, we can construct the product space where $X = X_1 \times X_2$, $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ and $\mu = \mu_1 \otimes \mu_2$. Let $f : X \to \bar{\mathbb{R}}$.

Lemma 4.5.1: Section of Measurable Function is Measurable

Suppose f is $\mathcal{A}$-measurable. Then, for all $x \in X_1$, the function $f_x : X_2 \to \bar{\mathbb{R}}$ defined by $y \mapsto f(x,y)$ is $\mathcal{A}_2$-measurable.

Proof. Suppose $B \in \mathcal{B}(\bar{\mathbb{R}})$. We need to prove that $f_x^{-1}(B) \in \mathcal{A}_2$ for all such B. To do this, we note that

$$f_x^{-1}(B) = \{y \in X_2 : f_x(y) \in B\} = \{y \in X_2 : f(x,y) \in B\} = \{y \in X_2 : (x,y) \in f^{-1}(B)\}$$

Now if we denote $E = f^{-1}(B)$. Then, $E \in \mathcal{A}$ since f is measurable. The last set in the equality becomes $\{y \in X_2 : (x,y) \in E\} = E_x$, the x-section of E. By Lemma 4.2.4, this set is measurable. $\square$

Remark: The above lemma is symmetric. We can also conclude on $f^y : X_1 \to \bar{\mathbb{R}}$.

Lemma 4.5.2: Measurability of Section Measures

Let $(X_1, \mathcal{A}_1, \mu_1), (X_2, \mathcal{A}_2, \mu_2)$ be two measure spaces with μ_1, μ_2 σ-finite. Let $E \in \mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$. Then,

1. $x \to \mu_2(E_x)$ is $\mathcal{A}_1$-measurable, and $y \to \mu_1(E^y)$ is $\mathcal{A}_2$-measurable.
2. $\int \mu_2(E_x)\,\mathrm{d}\mu_1 = \mu(E) = \int \mu_1(E^y)\,\mathrm{d}\mu_2$.

Proof. 1. **STEP I:** Let's first prove with assumption that $\mu_1(X_1), \mu_2(X_1) < \infty$.
(i) We first prove for rectangles. Suppose $E = A \times B$ where $A \in \mathcal{A}_1$ and $B \in \mathcal{A}_2$. Then,

$$E_x = \begin{cases} B, & x \in A, \\ \emptyset, & x \notin A \end{cases}$$

Then,

$$\mu_2(E_x) = \begin{cases} \mu_2(B), & x \in A \\ 0, & x \notin A \end{cases} = \mathbb{1}_A(x)\mu_2(B)$$

which is a simple function, and is measurable.
(ii) Now we prove on the algebra generated by rectangles. Let $\mathcal{C}$ denote the collection of all rectangles. Suppose $E \in \mathcal{R}(\mathcal{C})$. Then, $E = \bigsqcup_{j=1}^n E_j$, where $E_j \in \mathcal{C}$. Thus, $E_j = A_j \times B_j$, where $A_j \in \mathcal{A}_1$ and $B_j \in \mathcal{A}_2$. Then, by the previous observation,

$$E_x = \left(\bigsqcup_{j=1}^n E_j\right)_x = \bigsqcup_{j=1}^n (E_j)_x$$

Therefore,

$$\begin{aligned} \mu_2(E_x) &= \sum_{j=1}^n \mu_2((E_j)_x) && \text{(Finite additivity)} \\ &= \sum_{j=1}^n \mathbb{1}_{A_j}(x)\mu_2(B_j) && \text{(STEP I Part (i))} \end{aligned}$$

which is still a simple function and thus measurable.
(iii) Now we want to prove on the σ-algebra. But here is the difficulty: If we want to use our previous generating class argument, let

$$\mathcal{G} = \{E \in \mathcal{A} : \mu_2(E_x) \text{ is } \mathcal{A}_1\text{-measurable}\}$$

Then, we want to prove that $\mathcal{G}$ is a σ-algebra. However, it is hard to show that $\mathcal{M}$ is closed by finite union. Suppose $E, F \in \mathcal{M}$. Then, $\mu_2(E_x), \mu_2(F_x)$ are measurable. Then, we want to prove $\mu_2((E \cup F)_x) = \mu_2(E_x \cup F_x)$ is measurable. But this is difficult since E_x and F_x might have some intersection. Therefore, we turn to **monotone class theorem**. The claim is:

$$\mathcal{G} = \{E \in \mathcal{A} : \mu_2(E_x) \text{ is } \mathcal{A}_1\text{-measurable}\} \text{ is a monotone class}$$

To prove this, take $\{E^{(n)}\}_{n=1}^{\infty} \subseteq \mathcal{G}$, and $E^{(n)} \nearrow E$. Our goal is to prove that $E \in \mathcal{G}$. First of all, $\mu_2(E_x^{(n)})$ is $\mathcal{A}_1$-measurable for all n since $E^{(n)} \in \mathcal{G}$. Since $E^{(n)} \nearrow E$, we have $E_x^{(n)} \nearrow E_x$ for all $x \in X_1$. Since μ_2 is a measure, it is continuous from below. Therefore, $\mu_2(E_x) = \lim_{n\to\infty} \mu_2(E_x^{(n)})$. Since $\mu_2(E_x^{(n)})$ are $\mathcal{A}_1$-measurable, the limit $\lim_{n\to\infty} \mu_2(E_x^{(n)})$ is also $\mathcal{A}_1$-measurable, which means that $\mu_2(E_x)$ is measurable, and $E_x \in \mathcal{G}$.
Similarly, take $\{E^{(n)}\}_{n=1}^{\infty} \subseteq \mathcal{G}$ and $E^{(n)} \searrow E$. Using the argument above, we have $E_x^{(n)} \searrow E_x$. Since we assume that $\mu_2(X) < \infty$, and μ_2 is a measure, we can use continuous from above and get $\mu_2(E_x) = \lim_{n\to\infty} \mu_2(E_x^{(n)})$, which also shows that $E_x \in \mathcal{G}$ by the same reason.

Since $\mathcal{G}$ is a monotone class, and by previous argument (i) and (ii), we know that $\mathcal{G} \supseteq \mathcal{R}(\mathcal{C})$. By Monotone class theorem, $\mathcal{G} \supseteq \sigma(\mathcal{C}) = \mathcal{A}$. Therefore, $\mu_2(E_x)$ is measurable for all $E \in \mathcal{A}$. The argument for $\mu_1(E^y)$ is completely symmetric.

STEP II: Now we want to get rid of the assumptions that $\mu_1(X), \mu_2(X) < \infty$ by using the fact that μ_1, μ_2 are σ-finite. By σ-finiteness, we have:

$$\exists\, A^{(n)} \in \mathcal{A}_1,\ X_1 = \bigcup_{n=1}^{\infty} A^{(n)},\ A^{(n)} \text{ increasing},\ \mu_1(A^{(n)}) < \infty$$

$$\exists\, B^{(n)} \in \mathcal{A}_2,\ X_2 = \bigcup_{n=1}^{\infty} B^{(n)},\ B^{(n)} \text{ increasing},\ \mu_2(B^{(n)}) < \infty$$

Then, let $F^{(n)} = A^{(n)} \times B^{(n)} \in \mathcal{C}$. It has measure $\mu(F^{(n)}) = \mu_1(A^{(n)})\mu_2(B^{(n)}) < \infty$, and $\bigcup_{n=1}^{\infty} F^{(n)} = X_1 \times X_2$. Therefore, μ is also σ-finite.

Now, for $E \in \mathcal{A}$, take $E \cap F^{(n)}$. Then, $E \cap F^{(n)}$ has finite measure since $F^{(n)}$ has finite measure. By STEP I, we have that $\mu_2((E \cap F^{(n)})_x)$ is $\mathcal{A}_1$-measurable for any n. Since $E \cap F^{(n)} \nearrow E$, we have $(E \cap F^{(n)})_x \nearrow E_x$. By continuity from below, we know that $\mu_2(E_x)$ is $\mathcal{A}_1$-measurable. The argument for $\mu_1(E^y)$ is completely symmetric.

2. Similarly, we split it into three steps.
(i) For $E \in \mathcal{C}$, i.e., E is a rectangle, we have $E = A \times B$, where $A \in \mathcal{A}_1$, $B \in \mathcal{A}_2$. We have shown in the proof of 1 that

$$\mu_2(E_x) = \mathbb{1}_A(x)\mu_2(B)$$

Therefore,

$$\int \mu_2(E_x)\,\mathrm{d}\mu_1(x) = \int \mu_2(B)\mathbb{1}_A(x)\,\mathrm{d}\mu_1(x) = \mu_1(A)\mu_2(B) = \mu(E)$$

(ii) For $E \in \mathcal{R}(\mathcal{C})$. We have shown in the proof of 1 that

$$\mu_2(E_x) = \sum_{j=1}^{n} \mathbb{1}_{A_j}(x)\mu_2(B_j)$$

Therefore,

$$\begin{aligned}\int \mu_2(E_x)\,\mathrm{d}\mu_1(x) &= \int \sum_{j=1}^{n} \mathbb{1}_{A_j}(x)\mu_2(B_j)\,\mathrm{d}\mu_1(x) = \sum_{j=1}^{n} \int \mathbb{1}_{A_j}(x)\mu_2(B_j)\,\mathrm{d}\mu_1(x)\\ &= \sum_{j=1}^{n} \mu_1(A_j)\mu_2(B_j) = \sum_{j=1}^{n} \mu(E_j) = \mu(E)\end{aligned}$$

where $E = \bigsqcup_{j=1}^{n} E_j$, E_j are disjoint.

(iii) For $E \in \mathcal{A}$, we need to use monotone class theorem. Assume that $\mu_1(X_1), \mu_2(X_2) < \infty$. Let

$$\mathcal{G} = \left\{ E \in \mathcal{A} : \int \mu_2(E_x)\,\mathrm{d}\mu_1 = \mu(E) \right\}$$

We have shown that $\mathcal{G} \supseteq \mathcal{R}(\mathcal{C})$. The remaining part is to prove that $\mathcal{G}$ is a monotone class. Assume $E^{(n)} \in \mathcal{G}$ and $E^{(n)} \nearrow E$. We want to show that $E \in \mathcal{G}$. Since $E^{(n)} \in \mathcal{G}$, we have

$$\int \mu_2(E_x^{(n)})\,\mathrm{d}\mu_1 = \mu(E^{(n)}),\ E_x^{(n)} \nearrow E_x$$

By continuity from below, we know that $\mu(E^{(n)}) \nearrow \mu(E)$ and $\mu_2(E_x^{(n)}) \nearrow \mu_2(E_x)$. Then, by monotone convergence theorem, the LHS of the above equation has limit

$$\lim_{n\to\infty} \int \mu_2(E_x^{(n)})\,\mathrm{d}\mu_1 = \int \mu_2(E_x)\,\mathrm{d}\mu_1 = \lim_{n\to\infty} \mu(E^{(n)}) = \mu(E)$$

which shows that $E \in \mathcal{G}$. Similarly for the decreasing sequence, assume $E^{(n)} \in \mathcal{G}$ and $E^{(n)} \searrow E$. Using the same argument above, we can have

$$\lim_{n\to\infty} \int \mu_2(E_x^{(n)})\,\mathrm{d}\mu_1 = \lim_{n\to\infty} \mu(E^{(n)}) = \mu(E)$$

where the last equality comes from continuity from above (which can be used since we have finiteness assumption). Now, instead of using monotone convergence theorem, we use dominated convergence theorem to pass the limit into the integral. Since $\mu_2(E_x^{(n)}) \leqslant \mu_2(X_2)$, where $\mu_2(X_2)$ is an integrable function. Then, we can have

$$\lim_{n\to\infty} \int \mu_2(E_x^{(n)})\,\mathrm{d}\mu_1 = \int \mu_2(E_x)\,\mathrm{d}\mu_1 = \lim_{n\to\infty} \mu(E^{(n)}) = \mu(E)$$

Finally, we get rid of the finiteness assumption using the σ-finiteness. By σ-finiteness, we have:

$$\exists\, A^{(n)} \in \mathcal{A}_1,\, X_1 = \bigcup_{n=1}^{\infty} A^{(n)},\, A^{(n)} \text{ increasing},\, \mu_1(A^{(n)}) < \infty$$

$$\exists\, B^{(n)} \in \mathcal{A}_2,\, X_2 = \bigcup_{n=1}^{\infty} B^{(n)},\, B^{(n)} \text{ increasing},\, \mu_2(B^{(n)}) < \infty$$

Then, let $F^{(n)} = A^{(n)} \times B^{(n)} \in \mathcal{C}$. It has measure $\mu(F^{(n)}) = \mu_1(A^{(n)})\mu_2(B^{(n)}) < \infty$, and $\bigcup_{n=1}^{\infty} F^{(n)} = X_1 \times X_2$. Therefore, μ is also σ-finite.

Now, for $E \in \mathcal{A}$, take $E \cap F^{(n)}$. Then, $E \cap F^{(n)}$ has finite measure since $F^{(n)}$ has finite measure. By previous argument with finiteness assumption, we have

$$\int \mu_2([E \cap F^{(n)}]_x)\,\mathrm{d}\mu_1 = \mu(E \cap F^{(n)})$$

Since $E \cap F^{(n)} \nearrow E$, we have $(E \cap F^{(n)})_x \nearrow E_x$. By continuity from below, $\mu(E \cap F^{(n)}) \nearrow \mu(E)$, and by monotone convergence theorem, we have $\int \mu_2([E \cap F^{(n)}]_x)\,\mathrm{d}\mu_1 \nearrow \int \mu_2(E_x)\,\mathrm{d}\mu_1$. This shows that

$$\int \mu_2(E_x)\,\mathrm{d}\mu_1 = \mu(E)$$

The arugment for $\mu_1(E^y)$ is completely symmetric. $\square$

4.5.1 Tonelli's Theorem

Tonelli's theorem is about the criterion of interchanging integrals for non-negative functions.

Theorem 4.5.3: Tonelli's Theorem

Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be measure spaces. Let μ_1 and μ_2 be σ-finite. Let $(X, \mathcal{A}, \mu)$ be their product space. Let $f : X_1 \times X_2 \to \bar{\mathbb{R}}$ be a non-negative measurable function. Then,

$$\int \left[\int f_x(y)\,\mathrm{d}\mu_2\right]\mathrm{d}\mu_1 = \int f\,\mathrm{d}\mu = \int \left[\int f^y(x)\,\mathrm{d}\mu_1\right]\mathrm{d}\mu_2$$

Proof. (i) First assume that $f = c\mathbb{1}_E$, where $E \in \mathcal{A}$ and $c \geqslant 0$. Then,

$$f_x = (c\mathbb{1}_E)_x = c(\mathbb{1}_E)_x = c\mathbb{1}_{E_x}$$

We have proved that f_x is measurable in Lemma 4.5.1, so the inner integration $\int f_x(y)\,\mathrm{d}\mu_2$ makes sense. Therefore,

$$\int f_x(y)\,\mathrm{d}\mu_2(y) = \int c\mathbb{1}_{E_x}(y)\,\mathrm{d}\mu_2(y) = c\mu_2(E_x)$$

We have proved in Lemma 4.5.2 that $\mu_2(E_x)$ is measurable, so now the outer integration also makes sense. Therefore,

$$\int \left[\int f_x(y)\,\mathrm{d}\mu_2\right]\mathrm{d}\mu_1 = \int c\mu_2(E_x)\,\mathrm{d}\mu_1(x) = c\mu(E)$$

where the last equality comes from the 2nd property in Lemma 4.5.2. But note that $\int f \,\mathrm{d}\mu = \int c\mathbb{1}_E \,\mathrm{d}\mu = c\mu(E)$, so we have $\int \left[\int f_x(y)\,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1 = \int f \,\mathrm{d}\mu$. The argument for f^y is completely symmetric.

(ii) Assume $f = \sum_{j=1}^n c_j \mathbb{1}_{E_j}$ is a simple function, where $c_j \geqslant 0$ and $E_j \in \mathcal{A}$. Then,

$$f_x = \left(\sum_{j=1}^n c_j \mathbb{1}_{E_j}\right)_x = \sum_{j=1}^n (c_j \mathbb{1}_{E_j})_x$$

Then,

$$\int f_x(y)\,\mathrm{d}\mu_2(y) = \int \sum_{j=1}^n (c_j \mathbb{1}_{E_j})_x(y)\,\mathrm{d}\mu_2(y) = \sum_{j=1}^n \int (c_j \mathbb{1}_{E_j})_x(y)\,\mathrm{d}\mu_2(y) = \sum_{j=1}^n c_j \mu_2((E_j)_x)$$

and it is measurable w.r.t. μ_1, so we can further integrate

$$\int \left[\int f_x(y)\,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1 = \int \sum_{j=1}^n c_j \mu_2((E_j)_x)\,\mathrm{d}\mu_1 = \sum_{j=1}^n \int c_j \mu_2((E_j)_x)\,\mathrm{d}\mu_1 = \sum_{j=1}^n c_j \mu(E_j) = \int f \,\mathrm{d}\mu$$

(iii) Finally, assume that f is an arbitrary nonnegative function. Suppose f_j are simple functions and $f_j \nearrow f$. Then, it is clear that $(f_j)_x \nearrow f_x$. By monotone convergence theorem,

$$\lim_{n\to\infty} \int (f_j)_x \,\mathrm{d}\mu_2 = \int f_x \,\mathrm{d}\mu_2$$

Since we have proved that the integral on the LHS are measurable w.r.t. μ_1, we know that the limit is also measurable. Therefore, we can further integrate, and using monotone convergence theorem again to get

$$\lim_{n\to\infty} \int \left[\int (f_j)_x \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1 = \int \left[\int f_x \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1$$

But, since f_j's are simple, we can combine the properties we have proved before so that we get

$$\int f \,\mathrm{d}\mu = \lim_{n\to\infty} \int f_j \,\mathrm{d}\mu = \lim_{n\to\infty} \int \left[\int (f_j)_x \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1 = \int \left[\int f_x \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1$$

where the first equality is again by monotone convergence theorem. □

4.5.2 Fubini's Theorem

Fubini's theorem is about the criterion of interchanging integrals for integrable functions.

Theorem 4.5.4: Fubini's Theorem

Let $(X_1, \mathcal{A}_1, \mu_1)$ and $(X_2, \mathcal{A}_2, \mu_2)$ be measure spaces. Let μ_1 and μ_2 be σ-finite. Let $(X, \mathcal{A}, \mu)$ be their product space. Let $f : X_1 \times X_2 \to \bar{\mathbb{R}}$ be a measurable and integrable function. Then,

$$\int \left[\int f_x(y)\,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1 = \int f \,\mathrm{d}\mu = \int \left[\int f^y(x)\,\mathrm{d}\mu_1\right] \mathrm{d}\mu_2$$

Remark: Before proving it, there is some rigor problem of this theorem that needs to be claimed. For an integrable function, we can write $f = f^+ - f^-$, and $\int f^+ \,\mathrm{d}\mu < \infty$. Then, we have

$$\infty > \int f^+ \,\mathrm{d}\mu = \int \left[\int f_x^+ \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1$$

by Tonelli's theorem. This indicates that the function $\int f_x^+ \,\mathrm{d}\mu_2$ is finite a.e. (otherwise the integral of it would be infinite). Similarly for $\int f_x^- \,\mathrm{d}\mu_2$. We also have

$$\int f \,\mathrm{d}\mu = \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu = \int \left[\int f_x^+ \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1 - \int \left[\int f_x^- \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1 = \int \left[\int f_x^+ \,\mathrm{d}\mu_2 - \int f_x^- \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1$$

where the middle equality holds by Tonelli's theorem. However, there might be the case that at some point y, $\int f_x^+ \,\mathrm{d}\mu_2 = \int f_x^- \,\mathrm{d}\mu_2 = \infty$, in which case $\int f_x^+(y) \,\mathrm{d}\mu_2 - \int f_x^-(y) \,\mathrm{d}\mu_2$ is not defined. To solve this problem, we would replace these functions by functions that equal to them a.e.. Define

$$g^+(x) = \begin{cases} \int f_x^+(y) \,\mathrm{d}\mu_2, & \text{if } \int f_x^+(y) \,\mathrm{d}\mu_2 < \infty \\ 0, & \text{otherwise} \end{cases}, \quad g^-(x) = \begin{cases} \int f_x^-(y) \,\mathrm{d}\mu_2, & \text{if } \int f_x^-(y) \,\mathrm{d}\mu_2 < \infty \\ 0, & \text{otherwise} \end{cases}$$

It is easy to see that $g^+(x), g^-(x)$ are measurable functions, and $g(x) = g^+(x) - g^-(x)$ is always defined. Then, we are actually going to prove that:

$$\int g(x) \,\mathrm{d}\mu_1 = \int f \,\mathrm{d}\mu$$

Proof. Since f is integrable, we have f^+ and f^- are integrable. By Tonelli's theorem, we have

$$\iint f_x^+ \,\mathrm{d}\mu_2 \,\mathrm{d}\mu_1 = \int f^+ \,\mathrm{d}\mu < \infty$$

Thus, $\int f_x^+ \,\mathrm{d}\mu_2$ is also an integrable function. Define the function g as above. If we let

$$E = \left\{x \in X_1 : \int f_x^+ \,\mathrm{d}\mu_2 < \infty\right\}$$

Then, $g^+ = [\int f_x^+(y) \,\mathrm{d}\mu_2]\mathbb{1}_E$. Since E is a measurable set (inverse image set of a measurable function), g^+ is a measurable function. Moreover, since $\int f_x^+(y) \,\mathrm{d}\mu_2$ is finite a.e., $\mu_1(E^c) = 0$. Therefore,

$$\int g^+ \,\mathrm{d}\mu_1 = \iint f_x^+ \,\mathrm{d}\mu_2 \,\mathrm{d}\mu_1 < \infty$$

This shows that g^+ is an integrable function. This argument can also be applied to g^-. Then, we have

$$\begin{aligned} \int f \,\mathrm{d}\mu &= \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu \\ &= \int \left[\int f_x^+ \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1 - \int \left[\int f_x^- \,\mathrm{d}\mu_2\right] \mathrm{d}\mu_1 && \text{(Tonelli's Theorem)} \\ &= \int g^+ \,\mathrm{d}\mu_1 - \int g^- \,\mathrm{d}\mu_1 = \int g \,\mathrm{d}\mu_1 && (g^+ \text{ and } g^- \text{ are integrable}) \end{aligned}$$

which is the equality that we desire. □

Final Remark: If instead of integrable, we assume that

$$\int \left[\int |f_x| \, \mathrm{d}\mu_2 \right] \mathrm{d}\mu_1 < \infty$$

Then, the conclusions of Fubini's theorem are in force. This is because, $f_x^+, f_x^- \leqslant |f_x|$. Therefore,

$$\int \left[\int f_x^+ \, \mathrm{d}\mu_2 \right] \mathrm{d}\mu_1 < \infty, \quad \int \left[\int f_x^- \, \mathrm{d}\mu_2 \right] \mathrm{d}\mu_1 < \infty$$

By Tonelli's theorem, we have

$$\int f^+ \, \mathrm{d}\mu < \infty, \quad \int f^- \, \mathrm{d}\mu < \infty$$

which shows that f is integrable. Then, we can use Fubini's theorem to get the same conclusions.

4.6 Hahn-Jordan Decomposition

4.6.1 Signed Measure

In this section, we generalize the definition of measure to *signed measure*.

Definition 4.6.1: Signed Measure

Let $(X, \mathcal{A})$ be a measurable space. A **signed measure** on $(X, \mathcal{A})$ is a function $\nu : \mathcal{A} \to \bar{\mathbb{R}}$ such that

- $\nu(\emptyset) = 0$.
- ν is countably additive.

Remark: There are some implicit statements in this definition:

- If $E, F \in \mathcal{A}$ are disjoint, then $\nu(E \cup F) = \nu(E) + \nu(F)$ by definition. However, it might be the case that $\nu(E) = +\infty$ but $\nu(F) = -\infty$, in which case $\nu(E \cup F)$ is not defined. Therefore, for well-definedness, this situation can never happen.
- Suppose $\{E_j\}_{j=1}^{\infty} \subseteq \mathcal{A}$ is a sequence of disjoint sets. Then, we have $\nu\left(\bigsqcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \nu(E_j)$ by countable additivity. However, the sum on the RHS might be sensitive to the order of sum. Therefore, for well-definedness, we must have the latter sum *converges absolutely* if it is finite.

There are some really simple properties that we can expect for signed measure.

Property 4.6.2: Generalized Monotonicity

Let $(X, \mathcal{A})$ be a measurable space and ν be a signed measure on it. Let $E, F \in \mathcal{A}$ with $E \subseteq F$. Then,

(a) If $|\nu(E)| < \infty$, then $\nu(F \backslash E) = \nu(F) - \nu(E)$.

(b) If $\nu(E) = +\infty$, then $\nu(F) = +\infty$.

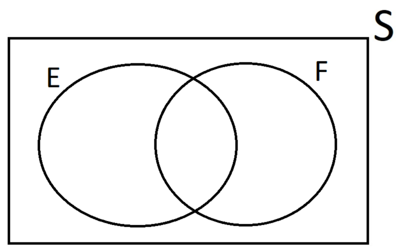
(c) If $\nu(E) = -\infty$, then $\nu(F) = -\infty$.

Proof. (a) We can write $F = E \cup (F \backslash E)$. Then, by additivity of signed measure, $\nu(F) = \nu(E) + \nu(F\backslash E)$. Since $|\nu(E)| < \infty$, we can subtract both sides by $\nu(E)$, and get $\nu(F\backslash E) = \nu(F) - \nu(E)$.
(b) We still have $\nu(F) = \nu(E) + \nu(F\backslash E)$, but with $\nu(E) = \infty$. Then, $\nu(F\backslash E) \neq -\infty$ by well-definedness. This leads to $\nu(F) = +\infty$.
(c) We still have $\nu(F) = \nu(E) + \nu(F\backslash E)$, but with $\nu(E) = -\infty$. Then, $\nu(F\backslash E) \neq +\infty$ by well-definedness. This leads to $\nu(F) = -\infty$. □

Property 4.6.3: $+\infty$ and $-\infty$ cannot be reached simultaneously

Let $(X, \mathcal{A})$ be a measurable space and ν be a signed measure on it. Let $E, F \in \mathcal{A}$ and $\nu(E) = +\infty$. Then, $\nu(F) > -\infty$.

Proof. $\nu(E) = +\infty = \nu(E \cap F) + \nu(E\backslash F)$. Note that $\nu(E \cap F) \neq -\infty$. We separate few cases to discuss.



- If $\nu(E \cap F) < \infty$, then $\nu(E\backslash F) = +\infty$. Then, we can't have $\nu(F) = -\infty$ since $E\backslash F$ and F are disjoint, for well-definedness, these two cannot be one $+\infty$ and one $-\infty$.

- If $\nu(E \cap F) = +\infty$. Then $\nu(F) = +\infty$ since $E \cap F \subseteq F$, and by Property 4.6.2.

□

Due to the above property, the range of ν is either $(-\infty, +\infty]$, or $[-\infty, +\infty)$. For simplicity, we assume, from now on, that:

$$\nu(F) \in (-\infty, +\infty], \text{ for all } F \in \mathcal{A}$$

Now we bring some previous propositions of measures onto signed measures.

Proposition 4.6.4: Continuous above and below

Let $(X, \mathcal{A})$ be a measurable space and ν be a signed measure on it. Then,

(a) ν is continuous from below. i.e., If $\{E_j\}_{j=1}^{\infty}$ is an increasing sequence on $\mathcal{A}$, then $\lim_{j\to\infty} \nu(E_j) = \nu\left(\bigcup_{j=1}^{\infty} E_j\right)$.

(b) ν is continuous from above. i.e., if $\{E_j\}_{j=1}^{\infty}$ is a decreasing sequence in $\mathcal{A}$ and $\nu(E_{n_0}) < \infty$ for some $n_0 \in \mathbb{N}$. Then, $\lim_{j\to\infty} \nu(E_j) = \nu\left(\bigcap_{j=1}^{\infty} E_j\right)$.

Proof. (a) Let $F_1 = E_1, F_2 = E_2\backslash E_1, \cdots, F_k = E_k\backslash E_{k-1}, \cdots$. Then, $F_k \in \mathcal{A}$ is a disjoint sequence and $\bigcup_{j=1}^{\infty} E_j =$

$\bigsqcup_{k=1}^{\infty} F_k$. Therefore,

$$\nu\left(\bigcup_{j=1}^{\infty} E_j\right) = \nu\left(\bigsqcup_{k=1}^{\infty} F_k\right) = \sum_{k=1}^{\infty} \nu(F_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \nu(F_k) = \lim_{n\to\infty} \nu\left(\bigsqcup_{k=1}^{n} F_k\right) = \lim_{n\to\infty} \nu(E_n)$$

(b) Let $F_i = E_1 \backslash E_i$. Then, $F_1 \subseteq F_2 \subseteq F_3 \subseteq \cdots$, and $\mu(E_1) = \mu(F_i) + \mu(E_i)$ by finite additivity. Also, $\bigcup_i^\infty F_i = E_1 \backslash (\bigcap_{i=1}^\infty E_i)$. By continuity from below, we have

$$\mu\left(E_1 \backslash \left(\bigcap_{i=1}^{\infty} E_i\right)\right) = \mu\left(\bigcup_{i=1}^{\infty} F_i\right) = \lim_{n\to\infty} \mu(F_n) = \lim_{n\to\infty} (\mu(E_1) - \mu(E_n)) = \mu(E_1) - \lim_{n\to\infty} \mu(E_n) \tag{4.3}$$

Note that, by countable additivity, we have

$$\mu\left(E_1 \backslash \left(\bigcap_{i=1}^{\infty} E_i\right)\right) + \mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \mu(E_1) \tag{4.4}$$

where we can subtract $\mu(\bigcap_{i=1}^\infty E_i)$ on both sides since $\mu(E_1) < \infty$ and the sequence is decreasing. Then, combining the equation 4.3 and 4.4, we have

$$\mu(E_1) - \mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \mu(E_1) - \lim_{n\to\infty} \mu(E_n)$$

Since $\mu(E_1) < +\infty$, we can substract this term from both sides to get our expected result:

$$\mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{n\to\infty} \mu(E_n)$$

which proves the result. □

4.6.2 Hahn Decomposition Theorem

This section is aimed to decompose any signed measure into two sets, where on one set, it always takes positive values, and on the other set, it always takes negative values.

Definition 4.6.5: Positive Set/Negative Set

Let $(X, \mathcal{A})$ be a measurable space and ν be a signed measure on it. Then,

- A set $E \in \mathcal{A}$ is called **positive** for ν if $\nu(F) \geqslant 0$ for all $F \in \mathcal{A}$ such that $F \subseteq E$.
- A set $E \in \mathcal{A}$ is called **negative** for ν if $\nu(F) \leqslant 0$ for all $F \in \mathcal{A}$ such that $F \subseteq E$.
- A set $E \in \mathcal{A}$ is called **null** for ν if $\nu(F) = 0$ for all $F \in \mathcal{A}$ such that $F \subseteq E$.

Theorem 4.6.6: Hahn Decomposition Theorem

Let $(X, \mathcal{A})$ be a measurable space and ν be a signed measure on it. Then, there exists a positive set P and a negative set N for ν such that $P \cup N = X$ and $P \cap N = \emptyset$. If P', N' is another such pair, then $P \Delta P' = N \Delta N'$ is null for ν.

Proof. Define

$$\alpha = \inf_{A \in \mathcal{A}} \nu(A)$$

We know that $\nu(A) > -\infty$. However, α can be $-\infty$ since it takes the infimum. The goal of the next step is to prove that α cannot be negative infinity by proving that, if so, there would be a set B such that $\nu(B) = -\infty$, which is a contradiction with our definition.

STEP I: We assume that $\alpha = -\infty$. Then, we will iterate to construct a set with negative infinite signed measure.

- ***ITERATE I:*** Let $A_0 = X$. Define

  $$\lambda(C) = \inf_{E \subseteq C} \nu(E), \quad E \in \mathcal{A}$$

  Then,

  $$\lambda(A_0) = \lambda(X) = \alpha = -\infty$$

  By the definition of infimum, there exists $B_1 \subseteq A_0$, such that $\nu(B_1) \leqslant -1$. Now, we define

  $$A_1 = \begin{cases} B_1, & \text{if } \lambda(B_1) = -\infty, \\ A_0 \backslash B_1, & \text{if } \lambda(B_1) > -\infty \end{cases}$$

  In the first case, we know that $\lambda(A_1) = \lambda(B_1) = -\infty$. In the second case, which we call a *bifurcation*, we will also have $\lambda(A_1) = -\infty$. To show this, suppose $\lambda(A_1) = \lambda(A_0 \backslash B_1) = c > -\infty$ and $\lambda(B_1) = d > -\infty$. Then, for any $E \subseteq A_0$, we have

  $$\nu(E) = \nu(E \cap B_1) + \nu(E \cap (A_0 \backslash B_1)) \geqslant \lambda(B_1) + \lambda(A_0 \backslash B_1)$$

  where the last inequality is by the definition of λ. Taking infimum on both sides, we then have

  $$\lambda(A_0) = \alpha \geqslant \lambda(B_1) + \lambda(A_0 \backslash B_1) = c + d > -\infty$$

  which is a contradiction. To conclude, we construct in this step

  $$A_1 \subseteq A_0, \quad \lambda(A_1) = -\infty, \quad B_1 \subseteq A_0, \quad \nu(B_1) \leqslant -1$$

- ***ITERATE II:*** Now we can continue the steps above again based on A_1. We can similarly take $B_2 \subseteq A_1$ such that $\nu(B_2) \leqslant -2$, and then take

  $$A_2 = \begin{cases} B_2, & \text{if } \lambda(B_2) = -\infty, \\ A_1 \backslash B_2, & \text{if } \lambda(B_2) > -\infty \end{cases}$$

  and then we will have

  $$A_2 \subseteq A_1, \quad \lambda(A_2) = -\infty, \quad B_2 \subseteq A_1, \quad \nu(B_2) \leqslant -2$$

- ***ITERATE k:*** After k steps, we will have

  $$A_{k+1} \subseteq A_k, \quad \lambda(A_k) = -\infty, \quad B_k \subseteq A_{k-1}, \quad \nu(B_k) \leqslant -k$$

Suppose that there is a bifurcation happened in the iteration step $n = k + 1$. Then, $A_{k+1} = A_k \backslash B_{k+1}$. Then, for all

$j \geqslant k+1$, we have

$$A_j \subseteq A_{k+1} \quad \Longrightarrow \quad B_{j+1} \subseteq A_j \subseteq A_{k+1}$$

Moreover, since it is a bifurcation,

$$A_{k+1} \cap B_{k+1} = \emptyset$$

Combining these two, we have $B_{j+1} \cap B_{k+1} = \emptyset$. Now we analyze two cases of bifurcation.

1. If there are only finitely many bifurcations. That is, there exists N such that, for all $n \geqslant N$, we have $B_n = A_n$. Then, since A_n is a decreasing sequence, B_n is also decreasing after $n \geqslant N$. Suppose $B_n \searrow B$. By continuity from above, since $\nu(B_n) \leqslant -n$, we have

$$\nu(B) = \lim_{n\to\infty} \nu(B_n) = -\infty$$

which is a contradiction since ν cannot take value $-\infty$.

2. If there are infinitely may bifurcations. Then, suppose $n_0, n_1, n_2, \cdots$ are the iteration steps where the bifurcation happens. Then, by the argument above about bifurcation, we have $B_{n_j} \cap B_{n_k} = \emptyset$ for all $n_j \neq n_k$. Let $B = \bigsqcup_{j=1}^{\infty} B_{n_j}$, then

$$\nu(B) = \sum_{j=1}^{\infty} \nu(B_{n_j}) = -\infty$$

Since each $\nu(B_{n_j}) \leqslant -n_j$, which is also a contradiction.

Therefore, $\alpha > -\infty$.

STEP II: In this step, we want to construct a set N such that $\nu(N) = \alpha$. By the definition of infimum, there exists a sequence of sets C_n, such that

$$\alpha \leqslant \nu(C_n) \leqslant \alpha + \frac{1}{2^n} \tag{4.5}$$

Then, we calculate

$$\begin{aligned}
\nu(C_n \cup C_{n+1}) &= \nu(C_n) + \nu(C_{n+1}) - \nu(C_n \cap C_{n+1}) \\
&\leqslant \alpha + \frac{1}{2^n} + \alpha + \frac{1}{2^{n+1}} - \nu(C_n \cap C_{n+1}) && \text{(Equation 4.5)} \\
&\leqslant \alpha + \frac{1}{2^n} + \frac{1}{2^{n+1}} && (\nu(C_n \cap C_{n+1}) \geqslant \alpha \text{ by the definition of } \alpha)
\end{aligned}$$

We can do this iteratively to get

$$\nu\left(\bigcup_{k=n}^{n+q} C_k\right) \leqslant \alpha + \frac{1}{2^n} + \cdots + \frac{1}{2^{n+q}}$$

Then, by continuity from below, we have

$$\nu\left(\bigcup_{k=n}^{\infty} C_k\right) = \lim_{q\to\infty} \nu\left(\bigcup_{k=n}^{n+q} C_k\right) = \alpha + \frac{2}{2^n} = \alpha + \frac{1}{2^{n-1}}$$

If we denote $D_n = \bigcup_{k=n}^{\infty} C_k$, then D_n is a decreasing sequence. Let $N = \bigcap_{n=1}^{\infty} D_n$. Then,

$$\nu(N) \leqslant \lim_{n\to\infty} \left(\alpha + \frac{1}{2^{n-1}}\right) = \alpha$$

by continuity from above. However, $\nu(N) \geqslant \alpha$ by definition of α. Therefore, we have $\nu(N) = \alpha$, which is what we want.

STEP III: We claim that this set N is what we want in the theorem.
1. Let $P = N^c$. Then, for some $E \subseteq P$, $E \in \mathcal{A}$, suppose that $\nu(E) < 0$. Then, we have

$$\nu(N \cup E) = \nu(N) + \nu(E) = \alpha + \nu(E) < \alpha$$

which is a contradiction since α is the minimum.
2. Similarly, for some $F \subseteq N$, suppose that $\nu(F) > 0$. We first claim that $\nu(F) \neq +\infty$. This is because, $F \subseteq N$, if $\nu(F) = +\infty$, then $\nu(N) = +\infty$ by Property 4.6.2. However, this cannot happen since $\nu(N) = \alpha \leqslant \nu(\emptyset) = 0$. Then, we have

$$\nu(N \backslash F) = \nu(N) - \nu(F) = \alpha - \nu(F) < \alpha$$

which is also a contradiction.

STEP IV: Finally, suppose that P', N' is another pair of sets that satisfy the properties in the theorem. Then, $P\backslash P' \subseteq P$ obviously. Moreover, $P\backslash P' \subseteq N$ since P'^c is a negative set, and $P\backslash P' \subseteq P'^c$. Therefore, $P\backslash P'$ is both positive and negative, hence a null set. Similarly, $P'\backslash P$ is also a null set. □

4.6.3 Jordan Decomposition Theorem

Besides decomposing sets, we can also decompose the signed measure into two parts.

Definition 4.6.7: Mutually Singular Measures

Let $(X, \mathcal{A})$ be a measurable space and let μ and ν be signed measures on it. The signed measures μ and ν are said to be **mutually singular**, or ν is **singular** with respect to μ, or vice versa, if there exists $E, F \in \mathcal{A}$ such that $E \cap F = \emptyset$, $E \cup F = X$, E is null for μ, and F is null for ν. We denote it by $\mu \perp \nu$.

Example 1: Let λ be the Lebesgue measure. Let $\mathbb{Q} = \{q_j, j \geqslant 1\}$ be the enumeration of all rational numbers. Let $\sum_{j=1}^{\infty} c_j < \infty$. We then let

$$\nu = \sum_{j=1}^{\infty} c_j \delta_{q_j}$$

where $\delta_x(A)$ is the Dirac measure on set A. Then, $\nu(\mathbb{Q}^c) = 0$, and $\lambda(\mathbb{Q}) = 0$. Therefore, $\lambda \perp \nu$.

Theorem 4.6.8: Jordan Decomposition Theorem

If ν is a signed measure on measurable space $(X, \mathcal{A})$, then there exists *unique* measures ν^+ and ν^- such that

$$\nu = \nu^+ - \nu^-, \quad \text{and} \quad \nu^+ \perp \nu^-$$

Proof. Let $X = P \cup N$ be a Hahn decomposition for ν, and define $\nu^+(E) = \nu(E \cap P)$ and $\nu^-(E) = -\nu(E \cap N)$. Then, for any set $A \in \mathcal{A}$, we have

$$\nu(A) = \nu(A \cap P) + \nu(A \cap N) = \nu^+(A) - \nu^-(A)$$

and they are mutually singular obviously. Moreover, if we also have $\nu = \mu^+ - \mu^-$ and $\mu^+ \perp \mu^-$, let $E, F \in \mathcal{A}$ such that $E \cap F = \emptyset$ and $E \cup F = X$, and $\mu^+(E) = \mu^-(F) = 0$. Then, $X = E \cup F$ is another Hahn decomposition for ν. By Theorem 4.6.6, $P \Delta E$ is null. Therefore, for any $A \in \mathcal{A}$,

$$\mu^+(A) = \mu^+(A \cap E) = \nu(A \cap E) = \nu(A \cap P) = \nu^+(A)$$

and similar for μ^-, which proves that this decomposition is unique. □

4.7 Radon-Nikodym Theorem

Definition 4.7.1: Absolute Continuity

Let $(X, \mathcal{A})$ be a measurable space. Let ν be a signed measure of the space and μ be a measure of the space. Then, ν is **absolutely continuous** with respect to μ, if

$$\mu(A) = 0 \quad \text{implies} \quad \nu(A) = 0$$

We write this as $\nu \ll \mu$.

Example 1: Let f be an integrable function in the measure space $(X, \mathcal{A}, \mu)$. Then, let

$$\nu(A) = \int_A f \, \mathrm{d}\mu$$

Then, ν is a signed measure. To prove this, first notice that $\nu(\emptyset) = 0$, trivially. Moreover, if sets $A_n \in \mathcal{A}$ are disjoint, then the series $\sum_{n=1}^{\infty} \mathbb{1}_{A_n}(x) f(x)$ converges for every x to $I_A(x) f(x)$. In addition, $|\sum_{i=1}^{N} I_{A_n}(x) f(x)| \leqslant I_A(x)|f(x)|$. Since $f(x)$ is integrable, by dominated convergence theorem, we have $\nu(\bigsqcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \nu(A_n)$.

This example shows that, if we have a measure and a function f, then we can always define the signed measure so that it is just the integral of f with respect to μ. However, we also want to ask the converse question: If we have ν and μ, when does the f exist? The *Radon-Nikodym theorem* answers this question, and it is very useful for defining the density function of a random variable, and proving the existence of conditional expectation.

Theorem 4.7.2: Radon-Nikodym Theorem

Let $(X, \mathcal{A}, \mu)$ be a measure space, where μ is a σ-finite measure. Let ν be a σ-finite signed measure defined on this space. Then,

(a) There exists *unique* signed measures ν_1, ν_2 such that

$$\nu = \nu_1 + \nu_2, \quad \text{where} \quad \nu_1 \leqslant \mu, \, \nu_2 \perp \mu$$

(b) There exists a measurable function f such that

$$\nu_1(A) = \int_A f \, \mathrm{d}\mu$$

Proof. **STEP I: With Two Assumptions** We first assume that ν is a measure, and ν, μ are all finite measures. Define

$$\mathcal{H} = \left\{ f \text{ measurable, } f \geqslant 0 : \int_A f \, \mathrm{d}\mu \leqslant \nu(A), \forall A \in \mathcal{A} \right\}$$

(The motivation to define this set is: since $\nu = \nu_1 + \nu_2$, and now we have assumption ν is a measure, so $\nu_1 \leqslant \nu$ obviously.) This set is nonempty since $f = 0 \in \mathcal{H}$. Let

$$\alpha = \sup_{f \in \mathcal{H}} \int_X f \, \mathrm{d}\mu$$

Now we start to have some claims:

Claim 1: $\alpha < \infty$. To prove this, just observe that

$$\int_X f \, \mathrm{d}\mu \leqslant \nu(X) < \infty$$

since we have assumption that ν is finite.

With this information in hand, by the definition of infimum, there exists a sequence of measurable functions $\{f_n\}_{n=1}^{\infty}$ such that

$$\alpha - \frac{1}{n} \leqslant \int_X f_n \, \mathrm{d}\mu \leqslant \alpha \tag{4.6}$$

Then, we can take $g_n = \max\{f_1, f_2, \cdots, f_n\}$, and $\{g_n\}$ would be an increasing sequence of measurable functions. Let $g_n \nearrow g$.

Claim 2: $g_n \in \mathcal{H}$. To prove this, denote the set

$$E_{n,k} = \{x \in X : g_n(x) = f_k(x)\}, \quad 1 \leqslant k \leqslant n$$

That is, $E_{n,k}$ indicates the portion of X where g_n takes the value of f_k. There might be some overlap between these sets, but without loss of generality, we assume that $\{E_{n,k}\}_{k=1}^n$ are disjoint by distributing overlapping part to one of the sets, and eliminate this part from others. Then,

$$\int_A g_n \, \mathrm{d}\mu = \sum_{k=1}^n \int_{A \cap E_{n,k}} g_n \, \mathrm{d}\mu = \sum_{k=1}^n \int_{A \cap E_{n,k}} f_k \, \mathrm{d}\mu \leqslant \sum_{k=1}^n \nu(A \cap E_{n,k}) = \nu(A)$$

where the second equation holds by definition of g_n, the third inequality holds by definition of $\mathcal{H}$, and final equality holds by finite additivity.

Claim 3: $g \in \mathcal{H}$. To prove this, just use the monotone convergence theorem, we have

$$\begin{aligned} \int_A g \, \mathrm{d}\mu &= \lim_{n \to \infty} \int_A g_n \, \mathrm{d}\mu && \text{(MCT)} \\ &\leqslant \nu(A) && \text{(By the last equation)} \end{aligned}$$

Claim 4: $\int g \, \mathrm{d}\mu = \alpha$. To prove this, first notice that $\int g \, \mathrm{d}\mu \leqslant \alpha$ by the definition of α. Then, notice that since g_n is an increasing sequence, we have

$$\int g \, \mathrm{d}\mu \geqslant \int g_n \, \mathrm{d}\mu \geqslant \int f_n \, \mathrm{d}\mu \geqslant \alpha - \frac{1}{n}$$

Taking limit on both sides, we have $\int g \, \mathrm{d}\mu \geqslant \alpha$.

Now, we define

$$\nu_1(A) = \int_A g \, \mathrm{d}\mu \leqslant \nu(A), \quad \nu_2(A) = \nu(A) - \nu_1(A) \geqslant 0$$

Then, $\nu_1 \ll \nu$ follows naturally by this definition. The only thing we need to prove is stated in the following claim:
Claim 5: $\nu_2 \perp \mu$. To prove this, let's consider the inequality $\frac{\nu_2}{\mu} \geqslant \frac{1}{n}$. When ν_2 and μ are indeed singular, then the subsets that make this inequality true would be on ν_2's nonzero region, and this might potentially separate the disjoint nonzero domain of ν_2 and μ.
So, we consider the signed measure $\sigma_n = \nu_2 - \frac{1}{n}\mu$. By Hahn decomposition, there exists P_n such that for all $E \subseteq P_n$, we have $\sigma_n(E) \geqslant 0$, and $N_n = P_n^c$, where for all $F \subseteq N_n$, we have $\sigma_n(F) \leqslant 0$. We will first show that

$$g + \frac{1}{n}\mathbb{1}_{P_n} \in \mathcal{H}$$

To show this, notice that

$$\int_A \left(g + \frac{1}{n}\mathbb{1}_{P_n}\right) \mathrm{d}\mu = \nu_1(A) + \frac{1}{n}\mu(P_n \cap A) \tag{4.7}$$

However, $P_n \cap A \subseteq P_n$, so $\sigma(P_n \cap A) \geqslant 0$, which indicates that

$$\nu_2(P_n \cap A) - \frac{1}{n}\mu(P_n \cap A) \geqslant 0 \quad \Longrightarrow \quad \frac{1}{n}\mu(P_n \cap A) \leqslant \nu_2(P_n \cap A) \tag{4.8}$$

Combining Equation 4.7 and 4.8, we have

$$\int_A \left(g + \frac{1}{n}\mathbb{1}_{P_n}\right) \mathrm{d}\mu \leqslant \nu_1(A) + \nu_2(P_n \cap A) \leqslant \nu_1(A) + \nu_2(A) = \nu(A)$$

which proves that $g + \frac{1}{n}\mathbb{1}_{P_n} \in \mathcal{H}$. Then, we have

$$\int \left(g + \frac{1}{n}\mathbb{1}_{P_n}\right) \mathrm{d}\mu = \alpha + \frac{1}{n}\mu(P_n)$$

However, by the definition of α, we must have $\frac{1}{n}\mu(P_n) \leqslant 0$, otherwise it is a contradiction. Therefore, $\mu(P_n) = 0$ for all n. This indicates that, if we denote $P = \bigcup_{n=1}^{\infty} P_n$, we will have

$$\mu(P) \leqslant \sum_{n=1}^{\infty} \mu(P_n) = 0 \quad \Longrightarrow \quad \mu(P) = 0$$

Then, let $N = P^c = \bigcap_{n=1}^{\infty} P_n^c = \bigcap_{n=1}^{\infty} N_n$. Then, $\nu_2(N) \leqslant \nu_2(N_n)$ by monotonicity. Moreover, since $\sigma(N_n) \leqslant 0$, we have $\nu_2(N_n) - \frac{1}{n}\mu(N_n) \leqslant 0$, which shows that $\nu_2(N_n) \leqslant \frac{1}{n}\mu(N_n)$. Combining these two, we have

$$\nu_2(N) \leqslant \nu_2(N_n) \leqslant \frac{1}{n}\mu(N_n) \leqslant \frac{1}{n}\mu(X) < \infty$$

Therefore, we can take $n \to \infty$ on both sides, and have $\nu_2(N) = 0$. This completes the proof that $\nu_2 \perp \mu$.

STEP II: Remove 2nd Assumption Assume μ and ν are σ-finite. Then,

$$\exists\, E_n \text{ increasing, } X = \bigcup_{n=1}^{\infty} E_n,\, \mu(E_n) < \infty$$

$$\exists\, F_n \text{ increasing, } X = \bigcup_{n=1}^{\infty} F_n,\, \nu(F_n) < \infty$$

Then, take $G_n = E_n \cap F_n$. G_n is also increasing, and $X = \bigcup_{n=1}^{\infty} G_n$. Moreover, $\mu(G_n) < \infty$ and $\nu(G_n) < \infty$ by monotonicity. Now, take $H_1 = G_1, H_2 = G_2 \backslash G_1, \cdots, H_k = G_k \backslash G_{k-1}$. Then, H_k's are disjoint. Moreover, $X = \bigsqcup_{j=1}^{\infty} H_j$, and $\mu(H_j) < \infty$, $\nu(H_j) < \infty$. With this in hand, define

$$\mu_j(A) = \mu(H_j \cap A), \quad \nu_j(A) = \nu(H_j \cap A)$$

By STEP I, we have that $\nu_j = \nu_j^1 + \nu_j^2$ where $\nu_j^1 \ll \mu_j$, and $\nu_j^2 \perp \mu_j$. Let f_j be the function such that $\nu_j^1(A) = \int_A f \, \mathrm{d}\mu_j$ on H_j, and pick $N_j \subset H_j$ such that $\mu(N_j) = 0$ and $\nu_j^2(H_j \setminus N_j) = 0$.
Define the global objects

$$f := \sum_{j=1}^{\infty} f_j \, \mathbf{1}_{H_j} \quad (\text{measurable on } X), \qquad N := \bigcup_{j=1}^{\infty} N_j \quad (\mu(N) = 0).$$

Now set

$$\nu^1(A) := \sum_{j=1}^{\infty} \nu_j^1(A) = \sum_{j=1}^{\infty} \int_{A \cap H_j} f_j \, d\mu = \int_A f \, d\mu, \qquad \nu^2(A) := \sum_{j=1}^{\infty} \nu_j^2(A), \qquad A \in \mathcal{A}.$$

(The sums are well-defined by countable additivity across the partition $\{H_j\}$.)

Absolute continuity of ν^1. If $\mu(A) = 0$, then $\mu_j(A) = \mu(A \cap H_j) = 0$ for all j, hence $\nu_j^1(A) = 0$ for all j because $\nu_j^1 \ll \mu_j$. Therefore

$$\nu^1(A) = \sum_{j=1}^{\infty} \nu_j^1(A) = 0,$$

so $\nu^1 \ll \mu$. Equivalently, $\nu^1(A) = \int_A f \, d\mu$ for all A.

Singularity of ν^2. Since $\mu(N_j) = 0$ for each j, we have $\mu(N) = 0$. For any $A \subset X \setminus N$,

$$A = \bigsqcup_{j=1}^{\infty} \big(A \cap (H_j \setminus N)\big) = \bigsqcup_{j=1}^{\infty} \big(A \cap (H_j \setminus N_j)\big),$$

and by the choice of N_j, $\nu_j^2\big(A \cap (H_j \setminus N_j)\big) = 0$ for all j. Hence

$$\nu^2(A) = \sum_{j=1}^{\infty} \nu_j^2\big(A \cap (H_j \setminus N_j)\big) = 0.$$

Thus $\nu^2 \perp \mu$.

STEP III: Remove 1st Assumption Suppose now, ν is a signed measure. Let $\nu = \nu^+ - \nu^-$ be the Jordan decomposition coming from a Hahn decomposition $X = P \sqcup N$ (so $\nu^+(A) = \nu(A \cap P)$ and $\nu^-(A) = -\nu(A \cap N)$). Because ν is σ-finite, $\nu^\pm \leq |\nu|$ are σ-finite *measures.*

Apply STEP II separately to the (positive) measures ν^+ and ν^-: there exist decompositions

$$\nu^+ = a^+ + s^+, \qquad \nu^- = a^- + s^-,$$

with $a^\pm \ll \mu$ and $s^\pm \perp \mu$. Moreover, there exist measurable densities $f^+, f^- \geq 0$ such that

$$a^+(A) = \int_A f^+ \, d\mu, \qquad a^-(A) = \int_A f^- \, d\mu, \qquad A \in \mathcal{A}.$$

Define signed measures

$$\nu_1 := a^+ - a^-, \qquad \nu_2 := s^+ - s^-,$$

and the (signed) measurable function

$$f := f^+ - f^-.$$

Then for every $A \in \mathcal{A}$,

$$\nu(A) = \nu^+(A) - \nu^-(A) = (a^+ + s^+)(A) - (a^- + s^-)(A) = \nu_1(A) + \nu_2(A),$$

and

$$\nu_1(A) = a^+(A) - a^-(A) = \int_A f^+ \, d\mu - \int_A f^- \, d\mu = \int_A (f^+ - f^-) \, d\mu = \int_A f \, d\mu.$$

Absolute continuity of ν_1. If $\mu(A) = 0$ then $a^\pm(A) = 0$ (since $a^\pm \ll \mu$), hence $\nu_1(A) = 0$. Thus $\nu_1 \ll \mu$. Equivalently, $\nu_1(A) = \int_A f \, d\mu$ for all A.

Singularity of ν_2. Pick $N^+, N^- \in \mathcal{A}$ with $\mu(N^\pm) = 0$ and $s^\pm(X \setminus N^\pm) = 0$. Set $N := N^+ \cup N^-$. Then $\mu(N) = 0$, and for every $A \subset X \setminus N$ we have $s^\pm(A) = 0$, hence $\nu_2(A) = s^+(A) - s^-(A) = 0$. Therefore $\nu_2 \perp \mu$.

This yields the Lebesgue decomposition for signed ν:

$$\nu = \underbrace{\nu_1}_{\ll \mu} + \underbrace{\nu_2}_{\perp \mu}, \qquad \text{with } \ \nu_1(A) = \int_A f \, d\mu \ \ (A \in \mathcal{A}).$$

Uniqueness. If $\nu = \eta_1 + \eta_2$ with $\eta_1 \ll \mu$ and $\eta_2 \perp \mu$, then $\tau := \eta_1 - \nu_1 = -(\eta_2 - \nu_2)$ is both $\ll \mu$ and $\perp \mu$, hence $\tau = 0$. Thus $\eta_1 = \nu_1$ and $\eta_2 = \nu_2$. The density f is unique μ-a.e., since $\int_A (f - g) \, d\mu = 0$ for all A forces $f = g$ μ-a.e. $\square$

Definition 4.7.3: Radon-Nikodym Derivative

The decomposition in Radon-Nikodym theorem is called a **Lebesgue decomposition** of ν with respect to μ. If $\nu \ll \mu$, the result says that there exists a function f such that $\nu(A) = \int_A f \, \mathrm{d}\mu$. This f is called the **Radon-Nikodym derivative** of ν with respect to μ. We denote it by $f = \mathrm{d}\nu/\mathrm{d}\mu$.

Remark: We sometimes write

$$\mathrm{d}\nu = f \, \mathrm{d}\mu$$

Chapter 5

L^p Spaces and Convergence

In this chapter we will talk about several modes of convergence and some related topics of L^p spaces. We will always assume that we are working in a measure space $(X, \mathcal{A}, \mu)$, where μ is σ-finite and $\mathcal{A}$ is complete (that is, if f is measurable, and $g = f$ a.e., then g is also measurable).

We will also define an equivalence relation: $f \sim g$ means that $\mu(f \neq g) = 0$. That is, f and g are equivalent if $f = g$ a.e.. Define the space

$$M = \{f : X \to \bar{\mathbb{R}} \mid f \text{ is measurable}\}$$

Then, we take $\mathcal{M} = M/\sim$, and we will often work on the space $\mathcal{M}$ later on.

5.1 Almost Sure Convergence

Definition 5.1.1: Pointwise Convergence

Let $f, f_n : X \to \bar{\mathbb{R}}$ be measurable functions. Then, we say f_n converges to f (written as $f_n \to f$) **pointwise** if for all $x \in E$, we have $f_n(x) \to f(x)$.

Definition 5.1.2: Almost Sure Convergence

Let $f, f_n : E \to \bar{\mathbb{R}}$ be measurable functions. Then, f_n converges to f **almost surely** (written as $f_n \overset{a.e.}{\to} f$ or $f_n \overset{a.s.}{\to} f$) if there exists $E \in \mathcal{A}$, such that $\mu(E^c) = 0$, and $f_n \to f$ on E pointwise.

Several simple but important properties of this convergence mode is listed below.

Property 5.1.3: Convergence on Quotient Space

Let $f, f_n : X \to \bar{\mathbb{R}}$ be measurable functions, and $f_n \overset{a.s.}{\to} f$. Let $g_n \sim f_n$ and $g \sim f$. Then, we have $g_n \overset{a.s.}{\to} g$.

Proof. Since $f_n \overset{a.s.}{\to} f$, $f_n \sim g_n$ and $f \sim g$, we can get, respectively,

$$\exists\, E \in \mathcal{A},\ \mu(E^c) = 0,\ f_n(x) \to f(x) \text{ for all } x \in E$$

$$\exists\, F_n \in \mathcal{A},\ \mu(F_n^c) = 0,\ f_n(x) = g_n(x) \text{ for all } x \in F_n$$

$$\exists\, F \in \mathcal{A},\ \mu(F^c) = 0,\ f(x) = g(x) \text{ for all } x \in F$$

Let $H = E \cap F \cap \bigcap_{n \geqslant 1} F_n$. Then, $H \in \mathcal{A}$, and $\mu(H^c) = 0$ because

$$\mu(H^c) = \mu\left(E^c \cup F^c \cup \bigcup_{n=1}^{\infty} F_n^c\right) \leqslant \mu(E^c) + \mu(F^c) + \sum_{n=1}^{\infty} \mu(F_n^c) = 0$$

If $x \in H$, then $g_n(x) = f_n(x)$ because $x \in F_n$ for all n. Then, $f_n(x) \to f$ since $x \in E$. Finally, $f(x) = g(x)$ since $x \in F$. This indicates that $g_n(x) \to g(x)$ on $x \in H$, which means that $g_n(x) \overset{a.e.}{\to} g$. □

Property 5.1.4: Unique Convergence on Quotient Space

Let $f_n, f, g : X \to \bar{\mathbb{R}}$ be measurable functions. Let $f_n \overset{a.s.}{\to} f$ and $f_n \overset{a.s.}{\to} g$. Then, $f = g$ a.s..

Proof. Since $f_n \overset{a.s.}{\to} f$ and $f_n \overset{a.s.}{\to} g$, we can get, respectively,

$$\exists\, E \in \mathcal{A},\ \mu(E^c) = 0,\ f_n(x) \to f(x) \text{ for all } x \in E$$

$$\exists\, F \in \mathcal{A},\ \mu(F^c) = 0,\ f_n(x) \to g(x) \text{ for all } x \in F$$

Then, let $H = E \cap F$. $\mu(H^c) = 0$ because $\mu(H^c) = \mu(E^c \cup F^c) \leqslant \mu(E^c) + \mu(F^c) = 0$. On H, $f_n \to f$ pointwise, and $f_n \to f$ pointwise. Therefore, $f(x) = g(x)$ on H, which means that $f = g$ a.e.. □

Example: On interval $[0, 1]$, we use Lebesgue measure space $([0, 1], \mathcal{L}, \lambda)$. Let $f_n(x) = x^n$. Then,

$$f_n(x) \to 0 \text{ for all } x \in [0, 1), \quad \text{but } f_n(1) = 1 \nrightarrow 0$$

Thus, $f_n \overset{a.s.}{\to} f$ but not pointwise.

5.2 Uniform Convergence a.e.

Definition 5.2.1: Uniform Convergence

Let $f, f_n : E \to \bar{\mathbb{R}}$ be measurable functions. Then, $f_n \to f$ **uniformly** if

$$\forall\, \epsilon > 0,\ \exists\, N,\ \text{s.t. } \forall\, n \geqslant N,\ \text{we have } \sup_{x \in E} |f_n(x) - f(x)| \leqslant \epsilon \text{ for all } x \in E$$

Definition 5.2.2: Uniform Convergence a.e.

Let $f, f_n : X \to \bar{\mathbb{R}}$ be measurable functions. Then, $f_n \to f$ **uniformly a.e.** if there exists $E \in \mathcal{A}$, such that $\mu(E^c) = 0$, and $f_n \to f$ uniformly on E.

Related with this definition is an important modification of what we called a supremum.

Definition 5.2.3: Essential Supremum

Let $f : X \to \bar{\mathbb{R}}$ be mearuable. The **essential supremum** of $|f|$, denoted by $\operatorname{ess\,sup}|f|$, is

$$\operatorname{ess\,sup}|f| = \inf\{a > 0 : \mu\{|f| > a\} = 0\}$$

Example 1: Suppose we are on the space $(\mathbb{R}, \mathcal{L}, \lambda)$. Let

$$f(x) = \begin{cases} 1, & x \neq 0, \\ +\infty, & \text{otherwise} \end{cases}$$

Then, $\sup|f| = +\infty$, but $\operatorname{ess\,sup}|f| = 1$.

Few properties related to ess sup is discussed below.

Property 5.2.4

If $\operatorname{ess\,sup}|f| = c$, then $\mu\{|f| > c\} = 0$.

Proof. Since $\operatorname{ess\,sup}|f| = c$, we have $\mu\{|f| > c + \frac{1}{n}\} = 0$. Then,

$$\mu\{|f| > c\} = \mu\left(\bigcup_{n\geqslant 1}\left\{|f| > c + \frac{1}{n}\right\}\right) \leqslant \sum_{n=1}^{\infty}\mu\left\{|f| > c + \frac{1}{n}\right\} = 0$$

□

Property 5.2.5: $\operatorname{ess\,sup}$ is a Class Property

If $f \sim g$, then $\operatorname{ess\,sup}|f| = \operatorname{ess\,sup}|g|$.

Proof. If $f \sim g$, then there exists $E \in \mathcal{A}$ such that $\mu(E^c) = 0$ and $f = g$ on E. Denote $c = \operatorname{ess\,sup}|f|$. Denote the set $A = \{|g| > c\}$. Then,

$$\mu(A) \leqslant \mu(A \cap E^c) + \mu(A \cap E) \leqslant \mu(E^c) + \mu(A \cap E) \leqslant 0 + \mu\{|f| > c\} = 0$$

This shows that $\operatorname{ess\,sup}|g| \leqslant c$. We can also go reversely to show that $\operatorname{ess\,sup}|g| \geqslant c$. □

Property 5.2.6: $\operatorname{ess\,sup}$ Defines a Metric

Define

$$d(f, g) = \operatorname{ess\,sup}|f - g|$$

Then, d is a distance on the space $\mathcal{M}$.

Proof. It is positive obviously. It is also symmetric by definition. Now, we need to prove that $d(f, g) = 0$ implies $f \sim g$. To do this, we need to prove that $\operatorname{ess\,sup}|h| = 0$ implies $h = 0$ a.e.. Note that if $\operatorname{ess\,sup}|h| = 0$, we have

$\mu(\{|h| > 1/n\}) = 0$. Then,

$$\mu\{|h| > 0\} = \mu\left(\bigcup_{n=1}^{\infty}\left\{|h| > \frac{1}{n}\right\}\right) \leqslant \sum_{n=1}^{\infty}\mu\{|h| > 1/n\} = 0$$

Finally, we need to prove the triangular inequality: $d(f,h) \leqslant d(f,g) + d(g,h)$. Let $d(f,g) = a$ and $d(g,h) = b$. Then, consider the set

$$\mu\{|f-h| > a+b\}$$

Since $|f-h| \leqslant |f-g| + |g-h|$, we need to have either $|f-g| > a$ or $|g-h| > b$ if we need to make $|f-h| > a+b$. Therefore,

$$\mu\{|f-h| > a+b\} = \mu(\{|f-g| > a\} \cup \{|g-h| > b\}) \leqslant \mu\{|f-g| > a\} + \mu\{|g-h| > b\} = 0$$

where the last equality holds by Property 5.2.4. This shows that $\operatorname{ess\,sup}|f-h| \leqslant a+b$. □

Property 5.2.7: Criterion of Uniform Convergence a.e.

$f_n \to f$ uniformly a.e *if and only if* $d(f_n, f) \to 0$.

Proof. ($\Longrightarrow$) Suppose $f_n \to f$ uniformly a.e. This means that we can find $E \in \mathcal{A}$, such that $\mu(E^c) = 0$, and have

$$\forall\,\epsilon > 0,\ \exists\,N,\ \text{s.t. } \forall\,n \geqslant N,\ \text{we have } \sup_{x\in E}|f_n(x) - f(x)| \leqslant \epsilon \text{ for all } x \in E$$

This means that

$$\mu\{|f_n - f| > \epsilon\} \leqslant \mu(E^c) = 0 \quad \Longrightarrow \quad d(f_n, f) \leqslant \epsilon$$

($\Longleftarrow$) Suppose $d(f_n, f) \to 0$. Then,

$$\forall\,\epsilon > 0,\ \exists\,N,\ \text{ s.t. } \forall\,n \geqslant N,\ \text{ we have } \operatorname{ess\,sup}|f_n - f| \leqslant \epsilon$$

In this condition, $\mu\{|f_n - f| > \epsilon\} = 0$. Therefore, we can find

$$k \geqslant 1,\ \exists\,n_k,\ \text{ s.t. } \forall\,n \geqslant n_k,\ \text{ we have } \mu\left\{|f_n - f| > \frac{1}{2^k}\right\} = 0$$

Using countable subadditivity, we know that

$$\mu\left(\underbrace{\bigcup_{k=1}^{\infty}\bigcup_{n=n_k}^{\infty}\left\{|f_n - f| > \frac{1}{2^k}\right\}}_{E^c}\right) \leqslant \sum_{k=1}^{\infty}\sum_{n=n_k}^{\infty}\mu\left\{|f_n - f| > \frac{1}{2^k}\right\} = 0$$

We will prove that $f_n \to f$ uniformly on the set $E = \bigcap_{k=1}^{\infty}\bigcap_{n=n_k}^{\infty}\{|f_n - f| \leqslant 1/2^k\}$. For all $\epsilon > 0$, we can find a k_0 such that $1/2^{k_0} \leqslant \epsilon$. Then, for all $n \geqslant n_{k_0}$, we have that, for $x \in E$, $|f_n(x) - f(x)| \leqslant 1/2^{k_0} \leqslant \epsilon$. Take supremum on both sides, we then prove the assertion. □

Example 2: On interval $[0,1]$, the previous example $f_n(x) = x^n$ does not converge uniformly a.e..

5.3 Almost Uniform Convergence

To fix the Example 2 condition above, we introduce the following definition.

Definition 5.3.1: Almost Uniform Convergence

Let $f_n, f : X \to \bar{\mathbb{R}}$ be measurable functions. Then, $f_n \to f$ **almost uniformly** if

$$\forall\, \epsilon > 0,\ \exists\, E_\epsilon \in \mathcal{A},\ \text{s.t. } \mu(E_\epsilon^c) \leqslant \epsilon,\ f_n \to f \text{ uniformly on } E_\epsilon$$

What is the relationship between these convergence modes? First, it is clear that uniform convergence a.e. implies both almost sure convergence and almost uniform convergence. Moreover, almost uniform convergence implies almost sure convergence.

Proposition 5.3.2: Almost Uniform implies Almost Sure

If $f_n \to f$ almost uniformly, then $f_n \to f$ a.e..

Proof. Suppose $f_n \to f$ almost uniformly. Then, there exists $k \geqslant 1$, and $E_k \in \mathcal{A}$, such that $\mu(E_k^c) \leqslant 1/k$ and $f_n \to f$ uniformly on E_k. Since uniform convergence implies pointwise convergence, $f_n \to f$ pointwise on E_k. Consider the set $E^c = \bigcap_{k=1}^\infty E_k^c$. Then,

$$\mu\left(\bigcap_{k=1}^\infty E_k^c\right) \leqslant \mu(E_{k_0}^c) \leqslant \frac{1}{k_0},\ \forall\, k_0 \quad \Longrightarrow \quad \mu\left(\bigcap_{k=1}^\infty E_k^c\right) = 0$$

Then, let $E = \bigcup_{k=1}^\infty E_k$. Then, for every $x \in E$, $x \in E_k$ for some k, thus $f_n(x) \to f(x)$ pointwise. This shows the almost sure convergence. □

The surprising fact is that, if $\mu(X) < \infty$, almost sure convergence implies almost uniform convergence, and this is called *Egorov's Theorem*.

5.3.1 Egorov's Theorem

Theorem 5.3.3: Egorov's Theorem

Let $\mu(X) < \infty$. If $f_n \to f$ a.s., then $f_n \to f$ almost uniformly.

Proof. Suppose $f_n \to f$ a.s.. Then,

$$\forall\, k \geqslant 1, \exists\, n \geqslant 1,\ \text{s.t. } \forall\, \ell \geqslant n,\ \text{s.t. } |f_\ell - f| \leqslant \frac{1}{k} \text{ on some set } E \text{ with } \mu(E^c) = 0$$

This implies that

$$\mu\left(\left[\bigcap_{k=1}^\infty \bigcup_{n=1}^\infty \bigcap_{\ell=n}^\infty \left\{x : |f_\ell(x) - f(x)| \leqslant \frac{1}{k}\right\}\right]^c\right) = \mu\left(\bigcup_{k=1}^\infty \bigcap_{n=1}^\infty \bigcup_{\ell=n}^\infty \left\{x : |f_\ell(x) - f(x)| > \frac{1}{k}\right\}\right) = 0$$

$$\implies \quad \mu\left(\bigcap_{n=1}^{\infty} \underbrace{\bigcup_{\ell=n}^{\infty}\left\{x : |f_\ell(x) - f(x)| > \frac{1}{k}\right\}}_{A_n}\right) = 0$$

Notice that $A_n \supseteq A_{n+1}$ is a decreasing set. Since $\mu(X) < \infty$, we can use continuity from above, to get that

$$\lim_{n\to\infty} \mu(A_n) = \lim_{n\to\infty} \mu\left(\bigcup_{\ell=n}^{\infty}\left\{x : |f_\ell(x) - f(x)| > \frac{1}{k}\right\}\right) = \mu\left(\bigcap_{n=1}^{\infty}\bigcup_{\ell=n}^{\infty}\left\{x : |f_\ell(x) - f(x)| > \frac{1}{k}\right\}\right) = 0$$

Fix $\epsilon > 0$. By the equation above, there exists $n_{\epsilon,k}$ such that for all $\ell > n_{\epsilon,k}$, we have

$$\mu\left(\bigcup_{\ell=n_{\epsilon,k}}^{\infty}\left\{x : |f_\ell(x) - f(x)| > \frac{1}{k}\right\}\right) \leqslant \frac{\epsilon}{2^k}$$

Summing over k, and by countable subadditivity, we have

$$\mu\left(\underbrace{\bigcup_{k=1}^{\infty}\bigcup_{\ell=n_{\epsilon,k}}^{\infty}\left\{x : |f_\ell(x) - f(x)| > \frac{1}{k}\right\}}_{E_\epsilon^c}\right) \leqslant \epsilon \quad \implies \quad E_\epsilon = \bigcap_{k\geqslant 1}\bigcap_{\ell\geqslant n_{\epsilon,k}}\left\{x : |f_\ell(x) - f(x)| \leqslant \frac{1}{k}\right\}$$

Now, fix $\delta > 0$. Then there exists some k such that $\frac{1}{k} \leqslant \delta$. Fix $\ell \geqslant n_{\epsilon,k}$ such that $|f_\ell - f| \leqslant \frac{1}{k} \leqslant \delta$. Since this is true for all $x \in E_\epsilon$ (by the definition of E_ϵ), we can take supremum on both sides, and get

$$\sup_{x\in E_\epsilon} |f_\ell(x) - f(x)| \leqslant \delta, \quad \forall\, \ell \geqslant n_{\epsilon,k}$$

which implies uniform convergence. $\square$

Below is a figure that explains the relations between convergence modes.

Chapter 6

Differentiation

PART II: Probability Theory

Chapter 7

Probability Spaces

PART III: Stochastic Processes

Chapter 8

Martingale