MATH 540 Notes

Jiantong Liu

Spring 2024

These notes were live-texed from a measure theory class (MATH 540) taught by Professor X. Li in Spring 2024 at University of Illinois. Any mistakes and inaccuracies would be my own. This course mainly follows Folland's *Real Analysis: Modern Techniques and Their Applications*.

Contents

1 Abstract Measure Theory **2**
- 1.1 Introduction . . . 2
- 1.2 Measures . . . 3
- 1.3 Outer Measure . . . 6
- 1.4 Borel Measure . . . 13

2 Integration **23**
- 2.1 Measurable Functions . . . 23
- 2.2 Integration of Non-negative Functions . . . 27
- 2.3 Integration of Complex-Valued Functions . . . 35
- 2.4 Modes of Convergences . . . 41
- 2.5 Product Measures . . . 47

3 Signed Measure and Differentiation **55**
- 3.1 Differentiation on Euclidean Space . . . 55
- 3.2 Functions of Bounded Variation . . . 59

4 L^p-spaces **63**
- 4.1 Basic Theory of L^p-spaces . . . 63
- 4.2 Distribution Functions and Weak L^p-space . . . 64
- 4.3 Some Useful Inequalities . . . 65
- 4.4 The Dual of L^p-space . . . 67

1 Abstract Measure Theory

1.1 Introduction

Definition 1.1. Let X be an (non-empty) underlying space we are working over. We denote $\mathcal{P}(X)$ to be the power set of X, i.e., the set of all subsets of X.

Example 1.2. Let $X = \{1, 2\}$, then $\mathcal{P}(X) = \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}$.

Remark 1.3. If X is a finite set of size n, then $\mathcal{P}(X)$ is a finite set of size 2^n.

We will consider a subcollection $\mathcal{A}$ of subsets of X, i.e., a subset of the power set. We will try to define this as an algebra. Note that an algebra is just a ring with a module structure with respect to some other ring.

Definition 1.4. $\mathcal{A} \subseteq \mathcal{P}(X)$ is an algebra on X if it is

a. closed under finite union, i.e., given $E_1, E_2 \in \mathcal{A}$, then $E_1 \cup E_2 \in \mathcal{A}$, and

b. closed under complements, i.e., if $E \in \mathcal{A}$, then the complement $E^c \in \mathcal{A}$ as well.

Remark 1.5. An algebra $\mathcal{A}$ would be closed under finite intersection. Indeed, for any $E_1, E_2 \in \mathcal{A}$, we have $E_1 \cap E_2 \in \mathcal{A}$ if and only if $(E_1 \cap E_2)^c \in \mathcal{A}$, if and only if $E_1^c \cup E_2^c \in \mathcal{A}$, which is true by definition.

Lemma 1.6. If $\mathcal{A}$ is an non-empty algebra on X, then $\varnothing \in \mathcal{A}$ and $X \in \mathcal{A}$.

Proof. Since $\mathcal{A}$ is non-empty, take $E \in \mathcal{A}$, then $\varnothing = E \cap E^c \in \mathcal{A}$ as well. Also, $X = E \cup E^c \in \mathcal{A}$. □

Example 1.7. Let X be a set, and let $\mathcal{A} = \{\varnothing, X\} \subseteq \mathcal{P}(X)$. It is easy to verify that $\mathcal{A}$ is an algebra.

Definition 1.8. Let $\varnothing \neq \mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, then we say $\mathcal{A}$ is a σ-algebra on X if

a. closed under countable union, i.e., if $E_j \in \mathcal{A}$ for all $j \in \mathbb{N}$, then $\bigcup\limits_{j=1}^{\infty} E_j \in \mathcal{A}$;

b. if $E \in \mathcal{A}$, then $E^c \in \mathcal{A}$.

Lemma 1.9. If $\mathcal{A} \neq \varnothing$ is a σ-algebra on X, then $\{\varnothing, X\} \subseteq \mathcal{A}$ is a σ-algebra.

Example 1.10. Let X be an uncountable set, let $\mathcal{A} = \{E \subseteq X : E \text{ is countable or } E^c \text{ is countable}\}$, then $\mathcal{A}$ is a σ-algebra on X.

Theorem 1.11. Suppose there is a non-empty algebra $\mathcal{A} \subseteq \mathcal{P}(X)$ such that, given pairwise disjoint subsets $E_j \in \mathcal{A}$ for all $j \in \mathbb{N}$, we have $\bigcup\limits_{j=1}^{\infty} E_j \in \mathcal{A}$, then $\mathcal{A}$ is a σ-algebra on X.

Proof. Take $E_j \in \mathcal{A}$ for all $j \in \mathbb{N}$, we will show that $\bigcup\limits_{j=1}^{\infty} E_j \in \mathcal{A}$. To do this, we will rearrange the sets. Let $F_1 = E_1$, let $F_2 = E_2 \backslash E_1$, let $F_3 = E_3 \backslash (E_1 \cup E_2)$, and so on, such that let $F_k = E_k \backslash \bigcup\limits_{i=1}^{k-1} E_i$. We note

$$\begin{aligned} F_k &= E_k \cap \left(\bigcup_{j=1}^{k-1} E_j \right)^c \\ &= E_k \cap \left(\bigcap_{j=1}^{k-1} E_j^c \right) \in \mathcal{A}. \end{aligned}$$

One can also verify that $\bigcup\limits_{j=1}^{\infty} E_j = \bigcup\limits_{k=1}^{\infty} F_k$, and that F_k's are disjoint from the definition. □

Definition 1.12. Let X be a non-empty space. A topology on X is a family $\mathcal{F}$ of subsets of X satisfying the following conditions:

i. $\varnothing, X \in \mathcal{F}$;

ii. $\mathcal{F}$ is closed under arbitrary union;

iii. $\mathcal{F}$ is closed under finite intersection.

Every member of $\mathcal{F}$ is now called an open subset of X. A complement of an open subset of X is called a closed subset.

Definition 1.13. Let $\mathcal{A}_1, \mathcal{A}_2$ be σ-algebras. We say $\mathcal{A}_1$ is smaller than $\mathcal{A}_2$ if $\mathcal{A}_1 \subseteq \mathcal{A}_2$, and equivalently $\mathcal{A}_2$ is larger than $\mathcal{A}_1$.

Definition 1.14. Let $\mathcal{F}$ be a family of subsets of X, the smallest σ-algebra containing $\mathcal{F}$ is called the σ-algebra generated by $\mathcal{F}$. This is denoted by $\mathcal{M}(\mathcal{F})$.

Lemma 1.15. Let $\mathcal{F}$ be a family of subsets of X. Suppose $\mathcal{F} \subseteq \mathcal{A}$ where $\mathcal{A}$ is a σ-algebra, then $\mathcal{M}(\mathcal{F}) \subseteq \mathcal{A}$.

Proof. Obvious. □

Definition 1.16. Let $\mathcal{F}$ be a topology on X, then we say $(X, \mathcal{F})$ is a topological space. We say $\mathcal{M}(\mathcal{F})$ is the Borel σ-algebra on X, denoted by $\mathcal{B}_X = \mathcal{B}_{X,\mathcal{F}}$. Any member of $\mathcal{B}_X$ is called a Borel set.

Example 1.17. Let $X = \mathbb{R}$, we denote the corresponding Borel σ-algebra to be $\mathcal{B}_{\mathbb{R}}$.

Definition 1.18. A G_δ-set is a countable intersection of open subsets of X. A F_σ-set is a countable union of closed subsets of X.

Theorem 1.19. Both G_δ-sets and F_σ-sets are Borel sets, that is, $G_\delta, F_\sigma \subseteq \mathcal{B}_X$.

Proof. We will prove that any G_δ-set E is a Borel set, and similarly any F_σ-set is a Borel set. By definition $E = \bigcap_{j=1}^{\infty} O_j$, where each O_j is an open subset. To show $E \in \mathcal{B}_X$, we show that $E^c \in \mathcal{B}_X$. Note that $E^c = \left(\bigcap_{j=1}^{\infty} O_j\right)^c = \bigcup_{j=1}^{\infty} O_j^c$. Since $O_j \in \mathcal{B}_X$ for all j, then $O_j^c \in \mathcal{B}_X$ as well. Therefore, $E^c \in \mathcal{B}_X$ since a σ-algebra $\mathcal{B}_X$ is closed under countable unions. □

Definition 1.20. Let $X_1, \ldots, X_n$ be non-empty spaces. The product space is $\prod_{j=1}^{n} X_j$. Define $\pi_j : \prod_{i=1}^{n} X_i \to X_j$ by $\pi_j(x_1, \ldots, x_n) = x_j$. Let $\mathcal{A}_j$ be a σ-algebra on X_j, the product σ-algebra on $\prod_{i=1}^{n} X_j$ is the σ-algebra generated by $\{\pi_j^{-1}(E_j) : E_j \in \mathcal{A}_j \ \forall j \in \{1, \ldots, n\}\}$. The product σ-algebra is denoted by $\bigotimes_{j=1}^{n} \mathcal{A}_j = \prod_{j=1}^{n} \mathcal{A}_j$.

Example 1.21. $\mathcal{B}_{\mathbb{R}^n} = \bigotimes_{j=1}^{n} \mathcal{B}_{\mathbb{R}}$.

1.2 Measures

Definition 1.22. Let $\mathcal{A}$ be a σ-algebra on X. A measure μ on X and $\mathcal{A}$ is a function $\mu : \mathcal{A} \to [0, \infty]$ such that

a. $\mu(\varnothing) = 0$;

b. if $E_j \in \mathcal{A}$ for all $j \in \mathbb{N}$ and E_j's are disjoint, then $\mu\left(\bigcup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} \mu(E_j)$.

We then say $(X, \mathcal{A})$ is a measureable space. A measureable space is a triple $(X, \mathcal{A}, \mu)$ with measure μ specified.

Definition 1.23. Let μ be a measure on $(X, \mathcal{A})$.

1. If $\mu(X) < \infty$, then we say μ is a finite measure. In particular, if $\mu(X) = 1$, this is a probability measure.

2. If $X = \bigcup_{j=1}^{\infty} E_j$ such that $\mu(E_j) < \infty$ for all $j \in \mathbb{N}$, then we say μ is σ-finite.

3. If for all $E \in \mathcal{A}$ with $\mu(E) = \infty$, there is $F \in \mathcal{A}$ such that $F \subseteq E$ and $0 < \mu(F) < \infty$, then we say μ is semi-finite.

Remark 1.24. A σ-finite measure is semi-finite. However, the converse is not true.

Example 1.25. Let $f : X \to [0, \infty]$ be a function. For any $E \subseteq \mathcal{P}(E)$, we can define a measure $\mu(E) = \sum_{x \in E} f(x)$. Note that the summation makes sense only when E is finite. In case E is infinite, we should define $\sum_{x \in E} f(x) = \sup\{\sum_{x \in F} f(x) : F \subseteq E \text{ for finite } F\}$. Let μ be a measure on $\mathcal{P}(X)$.

- If $f(x) \equiv 1$ for all $x \in X$, then $\mu(E) = \sum_{x \in E} 1 = \mathrm{Card}(E)$. In this case, μ is called a counting measure.
- Suppose $x_0 \in X$ is fixed. Define
$$f(x) = \begin{cases} 1, & \text{if } x = x_0 \\ 0, & \text{if } x \neq x_0 \end{cases}$$
then for any $E \in \mathcal{P}(X)$,
$$\mu(E) = \begin{cases} 1, & \text{if } x_0 \in E \\ 0, & \text{if } x_0 \notin E \end{cases}$$
This is called the Dirac measure of x_0.

Definition 1.26. Let $(X, \mathcal{A}, \mu)$ be a measure space. A set $E \subseteq \mathcal{A}$ is called a null set if $\mu(E) = 0$.

If a statement about points $x \in X$ is true except for null sets, then we say the statement is true almost everywhere.

Example 1.27. Suppose $f(x) \leqslant 1$ for all $x \in X$, then we say f is bounded above by 1 everywhere. If we want to weaken this statement, we can say $f(x) \leqslant 1$ almost everywhere $x \in X$, which is true if and only if $\mu(\{x \in X : f(x) > 1\} = 0$.

Theorem 1.28. Let $E, F \in \mathcal{A}$ be such that $E \subseteq F$, then $\mu(E) \leqslant \mu(F)$.

Proof. We can write $F = E \cup (E \backslash F)$, then
$$\begin{aligned} \mu(F) &= \mu(E) + \mu(F \backslash E) \\ &\geqslant \mu(E) \end{aligned}$$
since $\mu(F \backslash E) \geqslant 0$. □

Theorem 1.29 (Sub-additivity). Let $E_j \in \mathcal{A}$ for all $j \in \mathbb{N}$, then $\mu\left(\bigcup_{j=1}^{\infty} E_j\right) \leqslant \sum_{j=1}^{\infty} \mu(E_j)$.

Proof. Set $F_1 = E_1$ and let $F_k = E_k \backslash \left(\bigcup_{j=1}^{k-1} E_j\right)$ be defined inductively, then $\bigcup_{k \in \mathbb{N}} F_k = \bigcup_{j \in \mathbb{N}} E_j$. Since F_k's are disjoint, we have
$$\begin{aligned} \mu\left(\bigcup_{j \in \mathbb{N}} E_j\right) &= \mu\left(\bigcup_{k \in \mathbb{N}} F_k\right) \\ &= \sum_{k=1}^{\infty} \mu(F_k) \\ &= \sum_{k=1}^{\infty} \mu(E_k) \\ &= \sum_{j=1}^{\infty} \mu(E_j) \end{aligned}$$
by Theorem 1.28. □

Theorem 1.30. Let $E_j \in \mathcal{A}$ for all $j \in \mathbb{N}$.

a. (Continuity from below): If $E_1 \subseteq E_2 \subseteq \cdots E_j \subseteq \cdots$ for all j, then $\mu\left(\bigcup\limits_{j=1}^{\infty} E_j\right) = \lim\limits_{j\to\infty} \mu(E_j)$.

b. (Continuity from above): If $E_1 \supseteq E_2 \supseteq \cdots E_j \supseteq \cdots$ for all $j \in \mathbb{N}$, then $\mu\left(\bigcap\limits_{j=1}^{\infty} E_j\right) = \lim\limits_{j\to\infty} \mu(E_j)$ if $\mu(E_1) < \infty$.

In particular, the limits on the right exist on $\bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$.

Example 1.31. Let μ be the counting measure on $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. For each $j \in \mathbb{N}$, we define $E_j = \{n \in \mathbb{N} : n > j\}$. Therefore $E_1 \supseteq E_2 \supseteq \cdots$ is a decreasing sequence of sets. Note that $\mu(E_1) = \mu(\{n \in \mathbb{N}\}) = \mathbb{N} = \infty$, and $\lim\limits_{j\to\infty} \mu(E_j) = \lim\limits_{j\to\infty} \infty = \infty$, but $\mu\left(\bigcap\limits_{j=1}^{\infty} E_j\right) = \mu(\varnothing) = 0$.

Proof.

a. Set $E_0 = \varnothing$. Now

$$\bigcup_{j=1}^{\infty} E_j = \bigcup_{j=1}^{\infty} (E_j \backslash E_{j-1})$$

and therefore

$$\begin{aligned}
\mu\left(\bigcup_{j=1}^{\infty} E_j\right) &= \mu\left(\bigcup_{j=1}^{\infty} (E_j \backslash E_{j-1})\right) \\
&= \sum_{j=1}^{\infty} \mu(E_j \backslash E_{j-1}) \\
&= \lim_{k\to\infty} \sum_{j=1}^{k} \mu(E_j \backslash E_{j-1}) \\
&= \lim_{k\to\infty} \mu\left(\bigcup_{j=1}^{k} E_j \backslash E_{j-1}\right) \\
&= \lim_{k\to\infty} \mu(E_k) \\
&= \lim_{j\to\infty} \mu(E_j).
\end{aligned}$$

b. For any $j \in \mathbb{N}$, set $F_j = E_1 \backslash E_j$. Note that $F_j \subseteq F_{j+1}$ since $E_j \supseteq E_{j-1}$. This is now an increasing sequence as in part a. By part a., we know $\mu\left(\bigcup\limits_{j=1}^{\infty} F_j\right) = \lim\limits_{j\to\infty} \mu(F_j)$. Now note that

$$\begin{aligned}
\bigcup_{j=1}^{\infty} F_j &= \bigcup_{j=1}^{\infty} (E_1 \backslash E_j) \\
&= \bigcup_{j=1}^{\infty} (E_1 \cap E_j^c) \\
&= E_1 \cap \bigcup_{j=1}^{\infty} E_j^c \\
&= E_1 \cap \left(\bigcap_{j=1}^{\infty} E_j\right)^c
\end{aligned}$$

$$
\begin{aligned}
&= \left(\left(E_1 \cap \left(\bigcap_{j=1}^{\infty} E_j\right)^c\right) \cup \left(\bigcap_{j=1}^{\infty} E_j\right)\right) \cap \left(\bigcap_{j=1}^{\infty} E_j\right)^c \\
&= \left(E_1 \cap \left(\bigcap_{j=1}^{\infty} E_j\right)^c\right) \cup \left(\bigcap_{j=1}^{\infty} E_j\right).
\end{aligned}
$$

Note that $E_1 \cap \left(\bigcap_{j=1}^{\infty} E_j\right)^c$ and $\bigcap_{j=1}^{\infty} E_j$ are disjoint, therefore by property of measure we have

$$
\begin{aligned}
\mu(E_1) &= \mu\left(E_1 \cap \left(\bigcap_{j=1}^{\infty} E_j\right)^c\right) + \mu\left(\bigcap_{j=1}^{\infty} E_j\right) \\
&= \mu\left(\bigcup_{j=1}^{\infty} F_j\right) + \mu\left(\bigcap_{j=1}^{\infty} E_j\right) \\
&= \lim_{j\to\infty} \mu(F_j) + \mu\left(\bigcap_{j=1}^{\infty} E_j\right).
\end{aligned}
$$

Recall that $F_j = E_1 \backslash E_j$ for all j, therefore $E_1 = F_j \cup F_j^c = F_j \cup E_j$, where F_j and E_j are disjoint, therefore $\mu(E_1) = \mu(F_j) + \mu(E_j)$. Since $\mu(E_1) < \infty$, and F_j is a subset of E_1 and hence also a real number, then $\mu(E_1)$ is a sum of two real numbers. Therefore, we have $\mu(E_1) - \mu(E_j) = \mu(F_j)$. With this, we have

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

In particular, we get

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

□

1.3 Outer Measure

Definition 1.32. An outer measure μ^* on X (or $\mathcal{P}(X)$) is a function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ such that

i. $\mu^*(\varnothing) = 0$,

ii. $\mu^*(A) \leqslant \mu^*(B)$ for all $A \subseteq B \subseteq X$,

iii. σ-subaddivity: $\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) \leqslant \sum_{j=1}^{\infty} \mu^*(A_j)$.

Example 1.33. Let $\rho : \mathcal{A} \to [0, \infty]$ be such that $\rho(\varnothing) = 0$, where $\mathcal{A} \subseteq \mathcal{P}(X)$ is a subcollection (but not necessarily an algebra) such that $\varnothing, X \in \mathcal{A}$.

For all $A \in \mathcal{P}(X)$, i.e., $A \subseteq X$, we define

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

Theorem 1.34. μ^* defined in Example 1.33 is an outer measure.

Proof.

i. Let $E_j = \varnothing$ for all $j \in \mathbb{N}$, then $\varnothing \subseteq \bigcup_{j=1}^{\infty} E_j$, and so

$$\sum_{j=1}^{\infty} \rho(E_j) = \sum_{j=1}^{\infty} \rho(\varnothing) = \sum_{j=1}^{\infty} 0 = 0$$

and therefore $\mu^*(\varnothing) = 0$.

ii. Let $A \subseteq B \subseteq X$. If $B \subseteq \bigcup_{j=1}^{\infty} E_j$, we have $A \subseteq \bigcup_{j=1}^{\infty} E_j$, then

$$\left\{\sum_{j=1}^{\infty} \rho(E_j) : E_j \in \mathcal{A}, B \subseteq \bigcup_{j=1}^{\infty} E_j\right\} \subseteq \left\{\sum_{j=1}^{\infty} \rho(E_j) : E_j \in \mathcal{A}, A \subseteq \bigcup_{j=1}^{\infty} E_j\right\}.$$

In particular, given subsets $S_1 \subseteq S_2$, then $\inf S_2 \leqslant \inf S_1$ and $\sup S_1 \leqslant \sup S_2$. This implies $\mu^*(A) \leqslant \mu^*(B)$.

iii. We want to show $\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) \leqslant \sum_{j=1}^{\infty} \mu^*(A_j)$. Now for any $j \in \mathbb{N}$, we have

$$\mu^*(A_j) = \inf\left\{\sum_{k=1}^{\infty} \rho(E_k) : E_k \in \mathcal{A} \text{ and } A_j \subseteq \bigcup_{k=1}^{\infty} E_k\right\}.$$

For any $\varepsilon > 0$, we note that $\mu^*(A_j) + \varepsilon \cdot 2^{-j}$ is not a lower bound of $\left\{\sum_{k=1}^{\infty} \rho(E_k) : E_k \in \mathcal{A} \text{ and } A_j \subseteq \bigcup_{k=1}^{\infty} E_k\right\}$. Then there exists $E_k^{(j)} \in \mathcal{A}$ for $k \in \mathbb{N}$ such that $A_j \subseteq \bigcup_{k=1}^{\infty} E_k^{(j)}$ and $\sum_{k=1}^{\infty} \rho(E_k^{(j)}) \leqslant \mu^*(A_j) + \varepsilon \cdot 2^{-j}$. Summing with respec to j, we get

$$\begin{aligned}\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \rho(E_k^{(j)}) &\leqslant \sum_{j=1}^{\infty} \mu^*(A_j) + \sum_{j=1}^{\infty} \varepsilon \cdot 2^{-j} \\ &= \sum_{j=1}^{\infty} \mu^*(A_j) + \varepsilon.\end{aligned}$$

Note that

$$\bigcup_{j=1}^{\infty} A_j \subseteq \bigcup_{j=1}^{\infty}\bigcup_{k=1}^{\infty} E_k^{(j)}$$

is a countable union of subsets of $\mathcal{A}$. We will calculate the value over μ^*. By definition of μ^*, we have

$$\begin{aligned}\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) &\leqslant \sum_{j=1}^{\infty}\sum_{k=1}^{\infty} \rho(E_k^{(j)}) \\ &\leqslant \sum_{j=1}^{\infty} \mu^*(A_j) + \varepsilon.\end{aligned}$$

Since this is true for all $\varepsilon > 0$, then take $\varepsilon \to 0$, we are done.

□

Definition 1.35. Let μ^* be an outer measure on $(X, \mathcal{P}(X))$. A set $A \subseteq X$ is called μ^*-measurable if $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$ for all $E \subseteq X$.

Remark 1.36. First note that $\mu^*(E) = \mu^*((E \cap A) \cup (E \cap A^c))$, therefore $\mu^*(E) \leqslant \mu^*(E \cap A) + \mu^*(E \cap A^c)$ for all $E \subseteq X$.

Theorem 1.37 (Fundamental Theorem of Measure Theory)**.** Let μ^* be an outer measure on X. Let $\mathcal{A}$ be the collection of all μ^*-measurable set, then $\mathcal{A}$ is a σ-algebra, and $\mu^*|_{\mathcal{A}}$ is a measure on $\mathcal{A}$, i.e., $(X, \mathcal{A}, \mu^*)$ is a measure space.

Proof. We first prove that $\mathcal{A}$ is an algebra. To see $\mathcal{A}$ is closed under complement, we have $A \in \mathcal{A}$ if and only if $A^c \in \mathcal{A}$. by the definition of measurable set. To show $\mathcal{A}$ is closed under finite union, suppose $A, B \in \mathcal{A}$, and we want to show $A \cup B \in \mathcal{A}$, which is true if and only if $\mu^*(E) = \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$ for all $E \subseteq X$, hence it suffices to show that $\mu^*(E) \geqslant \mu^*(E \cap (A \cup B)) + \mu^*(E \cap (A \cup B)^c)$. We have

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

and

$$\begin{aligned}\mu^*(E \cap (A \cup B)^c) &= \mu^*(E \cap (A \cup B)^c \cap A) + \mu^*(E \cap (A \cup B)^c \cap A^c) \\ &= \mu^*(\varnothing) + \mu^*(E \cap A^c \cap B^c) \\ &= \mu^*(E \cap A^c \cap B^c).\end{aligned}$$

Therefore

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

where the last two steps follow from the fact that $A, B \in \mathcal{A}$ are μ^*-measurable. Therefore, $\mathcal{A}$ is an algebra. We now want to show that it is a σ-algebra. It suffices to prove that $\mathcal{A}$ is closed under disjoint σ-unions. Let $A_j \in \mathcal{A}$ for all $j \in \mathbb{N}$ where they are pairwise disjoint, and we want to show that $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$. That is,

$$\mu^*(E) = \mu^*\left(E \cap \left(\bigcup_{j=1}^{\infty} A_j\right)\right) + \mu^*\left(E \cap \left(\bigcup_{j=1}^{\infty} A_j\right)^c\right)$$

for all $E \subseteq X$.

Lemma 1.38. For a pairwise disjoint family $A_1, \ldots, A_n \in \mathcal{A}$,

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

Subproof. We proceed by induction. For $n = 1$, this is obviously true. Now suppose $n > 1$. To simplify the notation, let $B_n = \bigcup_{j=1}^{n} A_j$, and use the convention that $B_0 = \varnothing$. Now

$$\begin{aligned}\mu^*(E \cap B_n) &= \mu^*(E \cap B_n \cap A_n) + \mu^*(E \cap B_n \cap A_n^c) \\ &= \mu^*(E \cap A_n) + \mu^*(E \cap B_{n-1}) \\ &= \sum_{i=1}^{n}(E \cap A_i) + \mu^*(E \cap B_0) \\ &= \sum_{i=1}^{n}(E \cap A_i) \\ &= \sum_{i=1}^{n}(E \cap A_i)\end{aligned}$$

for all $n \in \mathbb{N}$. This finishes the proof. ■

Now for any $E \subseteq X$, we have

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

since $B_n = \bigcup_{j=1}^{n} A_j \subseteq \bigcup_{j=1}^{\infty} A_j$. Now take $n \to \infty$, we get

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

This forces all inequalities here to be equality, therefore

$$\mu^*(E) = \mu^*\left(E \cap \left(\bigcup_{j=1}^{\infty} A_j\right)\right) + \mu^*\left(E \cap \left(\bigcup_{j=1}^{\infty} A_j\right)^c\right)$$

as desired. Finally, we need to show that the restriction still gives a measure. We already know

$$\mu^*(E) = \sum_{j=1}^{\infty} \mu^*(E \cap A_j) + \mu^*\left(\left(\bigcup_{j=1}^{\infty} A_j\right)^c\right)$$

for any $E \subseteq X$, then in particular take $E = \bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$ to be the disjoint union, then this forces

$$\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu^*(E \cap A_j) + \mu^*(\varnothing) = \sum_{j=1}^{\infty} \mu^*(E \cap A_j).$$

Therefore $\mu^*|_{\mathcal{A}}$ is a measure. ☐

Definition 1.39. A measure μ is said to be complete if its domain contains all subsets of null sets.

Example 1.40. Let $X = \{a, b\}$, $\mathcal{A} = \{\varnothing, \{a, b\}\}$. Define $\mu : \mathcal{A} \to [0, \infty]$ by setting $\mu^*(X) = 0$, $\mu^*(\varnothing) = 0$. This is not a complete measure because $\{a\} \notin \mathcal{A}$.

Theorem 1.41. Let $\mathcal{A}$ be the collection of all μ^*-measurable sets, then the measure $\mu^*|_{\mathcal{A}}$ is complete.

Proof. Let N be any null set in $\mathcal{A}$, i.e., $\mu^*(N) = 0$. Take an arbitrary subset $A \subseteq N$, we need to show $A \in \mathcal{A}$. Since $\mu^*(N) = 0$, then $\mu^*(A) = 0$ as well. For any $E \subseteq X$, we prove $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$. It is clear that

$$\begin{aligned}\mu^*(E) &\leqslant \mu^*(E \cap A) + \mu^*(E \cap A^c)\\ &\leqslant \mu^*(A) + \mu^*(E \cap A^c)\\ &\leqslant \mu^*(N) + \mu^*(E \cap A^c)\\ &= \mu^*(E \cap A^c)\\ &= \mu^*(E).\end{aligned}$$

by the subadditivity of μ^*. ☐

Definition 1.42. Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra. A function $\mu_0 : \mathcal{A} \to [0, \infty]$ is a pre-measure if

i. $\mu_0(\varnothing) = 0$,

ii. if $A_j \in \mathcal{A}$ for all $j \in \mathbb{N}$ with $\bigcup\limits_{j=1}^{\infty} A_j \in \mathcal{A}$, and they are pairwise disjoint, then $\mu_0\left(\bigcup\limits_{j=1}^{\infty} A_j\right) = \sum\limits_{j=1}^{\infty} \mu_0(A_j)$.

Therefore, the difference of a pre-measure from a measure is that a pre-measure is not defined on a σ-algebra.

Theorem 1.43. Let μ_0 be a pre-measure, then $\mu_0(A) \leqslant \mu_0(B)$ if $A, B \in \mathcal{A}$ are such that $A \subseteq B$.

Proof. We write $B = (B \backslash A) \cup A$, where $B \backslash A = B \cap A^c \in \mathcal{A}$, therefore

$$\begin{aligned} \mu_0(B) &= \mu_0(B \backslash A) + \mu_0(A) \\ &\geqslant \mu_0(A). \end{aligned}$$

□

Definition 1.44. Given a pre-measure μ_0, we extend it to an outer measure as follows: for any $E \subseteq X$, define $\mu^*(E) = \inf\{\sum\limits_{j=1}^{\infty} \mu_0(A_j) : E \subseteq \bigcup\limits_{j=1}^{\infty} A_j, A_j \in \mathcal{A}\}$.

Theorem 1.45 (Carathéodory's Extension Theorem)**.** Let μ^* be the outer measure induced by μ_0 specified in Definition 1.44, then

i. $\mu^*|_{\mathcal{A}} = \mu_0$, or equivalently, for any $A \in \mathcal{A}$, we have $\mu^*(A) = \mu_0(A)$;

ii. if $A \in \mathcal{A}$, then A is μ^*-measurable.

Proof.

i. We want to show that for any $E \in \mathcal{A}$, $\mu^*(E) = \mu_0(E)$. To show $\mu^*(E) \leqslant \mu_0(E)$, we choose $A_1 = E \in \mathcal{A}$, and $A_j = \varnothing$ for all $j \geqslant 2$, then $E \subseteq \bigcup\limits_{j=1}^{\infty} A_j$, therefore

$$\begin{aligned} \mu^*(E) &\leqslant \sum_{j=1}^{\infty} \mu_0(A_j) \\ &= \mu_0(E). \end{aligned}$$

It now suffices to show that $\mu_0(E)$ is a lower bound of $\{\sum\limits_{j=1}^{\infty} \mu_0(A_j) : E \subseteq \bigcup\limits_{j=1}^{\infty}, A_j \in \mathcal{A}\}$. Let $A_j \in \mathcal{A}$ and $\bigcup\limits_{j=1}^{\infty} A_j \supseteq E$. We prove that $\mu_0(E) \leqslant \sum\limits_{j=1}^{\infty} \mu_0(A_j)$. For any $n \in \mathbb{N}$, define $B_n = E \cap \left(A_n \backslash \bigcup\limits_{j=1}^{n-1} A_j\right)$, therefore $\bigcup\limits_{n=1}^{\infty} B_n = E \cap \left(\bigcup\limits_{j=1}^{\infty} A_j\right) = E$ where B_n's are disjoint. We have

$$\begin{aligned} \mu_0(E) &= \mu_0\left(\bigcup_{n=1}^{\infty} B_n\right) \\ &= \sum_{n=1}^{\infty} \mu_0(B_n) \\ &\leqslant \sum_{n=1}^{\infty} \mu_0(A_n) \\ &= \sum_{j=1}^{\infty} \mu_0(A_j). \end{aligned}$$

ii. For any $A \in \mathcal{A}$, we want to prove that $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$ for all $E \subseteq X$. It suffices to show that for any $E \subseteq X$, we have $\mu^*(E) \geqslant \mu^*(E \cap A) + \mu^*(E \cap A^c)$.

Pick arbitrary $\varepsilon > 0$, then $\mu^*(E) + \varepsilon$ is not a lower bound of $\{\sum_{j=1}^{\infty} \mu_0(A_j) : E \subseteq \bigcup_{j=1}^{\infty}, A_j \in \mathcal{A}\}$. Therefore, there exists some $A_j \in \mathcal{A}$ such that $E \subseteq \bigcup_{j=1}^{\infty} A_j$ and $\sum_{j=1}^{\infty} \mu_0(A_j) \leqslant \mu^*(E) + \varepsilon$. Since $\mu_0(A_j) = \mu_0(A_j \cap A) + \mu_0(A_j \cap A^c)$, then

$$\begin{aligned}
\sum_{j=1}^{\infty} \mu_0(A_j) &= \sum_{j=1}^{\infty} \mu_0(A_j \cap A) + \sum_{j=1}^{\infty} \mu_0(A_j \cap A^c) \\
&= \sum_{j=1}^{\infty} \mu^*(A_j \cap A) + \sum_{j=1}^{\infty} \mu^*(A_j \cap A^c) \\
&\geqslant \mu^*\left(\left(\bigcup_{j=1}^{\infty} A_j\right) \cap A\right) + \mu^*\left(\left(\bigcup_{j=1}^{\infty} A_j\right) \cap A^c\right) \\
&\geqslant \mu^*(E \cap A) + \mu^*(E \cap A^c).
\end{aligned}$$

Let $\varepsilon \to 0$, then $\mu^*(E) \geqslant \mu^*(E \cap A) + \mu^*(E \cap A^c)$, as desired.

□

Theorem 1.46. Let $\mathcal{A} \subseteq \mathcal{P}(X)$ be an algebra, and let μ_0 be a pre-measure on $\mathcal{A}$. Define $\mathcal{M}(\mathcal{A})$ to be the σ-algebra generated by $\mathcal{A}$.

a. The outer measure μ^* induced by μ_0 defines a measure function on $\mathcal{M}(\mathcal{A})$, and $\mu^*|_{\mathcal{A}} = \mu_0$.

b. If $\tilde{\mu}$ is another measure on $\mathcal{M}(\mathcal{A})$ that extends μ_0, then $\tilde{\mu}(E) \leqslant \mu^*(E)$ for all $E \subseteq \mathcal{M}(\mathcal{A})$, with equality if and only if $\mu^*(E) < \infty$.

c. If μ_0 is σ-finite, i.e., $X = \bigcup_{j=1}^{\infty} A_j$ with $A_j \in \mathcal{A}$ and $\mu_0(A_j) < \infty$ for all j, then $\mu^*|_{\mathcal{M}(\mathcal{A})}$ is the unique extension of μ_0 to a measure on $\mathcal{M}(\mathcal{A})$.

Proof.

a. Let $\mathcal{B}$ be the set of all μ^*-measurable sets, then $\mu^*|_{\mathcal{B}}$ is a measure on $\mathcal{B}$ that extends μ_0. By the fundamental theorem of measure theory, we know $\mathcal{B}$ is a σ-algebra. In particular, $\mathcal{B} \supseteq \mathcal{A}$, therefore $\mathcal{B} \supseteq \mathcal{M}(\mathcal{A})$. That means $\mu^*|_{\mathcal{M}(\mathcal{A})}$ is a measure as well.

b. Let $\tilde{\mu}$ be any measure on $\mathcal{M}(\mathcal{A})$ that extends μ_0. We first show that for all $E \in \mathcal{M}(\mathcal{A})$, then $\tilde{\mu}(E) \leqslant \mu^*(E)$. Recall that $\mu^*(E) = \inf\{\sum_{j=1}^{\infty} \mu_0(A_j) : E \subseteq \bigcup_{j=1}^{\infty} A_j, A_j \in \mathcal{A}\}$. Given a cover $E \subseteq \bigcup_{j=1}^{\infty} A_j$ and fix $A_j \in \mathcal{A}$. Therefore,

$$\begin{aligned}
\tilde{\mu}(E) &\leqslant \tilde{\mu}\left(\bigcup_{j=1}^{\infty} A_j\right) \\
&\leqslant \sum_{j=1}^{\infty} \tilde{\mu}(A_j) \\
&= \sum_{j=1}^{\infty} \mu_0(A_j),
\end{aligned}$$

therefore $\tilde{\mu}(E) \leqslant \mu^*(E)$. Assume we have $\mu^*(E) < \infty$, and we want to show that $\tilde{\mu}(E) = \mu^*(E)$. It suffices to show $\mu^*(E) \leqslant \tilde{\mu}(E)$.

Claim 1.47. Let $A_j \in \mathcal{A}$ for all $j \in \mathbb{N}$, then $\mu^*\left(\bigcup\limits_{j=1}^{\infty} A_j\right) = \tilde{\mu}\left(\bigcup\limits_{j=1}^{\infty} A_j\right)$.

Subproof. Note that $\bigcup\limits_{j=1}^{\infty} A_j \in \mathcal{M}(\mathcal{A})$, then we can just work on $\mathcal{M}(\mathcal{A})$. Consider $\mu^*|_{\mathcal{M}(\mathcal{A})}$ and $\tilde{\mu}$ are measures on $\mathcal{M}(\mathcal{A})$. Let $E_n = \bigcup\limits_{j=1}^{n} A_j$ for all $n \in \mathbb{N}$, then we have a nested increasing sequence of E_n's. In particular, we know $\bigcup\limits_{n=1}^{\infty} E_n = \bigcup\limits_{j=1}^{\infty} A_j$. Therefore

$$\begin{aligned}
\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) &= \mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) \\
&= \lim_{n\to\infty} \mu^*(E_n) \\
&= \lim_{n\to\infty} \mu^*\left(\bigcup_{j=1}^{n} A_j\right) \\
&= \lim_{n\to\infty} \mu_0\left(\bigcup_{j=1}^{n} A_j\right) \\
&= \lim_{n\to\infty} \tilde{\mu}\left(\bigcup_{j=1}^{n} A_j\right) \\
&= \tilde{\mu}\left(\bigcup_{j=1}^{\infty} A_j\right)
\end{aligned}$$

by continuity from below and closure of finite union. ■

We know from the claim that

$$\begin{aligned}
\mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) &= \lim_{n\to\infty} \mu_0\left(\bigcup_{j=1}^{n} A_j\right) \\
&\leqslant \lim_{n\to\infty} \sum_{j=1}^{n} \mu_0(A_j) \\
&= \sum_{j=1}^{\infty} \mu_0(A_j).
\end{aligned}$$

Take arbitrary $\varepsilon > 0$, then consider $\mu^*(E) + \varepsilon$, which is not a lower bound of the set anymore. Therefore, there exists $A_j \in \mathcal{A}$ for each $j \in \mathbb{N}$ such that $E \subseteq \bigcup\limits_{j=1}^{\infty} A_j$ and that $\sum\limits_{j=1}^{\infty} \mu_0(A_j) \leqslant \mu^*(E) + \varepsilon$. In particular, this means $\mu^*\left(\bigcup\limits_{j=1}^{\infty} A_j\right) \leqslant \mu^*(E) + \varepsilon$. Since $\mu^*(E) < \infty$, then

$$\begin{aligned}
\mu^*\left(\bigcup_{j=1}^{\infty} A_j \backslash E\right) &= \mu^*\left(\bigcup_{j=1}^{\infty} A_j\right) - \mu^*(E) \\
&< \varepsilon.
\end{aligned}$$

Now that

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

$$
\begin{aligned}
&= \tilde{\mu}\left(\bigcup_{j=1}^{\infty} A_j\right) \\
&= \tilde{\mu}(E) + \tilde{\mu}\left(\bigcup_{j=1}^{\infty} A_j \backslash E\right) \\
&< \tilde{\mu}(E) + \varepsilon
\end{aligned}
$$

by the claim. Therefore, for any $\varepsilon > 0$, we have $\mu^*(E) \leqslant \tilde{\mu}(E) + \varepsilon$ whenever $\mu^*(E) < \infty$. Take $\varepsilon \to 0$, we get $\mu^*(E) \leqslant \tilde{\mu}(E)$.

c. Since μ_0 is σ-finite, then there exists a decomposition $X = \bigcup\limits_{j=1}^{\infty} A_j$ for $A_j \in \mathcal{A}$ and that $\mu_0(A_j) < \infty$. For any $E \in \mathcal{M}(\mathcal{A})$, then

$$
\begin{aligned}
E &= E \cap X \\
&= E \cap \left(\bigcup_{j=1}^{\infty} A_j\right) \\
&= \bigcup_{j=1}^{\infty} (E \cap A_j)
\end{aligned}
$$

and

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

since $\mu^*(E \cap A_j) \leqslant \mu^*(A_j) = \mu_0(A_j) < \infty$.

□

1.4 Borel Measure

Recall that the Borel σ-algebra $\mathcal{B}_{\mathbb{R}}$ is the smallest σ-algebra containing all open sets. Let $\mathcal{G}$ be the set of all open sets in $\mathbb{R}$ with respect to the standard topology. Therefore $\mathcal{B}_{\mathbb{R}} = \mathcal{M}(\mathcal{G})$. We can in fact use something smaller than $\mathcal{G}$.

Theorem 1.48. $\mathcal{B}_{\mathbb{R}}$ is a σ-algebra generated by

a. $\mathcal{A}_0 = \{(a, b) : a, b \in \mathbb{R}, a < b\}$, or by

b. $\mathcal{A}_1 = \{(a, b] : a, b \in \mathbb{R}, -\infty \leqslant a < b < \infty\} \cup \{(a, \infty) : -\infty \leqslant a < \infty\} \cup \{\varnothing\}$.

Any member in $\mathcal{A}_1$ is called an h-interval.

Proof.

a. We want to show that $\mathcal{B}_{\mathbb{R}} = \mathcal{M}(\mathcal{A})$. Obviously $\mathcal{A}_0 \subseteq \mathcal{G}$, then $\mathcal{M}(\mathcal{G})$ is a σ-algebra containing $\mathcal{A}_0$, then $\mathcal{M}(\mathcal{A}_0) \subseteq \mathcal{M}(\mathcal{G}) = \mathcal{B}_{\mathbb{R}}$. Conversely, recall that any open subset in $\mathbb{R}$ is a σ-union of open intervals, therefore $\mathcal{G} \subseteq \mathcal{M}(\mathcal{A})$, so $\mathcal{B}_{\mathbb{R}} = \mathcal{M}(\mathcal{G}) \subseteq \mathcal{M}(\mathcal{A}_0)$, therefore $\mathcal{B}_{\mathbb{R}} = \mathcal{M}(\mathcal{A}_0)$.

b. We first show that $\mathcal{M}(\mathcal{A}_1) \subseteq \mathcal{B}_{\mathbb{R}}$. Since $\mathcal{M}(\mathcal{A}_1)$ is the smallest σ-algebra containing $\mathcal{A}_1$, then it suffices to show that $\mathcal{A}_1 \subseteq \mathcal{B}_{\mathbb{R}}$. It is easy to see that $(a, b] = \bigcap_{n=1}^{\infty}(a, b+\frac{1}{n}) \in \mathcal{B}_{\mathbb{R}}$, and $(a, \infty) = \bigcup_{n=1}^{\infty}(a, n) \in \mathcal{B}_{\mathbb{R}}$.

We now verify that $\mathcal{B}_{\mathbb{R}} \subseteq \mathcal{M}(\mathcal{A}_1)$. By a. we know $\mathcal{B}_{\mathbb{R}} = \mathcal{M}(\mathcal{A}_0)$, so it suffices to show that $\mathcal{A}_0 \subseteq \mathcal{M}(\mathcal{A}_1)$. For $a < b$, we have $(a, b) = \bigcup_{n=1}^{\infty}(a, b-\frac{1}{n}]$, therefore the right-hand side is a σ-union of intervals, hence belongs to $\mathcal{M}(\mathcal{A}_1)$, and we are done.

□

Definition 1.49. We define $\mathcal{A}_2$ to be the collection of finite disjoint unions of h-intervals, e.g., $\bigcup_{j=1}^{n}(a_j, b_j]$, then $\mathcal{A}_2$ is an algebra.

Definition 1.50. A function on $\mathbb{R}$ is said to be right continuous if $\lim_{x \to x_0^+} F(x) = F(x_0)$.

Theorem 1.51. Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and right continuous. Let $I_j = (a_j, b_j]$ for $j = 1, \dots, n$ be disjoint h-intervals. We define the pre-measure μ_0 on $\mathcal{A}_2$ by $\mu_0(\varnothing) = 0$ and $\mu_0\left(\bigcup_{j=1}^{n}(a_j, b_j]\right) = \sum_{j=1}^{n}[F(b_j) - F(a_j)]$.

Proof. First one can check that μ_0 is well-defined, that is, given any partition of h-interval, the μ_0-measurements on the interval are the same.

Second, we need to show that μ_0 satisfies σ-additivity, that is, if $\bigcup_{j=1}^{\infty} I_j \in \mathcal{A}_2$ such that I_j's are disjoint, then $\mu_0\left(\bigcup_{j=1}^{\infty} I_j\right) = \sum_{j=1}^{\infty}\mu_0(I_j)$. It is easy to verify finite additivity, so we now assume

$$\bigcup_{j=1}^{\infty} I_j = I = (a, b] \in \mathcal{A}_2$$

for $-\infty \leqslant a < b < \infty$, then we will show that

$$F(b) - F(a) = \mu_0(I) = \sum_{j=1}^{\infty}\mu_0(I_j)$$

for $I_j = (a_j, b_j]$.

To show $\mu_0(I) \geqslant \sum_{j=1}^{\infty}\mu(I_j)$, we know $F(b) - F(a) \geqslant \sum_{j=1}^{n}[F(b_j) - F(a_j)]$, therefore taking the limit of $n \to \infty$ gives $F(b) - F(a) \geqslant \sum_{j=1}^{\infty}\mu_0(I_j)$.

To show $\mu_0(I) \leqslant \sum_{j=1}^{\infty}\mu(I_j)$, since F is right continuous, then for all $\varepsilon > 0$, there exist $\delta > 0$ such that $F(a+\delta) - F(a) < \varepsilon$. Therefore, for every $j > 0$, there exists $\delta_j > 0$ such that $F(b_j + \delta_j) - F(b_j) < 2^{-j}\varepsilon$, then

$$\begin{aligned}
[a+\delta, b] &\subseteq (a, b] \\
&= \bigcup_{j=1}^{\infty}(a_j, b_j] \\
&= \bigcup_{j=1}^{\infty}(a_j, b_j + \delta_j).
\end{aligned}$$

By compactness, there exists some $N \in \mathbb{N}$ such that $[a+\delta, b] \subseteq \bigcup_{j=1}^{N}(a_j, b_j + \delta_j)$. Assume $b_j + \delta_j \in (a_{j+1}, b_{j+1}]$, then

$$\mu_0(I) = \mu_0((a, b])$$

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

since F is increasing. Let $\varepsilon \to 0$ and we are done. □

Theorem 1.52. Let F be increasing and right-continuous, then

a. there is a unique measure μ_F on $\mathbb{R}$ such that $\mu_F((a,b]) = F(b) - F(a)$ for all $a, b \in \mathbb{R}$;

b. if G is another increasing and right-continuous function, then $\mu_F = \mu_G$ if and only if $F - G$ is a constant function;

c. if μ is a Borel measure on $\mathbb{R}$ that is finite on all bounded Borel sets, i.e., a set $S \subseteq \mathbb{R}$ contained in $[-M, M]$ for some $M \in \mathbb{R}$, then

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

is an increasing function and right-continuous, and $\mu_F = \mu$.

Proof.

a. Consider $\mathbb{R} = \bigcup_{j=-\infty}^{\infty} (j, j+1]$, then the pre-measure $\mu_0((j,j+1]) = F(j+1) - F(j) < \infty$ defined on h-intervals is σ-finite. Therefore there exists a unique extension of measure μ of μ_0 on $\mathcal{B}_{\mathbb{R}} = \mathcal{M}(\mathcal{A}_2)$ such that $\mu|_{\mathcal{A}_2} = \mu_0$.

b. We have $\mu_F((a,b]) = F(b) - F(a)$ and $\mu_G((a,b]) = G(b) - G(a)$, then

$$
\begin{aligned}
\mu_F((a,b]) = \mu_G((a,b]) &\iff F(b) - F(a) = G(b) - G(a) \\
&\iff F(b) - G(b) = G(a) - F(a) \\
&\iff F - G \text{ is constant.}
\end{aligned}
$$

c. First note that F is an increasing function since the measure function is increasing. Take any $x_0 \in \mathbb{R}$, we want to show that $\lim\limits_{x \to x_0^+} F(x) = F(x_0)$. We prove this by cases, either $x_0 = 0$, $x_0 > 0$, or $x_0 < 0$. We will only prove the first case, but the two other cases are analogous. Suppose $x_0 = 0$, take a nested sequence of intervals $E_n = (0, \frac{1}{n}]$, with $E_n \supseteq E_{n+1}$ for all $n \in \mathbb{N}$, then

$$\begin{aligned}
\lim_{x \to 0^+} F(x) &= \lim_{x \to 0^+} \mu((0, x]) \\
&= \lim_{n \to 0} \mu((0, \frac{1}{n}]) \\
&= \lim_{n \to \infty} \mu(E_n) \\
&= \mu\left(\bigcap_{n=1}^{\infty} E_n\right) \\
&= \mu(\varnothing) \\
&= 0 \\
&= F(0)
\end{aligned}$$

since $\mu(E_1) < \infty$.

$\square$

Definition 1.53. Suppose F is increasing and right-continuous, then by Theorem 1.51 we can use F to create μ_0 on $\mathcal{A}_2$, and get an outer measure μ^* induced by μ_0. Let $\mathcal{A}$ be the collection of all μ^*-measurable sets, then $\mu^*|_{\mathcal{A}}$ is a measure. Note that $\mathcal{A} \supseteq \mathcal{B}_{\mathbb{R}}$: since μ_F is only defined on $\mathcal{B}_{\mathbb{R}}$, then $\mu^*|_{\mathcal{A}}$ becomes the extension of μ_F on $\mathcal{A}$. We denote this measure to be $\bar{\mu}_F$, as the extension of μ_F, called the Lebesgue-Stieltjes measure.

Remark 1.54. In particular, if $F(x) = x$ for all $x \in \mathbb{R}$, then $\bar{\mu}_F$ is called a Lebesgue measure, denoted by $\mathfrak{m}$, with $\mathfrak{m}((a, b]) = F(b) - F(a) = b - a$.

Definition 1.55. Let μ be a Lebesgue-Stieltjes measure associated to an increasing and right-continuous function F. Let $\mathcal{M}_\mu$ be the domain of the measure μ, which gives the collection of measurable sets. For any measurable set $E \in \mathcal{M}_\mu$, we have

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

Theorem 1.56. For all $E \in \mathcal{M}_\mu$, we have

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

Proof. Let $\tilde{\mu}(E)$ be the right-hand side of this equation, so we will show that $\mu(E) = \tilde{\mu}(E)$. Note that we have a partition

$$(a_j, b_j) = \bigcup_{k=1}^{\infty} I_k^{(j)},$$

where $I_k^{(j)} = (b_j - \frac{1}{2^k}(b_j - a_j), b_j - \frac{1}{2^{k+1}}(b_j - a_j)]$. Now $E \subseteq \bigcup\limits_{j=1}^{\infty} (a_j, b_j)$, so $E \subseteq \bigcup\limits_{j=1}^{\infty} \bigcup\limits_{k=1}^{\infty} I_k^{(j)}$, and thus

$$\sum_{j=1}^{\infty} \mu((a_j, b_j)) = \sum_{j=1}^{\infty} \mu\left(\bigcup_{k=1}^{\infty} I_k^{(j)}\right)$$

$$= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \mu(I_k^{(j)}).$$

Because $\left\{ \sum_{j=1}^{\infty} \mu((a_j, b_j)) : E \subseteq \bigcup_{j=1}^{\infty} (a_j, b_j) \right\} \subseteq \left\{ \sum_{j=1}^{\infty} \mu((a_j, b_j]) : E \subseteq \bigcup_{j=1}^{\infty} (a_j, b_j] \right\}$, then $\tilde{\mu}(E) \geqslant \mu(E)$.

We now show that $\mu(E) \geqslant \tilde{\mu}(E)$. Pick arbitrary $\varepsilon > 0$, then we know $\mu(E) + \varepsilon$ is not a lower bound of the set $\left\{ \sum_{j=1}^{\infty} \mu((a_j, b_j]) : E \subseteq \bigcup_{j=1}^{\infty} (a_j, b_j] \right\}$, hence there exists $(a_j, b_j]$ for $j \geqslant 1$ such that $E \subseteq \bigcup_{j \geqslant 1} (a_j, b_j]$. Therefore $\sum_{j=1}^{\infty} \mu((a_j, b_j]) \leqslant \mu(E) + \varepsilon$. By the right continuity of F, for $\varepsilon \cdot 2^{-j} > 0$, there exists $\delta_j > 0$ such that $F(b_j + \delta_j) - F(b_j) < \varepsilon \cdot 2^{-j}$, then $E \subseteq \bigcup_{j=1}^{\infty} (a_j, b_j + \delta_j)$. We know

$$\begin{aligned}
\tilde{\mu}(E) &\leqslant \sum_{j=1}^{\infty} \mu((a_j, b_j + \delta_j)) \\
&= \sum_{j=1}^{\infty} [F(b_j + \delta_j) - F(a_j)] \\
&= \sum_{j=1}^{\infty} [F(b_j + \delta_j) - F(b_j) + F(b_j) - F(a_j)] \\
&\leqslant \sum_{j=1}^{\infty} [F(b_j + \delta_j) - F(b_j)] + \sum_{j=1}^{\infty} [F(b_j) - F(a_j)] \\
&< \sum_{j=1}^{\infty} \varepsilon \cdot 2^{-j} + \sum_{j=1}^{\infty} \mu((a_j, b_j]) \\
&< \varepsilon + \mu(E) + \varepsilon \\
&= \mu(E) + 2\varepsilon.
\end{aligned}$$

Taking small enough ε finishes the proof. □

Remark 1.57. The union of h-intervals may not be open, so often times we use the characterization in Theorem 1.56 instead.

Theorem 1.58. For any $E \subseteq \mathcal{M}_\mu$, we have

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

Proof. Let $\tilde{\mu}(E) = \inf\{\mu(U) : \text{ open } U \supseteq E\}$. First, $\mu(E) \leqslant \tilde{\mu}(E)$: since $E \subseteq U$, then $\mu(E) \leqslant \mu(U)$, therefore $\mu(E) \leqslant \tilde{\mu}(E)$. To see $\tilde{\mu}(E) \leqslant \mu(E)$, we have $\mu(E) + \varepsilon$ is not a lower bound of $\left\{ \sum_{j=1}^{\infty} \mu((a_j, b_j]) : E \subseteq \bigcup_{j=1}^{\infty} (a_j, b_j) \right\}$, then there exists (a_j, b_j) for each $j \in \mathbb{N}$ such that $E \subseteq \bigcup_{j=1}^{\infty} (a_j, b_j)$, and that $\sum_{j=1}^{\infty} \mu((a_j, b_j)) \leqslant \mu(E) + \varepsilon$. Therefore, take U to be the open set $\bigcup_{j=1}^{\infty} (a_j, b_j)$, then

$$\tilde{\mu}(E) \leqslant \mu(U) \leqslant \sum_{j=1}^{\infty} \mu((a_j, b_j)) \leqslant \mu(E) + \varepsilon$$

as desired.

Now let $\nu(E) = \sup\{\mu(K) : \text{ compact } K \subseteq E\}$. We note that if $K \subseteq E$, then $\mu(K) \leqslant \mu(E)$, therefore $\nu(E) \leqslant \mu(E)$. To prove the reverse inequality, we consider the following cases:

- E is bounded.
  - E is closed. Since E is bounded and closed, it is compact over $\mathbb{R}$, thus $\mu(E) \leqslant \nu(E)$.
  - E is bounded but not closed. We have $\mu(\bar{E}\backslash E) = \inf\{\mu(U) : \text{ open } U \supseteq \bar{E}\backslash E\}$. For any $\varepsilon > 0$, there exists an open set U such that $U \supseteq \bar{E}\backslash E$ and $\mu(U) \leqslant \mu(\bar{E}\backslash E) + \varepsilon$. Set $K = \bar{E}\backslash U$, then K is compact. Since all measures here are finite, we have

$$\begin{aligned}\mu(K) &= \mu(E) - \mu(E \cap U)\\ &= \mu(E) - [\mu(U) - \mu(U\backslash E)]\\ &\geqslant \mu(E) - \mu(U) + \mu(\bar{E}\backslash E)\\ &\geqslant \mu(E) - \varepsilon.\end{aligned}$$

    Therefore $\nu(E) \geqslant \mu(E) - \varepsilon$, and we are done by taking $\varepsilon \to 0$.

- E is not bounded. Suppose $E = \bigcup_{j=-\infty}^{\infty}((j, j+1] \cap E)$, then denote $E_j = E \cap (j, j+1]$, which is bounded. Therefore, we know the statement is true for each E_j for $j \geqslant 1$, thus $\mu(E_j) = \sup\{\mu(K) : \text{ compact } K \subseteq E_j\}$. Take arbitrary $\varepsilon > 0$, then $\mu(E_j) - \frac{1}{3}\varepsilon \cdot 2^{-|j|}$ is not the upper bound of $\{\mu(K) : \text{ compact } K \subseteq E_j\}$, then there exists a compact set $K_j \subseteq E_j$ such that $\mu(K_j) \geqslant \mu(E_j) - \frac{1}{3}\varepsilon \cdot 2^{-|j|}$. Since $K_j \subseteq E_j$ and E_j's are disjoint, then K_j's are disjoint. Therefore, for $n \in \mathbb{N}$, set $H_n = \bigcup_{j=-n}^{n} K_j$, which is a finite disjoint union of compact sets, so this is a compact set. But $H_n \subseteq E$, then

$$\begin{aligned}\mu(H_n) &= \mu\left(\bigcup_{j=-n}^{n} K_j\right)\\ &= \sum_{j=-n}^{n} \mu(K_j)\\ &\geqslant \sum_{j=-n}^{n} \mu(E_j) - \frac{\varepsilon}{3}\sum_{j=-n}^{n} 2^{-|j|}\\ &\geqslant \sum_{j=-n}^{n} \mu(E_j) - \frac{\varepsilon}{3}\sum_{j=-\infty}^{\infty} 2^{-|j|}\\ &\geqslant \sum_{j=-n}^{n} \mu(E_j) - \varepsilon.\end{aligned}$$

  Note that H_n still depends on n, so we should not take $n \to \infty$ here. Since $\nu(E)$ is the upper bound of $\mu(K)$'s for compact $K \subseteq E$, then $\nu(E) \geqslant \mu(H_n)$, therefore

$$\begin{aligned}\nu(E) &\geqslant \sum_{j=-n}^{n} \mu(E_j) - \varepsilon\\ &= \mu\left(\bigcup_{j=-n}^{n} E_j\right) - \varepsilon.\end{aligned}$$

  Take $n \to \infty$, then

$$\begin{aligned}\nu(E) &\geqslant \lim_{n\to\infty} \mu\left(\bigcup_{j=-n}^{n} E_j\right) - \varepsilon\\ &= \mu\left(\bigcup_{j=-\infty}^{\infty} E_j\right) - \varepsilon\end{aligned}$$

$$= \mu(E) - \varepsilon.$$

Let $\varepsilon \to 0$, we are done.

□

Theorem 1.59. Let $E \subseteq \mathbb{R}$, then the following are equivalent:

a. $E \in \mathcal{M}_\mu$;

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 a F_σ-set and $\mu(N_2) = 0$.

Proof.

- *b.* $\Rightarrow$ *a.*: note that $\mathcal{M}_\mu \supseteq \mathcal{B}_\mathbb{R}$, then both V and N_1 are measurable, therefore E is measurable, i.e., $E \in \mathcal{M}_\mu$.
- *c.* $\Rightarrow$ *a.*: similar to the case above.
- *a.* $\Rightarrow$ *b.*:
  - If $\mu(E) < \infty$, recall $\mu(E) = \inf\{\mu(U) : \text{ open } U \supseteq E\}$. For any $k \in \mathbb{N}$, consider $2^{-k} > 0$, then there exists open subset $U_k \supseteq E$ such that $\mu(U_k) \leqslant \mu(E) + 2^{-k}$. Let $V = \bigcap\limits_{k=1}^{\infty} U_k$ be a G_δ-set, then $V \supseteq E$ as well. It suffices to show that $V \backslash E$ is a null set. We know
  $$\begin{aligned} \mu(V) &= \mu\left(\bigcap_{k=1}^{\infty} U_k\right) \\ &\leqslant \mu(U_k) \\ &\leqslant \mu(E) + 2^{-k} \end{aligned}$$
  for all $k \in \mathbb{N}$. Since $\mu(V)$ and $\mu(E)$ are independent of k, then take $k \to \infty$, therefore $\mu(V) \leqslant \mu(E)$. But since $E \subseteq V$, then $\mu(E) \leqslant \mu(V)$, therefore this gives equality. Since $\mu(E) < \infty$, then $\mu(V) - \mu(E) = 0$, then $\mu(V \backslash E) = 0$ by additivity.
  - If $\mu(E) = \infty$, then the proof can be done using the previous case.
- *a.* $\Rightarrow$ *c.*: the proof is similar to the case above.

□

Theorem 1.60. Let $E \in \mathcal{M}_\mu$, and suppose $\mu(E) < \infty$. For any $\varepsilon > 0$, there exists some set A that is a finite union of open intervals such that $\mu(E \Delta A) = \mu((E \backslash A) \cup (A \backslash E)) < \varepsilon$.

Proof. Note that $\mu(E) = \sup\{\mu(K) : \text{ compact } K \subseteq E\}$. For any $\varepsilon > 0$, there exists compact $K \subseteq E$ such that $\mu(E) - \frac{\varepsilon}{2} < \mu(K)$, which is equivalent to having $\mu(E \backslash K) < \frac{\varepsilon}{2}$. Similarly, recall that $\mu(E) = \inf\{\mu(U) : \text{ open } U \supseteq E\}$, but open set U on $\mathbb{R}$ is characterized as a union of open intervals, therefore this is just $\mu(E) = \inf\{\sum\limits_{j=1}^{\infty} \mu((a_j, b_j)) : \bigcup\limits_{j=1}^{\infty} (a_j, b_j) \supseteq E\}$. Therefore, there exists $\bigcup\limits_{j=1}^{\infty} I_j \supseteq E$, where I_j is open interval for each j, such that $\mu\left(\bigcup\limits_{j=1}^{\infty} I_j\right) < \mu(E) + \frac{\varepsilon}{2}$. Since $\mu(E)$ is finite, then $\mu\left(\bigcup\limits_{j=1}^{\infty} I_j \backslash E\right) < \frac{\varepsilon}{2}$. Now $K \subseteq E \subseteq \bigcup\limits_{j=1}^{\infty} I_j$, but K is compact, so there exists $I_1, \ldots, I_n$ such that their union cover K. Set $A = \bigcup\limits_{j=1}^{m} I_j$, and we are done. □

Definition 1.61. Let $F(x) = x$ be a function for all $x \in \mathbb{R}$, then μ_F is called the Lebesgue measure defined by $\mathfrak{m}((a, b]) = b - a$. The domain of m is $\mathcal{L}$.

For $E \subseteq \mathbb{R}$ and $s, r \in \mathbb{R}$, we denote $E + s = \{x + s : x \in E\}$ and $rE = \{rx : x \in E\}$.

Theorem 1.62. If $E \in \mathcal{L}$, then $\mathfrak{m}(E+s) = \mathfrak{m}(E)$ and $\mathfrak{m}(rE) = |r|\mathfrak{m}(E)$.

Proof. We prove the first claim. For any $E \in \mathcal{L}$ and $s \in \mathbb{R}$, define $m_s = \mathfrak{m}(E+s)$, then this is a measure.

Claim 1.63. For any $E \in \mathcal{L}$, $m_s(E) = \mathfrak{m}(E)$.

Subproof. First note that this is true if E is a finite (disjoint) union of h-intervals of m_s, as $\mathfrak{m}$ extends the pre-measure μ_0. On $\mathcal{B}_{\mathbb{R}}$, the extension is unique, so $m_s(E) = \mathfrak{m}(E)$ if $E \in \mathcal{B}_{\mathbb{R}}$. Moreover, recall $E \in \mathcal{L}$ if and only if $E = V \backslash N_1$ for $V \in \mathcal{B}_{\mathbb{R}}$. Therefore this is true for all $E \in \mathcal{L}$. ■

□

Definition 1.64. The Cantor set $\mathscr{C}$ is constructed iteratively from the interval $[0,1]$, that for any remaining connected interval $[m,n]$, we delete the subinterval $(m + \frac{1}{3}(n-m), m + \frac{2}{3}(n-m))$ from $[m,n]$.

Remark 1.65. Note that

$$\begin{aligned}\mathfrak{m}(\mathscr{C}) &= \mathfrak{m}([0,1]) - \frac{1}{3} - \frac{2}{3^2} - \frac{2^2}{3^3} - \cdots \\ &= 1 - \sum_{j=0}^{\infty} \frac{2^j}{3^{j+1}} \\ &= 1 - 1 \\ &= 0.\end{aligned}$$

Remark 1.66. If E is countable, then

$$\begin{aligned}\mathfrak{m}(E) &= \sum_{j=1}^{\infty} \mathfrak{m}(\{a_j\}) \\ &= 0.\end{aligned}$$

Theorem 1.67. The Cantor set $\mathscr{C}$ is uncountable.

Proof. Alternatively, the Cantor set $\mathscr{C}$ can be represented as

$$\mathscr{C} = \{x \in [0,1] : x = \sum_{j=1}^{\infty} a_j 3^{-j}, a_j \in \{0,2\}\}.$$

To prove that $\mathscr{C}$ is uncountable, it suffices to build a surjection $f : \mathscr{C} \to [0,1]$. For $x \in \mathscr{C}$, we have $x = \sum\limits_{j=1}^{\infty} a_j 3^{-j}, a_j \in \{0,2\}$. Set $f(x) = \sum\limits_{j=1}^{\infty} \frac{a_j}{2} 2^{-j}$ for $\frac{a_j}{2} \in \{0,1\}$, therefore this gives a decimal representation with base 2, so any real number in $[0,1]$ can be represented in this form, therefore we have a surjection. □

Theorem 1.68. Let $F \subseteq \mathbb{R}$ be such that every subset of F is Lebesgue measurable, then $\mathfrak{m}(F) = 0$.

Corollary 1.69. If $\mathfrak{m}(F) > 0$, then there exists a subset S of F such that $S \notin \mathcal{L}$.

Remark 1.70 (Banach-Tarski Paradox)**.** Given a ball $B = S^2$, then there exists some $m \in \mathbb{N}$ such that $B = V_1 \cup \cdots \cup V_m$ is a union of subsets V_i that are not Lebesgue measurable and $\mathfrak{m}(B) \neq \mathfrak{m}(V_1 \cup \cdots \cup V_m)$.

Definition 1.71. For any $x \in \mathbb{R}$, we defined the cosets over $\mathbb{Q}$ to be $\mathbb{Q} + x = \{r + x : r \in \mathbb{Q}\}$ for any x. This is called the coset of an additive group $\mathbb{R}$.

Let E be the set that contains exactly one point from each coset of $\mathbb{Q}$ as representations, which requires the axiom of choice. Now E allows us make a partition on $\mathbb{R}$.

Lemma 1.72.

1. $(E+r_1)\cap(E+r_2)=\varnothing$ if $r_1\neq r_2$ and $r_1,r_2\in\mathbb{Q}$.
2. $\mathbb{R}=\bigcup_{r\in\mathbb{Q}}(E+r)$.

Proof.

1. Suppose $x\in(E+r_1)\cap(E+r_2)$, then $x=e_1+r_1=e_2+r_2$ for some $e_1,e_2\in E$. Therefore $e_1-e_2=r_2-r_1$, which is a non-zero rational number, therefore $0\neq e_1-e_2\in\mathbb{Q}$. Therefore e_1 and e_2 are in the same coset, so $e_1=e_2$, contradiction.
2. Obviously $\mathbb{R}\supseteq\bigcup_{r\in\mathbb{Q}}(E+r)$. Take any $x\in\mathbb{R}$, then E contains a point y from the coset $\mathbb{Q}+x$, therefore $y-x\in\mathbb{Q}$, so take $r=y-x$, then $x\in E+r$.

□

Proof of Theorem 1.68. We have

$$\begin{aligned}F&=F\cap\mathbb{R}\\&=F\cap\bigcup_{r\in\mathbb{Q}}(E+r)\\&=\bigcup_{r\in\mathbb{Q}}(F\cap(E+r)).\end{aligned}$$

Now let $F_r=F\cap(E+r)$ for all $r\in\mathbb{Q}$, then $F=\bigcup_{r\in\mathbb{Q}}F_r$ for $F_r\in\mathcal{L}$ by Lemma 1.72. It remains to verify that $\mathfrak{m}(F_r)=0$ for all $r\in\mathbb{Q}$. Recall

$$\mathfrak{m}(F_r)=\sup\{\mathfrak{m}(K):\text{ compact }K\subseteq F_r\},$$

then it suffices to show that

Claim 1.73. For any compact set $K\subseteq F_r$, $\mathfrak{m}(K)=0$.

Indeed, take the supremum over all compact subsets and we are done.

Subproof. Let $K_r=K+r$ for all $r\in\mathbb{Q}$.

First, we show that $K_{r_1}\cap K_{r_2}=\varnothing$ if $r_1\neq r_2$ for $r_1,r_2\in\mathbb{Q}$. Assume there exists $x\in K_{r_1}\cap K_{r_2}$, then $K\subseteq F_r\subseteq E+r$, so we know $K_{r_1}=K+r_1\subseteq E+r+r_1$ and $K_{r_2}=K+r_2\subseteq E+r+r_2$. Therefore, $x\in(E+r+r_1)\cap(E+r+r_2)$, but by Lemma 1.72 we know $(E+r+r_1)\cap(E+r+r_2)=\varnothing$, contradiction.

Set $H=\bigcup_{r\in\mathbb{Q}}K_r$ be a disjoint union. Since the right-hand side is a Borel set, then it is Lebesgue measurable, so by σ-additivity, we have

$$\begin{aligned}\mathfrak{m}(H)&=\mathfrak{m}\left(\bigcup_{r\in\mathbb{Q}}K_r\right)\\&=\sum_{r\in\mathbb{Q}}\mathfrak{m}(K_r)\\&=\sum_{r\in\mathbb{Q}}\mathfrak{m}(K)\\&=\mathfrak{m}(K)\sum_{r\in\mathbb{Q}}1.\end{aligned}$$

We need to bound the set, so instead of summation over $\mathbb{Q}$, we will sum over $\mathbb{Q}\cap[0,1]$ instead, so for $H=\bigcup_{r\in\mathbb{Q}\cap[0,1]}K_r$ we get

$$\mathfrak{m}(H)=\mathfrak{m}(K)\sum_{r\in\mathbb{Q}\cap[0,1]}1.$$

That is, $\mathfrak{m}(H)$ is just $\mathfrak{m}(K)$ times the number of rational numbers in $[0, 1]$, which are countably many, therefore $\mathfrak{m}(H) = \mathfrak{m}(K) \cdot \mathbb{N}$.

Assume, towards contradiction, that $\mathfrak{m}(K) \neq 0$, then we have $\mathfrak{m}(K) > 0$, so $\mathfrak{m}(H) = \infty$. But we know H is bounded by $[0, 1]$ already, therefore $\mathfrak{m}(H)$ is finite, contradiction. ■

□

Remark 1.74. Not every set is Lebesgue measurable.

2 Integration

2.1 Measurable Functions

Definition 2.1. Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces. A function $f : X \to Y$ is called $(\mathcal{A}, \mathcal{B})$-measurable if $f^{-1}(E) \in \mathcal{A}$ for any $E \in \mathcal{B}$. That is, the preimage of a measurable set is measurable.

Definition 2.2. Let $(X, \mathcal{A})$ be a measurable space.

a. If $f : X \to \mathbb{R}$ is $(\mathcal{A}, \mathcal{B}_{\mathbb{R}})$-measurable, then we say the function f is $\mathcal{A}$-measurable.

b. A complex-valued function $f : X \to \mathbb{C}$ is $\mathcal{A}$-measurable if $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are $\mathcal{A}$-measurable.

Definition 2.3. A function $f : \mathbb{R} \to \mathbb{C}$ is called Lebesgue measurable if it is $\mathcal{L}$-measurable (on both the real part and the imaginary part).

Lemma 2.4. Let $\mathcal{B}$ be a σ-algebra generated by $\mathcal{B}_0$, then $f : X \to Y$ is $(\mathcal{A}, \mathcal{B})$-measurable if and only if $f^{-1}(E) \in \mathcal{A}$ for all $E \in \mathcal{B}_0$.

Proof.

($\Rightarrow$): this is obvious by Definition 2.1.

($\Leftarrow$): let $M = \{E \subseteq Y : f^{-1}(E) \in \mathcal{A}\}$. Note that $\mathcal{M} \supseteq \mathcal{B}_0$ is a σ-algebra, and since $\mathcal{B}$ is the σ-algebra generated by $\mathcal{B}_0$, then $\mathcal{M} \supseteq \mathcal{B}$. Therefore, for all $E \in \mathcal{B}$, we have $f^{-1}(E) \in \mathcal{A}$.

□

Theorem 2.5. Let X and Y be topological spaces, then every continuous function $f : X \to Y$ is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable.

Proof. Note that f is continuous if and only if $f^{-1}(U)$ is open in X for any open subset U in Y, and since $\mathcal{B}_Y$ is the σ-algebra generated by all open subsets of Y, therefore by Lemma 2.4 we know f is $(\mathcal{B}_X, \mathcal{B}_Y)$-measurable. □

Theorem 2.6. Let $f : X \to \mathbb{R}$ be a function, then the following are equivalent:

a. f is $\mathcal{A}$-measurable;

b. $f^{-1}((a, \infty)) \in \mathcal{A}$ for all $a \in \mathbb{R}$;

c. $f^{-1}([a, \infty)) \in \mathcal{A}$ for all $a \in \mathbb{R}$;

d. $f^{-1}((-\infty, a)) \in \mathcal{A}$ for all $a \in \mathbb{R}$;

e. $f^{-1}((-\infty, a]) \in \mathcal{A}$ for all $a \in \mathbb{R}$;

Proof. Since the proofs will be analogous to one another, it suffices to show the equivalence between a. and b.

a. $\Rightarrow$ *b.*: since $(a, \infty) \in \mathcal{B}_{\mathbb{R}}$ is a Borel set, then $f^{-1}((a, \infty)) \in \mathcal{A}$ since f is $\mathcal{A}$-measurable.

b. $\Rightarrow$ *a.*: let $\mathcal{B}_0 = \{(a, \infty) : a \in \mathbb{R}\}$, then $\mathcal{B}_{\mathbb{R}}$ is a σ-algebra generated by $\mathcal{B}_0$. The statement then follows from Lemma 2.4. □

Theorem 2.7. If $f, g : X \to \mathbb{C}$ are $\mathcal{A}$-measurable, then so are $f + g$ and $f \cdot g$.

Proof. Assume, without loss of generality, that f and g are $\mathbb{R}$-valued functions.

First, we show that $f + g$ is $\mathcal{A}$-measurable. By Theorem 2.6, it suffices to show that $(f + g)^{-1}((-\infty, a)) \in \mathcal{A}$ for all $a \in \mathbb{R}$. Fix $a \in \mathbb{R}$, this is the set of elements $x \in X$ such that $(f + g)(x) < a$. Note that $x \in X$ satisfies $(f + g)(x) = f(x) + g(x) < a$ if and only if $f(x) < a - g(x)$, where both expressions are real numbers. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, there exists some $r \in \mathbb{Q}$ such that $f(x) < r < a - g(x)$. Therefore,

$$\begin{aligned}\{x \in X : f(x) + g(x) < a\} &= \bigcup_{r \in \mathbb{Q}} (\{x \in X : f(x) < r\} \cap \{x \in X : r < a - g(x)\}) \\ &= \bigcup_{r \in \mathbb{Q}} \left(f^{-1}((-\infty, r)) \cap g^{-1}((-\infty, a - r))\right) \in \mathcal{A}\end{aligned}$$

since $f^{-1}((-\infty, r)) \in \mathcal{A}$ and $g^{-1}((-\infty, a - r)) \in \mathcal{A}$.

Remark 2.8. Note that if f is $\mathcal{A}$-measurable, then $-f$ is $\mathcal{A}$-measurable. Therefore, the sum and the difference of two $\mathcal{A}$-measurable functions is still $\mathcal{A}$-measurable.

We now show that $f \cdot g$ is also $\mathcal{A}$-measurable.

Claim 2.9. If $f : X \to \mathbb{R}$ is $\mathcal{A}$-measurable, then f^2 is $\mathcal{A}$-measurable as well.

Subproof. By Theorem 2.6, it suffices to show $\{x \in X : f^2(x) > \alpha\} \in \mathcal{A}$ for all $\alpha \in \mathbb{R}$.

- If $\alpha < 0$, then $\{x \in X : f^2(x) > \alpha\} = X \in \mathcal{A}$.
- If $\alpha \geqslant 0$, then $\{x \in X : f^2(x) > \alpha\} = \{x \in X : f(x) > \sqrt{\alpha}\} \cup \{x \in X : f(x) < -\sqrt{\alpha}\}$. Since f is $\mathcal{A}$-measurable, then this is a union of two $\mathcal{A}$-measurable sets, which is still $\mathcal{A}$-measurable.

■

Now $fg = \frac{1}{2}\left((f+g)^2 - f^2 - g^2\right)$ which is $\mathcal{A}$-measurable. □

Definition 2.10. The extended real line is $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, and correspondingly $\mathcal{B}_{\bar{\mathbb{R}}} = \{E \subseteq \bar{\mathbb{R}} : E \cap \mathbb{R} \in \mathcal{B}_{\mathbb{R}}\}$. Any member in $\mathcal{B}_{\bar{\mathbb{R}}}$ is called a Borel set in $\bar{\mathbb{R}}$.

A function $f : X \to \bar{\mathbb{R}}$ is called $\mathcal{A}$-measurable if $f^{-1}(E) \in \mathcal{A}$ for all $E \in \mathcal{B}_{\bar{\mathbb{R}}}$.

We deduce results analogous to Theorem 2.6.

Theorem 2.11. Let $f : X \to \bar{\mathbb{R}}$ be a function, then the following are equivalent:

a. f is $\mathcal{A}$-measurable;
b. $f^{-1}((a, \infty]) \in \mathcal{A}$ for all $a \in \mathbb{R}$;
c. $f^{-1}([a, \infty]) \in \mathcal{A}$ for all $a \in \mathbb{R}$;
d. $f^{-1}([-\infty, a)) \in \mathcal{A}$ for all $a \in \mathbb{R}$;
e. $f^{-1}([-\infty, a]) \in \mathcal{A}$ for all $a \in \mathbb{R}$;

Theorem 2.12. Let $\{f_j\}_{j=1}^{\infty}$ be a sequence of $\bar{\mathbb{R}}$-valued measurable functions on $(X, \mathcal{A})$, then the functions

- $g_1(x) = \sup\limits_{j \in \mathbb{N}} f_j(x) = \sup\{f_j(x) : j \in \mathbb{N}\}$;
- $g_2(x) = \inf\limits_{j \in \mathbb{N}} f_j(x) = \inf\{f_j(x) : j \in \mathbb{N}\}$;
- $g_3(x) = \limsup\limits_{j \in \mathbb{N}} f_j(x) = \limsup\{f_j(x) : j \in \mathbb{N}\}$;
- $g_4(x) = \liminf\limits_{j \in \mathbb{N}} f_j(x) = \liminf\{f_j(x) : j \in \mathbb{N}\}$

are measurable.

Proof. We prove $g_1^{-1}((a, \infty]) \in \mathcal{A}$ for all $a \in \mathbb{R}$. Recall that $g_1^{-1}((a, \infty]) = \{x \in X : \infty \geqslant \sup\limits_j f_j(x) > a\} = \bigcup\limits_{j=1}^{\infty} \{x \in X : \infty \geqslant f_j(x) > a\}$. Since each f_j is $\mathcal{A}$-measurable, then each set is measurable, and so is the countable union of such functions. Therefore $g_1(x)$ is measurable. Similarly, we can show that $g_2(x)$ is measurable.

We also prove that g_3 is measurable. Recall that $\limsup\limits_{j \to \infty} f_j(x) = \inf\limits_{j \in \mathbb{N}} \sup\limits_{k > j} f_k(x)$, then it is measurable since supremum and infimum are measurable as functions. Similarly, we can show that $g_4(x)$ is measurable. □

Definition 2.13. Let $f : X \to \bar{\mathbb{R}}$ be a function, then define $f^+(x) = \max\{f(x), 0\}$ and $f^-(x) = \max\{-f(x), 0\}$.

Remark 2.14.

- $f^+ \geqslant 0$;

- $f^- \geqslant 0$;
- $f = f^+ - f^-$;
- $|f| = f^+ + f^-$;
- If f is measurable, then so are f^+, f^-, $|f|$.

Definition 2.15. Let $E \subseteq X$. The characteristic function or the indicator function is

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

Remark 2.16. If $E \in \mathcal{A}$, then χ_E is ($\mathcal{A}$-)measurable.

Definition 2.17. A simple function on X is a function that can be written as a finite $\mathbb{C}$-linear combination of characteristic functions of sets in $\mathcal{A}$.

Theorem 2.18. Any simple function f can be represented as a standard representation of the form

$$f(x) = \sum_{j=1}^{n} a_j \chi_{E_j}$$

where E_j's are disjoint, $a_j \in \mathbb{C}$ and $\bigcup_{j=1}^{n} E_j = X$.

Proof. We can write $f(x) = \sum_{k=1}^{m} a_k \chi_{E_k}(X)$ for some measurable sets $E_k \in \mathcal{A}$. Since each characteristic function takes only two values, then f takes finitely many valuers, say $z_1, \dots, z_m$. Now we can write $f(x) = \sum_{j=1}^{m} z_j \chi_{E_j}(x)$ where $E_j = \{x \in X : f(x) = z_j\} = f^{-1}(\{z_j\})$. In particular, E_j's are disjoint. However, these sets may not cover X. Let $E_{m+1} = X \setminus \bigcup_{j=1}^{m} E_j$, then $\bigcup_{j=1}^{m+1} E_j = X$, hence

$$f(x) = \sum_{j=1}^{m+1} z_j \chi_{E_j}(x)$$

where $z_{m+1} = 0$. $\square$

Remark 2.19. Equivalently, a function $f : X \to \mathbb{C}$ is simple if and only if f is measurable and the range of f is a finite subset of $\mathbb{C}$.

Theorem 2.20. Let $(X, \mathcal{A})$ be a measurable space.

a. If $f : X \to [0, \infty]$ is measurable, then there exists a sequence $\{\varphi_n\}_{n \geqslant 1}$ of simple functions such that
   - $0 \leqslant \varphi_1 \leqslant \varphi_2 \leqslant \cdots \leqslant f$,
   - $\lim_{n\to\infty} \varphi_n(x) = f(x)$ for all $x \in X$, and
   - $\varphi_n \rightrightarrows f$ converges uniformly on A, i.e., $\lim_{n\to\infty} \sup_{x \in A} |\varphi_n(x) - f(x)| = 0$, for any set A on which f is bounded.

b. If $f : X \to \mathbb{C}$ is measurable, then there exists a sequence $\{\varphi_n\}_{n \geqslant 1}$ of simple functions such that
   - $0 \leqslant |\varphi_1| \leqslant |\varphi_2| \leqslant \cdots \leqslant |f|$.
   - $\lim_{n\to\infty} \varphi_n(x) = f(x)$ for all $x \in X$.
   - $\varphi_n \rightrightarrows f$ converges uniformly on any set on which f is bounded.

Proof.

a. Take arbitrary $n \in \mathbb{N} \cup \{0\}$ and arbitrary $k \in \mathbb{Z}$. We define a dyadic interval to be

$$I_{k,n} = (k2^{-n}, (k+1)2^{-n}],$$

then let $\mathcal{I} = \{I_{k,n} : k, n\}$. For any $I, J \in \mathcal{I}$, we either have $I \subseteq J$, $J \subseteq I$, or $I \cap J = \varnothing$. That is, we have a graded structure on $\mathcal{I}$. Now define $E_{k,n} = \{x \in X : f(x) \in I_{k,n}\} = f^{-1}(I_{k,n})$ and $F_n = f^{-1}((2^n, \infty))$. Therefore, for a fixed n, the $I_{k,n}$'s give a partition of $(0, 2^n)$ on the y-axis, and $f(F_n)$ covers the rest of the y-axis. We define a simple function

$$\varphi_n(x) = \sum_{k=1}^{2^{2n}-1} k2^{-n}\chi_{E_{k,n}}(x) + 2^n\chi_{F_n}(x).$$

Claim 2.21. For any $n \in \mathbb{N}$, $\varphi_n(x) \leqslant \varphi_{n+1}(x)$.

Subproof. This follows from the definition. ■

Claim 2.22. We have $0 \leqslant f(x) - \varphi_n(x) \leqslant 2^{-n}$ for all $x \in F_n^c = \{x \in X : f(x) \leqslant 2^n\}$.

Subproof. We have

$$f(x) = \sum_{k=0}^{2^{2n}-1} f(x)\chi_{E_{k,n}}(x) + f(x)\chi_{F_n}(x)$$

which partitions $(0, \infty)$ to $\bigcup_{k=0}^{2^{2n}-1} I_{k,n}$ and $(2^n, \infty)$. Therefore

$$\begin{aligned} f(x) - \varphi_n(x) &= \sum_{k=0}^{2^{2n}-1} [f(x) - k \cdot 2^{-n}]\chi_{E_{k,n}}(x) + (f(x) - 2^n)\chi_{F_n}(x) \\ &= \sum_{k=0}^{2^{2n}-1} [f(x) - k \cdot 2^{-n}]\chi_{E_{k,n}}(x) \\ &\geqslant 0 \end{aligned}$$

if $x \in F_n^c$. We now bound the difference from above by enlarging it, and since $E_{k,n}$'s are disjoint, then

$$\begin{aligned} \sum_{k=0}^{2^{2n}-1} [f(x) - k \cdot 2^{-n}]\chi_{E_{k,n}}(x) &\leqslant \sum_{k=0}^{2^{2n}-1} [(k+1)2^{-n} - k2^{-n}]\chi_{E_{k,n}}(x) \\ &= \sum_{k=0}^{2^{2n}-1} 2^{-n}\chi_{E_{k,n}}(x) \\ &= 2^{-n}\sum_{k=0}^{2^{2n}-1} \chi_{E_{k,n}}(x) \\ &\leqslant 2^{-n} \end{aligned}$$

as desired. ■

Claim 2.23. $\lim_{n\to\infty} \varphi_n(x) = f(x)$ for all $x \in X$.

Subproof.

- Suppose $f(x) = \infty$, then recall $\varphi_n(x) = 2^n\chi_{F_n}(x) = 2^n$, so obviously both values equal to ∞.

- Suppose $0 \leqslant f(x) < \infty$, then for large enough n, we have $2^n > f(x)$, therefore $x \in F_n^c$ in this case. By Claim 2.22, $0 \leqslant f(x) - \varphi_n(x) \leqslant 2^{-n}$ for n large enough, so when we let $n \to \infty$, then

$$0 \leqslant \lim_{n\to\infty} [f(x) - \varphi_n(x)] \leqslant 0$$

and therefore by squeeze theorem the limit exists and must equal to 0, i.e., $\lim_{n\to\infty} \varphi_n(x) = f(x)$.

■

Claim 2.24. $\varphi_n \rightrightarrows f$ converges uniformly on any set on which f is bounded.

Subproof. Let A be a set on which f is bounded. For any $x \in A$, there exists some large enough n such that $0 \leqslant f(x) - \varphi_n(x) \leqslant 2^{-n}$ by Claim 2.22, so

$$0 \leqslant \sup_{x\in A} |f(x) - \varphi_n(x)| \leqslant 2^{-n},$$

so taking $n \to \infty$ gives

$$\lim_{n\to\infty} \sup_{x\in A} |f(x) - \varphi_n(x)| = 0,$$

i..e, $\varphi_n \rightrightarrows f$ on A. ■

b. Write $f = \mathrm{Re}(f) + i\,\mathrm{Im}(f)$, then both $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are measurable. Now write $\mathrm{Re}(f) = (\mathrm{Re}(f))^+ - (\mathrm{Re}(f))^-$ and $\mathrm{Im}(f) = (\mathrm{Im}(f))^+ - (\mathrm{Im}(f))^-$. By part a., we find a desirable sequence for each of these four parts of the function, then taking the sum/difference gives the desired sequence for f.

□

2.2 Integration of Non-negative Functions

Definition 2.25. Let $(X, \mathcal{A}, \mu)$ be a measure space, and let L^+ be the collection of all non-negative measurable functions on X, i.e., $f \in L^+$ if and only if $f : X \to [0, \infty]$.

Let $\varphi \in L^+$ be a simple function, then we can represent φ as

$$\varphi(x) = \sum_{j=1}^{n} a_j \chi_{E_j}(x)$$

for disjoint $E_j \in \mathcal{A}$ such that $\bigcup_{j=1}^{n} = X$.

We first define the integral for simple functions to be

$$\int_X \varphi d\mu = \sum_{j=1}^{n} a_j \mu(E_j).$$

Here we set $0 \cdot \infty = 0$. For any $A \subseteq X$, we define the integral to be

$$\int_A \varphi d\mu = \int_X \varphi\chi - Ad\mu.$$

To extend our definition to general non-negative functions, we need to define the following. For any $f \in L^+$, set

$$\int_X f d\mu = \sup \left\{ \int_X \varphi d\mu : 0 \leqslant \varphi \leqslant f \text{ for simple function } \varphi \right\}.$$

Since any non-negative measurable function is a limit of simple functions, then such simple functions exist, hence the supremum exists, which is either a real number or ∞.

Proposition 2.26. Let φ and ψ be simple functions in L^+, then

a. if $c \geqslant 0$, $\int_X c\varphi d\mu = c\int_X \varphi d\mu$;

b. $\int_X \varphi d\mu + \int_X \psi d\mu = \int_X (\varphi+\psi) d\mu$;

c. if $\varphi \leqslant \psi$ pointwise, then $\int_X \varphi d\mu \leqslant \int_X \psi d\mu$;

d. for any $A \in \mathcal{A}$, define $\nu : A \to \int_A \varphi d\mu$, then ν is a measure on $\mathcal{A}$.

Proof.

a. This follows from the definition.

b. Set $\varphi(X) = \sum_{j=1}^{n} a_j \chi_{E_j}(X)$ and $\psi(x) = \sum_{k=1}^{m} b_k \chi_{F_k}(x)$ as standard representations. To add the functions together, we need to refine the partition. Recall $X = \bigcup_{j=1}^{m} E_j = \bigcup_{k=1}^{m} F_k$, then we write

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

and similarly

$$F_k = F_k \cap X = F_k \cap \left(\bigcup_{j=1}^{n} E_j\right) = \bigcup_{j=1}^{n} (F_k \cap E_j).$$

Therefore

$$\begin{aligned} \varphi(x) &= \sum_{j=1}^{n} a_j \chi_{E_j} \\ &= \sum_{j=1}^{n} a_j \sum_{k=1}^{m} \chi_{E_j \cap F_k} \\ &= \sum_{j=1}^{n} \sum_{k=1}^{m} a_j \chi_{E_j \cap F_k} \end{aligned}$$

and similarly

$$\psi(x) = \sum_{j=1}^{n} \sum_{k=1}^{m} b_k \chi_{E_j \cap F_k}.$$

Therefore

$$\begin{aligned} (\varphi+\psi)(x) &= \varphi(x) + \psi(x) \\ &= \sum_{j,k} (a_j + b_k) \chi_{E_j \cap F_k}. \end{aligned}$$

Finally,

$$\begin{aligned} \int_X (\varphi+\psi) d\mu &= \sum_{j,k} (a_j + b_k) \mu(E_j \cap F_k) \\ &= \sum_{j,k} a_j \mu(E_j \cap F_k) + \sum_{j,k} b_k \mu(E_j \cap F_k) \\ &= \int_X \varphi d\mu + \int_X \psi d\mu. \end{aligned}$$

c. Using the same partition trick, since $\varphi \leqslant \psi$, then $a_j \leqslant b_k$ whenever $E_j \cap F_k \neq \varnothing$. Therefore,

$$\begin{aligned}\int_X \varphi d\mu &= \sum_{j,k} a_j \mu(E_j \cap F_k) \\ &\leqslant \sum_{j,k} b_k \mu(E_j \cap F_k) \\ &= \int_X \psi d\mu.\end{aligned}$$

d. It is easy to verify that

$$\nu(\varnothing) = \int_{\varnothing} \varphi d\mu = 0.$$

It remains to show that ν satisfies σ-additivity. Take a sequence $\{A_k\}_{k\geqslant 1} \subseteq \mathcal{A}$, such that A_k's are disjoint. Given a standard representation $\varphi = \sum\limits_{j=1}^{n} a_j \chi_{E_j}$, and we have

$$\begin{aligned}\nu\left(\bigcup_{k=1}^{\infty} A_k\right) &= \int_{\bigcup\limits_{k=1}^{\infty} A_k} \varphi d\mu \\ &= \int_X \varphi \chi_{\bigcup\limits_{k=1}^{\infty} A_k} d\mu \\ &= \int_X \sum_{j=1}^{n} a_j \chi_{E_j} \chi_{\bigcup\limits_{k=1}^{\infty} A_k} d\mu \\ &= \int_X \sum_{j=1}^{n} a_j \chi_{E_j \cap \left(\bigcup\limits_{k=1}^{\infty} A_k\right)} d\mu \\ &= \sum_{j=1}^{n} a_j \mu\left(E_j \cap \bigcup_{k=1}^{\infty} A_k\right) \\ &= \sum_{j=1}^{n} a_j \mu\left(\bigcup_{k=1}^{\infty} (E_j \cap A_k)\right) \\ &= \sum_{j=1}^{n} a_j \sum_{k=1}^{\infty} \mu(E_j \cap A_k) \\ &= \sum_{k=1}^{\infty} \sum_{j=1}^{n} a_j \mu(E_j \cap A_k) \\ &= \sum_{k=1}^{\infty} \int_{A_k} \varphi d\mu \\ &= \sum_{k=1}^{\infty} \nu(A_k).\end{aligned}$$

Note that we can only switch the summation because one of them is infinite while the other one is finite.

$\square$

Remark 2.27. Let φ, ψ be simple functions such that $\varphi \leqslant \psi$, then $\int\limits_X \varphi \leqslant \int\limits_X \psi$. Therefore, this is true for any functions $f, g \in L^+$ as well.

Theorem 2.28 (Monotone Convergence)**.** Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of functions in L^+ such that $f_j \leqslant f_{j+1}$ for all $j \in \mathbb{N}$, then

$$\lim_{n\to\infty}\int_X f_n d\mu = \int_X \lim_{n\to\infty} f_n d\mu$$

Remark 2.29. By Remark 2.27, the limit on the left-hand side exists.

Proof. Since the sequence $\{f_n\}_{n\in\mathbb{N}}$ is monotonely increasing, then $\lim\limits_{n\to\infty} f_n$ exists in $\bar{\mathbb{R}}$. Set $f = \lim\limits_{n\to\infty} f_n$, then $f \in L^+$ as well. In particular, $f = \sup\limits_{n\in\mathbb{N}} f_n$ as well, so $f_n \leqslant f$ for all $n \in \mathbb{N}$. Therefore,

$$\int_X f_n d\mu \leqslant \int_X f d\mu$$

for all $n \in \mathbb{N}$. Since $\{\int_X f_n d\mu\}_{n\geqslant 1}$ is a monotone sequence, the limit exists, therefore taking the limit $n \to \infty$ gives

$$\lim_{n\to\infty}\int_X f_n d\mu \leqslant \int_X \lim_{n\to\infty} f_n d\mu.$$

It remains to show

$$\lim_{n\to\infty}\int_X f_n d\mu \geqslant \int_X \lim_{n\to\infty} f_n d\mu.$$

Claim 2.30. Let φ be any simple function such that $0 \leqslant \varphi \leqslant f$. For any fixed $\alpha \in (0,1)$, let $E_n = \{x \in X : f_n(x) \geqslant \alpha\varphi(x)\}$, then

a. $E_n \subseteq E_{n+1}$ for all $n \in \mathbb{N}$, and $X = \bigcup\limits_{n=1}^{\infty} E_n$;

b. $\int_X \varphi d\mu = \lim\limits_{n\to\infty} \int_{E_n} \varphi d\mu$.

Subproof.

a. Since $f_{n+1} \geqslant f_n$, then $E_n \subseteq E_{n+1}$ for all $n \in \mathbb{N}$. To show $X = \bigcup\limits_{n=1}^{\infty} E_n$, we note that $E_n \subseteq X$ for all n implies $\bigcup\limits_{n=1}^{\infty} E_n \subseteq X$, and we claim that $X \subseteq \bigcup\limits_{n=1}^{\infty} E_n$. Take arbitrary $x \in X$,
   - if $\varphi(x) = 0$, then $f_n(x) \geqslant 0 = \varphi(x)$, so $x \in E_n$ for all n by definition;
   - if $\varphi(x) > 0$, recall $f(x) = \lim\limits_{n\to\infty} f_n(x)$, then there exists large enough $N \in \mathbb{N}$ such that $0 \leqslant f(x) - f_N(x) < (1-\alpha)\varphi(x)$, but $\varphi(x) \leqslant f(x)$, then $0 \leqslant f(x) - \varphi(x) < f_N(x) - \alpha\varphi(x)$. In particular, $x \in E_N$.

b. Recall from Proposition 2.26 that $\nu(A) = \int_A \varphi d\mu$ for all $A \in \mathcal{A}$ defines a measure. By the continuity from below for ν and part a., we know

$$\begin{aligned}\lim_{n\to\infty}\int_{E_n}\varphi d\mu &= \lim_{n\to\infty}\nu(E_n)\\ &= \nu\left(\bigcup_{n=1}^{\infty} E_n\right)\\ &= \nu(X)\\ &= \int_X \varphi d\mu.\end{aligned}$$

■

By Claim 2.30, we now have

$$\int_X f_n d\mu = \int_X f_n \chi_{E_n} d\mu = \int_X \alpha\varphi\chi_{E_n} d\mu = \alpha \int_X \varphi\chi_{E_n} d\mu.$$

Since this is true for all n, then taking $n \to \infty$ gives

$$\lim_{n\to\infty} \int_X f_n d\mu \geqslant \alpha \lim_{n\to\infty} \int_X \varphi\chi_{E_n} d\mu = \alpha \int_X \varphi d\mu$$

for any $\alpha \in (0, 1)$. Taking $\alpha \to 1$, we get

$$\lim_{n\to\infty} \int_X f_n d\mu \geqslant \int_X \varphi d\mu$$

for any function φ bounded by 0 and f. Taking the supremum over all such φ gives

$$\lim_{n\to\infty} \int_X f_n d\mu \geqslant \int_X f d\mu.$$

□

Theorem 2.31. Let $f_n \in L^+$ for all $n \in \mathbb{N}$, then

$$\int_X \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} \int_X f_n d\mu.$$

Proof.

Claim 2.32. Given any $f_1, f_2 \in L^+$,

$$\int_X (f_1 + f_2) d\mu = \int_X f_1 d\mu + \int_X f_2 d\mu.$$

Subproof. Since $f_1 \geqslant 0$, there exists simple functions φ_j's such that $0 \leqslant \varphi_j \leqslant f_1$ for all $\in \mathbb{N}$, $\varphi_j \leqslant \varphi_{j+1}$ for all j, and $\lim_{j\to\infty} \varphi_j = f_1$. Similarly, there are simple functions $0 \leqslant \psi_j \leqslant f_2$ for all $j \in \mathbb{N}$ with $\psi_j \leqslant \psi_{j+1}$ for all j, and that $\lim_{j\to\infty} \psi_j = f_2$. Therefore

$$\int_X (f_1 + f_2) d\mu = \int_X \lim_{j\to\infty} \varphi_j + \lim_{j\to\infty} \psi_j d\mu = \int_X \lim_{j\to\infty} (\varphi_j + \psi_j) d\mu.$$

Since $\varphi_j + \psi_j$ increases monotonically, so by Theorem 2.28, we have

$$\int_X (f_1 + f_2) d\mu = \int_X \lim_{j\to\infty} (\varphi_j + \psi_j) d\mu = \lim_{j\to\infty} \int_X \varphi_j + \psi_j d\mu$$

$$
\begin{aligned}
&= \lim_{j\to\infty}\left(\int_X \varphi_j d\mu + \int_X \psi_j d\mu\right)\\
&= \lim_{j\to\infty}\int_X \varphi_j d\mu + \lim_{j\to\infty}\int_X \psi_j d\mu\\
&= \int_X \lim_{j\to\infty}\varphi_j d\mu + \int_X \lim_{j\to\infty}\psi_j d\mu\\
&= \int_X f_1 d\mu + \int_X f_2 d\mu
\end{aligned}
$$

where we apply Theorem 2.28 at the last steps. ■

By Claim 2.32,

$$\int_X \sum_{n=1}^N f_n d\mu = \sum_{n=1}^N \int_X f_n d\mu$$

for all $n \in \mathbb{N}$. By Theorem 2.28,

$$
\begin{aligned}
\int_X \sum_{n=1}^\infty f_n d\mu &= \lim_{N\to\infty}\int_X \sum_{n=1}^N f_n d\mu\\
&= \lim_{N\to\infty}\sum_{n=1}^N \int_X f_n d\mu\\
&= \sum_{n=1}^N \int_X f_n d\mu.
\end{aligned}
$$

□

Theorem 2.33. Let $f \in L^+$, then $\int_X f d\mu = 0$ if and only if $f \equiv 0$ almost everywhere.

Proof.

($\Leftarrow$): Suppose $f \equiv 0$ almost everywhere, then for every choice of simple function φ such that $0 \leqslant \varphi \leqslant f$, $\varphi \equiv 0$ almost everywhere. Take the standard representation $\varphi = \sum_{j=1}^n a_j\chi_{E_j}$, then either $a_j = 0$ or $\mu(E_j) = 0$. Therefore,

$$
\begin{aligned}
\int_X \varphi d\mu &= \sum_{j=1}^n a_j\mu(E_j)\\
&= 0
\end{aligned}
$$

according to the convention that $0 \cdot \infty = 0$.

($\Rightarrow$): We claim that $\mu(\{x \in X : f(x) > 0\}) = 0$. To see this, note that

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

Denote $E_n = \{x \in X : f(x) > \frac{1}{n}\}$, then we just need to show that $\mu(E_n) = 0$ for all $n \in \mathbb{N}$. Note that

$$0 = \int_X f d\mu$$

$$\begin{aligned} &\geqslant \int_{E_n} f d\mu \\ &\geqslant \int_{E_n} \frac{1}{n} d\mu \\ &= \frac{1}{n} \times \mu(E_n), \end{aligned}$$

so $0 \leqslant \mu(E_n) \leqslant n \cdot 0 = 0$, hence $\mu(E_n) = 0$ for all $n \in \mathbb{N}$.

□

Corollary 2.34. If $f \in L^+$ and $\mu(E) = 0$, then

$$\int_E f d\mu = 0.$$

Proof. Note that

$$\int_E f d\mu = \int_X f \chi_E d\mu,$$

but $f\chi_E = 0$ almost everywhere since $\mu(E) = 0$, so by Theorem 2.33 we are done. □

Theorem 2.35. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in L^+. Suppose that $f_n \leqslant f_{n+1}$ for all $n \in \mathbb{N}$, and that $\lim\limits_{n\to\infty} f_n(x) = f(x)$ almost everywhere $x \in X$, then

$$\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu.$$

Proof. Let $E = \{x \in X : \lim\limits_{n\to\infty} f_n(x) = f(x)\}$, so E^c is a null set. Extend the function f to

$$f_{\text{ext}}(x) = \begin{cases} f(x), & \text{if } x \in E \\ 0, & \text{if } x \in E^c \end{cases}$$

then by Theorem 2.28 we have

$$\begin{aligned} \int_X f d\mu &= \int_E d\mu + \int_{E^c} 0 d\mu \\ &= \int_E f d\mu \\ &= \int_E \lim_{n\to\infty} f_n d\mu \\ &= \int_X \lim_{n\to\infty} f_n \chi_E d\mu \\ &= \lim_{n\to\infty} \int_X f_n \chi_E d\mu \\ &= \lim_{n\to\infty} \left(\int_E f_n d\mu + \int_{E^c} f_n d\mu \right) \\ &= \lim_{n\to\infty} \int_X f_n d\mu. \end{aligned}$$

□

Theorem 2.36 (Fatou's Lemma)**.** Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in L^+, then

$$\int_X \liminf_{n\to\infty} f_n d\mu \leqslant \liminf_{n\to\infty} \int_X f_n d\mu.$$

Remark 2.37. Note that Theorem 2.36 does not require Theorem 2.28, but we will use it to give a quick proof.

Proof. Note that for all $j \geqslant n$, we have

$$\inf_{k\geqslant n} f_k(x) \leqslant f_j(x).$$

Taking the integral, we have

$$\int_X \inf_{k\geqslant n} f_k d\mu \leqslant \int_X f_j d\mu$$

for all $j \geqslant n$. Therefore,

$$\int_X \inf_{k\geqslant n} f_k d\mu \leqslant \inf_{j\geqslant n} \int_X f_j d\mu$$

for all $n \in \mathbb{N}$. By definition,

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

By Theorem 2.28, taking the limit gives

$$\begin{aligned}\int_X \liminf_{n\to\infty} f_n d\mu &= \lim_{n\to\infty} \int_X \inf_{k\geqslant n} f_k d\mu \\ &\leqslant \lim_{n\to\infty} \inf_{j\geqslant n} \int_X f_j d\mu \\ &= \liminf_{n\to\infty} \int_X f_n d\mu.\end{aligned}$$

□

Remark 2.38. There is a different version of Theorem 2.36 concerning $\limsup$ instead.

Corollary 2.39. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence in L^+ and $\lim_{n\to\infty} f_n(x) = f(x)$ almost everywhere in $x \in X$, then

$$\int_X f d\mu \leqslant \liminf_{n\to\infty} \int_X f_n d\mu.$$

Theorem 2.40. Let $f \in L^+$ and $\int_X f d\mu < \infty$, then $\{x \in X : f(x) = \infty\}$ is a null set, and $\{x \in X : f(x) > 0\}$ is σ-finite.

Proof. We know that

$$\infty > \int_X f d\mu \geqslant \int_{\{x\in X : f(x)=\infty\}} f d\mu = \infty\mu(\{x \in X : f(x) = \infty\})$$

which forces $\mu(\{x \in X : f(x) = \infty\} = 0$. Also note that the level set

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

so we define $E_n = \{x \in X : f(x) > \frac{1}{n}\}$, so it remains to verify that $\mu(E_n) < \infty$ for all $n \in \mathbb{N}$. To see this,

$$\infty > \int_X f d\mu > \int_{E_n} f d\mu > \frac{1}{n}\mu(E_n),$$

therefore $\mu(E_n) < \infty$. □

2.3 Integration of Complex-Valued Functions

If f is a real-valued measurable function, we know $f = f^+ - f^-$ for $f^+, f^- \in L^+$. We know how to define $\int_X f^+ d\mu$ and $\int_X f^- d\mu$. To find the integral of f, we define

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu$$

if one of the two terms is not ∞. We need to resolve the issue when both of them are ∞.

Definition 2.41. Let f be a complex-valued measurable function, we say f is integrable if

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

that is, the L^1-norm $||f||_1 = \int_X |f| d\mu$ is finite. We define

$$L^1(X) = \left\{ f : \int_X |f| d\mu < \infty \right\}.$$

to be the set of L^1-integrable functions.

The following properties are obvious.

Theorem 2.42. Let $f, g \in L^1(X)$, then

a. $\int_X (\alpha f + \beta g) d\mu = \alpha \int_X f d\mu + \beta \int_X g d\mu$ for all $\alpha, \beta \in \mathbb{C}$;

b. if $|f| \leqslant |g|$ almost everywhere, then $\int_X |f| d\mu \leqslant \int_X |g| d\mu$;

c. let $\lambda(A) = \int_A |f| d\mu$ for all $A \in \mathcal{A}$, then λ is a measure on $\mathcal{A}$.

Theorem 2.43 (Triangle Inequality). Let $f \in L^1(X)$, then

$$\left| \int_X f d\mu \right| \leqslant \int_X |f| d\mu.$$

Proof.

- If f is real-valued, then

$$\left| \int_X f d\mu \right| = \left| \int_X f^+ d\mu - \int_X f^- d\mu \right| \leqslant \int_X f^+ d\mu + \int_X f^- d\mu = \int_X f^+ + f^- d\mu.$$

- If f is complex-valued, now we can just assume $\int_X f d\mu \neq 0$. Set

$$\alpha = \frac{\overline{\int_X f d\mu}}{\left| \int_X f d\mu \right|},$$

then we have $|\alpha| = 1$, and

$$\left|\int_X f d\mu\right| = \frac{\overline{\int_X f d\mu}\int_X f d\mu}{\left|\int_X f d\mu\right|} = \alpha \int_X f d\mu.$$

In particular, $\alpha \int_X f d\mu \in \mathbb{R}$. We know

$$\begin{aligned}\left|\int_X f d\mu\right| &= \operatorname{Re}\left(\alpha \int_X f d\mu\right)\\ &= \operatorname{Re}\left(\int_X \alpha f d\mu\right)\\ &= \int_X \operatorname{Re}(\alpha f) d\mu\\ &\leqslant \int_X |\operatorname{Re}(\alpha f)| d\mu\\ &\leqslant \int_X |\alpha f| d\mu\\ &= |\alpha| \int_X |f| d\mu\\ &= \int_X |f| d\mu.\end{aligned}$$

□

Theorem 2.44. Let $f, g \in L^1(X)$, then

a. $\int_X |f - g| d\mu = 0$ if and only if $f = g$ almost everywhere;

b. $\int_E f d\mu = \int_E g d\mu$ for all $E \in \mathcal{A}$ if and only if $f = g$ almost everywhere.

Proof.

a. We know $\int_X |f - g| d\mu = 0$ if and only if $|f - g| = 0$ almost everywhere, if and only if $f = g$ almost everywhere.

b. If $f = g$ almost everywhere, then obviously $\int_E f d\mu = \int_E g d\mu$ for all $E \in \mathcal{A}$. The other direction is left as an exercise.

□

By Theorem 2.44, we know if $f = g$ almost everywhere, then $\int_X f d\mu = \int_X g d\mu$.

Example 2.45. Let $X = [0, 1]$, set $f \equiv 1$ on X and

$$g(x) = \begin{cases} 1, & x \in [0, 1] \backslash \mathbb{Q} \\ 0, & x \in \mathbb{Q} \cap [0, 1] \end{cases}$$

on X, then $f = g$ almost everywhere. Therefore, in $L^1(X, \mathcal{A}, \mathcal{M})$, we say $f = g$. Note that in the sense of Riemann, they do not agree in terms of Riemann integrability, which is designed only for continuous functions in general.

Theorem 2.46 (Dominated Convergence Theorem)**.** Let $\{f_n\}_{n\geqslant 1}$ be a sequence in $L^1(X)$ such that

a. $\lim\limits_{n\to\infty} f_n = f$ almost everywhere,

b. there exists integrable function $g \in L^1$ such that $|f_n| \leqslant g$ for all $n \in \mathbb{N}$,

then $\int_X \lim\limits_{n\to\infty} f_n d\mu = \lim\limits_{n\to\infty} \int_X f_n d\mu$.

Proof. First, note that $f \in L^1$: since $|f| = \lim\limits_{n\to\infty} |f_n| \leqslant g \in L^1$, so $\int_X |f| d\mu \leqslant \int_X |g| d\mu < \infty$, hence $f \in L^1(X)$ by definition. Now note that $|f_n| \leqslant g$ if and only if $-g \leqslant f_n \leqslant g$ almost everywhere, then $f_n + g \in L^+$ for all $n \in \mathbb{N}$. By Theorem 2.36, we know

$$\begin{aligned}
\int_X \liminf_{n\to\infty} f_n d\mu + \int_X g d\mu &= \int_X \left(\liminf_{n\to\infty} f_n d\mu\right) + g \\
&= \int \liminf_{n\to\infty} (f_n + g) d\mu \\
&\leqslant \liminf_{n\to\infty} \int_X (f_n + g) d\mu \\
&= \liminf_{n\to\infty} \left(\int_X f_n d\mu + \int_X g d\mu\right) \\
&= \left(\liminf_{n\to\infty} \int_X f_n d\mu\right) + \int_X g d\mu,
\end{aligned}$$

therefore $\int_X \liminf\limits_{n\to\infty} f_n d\mu \leqslant \liminf\limits_{n\to\infty} \int_X f_n d\mu$. Since $g - f_n \in L^+$, then by Theorem 2.36 again, we know

$$\begin{aligned}
\int_X g d\mu - \int_X \limsup_{n\to\infty} f_n d\mu &= \int_X (g - \limsup_{n\to\infty} f_n) d\mu \\
&= \int_X (g + \liminf_{n\to\infty}(-f_n)) d\mu \\
&= \int_X \liminf_{n\to\infty} (g - f_n) d\mu \\
&\leqslant \liminf_{n\to\infty} \int_X (g - f_n) d\mu \\
&= \liminf_{n\to\infty} (\int_X g d\mu - \int_X f_n d\mu) \\
&= \int_X g d\mu - \limsup_{n\to\infty} \int_X f_n d\mu,
\end{aligned}$$

hence $\int_X \limsup\limits_{n\to\infty} f_n d\mu \geqslant \limsup\limits_{n\to\infty} \int_X f_n d\mu$. This gives

$$\int_X f d\mu = \int_X \limsup_{n\to\infty} f_n d\mu \geqslant \limsup_{n\to\infty} \int_X f_n d\mu \geqslant \liminf_{n\to\infty} \int_X f_n d\mu \geqslant \int_X \liminf_{n\to\infty} f_n d\mu = \int_X f d\mu$$

and forces

$$\limsup_{n\to\infty} \int_X f_n d\mu = \liminf_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu.$$

In particular, the limit exists, hence

$$\int_X \lim_{n\to\infty} f_n d\mu = \lim_{n\to\infty} \int_X f_n d\mu.$$

□

Theorem 2.47. Suppose that $\{f_j\}_{j\geqslant 1}$ is a sequence in L^1 such that $\sum_{j=1}^{\infty} \int_X |f_j| d\mu < \infty$, then $\sum_{j=1}^{\infty} f_j$ converges almost everywhere to a function in L^1 such that

$$\int_X \sum_{j=1}^{\infty} f_j d\mu = \sum_{j=1}^{\infty} \int_X f_j d\mu.$$

Proof. Let $g(x) = \sum_{j=1}^{\infty} |f_j(x)|$ for all $x \in X$, then

$$\int_X g d\mu = \int_X \sum_{j=1}^{\infty} |f_j| d\mu = \sum_{j=1}^{\infty} \int_X |f_j| d\mu < \infty.$$

Therefore $g \in L^1$. For all $n \in \mathbb{N}$, we set $g_n = \sum_{j=1}^{n} f_j$ and therefore $|g_n| \leqslant g$ for all $n \in \mathbb{N}$. Now by Theorem 2.46 we know

$$\begin{aligned}
\int_X \sum_{j=1}^{\infty} f_j d\mu &= \int_X \lim_{n\to\infty} g_n d\mu \\
&= \lim_{n\to\infty} \int_X g_n d\mu \\
&= \lim_{n\to\infty} \int_X \sum_{j=1}^{n} f_j d\mu \\
&= \lim_{n\to\infty} \sum_{j=1}^{n} \int_X f_j d\mu \\
&= \sum_{j=1}^{\infty} \int_X f_j d\mu.
\end{aligned}$$

□

Theorem 2.48. Let $f \in L^1$. For any $\varepsilon > 0$, there exists a simple function $\varphi \in L^1$ such that $||f - \varphi||_1 < \varepsilon$.

Proof. Note that $|f| \in L^+$, therefore there exists a sequence $\{\varphi_n\}_{n\geqslant 1}$ of simple functions such that $0 \leqslant |\varphi_1| \leqslant \cdots \leqslant |\varphi_n| \leqslant \cdots \leqslant |f|$ with $\lim_{n\to\infty} \varphi_n = f$. Therefore

$$|f - \varphi_n| \leqslant |f| + |\varphi_n| \leqslant 2|f| \in L^1.$$

By Theorem 2.46, we have

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

hence $\lim_{n\to\infty} \int_X |f - \varphi_n| d\mu = 0$. Now for any $\varepsilon > 0$, there exists some $N \in \mathbb{N}$ such that $\int_X |f - \varphi_N| < \varepsilon$. Take $\varphi = \varphi_N$, we have $||f - \varphi||_1 < \varepsilon$ as desired. □

Theorem 2.49. Let $f : [a, b] \to \mathbb{R}$ be a bounded function where $a, b \in \mathbb{R}$, then f is Riemann integrable if and only if the Lebesgue measure $\mathfrak{m}(\{x \in [a, b] : f \text{ is discontinuous}\}) = 0$.

Example 2.50. $\chi_{\mathbb{Q}}$ is not Riemann integrable on $[0,1]$ because it is discontinuous everywhere.

Example 2.51. Let $S = \{\frac{1}{n} : n \in \mathbb{N}\}$, then χ_S is Riemann integrable on $[0,1]$ because

$$\mathfrak{m}(\{x \in [0,1] : \chi_S \text{ is discontinuous at } x\} = \mathfrak{m}(S) = 0.$$

Example 2.52. Let $\mathscr{C}$ be the Cantor set, c.f., Definition 1.64, then $\chi_{\mathscr{C}}$ is Riemann integrable on $[0,1]$.

Proof. Given a partition $\mathcal{P}$ of $[a,b]$

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b,$$

recall that $||\mathcal{P}|| = \max\{|x_j - x_{j-1}| : 1 \leqslant j \leqslant n\}$, then we have two simple functions

$$U_{\mathcal{P}}(x) = \sum_{j=1}^{n} \sup_{x \in [x_{j-1},x_j)} f(x) \cdot \chi_{[x_{j-1},x_j)}(x)$$

and

$$L_{\mathcal{P}}(x) = \sum_{j=1}^{n} \inf_{x \in [x_{j-1},x_j)} f(x) \cdot \chi_{[x_{j-1},x_j)}(x).$$

We try to create a Riemann sum with respect to these two functions. We have

$$\begin{aligned}\int_{[a,b]} U_{\mathcal{P}} d\mathfrak{m} &= \sum_{j=1}^{n} \sup_{x \in [x_{j-1},x_j)} f(x)(x_j - x_{j-1}) \\ &:= U(f,\mathcal{P})\end{aligned}$$

and

$$\begin{aligned}\int_{[a,b]} L_{\mathcal{P}} d\mathfrak{m} &= \sum_{j=1}^{n} \inf_{x \in [x_{j-1},x_j)} f(x)(x_j - x_{j-1}) \\ &:= L(f,\mathcal{P}).\end{aligned}$$

Let us take a sequence of partitions $\{\mathcal{P}_n\}_{n \geqslant 1}$ such that

$$\mathcal{P}_1 \subseteq \mathcal{P}_2 \subseteq \cdots \subseteq \mathcal{P}_n \subseteq \cdots$$

and $\lim\limits_{n\to\infty} ||\mathcal{P}_n|| = 0$. Recall that f is Riemann integrable if and only if $L(f) =: \lim\limits_{n\to\infty} L(f,\mathcal{P}_n) = \lim\limits_{n\to\infty} U(f,\mathcal{P}_n) := U(f)$. We can bound f by the simple functions

$$L_{\mathcal{P}_1} \leqslant \cdots \leqslant L_{\mathcal{P}_n} \leqslant \cdots \leqslant f \leqslant \cdots \leqslant U_{\mathcal{P}_n} \leqslant \cdots \leqslant U_{\mathcal{P}_1}.$$

Therefore we get a monotone sequence and take the limit $n \to \infty$ since it exists in $\bar{\mathbb{R}}$, then $L := \lim\limits_{n\to\infty} L_{\mathcal{P}_n}$ and $U = \lim\limits_{n\to\infty} U_{\mathcal{P}_n}$ are $\bar{\mathbb{R}}$-valued functions, and are measurable. Since the limit preserves the order, we know that $L \leqslant f \leqslant U$. In particular, there exists some constant C such that

$$|U_{\mathcal{P}_n}| \leqslant \sup_{x\in[a,b]} |f(x)| \leqslant C$$

and

$$|L_{\mathcal{P}_n}| \leqslant \inf_{x\in[a,b]} |f(x)| \leqslant C$$

for all $n \in \mathbb{N}$. Therefore we get $|U| \leqslant C$ and $|L| \leqslant C$, where $C \in L^1([a,b])$. By Theorem 2.46, we have that

$$\int_{[a,b]} U d\mathfrak{m} = \int_{[a,b]} \lim_{n\to\infty} U_{\mathcal{P}_n} d\mathfrak{m}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \int\limits_{[a,b]} U_{\mathcal{P}_n} d\mathfrak{m} \\
&= \lim_{n\to\infty} U(f,\mathcal{P}_n) \\
&= U(f)
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\int\limits_{[a,b]} L d\mathfrak{m} &= \int\limits_{[a,b]} \lim_{n\to\infty} L_{\mathcal{P}_n} d\mathfrak{m} \\
&= \lim_{n\to\infty} \int\limits_{[a,b]} L_{\mathcal{P}_n} d\mathfrak{m} \\
&= \lim_{n\to\infty} L(f,\mathcal{P}_n) \\
&= L(f).
\end{aligned}
$$

Therefore, we know

$$
\begin{aligned}
f \text{ is Riemann integrable } &\iff U(f) = L(f) = \int\limits_a^b f dx \text{ in the Riemann sense} \\
&\iff \int\limits_{[a,b]} U d\mathfrak{m} = \int\limits_{[a,b]} L d\mathfrak{m} \\
&\iff \int\limits_{[a,b]} (U-L) d\mathfrak{m} = 0 \\
&\iff \mathfrak{m}(\{x \in [a,b] : U(x) > L(x)\}) = 0.
\end{aligned}
$$

Claim 2.53. If $f : [a,b] \to \mathbb{R}$ is a bounded Riemann integrable function, then f is Lebesgue integrable. Moreover,

$$
\int\limits_{[a,b]} f d\mathfrak{m} = \int\limits_a^b f dx.
$$

Subproof. We have

$$
\begin{aligned}
\{x \in [a,b] : f(x) \neq U(x)\} &\subseteq \{x \in [a,b] : L(x) \neq U(x)\} \\
&= \{x \in [a,b] : U(x) > L(x)\}
\end{aligned}
$$

and therefore

$$
\mathfrak{m}(\{x \in [a,b] : f(x) \neq U(x)\}) = 0.
$$

Hence,

$$
\begin{aligned}
\int\limits_{[a,b]} f d\mathfrak{m} &= \int\limits_{[a,b]} U d\mathfrak{m} \\
&= U(f) \\
&= \int\limits_a^b f dx.
\end{aligned}
$$

■

It now suffices to prove the following claim.

Claim 2.54. $\mathfrak{m}(\{x \in [a,b] : U(x) > L(x)\}) = 0$ if and only if $\mathfrak{m}(\{x \in [a,b] : f \text{ is discontinuous at } x\}) = 0$.

Subproof. For any $A \subseteq [a,b]$, we define the oscillation of f to be $\omega_f(A) = \sup\limits_{x\in A} f(x) - \inf\limits_{x\in A} f(x)$. Now f is continuous at x_0 if and only if the oscillation of f at x_0 is $\Omega_f(x_0) := \lim\limits_{\delta\to 0} \omega_f((x_0-\delta, x_0+\delta)) = 0$. Note that the function is monotone with respect to δ, therefore the limit exists. Let $x \in [a,b] \setminus \bigcup\limits_{n=1}^{\infty} \mathcal{P}_n$ with a zero-measure subset removed. Denote the subinterval in $\mathcal{P}_n$ containing x by I_n, then

$$\begin{aligned}\Omega_f(x) &= \lim_{n\to\infty} \omega_f(I_n)\\ &= \lim_{n\to\infty}[U_{\mathcal{P}_n}(x) - L_{\mathcal{P}_n}(x)]\\ &= U(x) - L(x).\end{aligned}$$

Therefore,

$$\begin{aligned} f \text{ is continuous at } x &\iff \Omega_f(x) = 0\\ &\iff U(x) = L(x)\\ &\iff U(x) = L(x),\end{aligned}$$

and we conclude that

$$\mathfrak{m}(\{x \in [a,b] : f \text{ is discontinuous at } x\} = \mathfrak{m}(\{x \in [a,b] : U(x) > L(x)\}$$

as desired. ■

□

2.4 Modes of Convergences

Definition 2.55. We say $\{f_n\}_{n\geqslant 1}$ converges to f uniformly on E if $\lim\limits_{n\to\infty} \sup\limits_{x\in E} |f_n(x) - f(x)| = 0$, and write $f_n \rightrightarrows f$ on E as $n \to \infty$.

Remark 2.56. If $f_n \rightrightarrows f$ on E, then $f_n \to f$ on E.

Definition 2.57. We say $\{f_n\}_{n\geqslant 1}$ converges to f in L^1 if $\lim\limits_{n\to\infty} ||f_n - f||_1 = 0$, and write $f_n \xrightarrow{L^1} f$ as $n \to \infty$.

Definition 2.58. We say that $\{f_n\}_{n\geqslant 1}$ converges to f in measure μ if for all $\varepsilon > 0$, $\lim\limits_{n\to\infty} \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) = 0$. We write $f_n \xrightarrow{\mu} f$ as $n \to \infty$.

We now study the relations between different types of convergence.

Theorem 2.59. If $f_n \xrightarrow{L^1} f$, then $f_n \xrightarrow{\mu} f$.

Proof. Pick $\varepsilon > 0$, and let $E_n = \{x \in X : |f_n(x) - f(x)| > \varepsilon\}$. Now

$$\begin{aligned}\varepsilon\mu(E_n) &= \int\limits_{E_n} \varepsilon d\mu\\ &\leqslant \int\limits_{E_n} |f_n - f| d\mu\\ &\leqslant \int\limits_X |f_n - f| d\mu\end{aligned}$$

$$= ||f_n - f||_1,$$

therefore $0 \leqslant \mu(E_n) \leqslant \frac{1}{\varepsilon}||f_n - f||_1$. Let $n \to \infty$, then $0 \leqslant \lim_{n\to\infty} \mu(E_n) \leqslant 0$ so by squeeze theorem we have $\lim_{n\to\infty} \mu(E_n) = 0$. By definition, $f_n \xrightarrow{\mu} f$. □

Example 2.60. Let $f_n = \frac{\chi_{(0,n)}}{n}$ be a function on $\mathbb{R}$, then $f_n \rightrightarrows 0$ on $\mathbb{R}$. Thus, $f_n \to 0$ on $\mathbb{R}$ pointwise. Moreover, $f_n \xrightarrow{u} 0$, but $f_n \not\xrightarrow{L^1} 0$, thus the converse of Theorem 2.59 is not true:

$$\begin{aligned}
\lim_{n\to\infty} \int_X |f_n - 0| d\mathfrak{m} &= \lim_{n\to 0} \int_X |f_n| d\mathfrak{m} \\
&= \frac{1}{n} \int_X \chi_{(0,n)} d\mathfrak{m} \\
&= \frac{n}{n} \\
&= 1.
\end{aligned}$$

Example 2.61. Let $f_n = \chi_{(n,n+1)}$ be a function on $\mathbb{R}$, then $f_n \to 0$ on $\mathbb{R}$ pointwise, but $f_n \not\xrightarrow{\mathfrak{m}} 0$ does not converge to 0 on measure $\mathfrak{m}$: for any $\varepsilon > 0$,

$$\mathfrak{m}(\{x \in X : |\chi_{(n,n+1)}(x)| > \varepsilon\}) = \mathfrak{m}(\{x \in (n, n+1) : \varepsilon < 1\}),$$

so for any $1 > \varepsilon > 0$, taking the limit $n \to 0$ gives

$$\lim_{n\to\infty} \mathfrak{m}(\{x \in X : |\chi_{(n,n+1)}(x)| > \varepsilon\}) = 1.$$

Definition 2.62. Let $\{f_n\}_{n\geqslant 1}$ be a sequence of measurable functions. We say the sequence is Cauchy in measure if for all $\sigma > 0$, for all $\varepsilon > 0$, there exists some $N \in \mathbb{N}$ such that $\mu(\{x \in X : |f_n(x) - f_m(x)| > \varepsilon\}) < \sigma$ for all $m, n \geqslant N$.

Equivalently, the sequence is Cauchy in measure if for any $\varepsilon > 0$,

$$\lim_{m,n\to\infty} \mu(\{x \in X : |f_n(x) - f_m(x)| > \varepsilon\}) = 0.$$

Theorem 2.63. Suppose $\{f_n\}_{n\geqslant 1}$ is Cauchy in measure, then there exists a subsequence $\{f_{n_j}\}_{j\geqslant 1}$ such that $f_{n_j} \to f$ almost everywhere as $j \to \infty$.

Proof. Let $\sigma = \varepsilon = 2^{-j}$ for all $j \in \mathbb{N}$, then there exists $n_j \in \mathbb{N}$ such that $\mu(\{x \in X : |f_{n_{j+1}}(x) - f_{n_j}(x)| > 2^{-j}\}) < 2^{-j}$, therefore we have choices $n_j < n_{j+1}$ for all J. Now we know $\{f_{n_j}\}_{j\geqslant 1}$ is a subsequence, so let $g_j = f_{n_j}$ for all $j \in \mathbb{N}$. Therefore,

$$\mu(\{x \in X : |g_{j+1}(x) - g_j(x)| > 2^{-j}\}) \leqslant 2^{-j}$$

for all j. Let $E_j = \{x \in X : |g_{j+1}(x) - g_j(x)| > 2^{-j}\}$, then $\mu(E_j) \leqslant 2^{-j}$.

Claim 2.64. For all $k \in \mathbb{N}$ and $F_k = \bigcup_{j=k}^{\infty} E_j$, then $\{g_j\}_{j\geqslant 1}$ is pointwise Cauchy on F_k^c.

Subproof. We show that for $x \in F_k^c$, we have $\lim_{m,n\to\infty} |g_m(x) - g_n(x)| = 0$, which is equivalent to saying for all $\varepsilon > 0$, for all $x \in F_k^c$, there exists $N \in \mathbb{N}$ such that $|g_m(x) - g_n(x)| < \varepsilon$ for all $m, n \geqslant N$. Since $x \in F_k^c$, then $x \in \left(\bigcup_{j=k}^{\infty} E_j\right)^c = \bigcap_{j=k}^{\infty} E_j^c$, so for all $j \geqslant k$ we know $x \in E_j^c$, which is equivalent to saying that for all $j \geqslant k$, $|g_{j+1}(x) - g_j(x)| < 2^{-j}$. Without loss of generality, take arbitrary $m > n \geqslant k$, we get

$$|g_m(x) - g_n(x)| = \left|\sum_{j=n}^{m-1} [g_{j+1}(x) - g_j(x)]\right|$$

$$\begin{aligned}
&\leqslant \sum_{j=n}^{m+1} |g_{j+1}(x) - g_j(x)| \\
&\leqslant \sum_{j=n}^{m+1} 2^{-j} \\
&\leqslant 2^{1-n}.
\end{aligned}$$

Taking $n \to \infty$, we forces $\lim\limits_{m,n\to\infty} |g_m(x) - g_n(x)| = 0$, as desired. ■

Claim 2.65. Let $F = \bigcap\limits_{k=1}^{\infty} F_k$, then $\mu(F) = 0$.

Subproof. We know that for all $n \in \mathbb{N}$,

$$\begin{aligned}
\mu(F) &\leqslant \mu(F_n) \\
&= \mu\left(\bigcup_{j=n}^{\infty} F_j\right) \\
&\leqslant \sum_{j=n}^{\infty} \mu(E_j) \\
&\leqslant \sum_{j=n}^{\infty} 2^{-j} \\
&\leqslant 2^{1-n},
\end{aligned}$$

so for $n \to \infty$, we forces $\mu(F) = 0$. ■

Claim 2.66. If $x \in F^c$, then $\{g_j(x)\}_{j\geqslant 1}$ is a pointwise Cauchy sequence.

Subproof. For any $x \in F^c$, we know $x \in (\bigcap\limits_{k=1}^{\infty} F_k)^c = \bigcup\limits_{k=1}^{\infty} F_k^c$, therefore $x \in F_k^c$ for some $k \in \mathbb{N}$. By Claim 2.64, we conclude that $\{g_j(x)\}_{j\geqslant 1}$ is a pointwise Cauchy sequence. ■

Therefore, for any $x \in F^c$, we know $\{g_j(x)\}$ is Cauchy, so $\lim\limits_{j\to\infty} g_j(x)$ exists in $\mathbb{R}$. Let f be given by

$$f(x) = \begin{cases} \lim\limits_{j\to\infty} g_j(x), & x \in F^c \\ 0, & x \in F \end{cases}$$

then $\{g_j\}$ converges to f almost everywhere. Consider $\{g_j\}_{j\geqslant 1}$ as the said subsequence $\{f_{n_j}\}_{j\geqslant 1}$ of $\{f_n\}_{n\geqslant 1}$, then we are done. □

Theorem 2.67 (Cauchy Criterion)**.** The sequence $\{f_n\}_{n\geqslant 1}$ is Cauchy in measure if and only if there is a measurable function f such that $f_n \xrightarrow{\mu} f$.

Proof.

($\Leftarrow$): Suppose $f_n \xrightarrow{\mu} f$, and set $\varepsilon > 0$, then we want to show that $\lim\limits_{m,n\to 0} \mu(\{x \in X : |f_m(x) - f_n(x)| > \varepsilon\}) = 0$. We know, for any $x \in X$ that lies in the given subset, that

$$\begin{aligned}
\varepsilon &< |f_m(x) - f_n(x)| \\
&= |(f_m(x) - f(x)) + (f(x) - f_n(x))| \\
&\leqslant |f_m(x) - f(x)| + |f_n(x) - f(x)|,
\end{aligned}$$

therefore either $|f_m(x) - f(x)| > \frac{\varepsilon}{2}$ or $|f_n(x) - f(x)| > \frac{\varepsilon}{2}$. Therefore,

$$\{x \in X : |f_m(x) - f_n(x)| > \varepsilon\} \subseteq \{x \in X : |f_m(x) - f(x)| > \frac{\varepsilon}{2}\} \cup \{x \in X : |f_n(x) - f(x)| > \frac{\varepsilon}{2}\}.$$

Hence,

$$\mu(\{x \in X : |f_m(x) - f_n(x)| > \varepsilon\}) \leqslant \mu(\{x \in X : |f_m(x) - f(x)| > \frac{\varepsilon}{2}\}) + \mu(\{x \in X : |f_n(x) - f(x)| > \frac{\varepsilon}{2}\}),$$

but as $m, n \to \infty$, the two measures of the right-hand side converges to 0, which forces the measure on the left also converges to 0.

($\Rightarrow$): Since $\{f_n\}_{n\geqslant 1}$ is Cauchy in measure, then there exists a subsequence $\{g_j\}_{j\geqslant 1} = \{f_{n_j}\}_{j\geqslant 1}$ such that $\lim\limits_{j\to\infty} f_{n_j} = \lim\limits_{j\to\infty} g_j = f$ almost everywhere.

Claim 2.68. $g_j \xrightarrow{\mu} f$.

Subproof. Again, let $E_j = \{x \in X : |g_{j+1}(x) - g_j(x)| > 2^{-j}\}$, and set $F_k = \bigcup\limits_{j=k}^{\infty} E_j$ as in Theorem 2.63, then we know for all $x \in F_k^c$, we have

$$|g_m(x) - g_j(x)| \leqslant 2^{1-j}$$

for all $m, j \geqslant k$. Now let $m \to \infty$, then

$$|f(x) - g_j(x)| \leqslant 2^{1-j}$$

for any $j \geqslant k$ and $x \in F_k^c$. Fix $\varepsilon > 0$. For large enough j, we know $2^{1-j} < \varepsilon$ and therefore satisfies

$$\{x \in X : |g_j(x) - f(x)| > \varepsilon\} = \{x \in F_j : |g_j(x) - f(x)| > \varepsilon\} \cup \{x \in F_j^c : |g_j(x) - f(x)| > \varepsilon\}.$$

But note that for any $x \in F_j^c$, $|g_j(x) - f(x)| \leqslant 2^{1-j} < \varepsilon$, which forces the second set to be empty, therefore we have

$$\{x \in X : |g_j(x) - f(x)| > \varepsilon|\} = \{x \in F_j : |g_j(x) - f(x)| > \varepsilon\} \subseteq F_j.$$

Taking the measure, we have

$$\begin{aligned}\mu(\{x \in X : |g_j(x) - f(x)| > \varepsilon\}) &\leqslant \mu(F_j) \\ &\leqslant 2^{1-j} \\ &\to 0\end{aligned}$$

as $j \to \infty$. Therefore, $g_j \xrightarrow{\mu} f$. ■

Claim 2.69. $f_n \xrightarrow{\mu} f$.

Subproof. We know that

$$\begin{aligned}\varepsilon &< |f_n(x) - f(x)| \\ &< |f_n(x) - g_j(x)| + |g_j(x) - f(x)| \\ &\leqslant |f_n(x) - g_j(x)| + |g_j(x) - f(x)|\end{aligned}$$

and therefore either $|f_n(x) - g_j(x)| > \frac{\varepsilon}{2}$ or $|g_j(x) - f(x)| > \frac{\varepsilon}{2}$. Therefore,

$$\{x \in X : |f_n(x) - f(x)| > \varepsilon\} \subseteq \{x \in X : |f_n(x) - g_j(x)| > \frac{\varepsilon}{2}\} \cup \{x \in X : |g_j(x) - f(x)| > \frac{\varepsilon}{2}\}.$$

Taking the measure, we know that

$$\mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) \leqslant \mu(\{x \in X : |f_n(x) - g_j(x)| > \frac{\varepsilon}{2}\}) + \mu(\{x \in X : |g_j(x) - f(x)| > \frac{\varepsilon}{2}\}).$$

Let $j, n \to \infty$, then $\mu(\{x \in X : |g_j(x) - f(x)| > \frac{\varepsilon}{2}\}) \to 0$ since $g_j \xrightarrow{\mu} f$, and $\mu(\{x \in X : |f_n(x) - g_j(x)| > \frac{\varepsilon}{2}\}) \to 0$ since $\{f_n\}_{n \geqslant 1}$ is Cauchy in measure. Therefore, $\mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) \to 0$ as $j, n \to \infty$. In particular, that means

$$\lim_{n \to \infty} \mu(\{x \in X : |f_n(x) - f(x)| > \varepsilon\}) = 0.$$

■

□

Theorem 2.70. Suppose $f_n \xrightarrow{\mu} f$ in measure, then there exists a subsequence $\{f_{n_j}\}_{j \geqslant 1}$ such that $f_{n_j} \to f$ almost everywhere.

Proof. Since $f_n \xrightarrow{\mu} f$, then $\{f_n\}_{n \geqslant 1}$ is Cauchy in measure, therefore by Theorem 2.63 there exists a subsequence $\{f_{n_j}\}_{j \geqslant 1}$ such that $f_{n_j} \to f$ almost everwhere. □

Corollary 2.71. If $\{f_n\}_{n \geqslant 1}$ converges to f in L^1, i.e., $||f_n - f||_1 \to 0$, then there exists a subsequence $\{f_{n_j}\}_{j \geqslant 1}$ such that $f_{n_j} \to f$ almost everywhere.

Proof. This is obvious from Theorem 2.70. □

Definition 2.72. We say $\{f_n\}_{n \geqslant 1}$ converges to f almost uniformly on X if for any $\varepsilon > 0$, there exists a subset $E \subseteq X$ such that $\mu(E) < \varepsilon$ and $f_n \rightrightarrows f$ on E^c.

Theorem 2.73 (Egoroff)**.** Suppose that $\mu(X) < \infty$ and $f_n \to f$ almost everywhere on X, then $\{f_n\}_{n \geqslant 1}$ converges to f almost uniformly.

Proof. Without loss of generality, suppose $f_n \to f$ for all $x \in X$. For any $k \in \mathbb{N}$ and $n \in \mathbb{N}$, we define

$$E_n(k) = \bigcup_{m=n}^{\infty} \{x \in X : |f_m(x) - f(x)| > \frac{1}{k}\}.$$

Claim 2.74. Given any k, $E_n(k) \supseteq E_{n+1}(k)$ for all $n \in \mathbb{N}$.

Subproof. This follows from the definition of $E_n(k)$. ■

Claim 2.75. $\bigcap\limits_{n \geqslant 1} E_n(k) = \varnothing$.

Subproof. Suppose not, then there exists $x \in \bigcap\limits_{n \geqslant 1} E_n(k)$, hence $x \in E_n(k)$ for all $n \in \mathbb{N}$. By definition, we know there is a subsequence $\{f_{n_j}\}_{j \geqslant 1}$ of $\{f_n\}_{n \geqslant 1}$ such that $|f_{n_j}(x) - f(x)| > \frac{1}{k}$ for any $j \in \mathbb{N}$. Let $j \to \infty$, we know $0 = \lim\limits_{j \to \infty} |f_{n_j}(x) - f(x)| \geqslant \frac{1}{k}$, contradiction. ■

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

$$\begin{aligned} \lim_{n \to \infty} \mu(E_n(k)) &= \mu\left(\bigcap_{n=1}^{\infty} E_n(k)\right) \\ &= \mu(\varnothing) \\ &= 0. \end{aligned}$$

For arbitrary $\varepsilon > 0$, there exists some $n_k \in \mathbb{N}$ such that $\mu(E_{n_k}(k)) < \varepsilon \cdot 2^{-k}$. Take $E = \bigcup\limits_{k \geqslant 1} E_{n_k}(k)$, then

$$\mu(E) \leqslant \sum_{k \geqslant 1} \mu(E_{n_k}(k)) < \sum_{k \geqslant 1} \varepsilon \cdot 2^{-k} \leqslant \varepsilon.$$

Finally, we need to show that $f_n \rightrightarrows f$ on E^c. Take $x \in E^c$, then $x \in \bigcap\limits_{k \geqslant 1} [E_{n_k}(k)]^c$, therefore $x \in E_{n_k}(k)^c$ for all $k \in \mathbb{N}$.

Recall that

$$(E_{n_k}(k))^c = \bigcap_{m \geqslant n_k} \{x \in X : |f_m(x) - f(x)| \leqslant \frac{1}{k}\},$$

Thus, if $x \in E^c$, we know $|f_n(x) - f(x)| \leqslant \frac{1}{k}$ for all $k \in \mathbb{N}$ and $n \geqslant n_k$, hence $\sup\limits_{x \in E^c} |f_n(x) - f(x)| \leqslant \frac{1}{k}$ for all $k \in \mathbb{N}$ and $n \geqslant n_k$, therefore

$$0 \leqslant \lim_{n\to\infty} \sup_{x\in E^c} |f_n(x) - f(x)| \leqslant \frac{1}{k}.$$

In particular, this limits tends to 0 when $k \to \infty$. This shows that $\lim\limits_{n\to\infty} \sup\limits_{x\in E^c} |f_n(x) - f(x)| = 0$, in other words $f_n \rightrightarrows f$ on E^c. Therefore, f_n converges almost uniformly to f on E^c. $\square$

Remark 2.76. If f_n converges to f almost uniformly on X, then $f_n \to f$ almost everywhere on X and $f_n \xrightarrow{\mu} f$ on X.

Remark 2.77. The condition $\mu(X) < \infty$ in Theorem 2.73 is necessary. To see this, consider the measure space $(\mathbb{R}, \mathcal{L}, \mathfrak{m})$, and consider $f_n = \chi_{[n,\infty)}$ for all $n \in \mathbb{N}$. Now $f_n \to 0$ converges, but f_n does not converge to 0 in measure $\mathfrak{m}$. Indeed,

$$\begin{aligned} \mathfrak{m}(\{x \in \mathbb{R} : |f_n(x)| > \frac{1}{2}\}) &= \mathfrak{m}(\{x \in [n,\infty)\}) \\ &= \infty \not\to 0. \end{aligned}$$

By Remark 2.76, $\{f_n\}_{n\geqslant 1}$ does not converge to 0 almost uniformly on $\mathbb{R}$.

Remark 2.78. The hypothesis $\mu(X) < \infty$ in Theorem 2.73 can be replaced by $|f_n| \leqslant g$ for all $n \in \mathbb{N}$ and $g \in L^1(X)$.

Theorem 2.79. Let f be any complex-valued measurable function on E with $\mu(E) < \infty$. Then for any $\varepsilon > 0$, there exist a simple function φ and a measurable set $F \subseteq E$ such that

1. $\mu(E \backslash F) < \varepsilon$, and
2. $|f(x) - \varphi(x)| < \varepsilon$ for all $x \in F$.

Proof. Without loss of generality, assume $f \in L^+$. Let $\varphi_n(x) = \sum\limits_{k=0}^{2^{2n}-1} k2^{-n}\chi_{E_{n,k}}(x) + 2^n\chi_{F_n}(k)$, where $E_{n,k} = \{x \in E : f(x) \in (k2^{-n}, (k+1)2^{-n}]\}$ and $F_n = \{x \in E : f(x) > 2^n\}$. Therefore, $F_n \supseteq F_{n+1}$ and $\mu(F_n) \leqslant \mu(E) < \infty$ for all $n \in \mathbb{N}$, so by continuity from above we have

$$\begin{aligned} \lim_{n\to\infty} \mu(F_n) &= \mu\left(\bigcap_{n\geqslant 1} F_n\right) \\ &= \mu(\varnothing) \\ &= 0. \end{aligned}$$

For any $\varepsilon > 0$, there exists $N_1 \in \mathbb{N}$ such that $\mu(F_n) < \varepsilon$ for all $n \geqslant N_1$. Recall that $|\varphi_n(x) - f(x)| \leqslant 2^{-n}$ for all $x \notin F_n$, then $\sup\limits_{x\in F_n^c} |\varphi_n(x) - f(x)| \leqslant 2^{-n}$, then by squeeze theorem we have $\lim\limits_{n\to\infty} \sup\limits_{x\in F_n^c} |\varphi_n(x) - f(x)| = 0$. Hence, for any $\varepsilon > 0$, there exists $N_2 \in \mathbb{N}$ such that $\sup\limits_{x\in F_n^c} |\varphi_n(x) - f(x)| < \varepsilon$ for all $n \geqslant N_2$. Let $N = \max\{N_1, N_2\}$, then $|\varphi_N(x) - f(x)| < \varepsilon$ for all $x \notin F_N$, and $\mu(F_N) < \varepsilon$. Define $\varphi = \varphi_N$ to be the said simple function, and let $F = E \backslash F_N$. $\square$

Theorem 2.80. Let $\mu(X) < \infty$ and f be a complex-valued measurable function on X. For any $\varepsilon > 0$, there exists $0 < M \in \mathbb{R}$ and a measurable set $E \subseteq X$ such that $|f(x)| < M$ for all $x \in E$ and $\mu(E^c) < \varepsilon$.

Proof. By Theorem 2.79, for any $\varepsilon > 0$, there exists a simple function φ and a measurable set $E \subseteq X$ such that $\mu(E^c) < \varepsilon$ and $|f(x) - \varphi(x)| < \varepsilon$ for all $x \in E$. Using the triangle inequality and the fact that φ is a simple function on E, we know for any $x \in E$ that

$$\begin{aligned} |f(x)| &\leqslant |f(x) - \varphi(x)| + |\varphi(x)| \\ &< \varepsilon + |\varphi(x)| \\ &< \varepsilon + \sup_{x\in E} |\varphi(x)| \\ &=: M \in \mathbb{R}. \end{aligned}$$

$\square$

Theorem 2.81. For any $f \in L^1(\mathbb{R}, \mathcal{A}, \mu)$ where μ is a Lebesgue-Stieltjes measure, then for any $\varepsilon > 0$, there exists a continuous function g on $\mathbb{R}$ such that $||f - g||_1 < \varepsilon$.

Proof. For any $\varepsilon > 0$, there exists a simple function $\varphi \in L^1$ such that $||f - \varphi||_1 < \varepsilon$. Let us write $\varphi(x) = \sum_{j=1}^{n} a_j \chi_{E_j}$, where each $a_j \neq 0$, and each $\mu(E_j) < \infty$ for all j. We can replace E_j by a finite union of disjoint open intervals $I_k^{(j)}$ for each j, then $\mu\left(E_j \Delta \left(\bigcup_{k=1}^{K} I_k^{(j)}\right)\right) < \frac{\varepsilon}{2^j |a_j|}$. Therefore, χ_{E_j} can be replaced by $\chi_{\bigcup_{j=1}^{K} I_k^{(j)}}$, which can then be replaced by continuous functions g_j, where we replace the function upon intervals on $I_k^{(j)}$ for each k, such that $g = \sum_{i=1}^{n} g_j$. This gives the desired function g. $\square$

Theorem 2.82 (Lusin)**.** Let μ be a Lebesgue-Stieltjes measure on $\mathbb{R}$, and let f be any complex-valued function measurable function on E with $\mu(E) < \infty$, then f is almost a continuous function on E in the following sense: for any $\varepsilon > 0$, there exists a function g on E and a measurable set $F \subseteq E$ such that

1. g is continuous on E,
2. $\mu(E \backslash F) < \varepsilon$, and
3. $|f(x) - g(x)| < \varepsilon$ for all $x \in F$.

Proof Sketch.

- By Theorem 2.79, we know any complex-valued function is "almost simple", i.e., close to a simple function $\varphi \in L^1$ on E.
- Since φ is integrable, then by Theorem 2.81, we know continuous functions are dense in L^1, i.e., there exists a sequence $\{g_j\}_{j \geqslant 1}$ of continuous functions such that $||g_j - \varphi||_1 \to 0$ as $j \to \infty$. Here we can replace $|| \cdot ||_1$ by $|| \cdot ||_{L^1(E)}$.
- We can now find a subsequence $\{g_{n_j}\}_{j \geqslant 1}$ of $\{g_j\}_{j \geqslant 1}$ such that $g_{n_j} \to \varphi$ almost everywhere as $j \to \infty$.
- Note that limit of continuous functions may not be continuous, but the limit of uniform continuous functions is continuous, so we can find the continuous function g after applying Theorem 2.73 to $\{g_{n_j}\}_{j \geqslant 1}$.

$\square$

Remark 2.83 (Littlewood's Three Principles on $\mathbb{R}$)**.**

- Every (finite) measurable set in $\mathbb{R}$ is nearly a finite union of intervals.
- Every measurable (complex-valued) function on $\mathbb{R}$ is nearly continuous, c.f., Theorem 2.82.
- Every convergent sequence of measurable functions on a finite measure set is nearly uniformly convergent, c.f., Theorem 2.73.

2.5 Product Measures

We want to define a product measure on the product space $X \times Y = \{(x, y) : x \in X, y \in Y\}$.

Definition 2.84. Let $(X, \mathcal{A}, \mu_1)$ and $(Y, \mathcal{B}, \mu_2)$ be two measure spaces. For $A \in \mathcal{A}$ and $B \in \mathcal{B}$, we can define a rectangle $A \times B = \{(x, y) : x \in A, y \in B\}$.

Definition 2.85. The product σ-algebra of $\mathcal{A}$ and $\mathcal{B}$, denoted by $\mathcal{A} \otimes \mathcal{B}$, is the σ-algebra generated by rectangles $A \times B$ for $A \in \mathcal{A}$ and $B \in \mathcal{B}$. Therefore, it is the smallest σ-algebra containing all rectangles.

The goal is now to define a product measure $\mu_1 \times \mu_2$ on $\mathcal{A} \otimes \mathcal{B}$, such that $(\mu_1 \times \mu_2)(A \times B) = \mu_1(A)\mu_2(B)$ for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$. To do so, we create a pre-measure on the product algebra, and then get an outer measure, so by Theorem 1.37 we get a desired measure by restriction.

Lemma 2.86. Let $\mathcal{R}_0$ be the collection of finite disjoint unions of rectangles, then $\mathcal{R}_0$ is an algebra.

Proof. Recall that $(A \times B)^c = (X \times B^c) \cup (A^c \times Y)$, which is a union of two rectangles, therefore $\mathcal{R}_0$ is closed under complements if it is closed under finite union. Note that $(A \times B) \cap (E \times F) = (A \cap E) \times (B \cap F)$, therefore $\mathcal{R}_0$ is closed under finite intersection. This shows that $\mathcal{R}_0$, as a family of finite disjoint union of rectangles, is an algebra. □

Definition 2.87. Let $E \in \mathcal{R}_0$, then we can write $E = \bigcup_{j=1}^{n} (A_j \times B_j)$ for $A_j \in \mathcal{A}$ and $B_j \in \mathcal{B}$ such that A_j's and B_j's are disjoint. Now define $\pi(E) = \sum_{j=1}^{n} \mu_1(A_j)\mu_2(B_j)$. In this definition, we set $0 \cdot \infty = 0$.

Lemma 2.88. π is a pre-measure on $\mathcal{R}_0$.

Proof. Left as an exercise. □

For any $E \subseteq X \times Y$, we define $\pi^*(E) = \inf\{\sum_{j=1}^{\infty} \pi(R_j) : R_j \in \mathcal{R}_0, E \subseteq \bigcup_{j=1}^{\infty} R_j\}$, then π^* is the induced outer measure of π on $\mathcal{P}(X \times Y)$.

Definition 2.89. The product measure is defined by $\mu_1 \times \mu_2 = \pi^*|_{\mathcal{A}\otimes\mathcal{B}}$. That is, for any $E \in \mathcal{A}\otimes\mathcal{B}$, we set $(\mu_1 \times \mu_2)(E) = \pi^*(E)$.

Theorem 2.90. Let μ_1, μ_2 be σ-finite, then

1. $\mu_1 \times \mu_2$ is σ-finite,
2. $\mu_1 \times \mu_2$ is the unique measure on $\mathcal{A} \otimes \mathcal{B}$ such that $(\mu_1 \times \mu_2)(A \times B) = \mu_1(A)\mu_2(B)$ for any $A \in \mathcal{A}$ and $B \in \mathcal{B}$.

Proof.

1. Since μ_1 and μ_2 are σ-finite, then we can write $X = \bigcup_{j=1}^{\infty} A_j$ such that $A_j \in \mathcal{A}$ and $\mu_1(A_j) < \infty$ for all j, and similarly $Y = \bigcup_{k=1}^{\infty} B_k$ such that $B_k \in \mathcal{B}$ and $\mu_2(B_k) < \infty$ for all k. Now we know $X \times Y = \bigcup_{j,k}(A_j \times B_k)$. It suffices to show that $A_j \times B_k$ has finite measure over the product measure. By restricting to $\mathcal{R}_0$, we have
$$\begin{aligned}(\mu_1 \times \mu_2)(A_j \times B_k) &= \pi(A_j \times B_k) \\ &= \mu_1(A_j)\mu_2(B_k) \\ &< \infty\end{aligned}$$
for all j, k. Hence, $\mu_1 \times \mu_2$ is σ-finite.
2. This is obvious from properties of σ-finite measures.

□

Given $f : X \times Y \to \mathbb{C}$, we may want to compare $\int_Y \int_X f(x,y) d\mu_1 d\mu_2$, $\int_X \int_Y f(x,y) d\mu_2 d\mu_1$, and $\int_{X \times Y} f d(\mu_1 \times \mu_2)$.

Definition 2.91. Let $E \subseteq X \times Y$, for all $x \in X$ and $y \in Y$, we define the x-section of E to be $E_x = \{y \in Y : (x, y) \in E\}$. Similarly, the y-section of E is $E^y = \{x \in X : (x, y) \in E\}$.

Definition 2.92. Fix $f : X \times Y \to \mathbb{C}$. For any $x \in X$, the x-section of f is defined by $f_x(y) = f(x, y)$ for all $y \in Y$, hence we obtain a function $f_x : Y \to \mathbb{C}$. Similarly, for any $y \in Y$, the y-section of f is defined by $f^y(x) = f(x, y)$ for all $x \in X$, hence we obtain a function $f^y : X \to \mathbb{C}$.

Theorem 2.93.

a. If $E \in \mathcal{A} \otimes \mathcal{B}$, then $E_x \in \mathcal{B}$ and $E^y \in \mathcal{A}$ for all $x \in X$ and $y \in Y$.

b. If f is $\mathcal{A} \otimes \mathcal{B}$-measurable, then f_x is $\mathcal{B}$-measurble and f^y is $\mathcal{A}$-measurable for all $x \in X$ and $y \in Y$.

Proof.

a. Let $\mathcal{R} = \{E \subseteq X \times Y : E \in \mathcal{A} \otimes \mathcal{B} \text{ and } E_x \in \mathcal{B}\ \forall x \in X\}$. We prove that $\mathcal{R} = \mathcal{A} \otimes \mathcal{B}$. We know $\mathcal{R} \subseteq \mathcal{A} \otimes \mathcal{B}$ by definition, so it suffices to show that $\mathcal{R}$ is a σ-algebra, and that $\mathcal{R}$ contains all rectangles.

   First, we have
$$\left(\bigcup_{j \geqslant 1} E_j\right)_x = \bigcup_{j \geqslant 1} (E_j)_x$$
   and $(E^c)_x = (E_x)^c$, therefore $\mathcal{R}$ is a σ-algebra. Second, take any rectangle $A \times B$ with $A \in \mathcal{A}$ and $B \in \mathcal{B}$, then
$$(A \times B)_x = \begin{cases} \varnothing, & x \notin A \\ B, & x \in A, \end{cases}$$
   therefore $(A \times B)_x \in \mathcal{B}$ for all $x \in X$, thus $A \times B \in \mathcal{R}$.

b. This simply follows from part a.

□

Definition 2.94. Let $\mathcal{C} \subseteq \mathcal{P}(X)$, we say $\mathcal{C}$ is closed under countable increasing unions if given $E_j \in \mathcal{C}$ for all $j \in \mathbb{N}$ and $E_1 \subseteq E_2 \subseteq \cdots$ is an increasing sequence, then $\bigcup\limits_{j \geqslant 1} E_j \in \mathcal{C}$. Similarly, we say $\mathcal{C}$ is closed under countable decreasing intersections if for $E_j \in \mathcal{C}$ where $j \in \mathbb{N}$ and $E_1 \supseteq E_2 \supseteq \cdots$, then $\bigcap\limits_{j \geqslant 1} E_j \in \mathcal{C}$.

We say $\mathcal{C}$ is a monotone class if it is closed under countable increasing union and closed under countable decreasing intersection.

Remark 2.95. If $\mathcal{C}$ is a σ-algebra, then $\mathcal{C}$ is a monotone class. However, the converse may not be true. For instance, given $X = \{0, 1\}$ and $\mathcal{C} = \{\{0\}\}$, we know $\mathcal{C}$ is a monotone class but not an algebra.

Definition 2.96. Let $\mathcal{A} \subseteq \mathcal{P}(X)$, then $\mathcal{C}(\mathcal{A})$ denotes the smallest monotone class containing $\mathcal{A}$.

Lemma 2.97. Let $\mathcal{A}$ be an algebra, then $\mathcal{C}(\mathcal{A}) = \mathcal{M}(\mathcal{A})$, where $\mathcal{M}(\mathcal{A})$ is the smallest σ-algebra containing $\mathcal{A}$.

Proof. Left as an exercise. □

Lemma 2.97 can be applied to prove Theorem 2.98, known as a baby version of Fubini theorem.

Theorem 2.98. Suppose that $(X, \mathcal{A}, \mu_1)$ and $(Y, \mathcal{B}, \mu_2)$ are measure spaces. Let $E \in \mathcal{A} \otimes \mathcal{B}$, then $f(x) = \mu_2(E_x)$ for all x and $g(y) = \mu_1(E^y)$ are measurable functions. Moreover, $(\mu_1 \times \mu_2)(E) = \int\limits_X \mu_2(E_x) d\mu_1 = \int\limits_Y \mu_1(E^y) d\mu_2$.

Proof. Let $\mathcal{C}$ be the collection of $E \in \mathcal{A} \otimes \mathcal{B}$ such that $(\mu_1 \times \mu_2)(E) = \int\limits_X \mu_2(E_x) d\mu_1 = \int\limits_Y \mu_1(E^y) d\mu_2$, so it suffices to show $\mathcal{C} = \mathcal{A} \otimes \mathcal{B}$. Recall that $\mathcal{R}_0$ is the collection of finite disjoint unions of rectangles, then by Lemma 2.97, we know $\mathcal{A} \otimes \mathcal{B} = \mathcal{M}(\mathcal{R}_0) = \mathcal{C}(\mathcal{R}_0)$, the smallest monotone class containing $\mathcal{R}_0$. It suffices to show that $\mathcal{C} = \mathcal{C}(\mathcal{R}_0)$, and to conclude the proof we need to show

- $\mathcal{C} \supseteq \mathcal{R}_0$, and
- $\mathcal{C}$ is a monotone class.

Claim 2.99. For any $A \in \mathcal{A}$ and $B \in \mathcal{B}$, the rectangle $A \times B \in \mathcal{C}$.

Subproof. Let $E = A \times B$, then
$$\begin{aligned} (\mu_1 \times \mu_2)(E) &= (\mu_1 \times \mu_2)(A \times B) \\ &= \mu_1(A) \times \mu_2(B). \end{aligned}$$

We will show that $(\mu_1 \times \mu_2)(E) = \int\limits_X \mu_2(E_x) d\mu_1$, then the other equality follows similarly. We have

$$E_x = (A \times B)_x = \begin{cases} B, & x \in A \\ \varnothing, & x \notin A \end{cases}$$

and therefore

$$\begin{aligned} \mu_2(E_x) &= \begin{cases} \mu_2(B), & x \in A \\ 0, & x \notin A \end{cases} \\ &= \mu_2(B)\chi_A(x) \end{aligned}$$

so

$$\begin{aligned} \int\limits_X \mu_2(E_x) d\mu_1 &= \int\limits_X \mu_2(B)\chi_A(x) d\mu_1 \\ &= \mu_2(X) \int\limits_X \chi_A d\mu_1 \\ &= \mu_2(B)\mu_1(A) \\ &= (\mu_1 \times \mu_2)(A \times B) \\ &= (\mu_1 \times \mu_2)(E). \end{aligned}$$

■

Therefore, $\mathcal{R}_0 \subseteq \mathcal{C}$ by the finite additivity of measures.

To show $\mathcal{C}$ is a monotone class, so it suffices to show $\mathcal{C}$ is closed under increasing σ-unions and under decreasing σ-intersections. We show that $\mathcal{C}$ is closed under increasing σ-unions, and the other part can be proven analogously. For any $n \in \mathbb{N}$, let $E_n \in \mathcal{C}$ be such that $E_n \subseteq E_{n+1}$, so we want to show $E := \bigcup\limits_{n \geqslant 1} E_n \in \mathcal{C}$. First we show that $(\mu_1 \times \mu_2)(E) = \int\limits_Y \mu_1(E^y) d\mu_2$, and similarly we can show the other equality. By continuity from below, we have

$$\begin{aligned} \mu_1(E^y) &= \mu_1\left(\left(\bigcup_{n \geqslant 1} E_n\right)^y\right) \\ &= \mu_1\left(\bigcup_{n \geqslant 1} E_n^y\right) \\ &= \lim_{n \to \infty} \mu_1(E_n^y). \end{aligned}$$

By Theorem 2.28 and continuity of $\mu_1 \times \mu_2$, we know

$$\begin{aligned} \int\limits_Y \lim_{n \to \infty} \mu_1(E_n^y) d\mu_2 &= \int\limits_Y \mu_1(E^y) d\mu_2 \\ &= \lim_{n \to \infty} \int\limits_Y \mu_1(E_n^y) d\mu_2 \\ &= \lim_{n \to \infty} (\mu_1 \times \mu_2)(E_n) \\ &= (\mu_1 \times \mu_2)\left(\bigcup_{n \geqslant 1} E_n\right) \\ &= (\mu_1 \times \mu_2)(E). \end{aligned}$$

Therefore, $\int\limits_Y \mu_1(E^y) d\mu_2 = (\mu_1 \times \mu_2)(E)$. □

Theorem 2.100. Suppose $(X, \mathcal{A}, \mu_1)$ and $(Y, \mathcal{B}, \mu_2)$ are σ-finite measure spaces, then for any $E \in \mathcal{A} \otimes \mathcal{B}$, we know that functions $f(x) = \mu_2(E_x)$ and $g(x) = \mu_1(E^y)$ are measurable. Moreover, $(\mu_1 \times \mu_2)(E) = \int\limits_X \mu_2(E_x) d\mu_1 = \int\limits_Y \mu_1(E^y) d\mu_2$.

Proof. We can write $X = \bigcup\limits_{j \geqslant 1} A_j$ where $A_j \in \mathcal{A}$ are pairwise disjoint, and that $\mu_1(A_j) < \infty$ for all j. Similarly, we have $Y = \bigcup\limits_{k \geqslant 1} B_k$ where $B_k \in \mathcal{B}$ are pairwise disjoint, and that $\mu_2(B_k) < \infty$ for all k. Therefore,

$$\begin{aligned} X \times Y &= \left(\bigcup_{j \geqslant 1} A_j \right) \times \left(\bigcup_{k \geqslant 1} B_k \right) \\ &= \bigcup_{j,k} (A_j \times B_k) \end{aligned}$$

which is a σ-union of pairwise disjoint rectangles. Therefore, $X \times Y = \bigcup\limits_{i \geqslant 1} (X_i \times Y_i)$ such that $X_i \times Y_i$'s are disjoint, and so

$$(\mu_1 \times \mu_2)(X_i \times Y_i) = \mu_1(X_i)\mu_2(Y_i) < \infty.$$

Take a measurable set $E \in \mathcal{A} \otimes \mathcal{B}$ in the product space, then

$$\begin{aligned} E &= E \cap (X \times Y) \\ &= E \cap \left(\bigcup_{i \geqslant 1} (X_i \times Y_i) \right) \\ &= \bigcup_{i \geqslant 1} (E \cap (X_i \times Y_i)) \end{aligned}$$

which is an infinite union of finite measure sets. For each finite measure set $E \cap (X_i \times Y_i)$, we apply Theorem 2.98, and we know

$$\begin{aligned} (\mu_1 \times \mu_2)(E \cap (X_i \times Y_i)) &= \int\limits_{X_i} \mu_2((E \cap (X_i \times Y_i))_x) d\mu_1 \\ &= \int\limits_{X_i} \mu_2(E_x \cap Y_i) d\mu_1, \end{aligned}$$

and similarly $(\mu_1 \times \mu_2)(E \cap (X_i \times Y_i)) = \int\limits_{Y_i} \mu_1(E^y \cap X_i) d\mu_2$.

Now by Theorem 2.28 we know

$$\begin{aligned} (\mu_1 \times \mu_2)(E) &= \sum_{i \geqslant 1} (\mu_1 \times \mu_2)(E \cap (X_i \times Y_i)) \\ &= \sum_{i \geqslant 1} \int\limits_{X_i} \mu_2(E_x \cap Y_i) d\mu_1 \\ &= \int\limits_{X_i} \sum_{i \geqslant 1} \mu_2(E_x \cap Y_i) d\mu_1 \\ &= \int\limits_{X} \sum_{i \geqslant 1} \mu_2(E_x \cap Y_i) \chi_{X_i} d\mu_1 \end{aligned}$$

and then Claim 2.101 gives us the desired equality.

Claim 2.101. For all $x \in X$, $\sum\limits_{i \geqslant 1} \mu_2(E_x \cap Y_i) \chi_{X_i}(x) = \mu_2(E_x)$.

The other equality follows similarly.

Subproof. By Theorem 2.28,

$$\begin{aligned}
\sum_{i\geqslant 1}\mu_2(E_x\cap Y_i)\chi_{X_i}(x) &= \sum_{i\geqslant 1}\int\limits_{E_X}\chi_{Y_i}(y)\chi_{X_i}(x)d\mu_2\\
&= \int\limits_{E_x}\sum_{i\geqslant 1}\chi_{Y_i}(y)\chi_{X_i}(x)d\mu_2\\
&= \int\limits_{E_x}\sum_{i\geqslant 1}\chi_{X_i\times Y_i}(x,y)d\mu_2\\
&= \int\limits_{E_x}\chi_{X\times Y}(x,y)d\mu_2\\
&= \int\limits_{E_x}1d\mu_2\\
&= \mu_2(E_x).
\end{aligned}$$

■

□

Theorem 2.102 (Fubini-Tonelli)**.** Suppose that $(X,\mathcal{A},\mu_1)$ and $(Y,\mathcal{B},\mu_2)$ are σ-finite measure spaces.

a. (Tonelli) If $f\in L^+(X\times Y)$, then both $g(x)=\int\limits_Y f_x d\mu_2$ and $h(y)=\int\limits_X f^y d\mu_1$ are well-defined non-negative measurable functions. Moreover,

$$\int\limits_{X\times Y} fd(\mu_1\times\mu_2)=\int\limits_X\left(\int\limits_Y fd\mu_2\right)d\mu_1=\int\limits_Y\left(\int\limits_X fd\mu_1\right)d\mu_2.$$

b. (Fubini) If $f\in L^1(\mu_1\times\mu_2)$ over $(X\times Y,\mathcal{A}\otimes\mathcal{B},\mu_1\times\mu_2)$, then $f_x\in L^1(\mu_2)$ almost everywhere on X and $f^y\in L^1(\mu_1)$ almost everywhere on Y. Moreover,

$$\int\limits_{X\times Y} fd(\mu_1\times\mu_2)=\int\limits_X\left(\int\limits_Y fd\mu_2\right)d\mu_1=\int\limits_Y\left(\int\limits_X fd\mu_1\right)d\mu_2$$

Proof.

a. Since μ_1 and μ_2 are σ-finite, then for any $E\in\mathcal{A}\otimes\mathcal{B}$, we have $(\mu_1\times\mu_2)(E)=\int\limits_X\mu_2(E_x)d\mu_1=\int\limits_Y\mu_1(E^y)d\mu_2$. Therefore, if $f=\chi_E$, we know part a. is true. Recall that a simple function is a combination of those indicator functions, so part a. holds for simple functions. For any $f\in L^+(X\times Y)$, we know $f=\lim\limits_{n\to\infty}\varphi_n$, where $\{\varphi_n\}_{n\geqslant 1}$ is an increasing sequence of measurable simple functions in L^+. We now know

$$\begin{aligned}
\int\limits_{X\times Y} fd(\mu_1\times\mu_2) &= \int\limits_{X\times Y}\lim_{n\to\infty}\varphi_n d(\mu_1\times\mu_2)\\
&= \lim_{n\to\infty}\int\limits_{X\times Y}\varphi_n d(\mu_1\times\mu_2)\\
&= \int\limits_X\left(\int\limits_Y\lim_{n\to\infty}\varphi_n d\mu_2\right)d\mu_1
\end{aligned}$$

$$= \int_X \left(\int_Y f d\mu_2 \right) d\mu_1$$

by Theorem 2.28. We can show the other equality in a similar fashion.

b. Since f is integrable, then it has values in $\mathbb{C}$ almost everywhere. So without loss of generality, we may assume f to be a real-valued function. We now write $f = f^+ + f^-$, then we reduce the problem to non-negative measurable functions, which is illustrated in part a.

□

Example 2.103.

$$\int_0^\infty \frac{\sin(x)}{x} dx = \lim_{n\to\infty} \int_0^n \frac{\sin(x)}{x} dx = \frac{\pi}{2}.$$

Proof. This can be done by Fourier series. However, we will prove this by using Theorem 2.102. Note that

$$\frac{1}{x} = \int_0^\infty e^{-xt} dt$$

for all $x > 0$. Set $F(x,t) = e^{-xt}\sin(x)$ for $0 < x < \infty$ and $0 < t < \infty$.

Claim 2.104. $F \in L^1([0,n) \times [0,\infty))$.

Subproof. We know

$$|F(x,t)| \leqslant G(x,t) := \begin{cases} e^{-xt}x, & 0 \leqslant x < 1 \\ e^{-t}, & x \geqslant 1 \end{cases}$$

Obviously $G(x,t) \in L^+$ since it is a non-negative measurable function. By applying part a. of Theorem 2.102 to $G(x,t)$, we know

$$\begin{aligned} \int_{[0,n]\times[0,\infty)} G dx dt &= \int_0^n \left(\int_0^\infty G(x,t) dt \right) dx \\ &= \int_0^1 \left(\int_0^\infty F(x,t) dt \right) dx + \int_1^n \left(\int_0^\infty F(x,t) dt \right) dx \\ &< \infty \end{aligned}$$

and so $G \in L^1([0,n] \times [0,\infty))$. ■

Since $F(x,t)$ is integrable, then by part b. of Theorem 2.102, we have

$$\begin{aligned} \int_0^\infty \frac{\sin(x)}{x} dx &= \int_0^\infty \left(\int_0^\infty e^{-xt} dt \right) \sin(x) dx \\ &= \int_0^\infty \left(\int_0^\infty e^{-xt}\sin(x) dt \right) dx \\ &= \lim_{n\to\infty} \int_0^n \left(\int_0^\infty F(x,t) dt \right) dx \end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} \int_0^\infty \left(\int_0^n e^{-xt}\sin(x)dx \right) dt \\
&= \lim_{n\to\infty} \int_0^\infty \frac{1}{1+t^2}(1 - e^{-nt}\cos(n) - te^{-nt}\sin(n))dt.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\int_0^\infty \frac{\sin(x)}{x}dx &= \int_0^\infty \frac{1}{1+t^2}dt - \int_0^n \frac{e^{-nt}\cos(n)}{1+t^2}dt - \int_0^\infty \frac{te^{-nt}\sin(n)}{1+t^2}dt \\
&= \frac{\pi}{2} - \int_0^n \frac{e^{-nt}\cos(n)}{1+t^2}dt - \int_0^\infty \frac{te^{-nt}\sin(n)}{1+t^2}dt.
\end{aligned}
$$

Assuming the limit exists, then we have

$$
\int_0^\infty \frac{\sin(x)}{x}dx = \frac{\pi}{2} - \lim_{n\to\infty}\int_0^n \frac{e^{-nt}\cos(n)}{1+t^2}dt - \lim_{n\to\infty}\int_0^\infty \frac{te^{-nt}\sin(n)}{1+t^2}dt.
$$

We claim that both limits here are 0. We have

$$
\left| \frac{e^{-nt}\cos(n)}{1+t^2} \right| \leqslant \frac{e^{-nt}}{1+t^2} \leqslant \frac{1}{1+t^2}
$$

for any $n \in \mathbb{N}$. Note that the rightmost function is integrable on $(0,\infty)$. By Theorem 2.46,

$$
\begin{aligned}
\lim_{n\to\infty}\int_0^n \frac{e^{-nt}\cos(n)}{1+t^2}dt &= \int_0^\infty \lim_{n\to\infty} \frac{e^{-nt}\cos(n)}{1+t^2}dt \\
&= \int_0^\infty 0 \\
&= 0.
\end{aligned}
$$

Using similar techniques, we can verify that

$$
\lim_{n\to\infty}\int_0^\infty \frac{te^{-nt}\sin(n)}{1+t^2}dt = 0
$$

as well. $\square$

3 Signed Measure and Differentiation

We now study $(X, \mathcal{M}, \mu) = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \mathfrak{m})$.

3.1 Differentiation on Euclidean Space

Definition 3.1. Let $f : \mathbb{R}^n \to \mathbb{C}$ be a measurable function. If $\int_K |f| d\mathfrak{m} < \infty$ for all compact sets $K \subseteq \mathbb{R}^n$, then f is called locally integrable. Moreover, we denote $L^1_{\text{Loc}}(\mathbb{R}^n)$ to be the collection of functions $f : \mathbb{R}^n \to \mathbb{C}$ where f is locally integrable.

Definition 3.2. Let $B(x, r) = \{y \in \mathbb{R}^n : |y - x| < r\}$ be the open ball at x of radius r with respect to the 2-norm, i.e., Euclidean distance. Let $f \in L^1_{\text{Loc}}(\mathbb{R}^n)$, then the Hardy-Littlewood maximal function is defined by

$$Mf(x) = \sup_{r>0} \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f(y) d\mathfrak{m}$$

for almost every $x \in \mathbb{R}^n$.

Remark 3.3. Note that $\mathfrak{m}(B(x, r)) = \mathfrak{m}(B(0, r)) = C_n r^n$ since Lebesgue measure is invariant under translation, c.f., Theorem 1.62, where $C_n = \mathfrak{m}(B(0, 1))$.

Theorem 3.4 (Lebesgue Differentiation)**.** Let $f \in L^1_{\text{Loc}}(\mathbb{R}^n)$, then

$$\lim_{r\to 0} \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f d\mathfrak{m} = f(x)$$

for almost every $x \in \mathbb{R}^n$.

Remark 3.5. If f is continuous, then the statement is true for all $x \in \mathbb{R}^n$.

Remark 3.6. When $n = 1$, then Theorem 3.4 is an analogue of the fundamental theorem of calculus:

$$\frac{d}{dx} \int_0^x f(t) dt = f(x)$$

for almost every $x \in \mathbb{R}$.

Remark 3.7. Alternatively, we can define the uncentered maximal function

$$\tilde{M} f(x) = \sup_{B \ni x} \frac{1}{\mathfrak{m}(B)} \int_B f d\mathfrak{m}.$$

Note $Mf(x) \leqslant \tilde{M} f(x) \leqslant 2^n Mf(x)$. Then there is a version of Lebesgue Differentiation Theorem for uncentered maximal function:

$$\lim_{\substack{r(B)\to 0 \\ B \ni x}} \frac{1}{\mathfrak{m}(B)} \int_B f d\mathfrak{m} = f(x) \tag{3.8}$$

for almost every $x \in \mathbb{R}^n$.

Remark 3.9. Equation (3.8) holds if the ball B is replaced by a cube Q. However, Equation (3.8) is not true if balls are replaced by rectangles, i.e., pointing to many directions.

Remark 3.10 (Kakeya Needle Problem)**.** Suppose we have a unit line segment on a plane, and we move the segment continuously on the plane until it points towards the opposite direction. What is the smallest possible area covered by the continuous movement of the segment? In fact, there is no such minimal area: the area can be arbitrary small. This is due to the existence of Besicovitch sets.

Definition 3.11. A set in $\mathbb{R}^n$ is called a Besicovitch set if it contains unit line segments pointing to every possible directions, but its Lebesgue measure is zero.

Lemma 3.12 (Vitali Covering)**.** Let E be any Lebesgue measurable set in $\mathbb{R}^n$, and suppose $E \subseteq \bigcup_{\alpha\in A} B_\alpha$, where each B_α is a ball in $\mathbb{R}^n$. Moreover, suppose $\sup_{\alpha\in A} r(B_\alpha) < \infty$, where $r(B)$ is the radius of the ball B, then there exists a countable subcollection of disjoint subsets $\{B_{\alpha_k}\}_{k\in\mathbb{N}}$ of $\{B_\alpha\}_{\alpha\in A}$ such that $\mathfrak{m}(E) \leqslant C_n \sum_{k\geqslant 1} \mathfrak{m}(B_{\alpha_k})$.[1]

Proof. Take B_{α_1} such that $r(B_{\alpha_1}) > \frac{1}{2}\sup_{\alpha\in A} r(B_\alpha)$. Remove those balls B in $\mathcal{B}_0 = \{B_\alpha\}_{\alpha\in A}$ such that $B \cap B_{\alpha_1} \neq \varnothing$. Let $\mathcal{B}_1 = \{B_\alpha\}_{\alpha\in A}\backslash\{B_{\alpha_1} \cup \{B \in \mathcal{B}_0 : B \cap B_{\alpha_1} \neq \varnothing\}\}$. Suppose that we have disjoint balls $B_{\alpha_1}, \dots, B_{\alpha_k}$ chosen, then we need to choose $B_{\alpha_{k+1}}$ from the remaining balls such that

$$r(B_{\alpha_{k+1}}) \geqslant \frac{1}{2}\sup\left\{r(B_\alpha) : B_\alpha \cap \left(\bigcup_{j=1}^k B_{\alpha_j}\right) = \varnothing\right\}.$$

This gives us a desired sequence of disjoint balls. Suppose $\sum_{k\geqslant 1} \mathfrak{m}(B_{\alpha_k}) = \infty$, then $\mathfrak{m}(E) \leqslant \infty = C_n \sum_{k\geqslant 1} \mathfrak{m}(B_{\alpha_k})$. Therefore, we may assume that $\sum_{k\geqslant 1} \mathfrak{m}(B_{\alpha_k}) < \infty$.

Claim 3.13. $E \subseteq \bigcup_{k\geqslant 1} 5B_{\alpha_k}$, where $5B_{\alpha_k}$ is the induced ball from B_{α_k}, dilated by 5.

By Claim 3.13, we have

$$\begin{aligned}
\mathfrak{m}(E) &\leqslant \mathfrak{m}\left(\bigcup_{k\geqslant 1} 5B_{\alpha_k}\right)\\
&= \sum_{k\geqslant 1} \mathfrak{m}(5B_{\alpha_k})\\
&= \sum_{k\geqslant 1} 5^n \mathfrak{m}(B_{\alpha_k})\\
&= 5^n \sum_{k\geqslant 1} \mathfrak{m}(B_{\alpha_k}).
\end{aligned}$$

Proof of Claim 3.13. One can show that there exists $j_0 \in \mathbb{N}$ such that $B \cap B_{\alpha_{j_0}} \neq \varnothing$ but $r(B_{\alpha_{j_0}}) > \frac{1}{2}r(B)$. ■

□

Theorem 3.14. There exists a constant $C_n > 0$ such that, for any $\lambda > 0$ and $f \in L^1(\mathbb{R}^n)$,

$$\mathfrak{m}(\{x \in \mathbb{R}^n : Mf(x) > \lambda\}) \leqslant \frac{C_n||f||_1}{\lambda}.$$

We say M is of weak-$(1, 1)$.

Remark 3.15. We say M is of strong-$(1, 1)$ if for any $f \in L^1$, $||Mf||_1 \leqslant C||f||_1$.

Proof. For any $\lambda > 0$, we denote $E_\lambda = \{x \in \mathbb{R}^n : Mf(x) > \lambda\}$. For any $x \in E_\lambda$, let B_x be a ball centered at x such that

$$\frac{1}{\mathfrak{m}(B_x)}\int_{B_x} |f| d\mathfrak{m} > \lambda.$$

Such ball exists because $Mf(x) > \lambda$. Therefore, $E_\lambda \subseteq \bigcup_{x\in E_\lambda} B_x$. By Lemma 3.12, we have

$$\mathfrak{m}(B_x) > \frac{1}{\lambda}\int_{B_x} |f| d\mathfrak{m} \leqslant \frac{1}{\lambda}||f||_1 < \infty \tag{3.16}$$

[1]Here $C_n = \mathfrak{m}(B(0, 1))$ is the measure of the unit n-ball, which can be bounded by 5^n.

since $f \in L^1$ and $\lambda > 0$. Therefore, $\sup\limits_{x \in E_\lambda} \mathfrak{m}(B_x) < \infty$, so $\sup\limits_{x \in E_\lambda} r(B_x) < \infty$. By Lemma 3.12, we know

$$\mathfrak{m}(E_\lambda) \leqslant C_n \sum_{k \geqslant 1} \mathfrak{m}(B_{x_k}),$$

where $\{B_{x_k}\}_{k \geqslant 1}$ is a sequence of disjoint balls. But $\mathfrak{m}(E_\lambda) = \mathfrak{m}(\{x : Mf(x) > \lambda\})$. By Equation (3.16),

$$\begin{aligned}\sum_{k \geqslant 1} \mathfrak{m}(B_{x_k}) &\leqslant \frac{1}{\lambda} \sum_{k \geqslant 1} \int_{B_{x_k}} |f| d\mathfrak{m} \\ &= \frac{1}{\lambda} \int_{k \geqslant 1} B_{x_k} |f| d\mathfrak{m} \\ &\leqslant \frac{||f||_1}{\lambda}.\end{aligned}$$

Therefore, $\mathfrak{m}(E_\lambda) \leqslant \frac{C_n ||f||_1}{\lambda}$. □

Proof of Theorem 3.4. We first show that the statement is true if f is continuous. To prove the statement over $L^1_{\text{Loc}}(\mathbb{R}^n)$, recall that we know continuous functions are dense in the L^1-space, so we have

$$\sup_{r>0} \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f d\mathfrak{m} = Mf(x)$$

and recall that M is of weak-$(1,1)$ estimate, therefore the statement is true for any function f in the L^1-space.

Claim 3.17. The statement is true if $f \in C(\mathbb{R}^n)(\cap L^1(\mathbb{R}))$.[2]

Subproof. Let $f \in C(\mathbb{R}^n)$, then let

$$D_{x,r} := \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f d\mathfrak{m} - f(x),$$

and we will show that $\lim\limits_{r \to 0} D_{x,r} = 0$ for all $x \in \mathbb{R}^n$. We know

$$\begin{aligned}D_{x,r} &= \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f d\mathfrak{m} - \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f(x) d\mathfrak{m} \\ &= \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f(y) - f(x) dy.\end{aligned}$$

Since f is continuous, then for any $\delta > 0$, there exists $r_\delta > 0$ such that $|f(y) - f(x)| < \delta$ whenever $|y - x| < r_\delta$. For any $r < r_\delta$, if $y \in B(x,r)$, then $|y - x| < r < r_\delta$, therefore $|f(y) - f(x)| < \delta$ if $y \in B(x,r)$ for any $r < r_\delta$. For any $r < r_\delta$, we have

$$\begin{aligned}|D_{x,r}| &\leqslant \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} |f(y) - f(x)| dy \\ &< \delta.\end{aligned}$$

Therefore $|D_{x,r}| < \delta$ whenever $r < r_\delta$. Let $r \to 0$, then this gives $\lim\limits_{r \to 0} |D_{x,r}| < \delta$ for all $\delta > 0$. Let $\delta \to 0$, then $\lim\limits_{r \to 0} |D_{x,r}| = 0$. ■

[2] We can localize to make sure this is true.

Claim 3.18. For any function $f \in L^1(\mathbb{R}^n)$, the value

$$\lim_{r\to 0} \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f d\mathfrak{m}$$

exists in $\mathbb{R}$ for almost all $x \in \mathbb{R}$.

Subproof. Define

$$\Theta(f)(x) := \limsup_{r\to 0} \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f d\mathfrak{m} - \liminf_{r\to 0} \frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f d\mathfrak{m}.$$

We will show that for almost all $x \in X$, $\Theta(f)(x) = 0$. For any $\varepsilon > 0$, there exists $g \in C(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$ such that $||f-g||_1 < \varepsilon$. By Claim 3.17, $\Theta(g)(x) = 0$ for all $x \in \mathbb{R}^n$. Now

$$\begin{aligned}\Theta(f)(x) &= \Theta(f)(x) - \Theta(g)(x)\\ &\leqslant |\Theta(f-g)(x)|\\ &\leqslant M(f-g)(x).\end{aligned}$$

We know for $\lambda > 0$, the level set

$$\begin{aligned}\{x \in \mathbb{R}^n : |\Theta(f)| > \lambda\} &\subseteq \{x \in \mathbb{R}^n : |\Theta(f-g)(x)| > \lambda\}\\ &\subseteq \{x \in \mathbb{R}^n : M(f-g)(x) > \lambda\},\end{aligned}$$

hence the measure

$$\begin{aligned}\mathfrak{m}(\{x \in \mathbb{R}^n : |\Theta(f)| > \lambda\}) &\leqslant \mathfrak{m}(\{x \in \mathbb{R}^n : M(f-g)(x) > \lambda\})\\ &\leqslant \frac{C_n ||f-g||_1}{\lambda}\\ &< \frac{C_n \varepsilon}{\lambda}\end{aligned}$$

for any $\varepsilon > 0$, by the weak-$(1,1)$ estimate. Let $\varepsilon \to 0$, we have

$$\mathfrak{m}(\{x \in \mathbb{R}^n : |\Theta(f)| > \lambda\}) = 0$$

for all $\lambda > 0$. Now the set

$$\{x \in \mathbb{R}^n : |\Theta(f)| \neq 0\} = \bigcup_{n\geqslant 1} \{x \in \mathbb{R}^n : |\Theta(f)(x)| > \frac{1}{n}\}$$

is a union of null sets, therefore $\mathfrak{m}(\{x \in \mathbb{R}^n : |\Theta(f)| \neq 0\}) = 0$, hence $\Theta(f)(x) = 0$ for almost all $x \in X$. ■

It remains to show that

$$\lim_{r\to 0}\left(\frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f d\mathfrak{m} - f(x)\right) = 0$$

for almost all $x \in X$. For any $\varepsilon > 0$, there exists $g \in C(\mathbb{R}^n)$ such that $||f-g||_1 < \varepsilon$. By telescoping, we know

$$\begin{aligned}\left|\lim_{r\to 0}\left(\frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} f d\mathfrak{m} - f(x)\right)\right| &= \left|\lim_{r\to 0}\left(\frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} (f-g) d\mathfrak{m} - (f-g)(x)\right)\right|\\ &\leqslant \lim_{r\to 0}\left(\frac{1}{\mathfrak{m}(B(x,r))} \int_{B(x,r)} |f-g| d\mathfrak{m} + (f-g)(x)\right)\end{aligned}$$

$$\leqslant 2M(f-g)(x).$$

Therefore, the level set has measure

$$\begin{aligned}\mathfrak{m}\left(\left\{x\in\mathbb{R}^n : \lim_{r\to 0}\left|\frac{1}{\mathfrak{m}(B(x,r))}\int\limits_{\mathfrak{m}(B(x,r))} f d\mathfrak{m} - f(x)\right| > \lambda\right\}\right) &\leqslant \mathfrak{m}(\{x\in\mathbb{R}^n : 2M(f-g)(x) > \lambda\})\\ &\leqslant \frac{2C_n||f-g||_1}{\lambda}\\ &< \frac{2C_n\varepsilon}{\lambda}\end{aligned}$$

for all $\varepsilon > 0$. Let $\varepsilon \to 0$, we know

$$\mathfrak{m}\left(\left\{x\in\mathbb{R}^n : \lim_{r\to 0}\left|\frac{1}{\mathfrak{m}(B(x,r))}\int\limits_{\mathfrak{m}(B(x,r))} f d\mathfrak{m} - f(x)\right| > \lambda\right\}\right) = 0$$

for all $\lambda > 0$. Using the same argument as in Claim 3.18, we know

$$\lim_{r\to 0}\left|\frac{1}{\mathfrak{m}(B(x,r))}\int\limits_{B(x,r)} f - f(x)\right| = 0$$

for almost all $x \in X$. □

3.2 Functions of Bounded Variation

Theorem 3.19. Let $F : \mathbb{R} \to \mathbb{R}$ be an increasing function, then F is continuous almost everywhere.

Proof. Let $D = \{x \in \mathbb{R} : F \text{ is discontinuous at } x\}$, then it suffices to show that $\mathfrak{m}(D) = 0$. For any $x \in D$, we know the one-side limits do not agree:

$$\lim_{y\to x^+} F(y) = F(x^+) \neq F(x^-) = \lim_{y\to x^-} F(y)$$

Since F is an increasing function, then $F(x^+) > F(x^-)$. Let $I_x = (F(x^-), F(x^+))$.

Claim 3.20. $\{I_x : x \in D\}$ is a collection of disjoint open intervals.

Subproof. Let $x_1 < x_2$ be points in D, then we need to show that $I_{x_1} \cap I_{x_2} = \varnothing$. By the denseness, there exists $y \in \mathbb{R}$ such that $x_1 < y < x_2$, therefore

$$\begin{aligned}F(x_1^+) &= \inf\{F(x) : x > x_1\}\\ &\leqslant F(y)\\ &\leqslant \sup\{F(x) : x < x_2\}\\ &= F(x_2^-).\end{aligned}$$

Therefore $I_{x_1} \cap I_{x_2} = \varnothing$. ■

Moreover, D is a countable set, since there is a correspondence between discontinuous points $x \in D$ and bounded intervals I_x. Now for arbitrary I_x where $x \in D$, we can take $r_x \in I_x \cap \mathbb{Q}$, which exists since $\mathbb{Q}$ is dense in $\mathbb{R}$. Set $\mathcal{R} = \{r_x : x \in D\} \subseteq \mathbb{Q}$. Therefore, there is a correspondence between I_x's and r_x's. In particular, the cardinality of $\mathcal{R}$ is at most $\mathbb{Q}$, which is countable, therefore D is countable, hence $\mathfrak{m}(D) = 0$. □

Theorem 3.21. Let $f : [a,b] \to \mathbb{R}$ be an increasing function where $a < b \in \mathbb{R}$, then f is differentiable almost everywhere in $[a,b]$, and the Lebesgue measure

$$\int\limits_a^b f'(t)dt \leqslant f(b) - f(a).$$

Proof Sketch. One can prove differentiability by Lemma 3.12, and the inequality comes from Theorem 2.36. $\square$

Definition 3.22. Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function and fix $x \in \mathbb{R}$. We define the (total) variation of f at x to be $T_f(x) = \sup_{n\in\mathbb{N}}\{\sum_{i=1}^{n} |f(x_j) - f(x_{j-1})| : -\infty < x_0 < \cdots < x_n = x\}$. Moreover, we define $T_f(\infty) = \lim_{x\to\infty} T_f(x) = \sup_{n\in\mathbb{N}}\{\sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| : -\infty < x_0 < \cdots < x_n < \infty\}$.

We say f is of bounded variation on $\mathbb{R}$ if $T_f(\infty) < \infty$, and denote $\mathrm{BV}(\mathbb{R})$ to be the set of functions f of bounded variation on $\mathbb{R}$.

Alternatively, we may use the notation Var_f in place of T_f.

Definition 3.23. Let $a, b \in \mathbb{R}$, then we define the (total) variation of f on $[a, b]$ to be $\mathrm{Var}_f([a, b]) = \sup_{n\in\mathbb{N}}\{\sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| : a = x_0 < \cdots < x_n = b\}$.

If $\mathrm{Var}_f([a, b]) < \infty$, then we say f is of bounded variation on $[a, b]$. We denote $\mathrm{BV}([a, b])$ to be the set of functions f of bounded variation on $[a, b]$.

Example 3.24. Let $f : \mathbb{R} \to \mathbb{R}$ be an increasing and bounded function, then $f \in \mathrm{BV}(\mathbb{R})$. Indeed, $T_f(\infty) = f(\infty) - f(-\infty) = \lim_{x\to\infty} f(x) - \lim_{x\to-\infty} f(x) < \infty$.

Example 3.25. Let

$$f(x) = \begin{cases} x \sin\left(\frac{1}{x}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

be a function on $\mathbb{R}$, then $f \notin \mathrm{BV}(\mathbb{R})$. In fact, $f \notin \mathrm{BV}([a, b])$ whenever $a \leqslant 0 \leqslant b$. For instance, $\mathrm{Var}_f([0, 1]) = \infty$.

Theorem 3.26. Given a function $f : \mathbb{R} \to \mathbb{R}$, then $f \in \mathrm{BV}(\mathbb{R})$ (or $\mathrm{BV}([a, b])$) if and only if $f = g_1 - g_2$ where g_1, g_2 are bounded, increasing functions.

Proof.

($\Leftarrow$): if f can be written as $g_1 - g_2$, a difference of bounded, increasing functions, then $T_f(\infty) \leqslant T_{g_1}(\infty) + T_{g_2}(\infty) < \infty$, therefore $f \in \mathrm{BV}(\mathbb{R})$.

($\Rightarrow$): let us write $f = \frac{1}{2}(T_f + f) - \frac{1}{2}(T_f - f)$, then set $g_1 = \frac{1}{2}(T_f + f)$ and $g_2 = \frac{1}{2}(T_f - f)$.

Claim 3.27. Both g_1 and g_2 are bounded, increasing functions.

Subproof. Since f is totally bounded, i.e., $|f| \leqslant T_f(\infty)$, then f is bounded. Therefore, for $j = 1, 2$, we have $|g_j(x)| \leqslant \frac{1}{2}(T_f(\infty) + f(x)) < \infty$ since $T_f(\infty) < \infty$. Therefore, g_1 and g_2 are bounded functions.

We now prove that g_1 is an increasing function, then a similar argument shows g_2 is also increasing. Set $x < y$, then we want to show that $g_1(x) < g_1(y)$. By the definition of $T_f(x)$, for any $\varepsilon > 0$, there exists a sequence $x_0 < \cdots < x_n = x$ such that $T_f(x) - \varepsilon \leqslant \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})|$. Therefore, we know

$$\begin{aligned} T_f(y) &\geqslant \sum_{j=1}^{n} |f(x_j) - f(x_{j-1})| + |f(x) - f(y)| \\ &\geqslant T_f(x) - \varepsilon + |f(x) - f(y)|. \end{aligned}$$

Note that

$$\begin{aligned} g_1(y) &= \frac{1}{2}(T_f + f) \\ &\geqslant \frac{1}{2}(T_f(x) - \varepsilon + |f(x) - f(y)|) \end{aligned}$$

$$= \frac{1}{2}(T_f(x) + f(x) - \varepsilon + |f(x) - f(y)| + f(y) - f(x))$$
$$\geqslant \frac{1}{2}((T_f + f)(x) - \varepsilon + |f(x) - f(y)| - (f(x) - f(y)))$$
$$\geqslant \frac{1}{2}(T_f + f)(x) - \frac{1}{2}\varepsilon.$$

Take $\varepsilon \to 0$, we get $g_1(y) \geqslant g_1(x)$. ■

□

Theorem 3.28. Any $f \in \mathrm{BV}(\mathbb{R})$ is differentiable almost everywhere. Similar conclusion holds for $f \in \mathrm{BV}([a, b])$.

Definition 3.29. Let $f : [a, b] \to \mathbb{R}$ or $f : [a, b] \to \mathbb{C}$ be a function where $a, b \in \mathbb{R}$. We say f is absolutely continuous on $[a, b]$ if for any $\varepsilon > 0$, there exists $\delta > 0$ such that $\sum_{j=1}^{n} |f(b_j) - f(a_j)| < \varepsilon$ whenever $\sum_{j=1}^{n} (b_j - a_j) < \delta$ where $(a_1, b_1), \ldots, (a_n, b_n)$ are disjoint pairs. We denote $\mathrm{AC}([a, b])$ to be the collection of all absolutely continuous functions.

Remark 3.30. Any absolutely continuous function is a uniformly continuous function.

Corollary 3.31. If $f \in \mathrm{AC}([a, b])$, then $f \in \mathrm{BV}([a, b])$.

Lemma 3.32. Let $f \in \mathrm{AC}([a, b])$. Suppose $f'(x) = 0$ almost everywhere, then $f(x)$ is a constant function on $[a, b]$.

Example 3.33. We can define a function using the Cantor set that is not constant, but the derivative is zero almost everywhere. Set $f(0) = 0$ and $f(1) = 1$. We define the function on the removed intervals in the procedure of the Cantor set. For the first time, we remove $(\frac{1}{3}, \frac{2}{3})$, we define the function to be constant as $\frac{1}{2}$. We then need to remove $(\frac{1}{9}, \frac{2}{9})$ and $(\frac{7}{9}, \frac{8}{9})$, where we define the function to be constant as $\frac{1}{4}$ and $\frac{3}{4}$, respectively. We continue the process iteratively, so we define it to be constant function on each removed interval. Finally, we define the function to be 0 on the Cantor set.

Theorem 3.34. Let $F : [a, b] \to \mathbb{C}$ be a function, then the following are equivalent:

a. $F \in \mathrm{AC}([a, b])$;

b. there exists $f \in L^1([a, b], \mathcal{L}, \mathfrak{m})$ such that $F(x) - F(a) = \int_{[a,x)} f d\mathfrak{m}$;

c. F is differentiable for almost all $x \in [a, b]$, $F' \in L^1$, and $F(x) - F(a) = \int_a^x F'(t)dt$ for all $x \in [a, b]$.

Proof.

b. ⇒ a.: by part b., we know $F(x) = F(a) + \int_a^x f(t)dt$ for any $x \in [a, b]$, where $f \in L^1$. Set $g(x) = \int_a^x f(t)dt$, then it suffices to show $g \in \mathrm{AC}([a, b])$.

Claim 3.35. Let $f \in L^1$, then for any $\varepsilon > 0$, there exists $\delta > 0$ such that $\int_E |f| d\mathfrak{m} < \varepsilon$ whenever $\mathfrak{m}(E) < \delta$.

For any $\varepsilon > 0$, there exists $\delta > 0$ such that $\int_E |f| d\mathfrak{m} < \varepsilon$ whenever $\mathfrak{m}(E) < \delta$ because $f \in L^1$. For any disjoint intervals $(a_1, b_1), \ldots, (a_n, b_n)$ such that $\sum_{j=1}^{n} (b_j - a_j) < \delta$, we know $\mathfrak{m}\left(\bigcup_{j\geqslant 1} (a_j, b_j)\right) < \delta$, and thus

$$\sum_{j=1}^{n} |g(b_j) - g(a_j)| = \sum_{j=1}^{n} \int_{a_j}^{b_j} f d\mathfrak{m}$$
$$= \int_{\bigcup_{j=1}^{n} (a_j, b_j)} f d\mathfrak{m}$$

$$
\begin{aligned}
&\leqslant \int\limits_{\bigcup_{j=1}^{n}(a_j,b_j)} |f| d\mathfrak{m} \\
&< \varepsilon.
\end{aligned}
$$

a. ⇒ *c.*: we may assume F to be real-valued. Since $F \in \mathrm{AC}([a,b])$, then $F \in \mathrm{BV}([a,b])$, therefore $F = F_1 - F_2$ where F_1 and F_2 are bounded and increasing functions. Therefore, $\int_a^b F_1' d\mathfrak{m} \leqslant F_1(b) - F_1(a)$ and similarly $\int_a^b F_2' d\mathfrak{m} \leqslant F_2(b) - F_2(a)$. Now

$$
\begin{aligned}
\int_a^b |F'| d\mathfrak{m} &\leqslant \int_a^b F_1' d\mathfrak{m} + \int_a^b F_2' d\mathfrak{m} \\
&\leqslant F_1(b) - F_1(a) + F_2(b) - F_2(a) \\
&< \infty,
\end{aligned}
$$

therefore $F' \in L^1$. Now define $G(x) = \int_a^x F' d\mathfrak{m}$ to be a real-valued function since $F' \in L^1$, then by Claim 3.35, we know $G \in \mathrm{AC}([a,b])$. Therefore, $F - G \in \mathrm{AC}([a,b])$ as well.

Claim 3.36. $G'(x) = F'(x)$ for almost all $x \in [a,b]$.

Subproof. Recall that

$$
\begin{aligned}
G'(x) &= \lim_{h \to 0} \frac{G(x+h) - G(x)}{h} \\
&= \lim_{h \to 0} \frac{\int_x^{x+h} F'(t) dt}{h},
\end{aligned}
$$

which is $F'(x)$ for almost all $x \in [a,b]$ by Theorem 3.4 since $F' \in L'([a,b], \mathcal{L}, \mathfrak{m})$. ■

By Claim 3.36, $(F-G)' = 0$ almost everywhere on $[a,b]$. By Lemma 3.32, $F - G$ is a constant function, and this means $F(x) = G(x) + F(a)$ for all $x \in [a,b]$. Therefore,

$$
F(x) = F(a) + \int_a^x F' d\mathfrak{m}.
$$

c. ⇒ *b.*: take $f = F'$.

□

4 L^p-spaces

4.1 Basic Theory of L^p-spaces

Definition 4.1. Let $f : X \to \mathbb{C}$ be a measurable function on $(X, \mathcal{A}, \mu)$. Let $0 < p < \infty$, then the L^p-norm of f is defined by

$$||f||_p = \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}}.$$

Remark 4.2. This is a norm only when $1 \leqslant p < \infty$, since it satisfies Minkowski inequality

$$||f + g||_p \leqslant ||f||_p + ||g||_p.$$

When $0 < p < 1$, it is not an actual norm, but we recover a similar inequality

$$||f + g||_p \leqslant C_p(||f||_p + ||g||_p)$$

where $C_p > 1$.

Among $0 < p < \infty$, only $p = 2$ gives a Hilbert space, with a standard inner product structure on it.

Definition 4.3. For $0 < p < \infty$, we define the L^p-space to be the collection of functions $f : X \to \mathbb{C}$ where $||f||_p < \infty$, and denote it by $L^p(X, \mathcal{A}, \mu) = L^p(X) = L^p(\mu) = L^p$.

We can also define an L^p-space where $p = \infty$.

Definition 4.4. We define the L^∞-norm to be

$$||f||_\infty = \inf\{M \geqslant 0 : \mu(\{x \in X : |f(x)| > M\}) = 0\}.$$

The norm behaves very much like a maximal function, where we ignore the null sets. Therefore, we also call this the essential norm of $f(x)$, denoted $\operatorname{esssup}_{x\in X} |f(x)|$.

Remark 4.5. From the discussion above, we know $||f||_\infty$ behaves approximately like $\sup\limits_{x\in X} |f(x)|$.

Lemma 4.6. $|f(x)| \leqslant ||f||_\infty$ almost everywhere on X.

Proof. It suffices to show that $\mu(\{x \in X : |f(x)| > ||f||_\infty\}) = 0$. Let us write

$$\{x \in X : |f(x)| > ||f||_\infty\} = \bigcup_{n\geqslant 1}\{x \in X : |f(x)| > ||f||_\infty + \frac{1}{n}\},$$

then

$$\mu(\{x \in X : |f(x)| > ||f||_\infty\}) \leqslant \sum_{n\geqslant 1} \mu(\{x \in X : |f(x)| > ||f||_\infty + \frac{1}{n}\}).$$

It remains to show that for any $n \in \mathbb{N}$,

$$\mu(\{x \in X : |f(x)| > ||f||_\infty + \frac{1}{n}\}) = 0.$$

This is true by definition, otherwise there exists some $n \in \mathbb{N}$ such that $||f||_\infty \geqslant ||f||_\infty + \frac{1}{n}$, contradiction. $\square$

Definition 4.7. The L^∞-space is defined as the collection of measurable functions $f : X \to \mathbb{C}$ such that $||f||_\infty < \infty$.

Remark 4.8. Obviously $||f + g||_\infty \leqslant ||f||_\infty + ||g||_\infty$.

4.2 Distribution Functions and Weak L^p-space

Definition 4.9. The distribution function is given by

$$\begin{aligned}\lambda_f : (0,\infty) &\to [0,\infty]\\ \alpha &\mapsto \mu(\{x \in X : |f(x)| > \alpha\})\end{aligned}$$

Lemma 4.10. λ_f is a decreasing and right-continuous function.

Proof. To see λ_f is decreasing, note that $\{x \in X : |f(x)| > \alpha\} \subseteq \{x \in X : |f(x)| > \beta\}$ whenever $\alpha \geqslant \beta$. To prove λ_f is a right-continuous function, we can write

$$\begin{aligned}\{x \in X : |f(x)| > \alpha\} &= \bigcup_{n\geqslant 1}\{x \in X : |f(x)| > \alpha + \frac{1}{n}\}\\ &=: \bigcup_{n\geqslant 1} E_n.\end{aligned}$$

Note that $E_n \subseteq E_{n+1}$ whenever $n \geqslant 1$, so by continuity from below we know

$$\begin{aligned}\lim_{n\to\infty}\lambda_f(\alpha+\frac{1}{n}) &= \lim_{n\to\infty}\mu(E_n)\\ &= \mu\left(\bigcup_{n\geqslant 1} E_n\right)\\ &= \lambda_f(\alpha).\end{aligned}$$

Therefore, $\lim\limits_{t\to 0^+}\lambda_f(\alpha+t) = \lambda_f(\alpha)$. □

Theorem 4.11. For any $0 < p < \infty$,

$$||f||_p^p = \int_X |f|^p d\mu = p\int_0^\infty \alpha^{p-1}\lambda_f(\alpha)d\alpha.$$

Proof. We prove the case where μ is σ-finite. Let $E_\alpha = \{x \in X : |f(x)| > \alpha\}$, then $\lambda_f(\alpha) = \mu(E_\alpha) = \int_X \chi_{E_\alpha}(x)d\mu$. Now by Theorem 2.102,

$$\begin{aligned}p\int_0^\infty \alpha^{p-1}\lambda_f(\alpha)d\alpha &= p\int_0^\infty \alpha^{p-1}\int_X \chi_{E_\alpha}(x)d\mu d\alpha\\ &= p\int_X\int_0^\infty \alpha^{p-1}\chi_{E_\alpha}(x)d\alpha d\mu\\ &= p\int_X\int_0^{|f(x)|}\alpha^{p-1}d\alpha d\mu\\ &= \int_X |f(x)|^p d\mu\\ &= ||f||_p^p.\end{aligned}$$

To prove this in general, we run the usual argument: first prove it on simple functions, then by the denseness of simple functions in L^p-space to pass the result to $f \in L^p$ by the monotone convergence theorem. □

Definition 4.12. For $0 < p < \infty$, we define the weak L^p-norm to be

$$||f||_{p,\infty} = \left(\sup_{\alpha>0} \alpha^p \lambda_f(\alpha)\right)^{\frac{1}{p}}.$$

Definition 4.13. For $0 < p < \infty$, the weak L^p-space is defined to be

$$L^{p,\infty} = \{f : X \to \mathbb{C} : ||f||_{p,\infty} < \infty\}.$$

4.3 Some Useful Inequalities

Theorem 4.14 (Chebyshev Inequality). For any $\alpha > 0$ and any $f \in L^p$,

$$\lambda_f(\alpha) \leqslant \frac{||f||_p^p}{\alpha^p}.$$

Proof. We have

$$\begin{aligned} ||f||_p^p &= \int\limits_X |f|^p d\mu \\ &\geqslant \int\limits_{\{x\in X:|f(x)|>\alpha\}} |f|^p d\mu \\ &\geqslant \int\limits_{\{x\in X:|f(x)|>\alpha\}} \alpha^p d\mu \\ &= \alpha^p \mu(\{x \in X : |f(x)| > \alpha\}) \\ &= \alpha^p \lambda_f(\alpha). \end{aligned}$$

Since $\alpha > 0$, then $\lambda_f(\alpha) \leqslant \frac{||f||_p^p}{\alpha^p}$ as desired. ☐

Corollary 4.15. For any $\alpha > 0$ and any $f \in L^p$,

$$||f||_{p,\infty} \leqslant ||f||_p.$$

Therefore, a function in the L^p-space is a function in the weak L^p-space.

Lemma 4.16. Let $a, b \geqslant 0$ and $0 < \theta < 1$, then

$$a^\theta b^{1-\theta} \leqslant \theta a + (1-\theta)b \tag{4.17}$$

where the equality holds if and only if $a = b$.

Proof. First note that Equation (4.17) is trivial if $b = 0$. Now suppose $b \neq 0$, therefore $b > 0$, then Equation (4.17) is equivalent to

$$\left(\frac{a}{b}\right)^\theta \leqslant \theta\left(\frac{a}{b}\right) + (1-\theta).$$

Set $t = \frac{a}{b} \geqslant 0$, then we just need to prove that

$$t^\theta \leqslant \theta t + (1-\theta)$$

with equality if and only if $t = 1$.

Let $f(t) = t^\theta - \theta t$ be defined for $t \in [0, \infty)$. The derivative is $f'(t) = \theta t^{\theta-1} - \theta$. Note that

$$\begin{aligned} f'(t) = 0 &\iff \theta(t^{\theta-1} - 1) = 0 \\ &\iff t = 1. \end{aligned}$$

Therefore $t = 1$ gives the unique critical point, therefore it is a global extremum point. We find the second derivative to be $f''(t) < 0$ for any $t \in [0, \infty)$. In particular, $f'(t) < 0$ if $t > 1$ and $f'(t) > 0$ if $t < 1$. Therefore $f(t) \leqslant f(1)$, which indicates $t^\theta \leqslant \theta t + (1-\theta)$. ☐

Theorem 4.18 (Hölder Inequality)**.** Let $1 \leqslant p \leqslant \infty$, then

$$||fg||_1 \leqslant ||f||_p ||g||_{p'} \tag{4.19}$$

where p' is the conjugate of p, i.e., $\frac{1}{p} + \frac{1}{p'} = 1$.

Remark 4.20. If $p = \infty$, then $p' = 1$; if $p = 1$, then $p' = \infty$.

Proof.

- If $p = 1$, then $p' = \infty$, then by Lemma 4.6,

$$\begin{aligned} ||fg||_1 &= \int_X |fg| d\mu \\ &\leqslant \int_X |f| \cdot ||g||_\infty \\ &= ||f||_1 ||g||_\infty. \end{aligned}$$

- If $p = \infty$, then $p' = 1$, now the argument is the same as the previous case.
- If $1 < p < \infty$, then $1 < p' < \infty$.
  - Suppose $||f||_p = 0$ or $||g||_{p'} = 0$, then $f \equiv 0$ almost everywhere or $g \equiv 0$ almost everywhere. This means $||fg||_1 = 0$, hence Equation (4.19) holds.
  - If $||f||_p = \infty$ or $||g||_{p'} = \infty$, then we may assume that $||f||_p = \infty$ and $||g||_{p'} \neq 0$, then $||f||_p ||g||_{p'} = \infty$, which implies Equation (4.19).
  - This reduces the problem to the following case: suppose $0 < ||f||_p < \infty$ and $0 < ||g||_{p'} < \infty$. Therefore, Equation (4.19) is equivalent to
$$\left|\left| \frac{f}{||f||_p} \frac{g}{||g||_{p'}} \right|\right|_1 \leqslant 1$$
by normalization. It remains to show that $||fg||_1 \leqslant 1$ for any f, g with $||f||_p = 1$ and $||g||_{p'} = 1$. Let $a = |f(x)|^p$ and $b = |g(x)|^{p'}$, and set $\theta = \frac{1}{p}$. By Lemma 4.16,
$$|f(x)g(x)| \leqslant \frac{1}{p}|f(x)|^p + \frac{1}{p'}|g(x)|^{p'},$$
therefore
$$\begin{aligned} ||fg||_1 &\leqslant \frac{1}{p}\int_X |f|^p d\mu + \frac{1}{p'}\int_X |g|^{p'} d\mu \\ &= \frac{1}{p} + \frac{1}{p'} \\ &= 1. \end{aligned}$$

$\square$

Theorem 4.21. Suppose $1 < p < \infty$, then

$$||fg||_1 = ||f||_p ||g||_{p'}$$

if and only if

$$\alpha|f(x)|^p = \beta|g(x)|^{p'}$$

for almost all $x \in X$ and some constant α, β where $\alpha, \beta \neq 0$.

Theorem 4.22 (Minkowski). Let $1 \leqslant p \leqslant \infty$, then for any $f, g \in L^p$,

$$||f+g||_p \leqslant ||f||_p + ||g||_p.$$

Proof.

- If $p = \infty$ or $p = 1$, the proof is trivial.
- Now suppose $1 < p < \infty$, then for $\frac{1}{p} + \frac{1}{p'} = 1$, we have

$$\begin{aligned}|f+g|^p &= |f+g||f+g|^{p-1} \\ &\leqslant |f||f+g|^{p+1} + |g||f+g|^{p+1}.\end{aligned}$$

Therefore

$$\begin{aligned}\int_X |f+g|^p d\mu &\leqslant \int_X |f||f+g|^{p-1} d\mu + \int_X |g||f+g|^{p-1} d\mu \\ &\leqslant ||f||_p \left(\int_X |f+g|^{(p-1)p'} d\mu\right)^{\frac{1}{p'}} + ||g||_p \left(\int_X |f+g|^{(p-1)p'}\right)^{\frac{1}{p'}} \\ &= (||f||_p + ||g||_p) \left(\int_X |f+g|^p d\mu\right)^{\frac{1}{p'}}\end{aligned}$$

by Theorem 4.18. We may assume that $||f+g||_p \neq 0$, then this gives $||f+g||_p \leqslant ||f||_p + ||g||_p$.

$\square$

Theorem 4.23. Let $1 < p < \infty$ and $f, g \in L^p$, then $||f+g||_p = ||f||_p + ||g||_p$ if and only if there exists constant $C \geqslant 0$ such that either $f = Cg$ almost everywhere or $g = Cf$ almost everywhere.

4.4 The Dual of L^p-space

Definition 4.24. Let X be a vector space over $\mathbb{C}$ (or $\mathbb{R}$). A linear map $X \to \mathbb{C}$ is called a linear functional on X.

Definition 4.25. Let T be a linear functional on the normed space X, i.e., a vector space equipped with a norm function. We say T is a bounded linear functional if there exists $0 < C \in \mathbb{R}$ such that T satisfies

$$|T(x)| \leqslant C||x||$$

for every $x \in X$.

Definition 4.26. The dual space of X, denoted X^*, is the collection of all bounded linear functionals on X.

Theorem 4.27. Let $1 \leqslant p < \infty$, then for any $f \in L^p$,

$$||f||_p = \sup\left\{\left|\int_X fg d\mu\right| : ||g||_{p'} = 1\right\}.$$

Moreover, if μ is semi-finite, then

$$||f||_\infty = \sup\left\{\left|\int_X fg d\mu\right| : ||g||_1 = 1\right\}.$$

Theorem 4.28. If $1 < p < \infty$, then $(L^p)^* \cong L^{p'}$ is an isometric isomorphism, i.e., preserving the L^p-norm. Moreover, assuming μ to be σ-finite, now if $p = \infty$, then $(L^1)^* \cong L^\infty$.

Remark 4.29. We know $(L^\infty)^* \supseteq L^1$, but not necessarily an isometric isomorphism.

Theorem 4.30. The collection of continuous functions $C(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ if $1 \leqslant p < \infty$.

Theorem 4.31. The collection of continuous functions with compact support $C_c(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $1 \leqslant p < \infty$. That is, for any $f \in L^p$ where $1 \leqslant p < \infty$ and any $\varepsilon > 0$, there exists a function $g \in C_c(\mathbb{R}^n)$ such that $||f - g||_p < \varepsilon$.

Corollary 4.32. For $1 \leqslant p < \infty$, then

$$\lim_{t\to 0} \int_{\mathbb{R}} |f(x+t) - f(x)|^p dx = 0.$$

Proof. By Theorem 4.31, for any $\varepsilon > 0$, there exists $g \in C_c(\mathbb{R})$ such that $||f - g||_p < \varepsilon$.

Claim 4.33.

$$\lim_{t\to 0} \int_{\mathbb{R}} |g(x+t) - g(x)|^p dx = 0.$$

Subproof. Let $|t| \leqslant 1$, then there exists a compact set K such that

$$\int_{\mathbb{R}} |g(x+t) - g(x)|^p dx = \int_K |g(x+t) - g(x)|^p dx.$$

In particular, since g has compact support, so there exists $0 < M \in \mathbb{R}$ such that $|g(x)| \leqslant M$ for all $x \in K$, and $|g(x+t)| \leqslant M$ for all $x \in K$ and $|t| \leqslant 1$. Therefore, $|g(x+t) - g(x)|^p \leqslant 2^p M^p$ for all $x \in K$ and $|t| \leqslant 1$. Note that $2^p M^p \in L^1(K)$: the function is bounded since it is continuous. Therefore, by Theorem 2.46, we have

$$\begin{aligned}
\lim_{t\to 0} \int_{\mathbb{R}} |g(x+t) - g(x)|^p dx &= \int_{\mathbb{R}} \lim_{t\to 0} |g(x+t) - g(x)|^p dx \\
&= \int_{\mathbb{R}} \left(\lim_{t\to 0} |g(x+t) - g(x)|\right)^p dx \\
&= 0
\end{aligned}$$

by continuity. ■

Moreover, note that

$$\begin{aligned}
\left(\int_{\mathbb{R}} |f(x+t) - f(x)|^p dx\right)^{\frac{1}{p}} &= \left(\int_{\mathbb{R}} |f(x+t) - g(x+t) + g(x+t) - g(x) + g(x) - f(x)|^p dx\right)^{\frac{1}{p}} \\
&\leqslant \left(\int_{\mathbb{R}} |f(x+t) - g(x+t)|^p dx\right)^{\frac{1}{p}} + \left(\int_{\mathbb{R}} |g(x+t) - g(x)|^p dx\right)^{\frac{1}{p}} \\
&+ \left(\int_{\mathbb{R}} |g(x) - f(x)|^p dx\right)^{\frac{1}{p}} \\
&= 2||f - g||_p + \left(\int_{\mathbb{R}} |g(x+t) - g(x)|^p dx\right)^{\frac{1}{p}}.
\end{aligned}$$

Let $t \to 0$, then

$$\left(\limsup_{t\to 0} \int_{\mathbb{R}} |f(x+t) - f(x)|^p dx\right)^{\frac{1}{p}} \leqslant 2||f - g||_p$$

$$\leqslant 2\varepsilon$$

for any $\varepsilon > 0$ by Claim 4.33. Take $\varepsilon \to 0$, this forces limit to exists, and in particular $\lim_{t\to 0} \int_{\mathbb{R}} |f(x+t) - f(x)|^p dx = 0$. □

Remark 4.34. Note that $(L^p)^* \cong L^p$, i.e., L^p-space being self-adjoint, is true if and only if $p = 2$.