

Lecture Notes
MATH205A: Real Analysis I (Autumn 2020)
@ Stanford University

Tianyu Du

October 12, 2020

1 Measures

1.1 Motivation

Motivation of this course is to define a notion of *length* on subsets of \mathbb{R} such that

1. $length([a, b]) = b - a$.
2. (countable additivity) $length(\bigcup^\infty A_i) = \sum^\infty length(A_i)$ where A_i 's are disjoint.
3. (translation invariance) for all $a \in \mathbb{R}$, $length(A + a) = length(A)$.

Fact 1.1. it is impossible to construct such length for all subsets of \mathbb{R} .

Proof. This proof shows it is impossible to construct a notion of length on $[0, 1]$ with desired properties.

For $x, y \in [0, 1]$, define an equivalence relation as $x \sim y \iff x - y \in \mathbb{Q}$. By the axiom of choice, we may construct a set A containing exactly one element from each equivalence class of $x \in [0, 1]$. Obviously, $A \subseteq [0, 1]$.

For each $r \in [-1, 1] \cap \mathbb{Q}$, let $A_r := A + r$, and $A_r \subseteq [-1, 2]$. By translation invariance, $length(A_r) = length(A)$. Note that for any $y \in [0, 1]$, there exists some $x \in A$ such that $x \sim y$, therefore, $y \in A_{y-x} \subseteq \bigcup_r A_r$. Hence, $[0, 1] \subseteq \bigcup_r A_r$.

If the notion of length satisfies countable additivity, $length(\bigcup_r A_r)$ is either zero or infinity, which leads to a contradiction. ■

Lebesgue's Resolution: we only defines length for a subset of $\mathcal{P}(\mathbb{R})$, which contains *everything that may ever arrive in practice*, i.e., σ -algebras.

1.2 Algebras and σ -algebra

Definition 1.1. Let X be a set, a collection \mathcal{A} of subsets of X is called an **algebra** if

1. $X \in \mathcal{A}$,

$$2. A \in \mathcal{A} \implies A^c \in \mathcal{A},$$

$$3. A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}.$$

Consequently: (1) $A, B \in \mathcal{A} \implies A \cap B \in \mathcal{A}$; (2) $A_1, \dots, A_n \in \mathcal{A} \implies \bigcup_i A_i, \bigcap_i A_i \in \mathcal{A}$ (easily shown by induction); (3) $\emptyset \in \mathcal{A}$.

Definition 1.2. Let X be a set, a collection \mathcal{A} of subsets of X is called a σ -algebra if

$$1. X \in \mathcal{A},$$

$$2. A \in \mathcal{A} \implies A^c \in \mathcal{A},$$

$$3. A_1, A_2, \dots \in \mathcal{A} \implies \bigcup_i^\infty A_i \in \mathcal{A}.$$

Example 1.1 (trivial examples). The power set of X is a σ -algebra on X ; $\{\emptyset, X\}$ is a σ -algebra on X .

Example 1.2 (finite/co-finite algebra). Let X be an infinite set and \mathcal{A} be the collection of subsets A such that either A is finite or A^c is finite. \mathcal{A} is an algebra.

Proof. $X \in \mathcal{A}$ since $X^c = \emptyset$ is finite. For a $X \in \mathcal{A}$, if X is finite, then $X^c \in \mathcal{A}$. If X is infinite, X^c is finite and $X^c \in \mathcal{A}$. Let $A, B \in \mathcal{A}$, if both A and B are finite, $A \cup B$ is finite and in \mathcal{A} . If A is finite and B is co-finite, then $(A \cup B)^c = A^c \cap B^c \subseteq B^c$ is finite. If both A and B are co-finite, $(A \cup B)^c$ is finite so that $A \cup B \in \mathcal{A}$. ■

Note the \mathcal{A} is not a σ -algebra if X is infinite: take distinct points $x_1, x_2, \dots \in \mathcal{A}$, then the union of them is neither finite or co-finite, and therefore not in \mathcal{A} .

Example 1.3 (countable/co-countable σ -algebra). The collection of subsets $A \subseteq X$, such that either A is countable or A^c is countable, forms a σ -algebra.

Example 1.4. Let $X = \mathbb{R}$ and \mathcal{A} be the collection of all finite disjoint unions of half-open intervals (i.e., sets like $(a, b], (-\infty, b], (a, \infty)$), \mathcal{A} is an algebra. (Not working for open intervals).

Example 1.5 (counter example). Let X be an infinite set, \mathcal{A} be the collection of finite subsets of X . Then, \mathcal{A} is not an algebra.

Proposition 1.1. Let X be a set and $\{\mathcal{A}_i\}_{i \in \mathcal{I}}$ be an arbitrary (not necessarily countable) collection of σ -algebras, then $\bigcap_{i \in \mathcal{I}} \mathcal{A}_i$ is a σ -algebra.

Proof. Since $X \in \mathcal{A}_i$ for all $i \in \mathcal{I}$ ■

Corollary 1.1. Let X be a set, and \mathcal{P} is an arbitrary collection of subsets of X , then $\exists!$ smallest σ -algebra \mathcal{A} containing \mathcal{P} . That is, for any σ -algebra $\mathcal{B} \supseteq \mathcal{P}$, $\mathcal{A} \subseteq \mathcal{B}$. \mathcal{A} is defined as the σ -algebra **generated by** \mathcal{P} , denoted as $\sigma(\mathcal{P})$.

Proof. For any \mathcal{P} , the power set of X is obviously a σ -algebra containing \mathcal{P} . Then we can take \mathcal{A} as the intersection of all σ -algebras containing \mathcal{P} . ■

1.3 Borel σ -algebra

Definition 1.3. The **Borel σ -algebra** of \mathbb{R} , denoted as $\mathcal{B}(\mathbb{R})$, is the σ -algebra generated by the set of open intervals in \mathbb{R} .

Fact 1.2. $\mathcal{B}(\mathbb{R})$ is generated by the collection of all closed intervals as well.

Proof. Let \mathcal{F} denote the σ -algebra generated by all closed intervals. Any open interval can be written as a countable union of closed sets: $(a, b) = \bigcup_{n=1}^{\infty} [a + 1/n, b - 1/n]$, therefore $(a, b) \in \mathcal{F}$ and $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{F}$.

Similarly, $[a, b] = \bigcap_{n=1}^{\infty} (a - 1/n, b + 1/n)$, hence $\mathcal{B}(\mathbb{R})$ is a σ -algebra contains all closed sets. Therefore, $\mathcal{F} \subseteq \mathcal{B}(\mathbb{R})$. ■

Fact 1.3. $\mathcal{B}(\mathbb{R})$ is generated by

1. all open sets,
2. all closed sets,
3. all half-open intervals.

Example 1.6 (counter example). $\mathcal{B}(\mathbb{R})$ is not generated by the collection of singletons.

Proof. ■

Definition 1.4. The Borel algebra of \mathbb{R}^d , $\mathcal{B}(\mathbb{R}^d)$, is the σ -algebra generated by

1. all open sets in \mathbb{R}^d ,
2. all closed sets in \mathbb{R}^d ,
3. all closed cubes (regions) in \mathbb{R}^d : $\prod_{i=1}^d [a_i, b_i]$.

1.4 Measures

Definition 1.5. For a set X and a σ -algebra \mathcal{A} of X , the pair (X, \mathcal{A}) is called a **measurable space**.

Definition 1.6. A **measure** μ on a measurable space (X, \mathcal{A}) is a map $\mu : \mathcal{A} \rightarrow [0, \infty]$ such that

1. $\mu(\emptyset) = 0$,
2. $\mu(\bigcup_i^{\infty} A_i) = \sum_i^{\infty} \mu(A_i)$ for disjoint sequence (A_i)

For now, we don't require the translation invariance property.

The triple (X, \mathcal{A}, μ) is called a **measure space**.

Example 1.7 (counting measure).

Example 1.8 (point-mass measure).

Proposition 1.2. A measure μ possesses the following basic properties:

1. (Monotonicity) $A \subseteq B \implies \mu(A) \leq \mu(B)$.
2. (Sub-additivity) $\mu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$.
3. Let $A_1 \subseteq A_2 \subseteq \dots$ be an increasing set, let $\bigcup_{i=1}^{\infty} A_i$ denoted $A_i \nearrow A$, $\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$.
4. If $A_1 \searrow A \equiv \bigcap_{i=1}^{\infty} A_i$, and **there exists** $\mu(A_i) < \infty$, then $\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n)$.

Proof. ■

Example 1.9 (counter example). Let $X = \mathbb{Z}$, $\mathcal{A} = 2^{\mathbb{Z}}$ and μ be the counting measure. Define $A_i = \{i, i+1, \dots\}$, then $A_i \searrow A = \emptyset$, but $\lim_{n \rightarrow \infty} \mu(A_n) = \infty \neq \mu(\emptyset)$.

1.5 Outer Measure

Definition 1.7. Let X be a set, $\mu^* : 2^X \rightarrow [0, \infty]$ is an **outer measure** if

1. $\mu^*(\emptyset) = 0$.
2. $\mu^*(A) \leq \mu^*(B)$ whenever $A \subseteq B$.
3. (countable sub-additivity) $\mu^*(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$.

Key difference between outer measure and measure:

1. Outer measure does not require countable additivity,
2. outer measure is defined on 2^X instead of a σ -algebra .

Example 1.10.

1.6 Lebesgue Measure on \mathbb{R}

Definition 1.8. Let $A \subseteq \mathbb{R}$, define the **Lebesgue outer measure**:

$$\lambda^*(A) = \inf \left\{ \sum_{i \in \mathbb{N}} b_i - a_i : A \subseteq \bigcup_{i \in \mathbb{N}} (a_i, b_i) \right\} \quad (1)$$

The Lebesgue outer measure of a set A is simply in the infimum of total lengths (the conventional notion of length) of open intervals cover A .

Proposition 1.3. The Lebesgue measure satisfies the following properties:

1. λ^* is an outer measure on \mathbb{R} ,
2. $\lambda^*([a, b]) = b - a$ for all $a < b$.

Proof. (1.1) $\lambda^*(\emptyset) = 0$ since $(-\varepsilon, \varepsilon)$ covers \emptyset for arbitrarily small ε .

(1.2) Let $A \subseteq B$, Ω_A and Ω_B be collection of sequences of open intervals covering A and B respectively. Then, any cover of B must be a cover of A , that is, $\Omega_A \subseteq \Omega_B$. Therefore, $\lambda^*(A) \leq \lambda^*(B)$.

(1.3) Let $A_1, A_2, \dots \subseteq \mathbb{R}$ and $A = \bigcup_{i=1}^{\infty} A_i$. For all i , we may find (a_{ij}, b_{ij}) covers A_i such that

$$\sum_{j=1}^{\infty} (b_{ij} - a_{ij}) \leq \lambda^*(A_i) + \varepsilon 2^{-i} \quad (2)$$

Also, $\{(a_{ij}, b_{ij})\}_{i,j}$ is a countable union of open intervals that covers A .

$$\lambda^*(A) \leq \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (b_{ij} - a_{ij}) \quad (3)$$

$$\leq \sum_{i=1}^{\infty} (\lambda^*(A_i) + \varepsilon 2^{-i}) \quad (4)$$

$$= \sum_{i=1}^{\infty} \lambda^*(A_i) + \varepsilon \quad (5)$$

Therefore, $\lambda^*(A) \leq \sum_{i=1}^{\infty} \lambda^*(A_i)$.

(2) Note that $[a, b] \subseteq (a - \varepsilon, b + \varepsilon)$ for all $\varepsilon > 0$. Therefore,

$$\lambda^*([a, b]) \leq \inf_{\varepsilon > 0} \lambda^*(a - \varepsilon, b + \varepsilon) = b - a \quad (6)$$

Now show $\lambda^*([a, b]) \geq b - a$. We want to show that $\sum_{i=1}^{\infty} (b_i - a_i) \geq b - a$ for all possible covering of $[a, b]$, which implies the infimum of them is at least $b - a$.

Take an arbitrary covering $\{(a_i, b_i)\}_i$ of $[a, b]$. Since $[a, b]$ is compact, there exists a finite covering $[a, b] \subseteq \bigcup_{i=1}^n (a_i, b_i)$ (reindexed), it suffices to show the finite sum $\sum_{i=1}^n (b_i - a_i) \geq b - a$.

(1) We firstly define an *interval* to be any open, closed or half-open intervals. The *length* of an interval is the difference between two end points.

Note that if an interval I contains a finite collection of disjoint sub-intervals, then the length of I is at least the sum of lengths of sub-intervals. The equality holds when I is exactly finite union of disjoint sub-intervals.

(2) Suppose $[a, b] \subseteq \bigcup_{i=1}^n (a_i, b_i)$, let $I_i = [a, b] \cap (a_i, b_i)$. Easy to verify that the length of $I_i \leq$ length of $(a_i, b_i) = b_i - a_i$. Moreover, $\bigcup_{i=1}^n I_i = [a, b] \cup \bigcup_{i=1}^n (a_i, b_i) = [a, b]$.

(3) For all i , define $I'_i = I_i \setminus (I_1 \cup I_2 \cup \dots \cup I_{i-1})$. This procedure allows us to express $[a, b]$ as a finite union of disjoint sub-intervals: $[a, b] = \bigcup_{i=1}^n I'_i$. Each I'_i is a finite union of disjoint intervals as well, the conventional notion of I'_i is well-defined. Then $b - a =$ sum of lengths of I'_i .

However, $\ell(I'_i) \leq \ell(I_i) \leq b_i - a_i$ and sum of lengths of $I'_i \leq$ sum of lengths of $I_i \leq \sum_{i=1}^n b_i - a_i$. Therefore, $b - a \leq \sum_{i=1}^n b_i - a_i \leq \sum_{i=1}^{\infty} b_i - a_i$. Hence, $b - a = \sum_{i=1}^{\infty} b_i - a_i$ and $\lambda^*[a, b] = b - a$ consequently. ■

1.7 Construct Lebesgue Measure

Definition 1.9. Let X be a set with outer measure μ^* . A set $B \subseteq X$ is μ^* -**measurable** if

$$\forall A \subseteq X, \mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c) \quad (7)$$

Theorem 1.1. For any set X with outer measure μ^* on it, let \mathcal{M}_{μ^*} denote the set of all μ^* -**measurable** sets. Then, \mathcal{M}_{μ^*} is a σ -algebra and $\mu^*|_{\mathcal{M}_{\mu^*}}$ (μ^* restricted to \mathcal{M}_{μ^*}) is a measure.

Proof. To show B is μ^* -measurable, it suffices to show that $\forall A \subseteq X, \mu^*(A) \geq \mu^*(A \cap B) + \mu^*(A \cap B^c)$, because the opposite inequality always holds by sub-additivity.

(1.1) Let $A \subseteq X$, $\mu^*(A \cap \emptyset) + \mu^*(A \cap \emptyset^c) = \mu^*(A \cap \emptyset^c) = \mu^*(A)$, therefore, $\emptyset \in \mathcal{M}_{\mu^*}$.

(1.2) Let $A \subseteq X$ and $B \in \mathcal{M}_{\mu^*}$, $\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c) = \mu^*(A \cap (B^c)^c) + \mu^*(A \cap B^c)$.

Hence, $B^c \in \mathcal{M}_{\mu^*}$.

(1.3.1) Let $B_1, B_2 \in \mathcal{M}_{\mu^*}$, we are going to show $B_1 \cup B_2 \in \mathcal{M}_{\mu^*}$. Fix any $A \subseteq X$,

$$\mu^*(A \cap (B_1 \cup B_2)) = \mu^*(A \cap (B_1 \cup B_2) \cap B_1) + \mu^*(A \cap (B_1 \cup B_2) \cap B_1^c) \quad (8)$$

$$= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c \cap B_2) \quad (9)$$

Moreover,

$$\mu^*(A \cap (B_1 \cup B_2)^c) = \mu^*(A \cap B_1^c \cap B_2^c) \quad (10)$$

Therefore,

$$\mu^*(A \cap (B_1 \cup B_2)) + \mu^*(A \cap (B_1 \cup B_2)^c) = \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c \cap B_2) + \mu^*(A \cap B_1^c \cap B_2^c) \quad (11)$$

$$= \mu^*(A \cap B_1) + \mu^*(A \cap B_1^c) \text{ since } B_2 \in \mathcal{M}_{\mu^*} \quad (12)$$

$$= \mu^*(A) \text{ since } B_1 \in \mathcal{M}_{\mu^*} \quad (13)$$

Therefore, \mathcal{M}_{μ^*} is an algebra.

(1.3.2) Now show that \mathcal{M}_{μ^*} is a σ -algebra. For any sequence of sets $A_i \in \mathcal{M}_{\mu^*}$, we can define $B_n := A_n \setminus \bigcup_{j=1}^{n-1} A_j$ such that $\bigcup B_i = \bigcup A_i$. Therefore, it suffices to show \mathcal{M}_{μ^*} is closed under countable disjoint unions.

We are going to show the union $\bigcup B_i$ is μ^* -measurable for any disjoint sequence of μ^* -measurable B_i 's.

Claim: let $A \subseteq X$, $\mu^*(A) = \sum_{i=1}^n \mu^*(A \cap B_i) + \mu^*(A \cap (\bigcup_{i=1}^n B_i)^c)$. The claim can be proved by induction on n .

When $n = 1$, $\mu^*(A) = \mu^*(A \cap B) + \mu^*(A \cap B^c)$ because B_1 is μ^* -measurable.

Suppose the claim holds for n , then

$$\mu^*(A \cap (\bigcup_{i=1}^n B_i)^c) = \mu^*(A \cap (\bigcup_{i=1}^n B_i)^c \cap B_{n+1}) + \mu^*(A \cap (\bigcup_{i=1}^n B_i)^c \cap B_{n+1}^c) \quad (14)$$

because $B_{n+1} \in \mathcal{M}_{\mu^*}$. Moreover, since all B_i 's are disjoint, $B_{n+1} \subseteq B_i^c$ for all i . Hence,

$$B_{n+1} \subseteq \cap_{i=1}^n B_i^c = (\cup_{i=1}^n B_i)^c \quad (15)$$

Also,

$$(\cup_{i=1}^n B_i)^c \cap B_{n+1}^c = \cap_{i=1}^{n+1} B_i^c \quad (16)$$

Consequently,

$$\mu^*(A \cap (\cup_{i=1}^n B_i)^c) = \mu^*(A \cap B_{n+1}) + \mu^*(A \cap (\cup_{i=1}^{n+1} B_i)^c) \quad (17)$$

Hence,

$$\mu^*(A) = \sum_{i=1}^n \mu^*(A \cap B_i) + \mu^*(A \cap (\cap_{i=1}^n B_i^c)) \quad (18)$$

$$\geq \sum_{i=1}^n \mu^*(A \cap B_i) + \mu^*(A \cap (\cap_{i=1}^\infty B_i^c)) \quad (19)$$

$$= \sum_{i=1}^n \mu^*(A \cap B_i) + \mu^*(A \cap (\cup_{i=1}^\infty B_i)^c) \quad (20)$$

Take $n \rightarrow \infty$

$$\mu^*(A) \geq \sum_{i=1}^\infty \mu^*(A \cap B_i) + \mu^*(A \cap (\cup_{i=1}^\infty B_i)^c) \quad (21)$$

$$\geq \mu^*(A \cap \cup_{i=1}^\infty B_i) + \mu^*(A \cap (\cup_{i=1}^\infty B_i)^c) \quad (22)$$

Therefore, $\cup_{i=1}^\infty B_i$ is μ^* -measurable.

(2) Let B_1, B_2, \dots be a sequence of disjoint sets from \mathcal{M}_{μ^*} . Using the above fact and take $A = \cup_{i=1}^\infty B_i$,

$$\mu^*(A) \geq \mu^*(\cup_{i=1}^\infty B_i) + \mu^*(\emptyset) = \mu^*(\cup_{i=1}^\infty B_i) \quad (23)$$

The opposite inequality holds by sub-additivity. Therefore, μ^* is a measure on \mathcal{M}_{μ^*} . ■

Definition 1.10. Let λ^* be the Lebesgue outer measure on \mathbb{R} , then the collection \mathcal{M}_{λ^*} of λ^* -measurable sets is called the **Lebesgue σ -algebra**. The restriction $\lambda = \lambda^*|_{\mathcal{M}_{\lambda^*}}$, which is a measure on \mathcal{M}_{λ^*} , is called the **Lebesgue measure**. Any set in \mathcal{M}_{λ^*} is called a **Lebesgue measurable set**.

Theorem 1.2. $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{M}_{\lambda^*}$.

Proof. Note that $\{(-\infty, b] : b \in \mathbb{R}\}$ generates $\mathcal{B}(\mathbb{R})$, it suffices to show $\{(-\infty, b] : b \in \mathbb{R}\} \subseteq \mathcal{M}_{\lambda^*}$.

Let $B = (-\infty, b]$, we are going to show B is λ^* -measurable. Let $A \subseteq \mathbb{R}$ and (a_n, b_n) be a

sequence of open intervals covers A . For every $n \in \mathbb{N}$,

$$\lambda^*((a_n, b_n) \cap B) + \lambda^*((a_n, b_n) \cap B^c) = \lambda^*((a_n, b_n) \cap (-\infty, b]) + \lambda^*((a_n, b_n) \cap (b, \infty)) \quad (24)$$

Three cases follow:

1. $b > b_n$: $\lambda^*((a_n, b_n) \cap B) + \lambda^*((a_n, b_n) \cap B^c) = \lambda^*((a_n, b_n)) = b_n - a_n$.
2. $b_n > b > a_n$: $\lambda^*((a_n, b_n) \cap B) + \lambda^*((a_n, b_n) \cap B^c) = \lambda^*((a_n, b]) + \lambda^*((b, b_n]) = b_n - a_n$.
3. $a_n > b$: $\lambda^*((a_n, b_n) \cap B) + \lambda^*((a_n, b_n) \cap B^c) = \lambda^*((a_n, b_n)) = b_n - a_n$.

Therefore,

$$\lambda^*((a_n, b_n) \cap B) + \lambda^*((a_n, b_n) \cap B^c) = b_n - a_n \quad (25)$$

By monotonicity and sub-additivity:

$$\lambda^*(A \cap B) + \lambda^*(A \cap B^c) \leq \lambda^*(\cup(a_n, b_n) \cap B) + \lambda^*(\cup(a_n, b_n) \cap B^c) \quad (26)$$

$$\leq \sum \lambda^*((a_n, b_n) \cap B) + \lambda^*((a_n, b_n) \cap B^c) \quad (27)$$

$$= \sum_{n=1}^{\infty} b_n - a_n \quad (28)$$

Take the infimum of all such covering, we can show

$$\lambda^*(A \cap B) + \lambda^*(A \cap B^c) \leq \lambda^*(A) \quad (29)$$

Therefore, B is μ^* -measurable and \mathcal{M}_{λ^*} is a σ -algebra containing all such intervals and $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{M}_{\lambda^*}$. ■

1.8 Lebesgue Measure on \mathbb{R}^d

Definition 1.11. Steps to construct Lebesgue measure on \mathbb{R}^d :

1. Define open cubes on \mathbb{R}^d as a Cartesian product of open intervals: $Q := \prod_{i=1}^d (a_i, b_i)$. Define Lebesgue outer measure:

$$\lambda^*(A) := \inf \left\{ \sum_{n=1}^{\infty} \prod_{i=1}^d (b_{ni} - a_{ni}) : A \subseteq \bigcup_{n=1}^{\infty} Q_n \right\} \quad (30)$$

2. Show λ^* is an outer measure and $\lambda^*(Q) = \prod_{i=1}^d (b_i - a_i)$.
3. \mathcal{M}_{λ^*} is the Lebesgue σ -algebra on \mathbb{R}^d . Restricting λ^* on \mathcal{M}_{λ^*} defines the Lebesgue measure.
4. Show that any Borel set in \mathbb{R}^d is Lebesgue measurable by showing that there is a generating set of $\mathcal{B}(\mathbb{R}^d)$ is in \mathcal{M}_{λ^*} .

1.9 Uniqueness of the Lebesgue Measure

The next goal is to prove the uniqueness of Lebesgue measure on $\mathcal{B}(\mathbb{R}^d)$ subject to the criterion that the measure of any interval (cube) is the volume in the usual sense (product of side lengths).

Theorem 1.3. Let λ be the Lebesgue measure on \mathbb{R}^d , then for any Lebesgue measurable set A ,

1. $\lambda(A) = \inf\{\lambda(U) : \text{open } U \supseteq A\},$
2. $\lambda(A) = \sup\{\lambda(K) : \text{compact } K \subseteq A\}.$

Proof. (1.1) WLOG $\lambda(A) < \infty$, by monotonicity, $\lambda(A) \leq \lambda(U)$ for any open cover, $\lambda(A) \leq \inf\{.. \}$.

(1.2) Let $\varepsilon > 0$, \exists a sequence of open intervals (R_i) such that

$$\lambda(A) \leq \sum_{i=1}^{\infty} \lambda(R_i) \leq \lambda(A) + \varepsilon \quad (31)$$

Let $U := \cup R_i$ open, hence $\inf\{.. \} \leq \lambda(U) \leq \sum_{i=1}^{\infty} \lambda(R_i) \leq \lambda(A) + \varepsilon$. Since this ε can be arbitrarily small, we conclude $\inf\{.. \} \leq \lambda(A)$.

(2.1) let A be a Lebesgue measurable set, assume A is bounded, so that $\lambda(A) < \infty$. Then there exists a compact $C \supseteq A$. $C \setminus A$ is Lebesgue measurable as well.

By conclusion of part (1), there exists a open set $U \supseteq C \setminus A$ such that

$$\lambda(C \setminus A) \leq \lambda(U) \leq \lambda(C \setminus A) + \varepsilon \quad (32)$$

Let $K = C \setminus U$, K is compact. Moreover, let $a \in K$, then $a \in C$ and $a \notin U$. Therefore, $a \notin C \setminus A$, it must be $x \in A$. Hence, $K \subseteq A$.

$$\lambda(K) = \lambda(C \setminus U) \quad (33)$$

$$\geq \lambda(C) - \lambda(U) \quad (34)$$

$$\geq \lambda(C) - (\lambda(C \setminus A) + \varepsilon) \quad (35)$$

$$= \lambda(C) - \lambda(C) + \lambda(A) - \varepsilon \quad (36)$$

$$= \lambda(A) - \varepsilon \quad (37)$$

Take $\varepsilon \rightarrow 0$ and $\lambda(A) \leq \sup\{.. \}$. By monotonicity, $\lambda(A) \geq \sup\{.. \}$.

(2.2) Other cases: suppose A is unbounded and $\lambda(A) > 0$. Take an arbitrary $b < \lambda(A)$. We will show that $\sup\{.. \} \geq b$, this will prove that $\lambda(A) \leq \sup\{.. \}$.

To show $\sup\{.. \} \geq b$, it suffices to show that there exists a compact set $K \subseteq A$ such that $\lambda(K) \geq b$.

Let $\{C_j\}_{j=1}^{\infty}$ be a sequence of compact sets increasing to \mathbb{R}^d .

Then $A \cap C_j \uparrow A$ and $\lambda(A \cap C_1) < \infty$, which implies $\lambda(A) = \lim_{j \rightarrow \infty} \lambda(A \cap C_j)$. Since $b < \lambda(A)$, there exists j such that $\lambda(A \cap C_j) \geq b$, where $A \cap C_j$ is compact. Hence, $b \leq \sup\{.. \}$ and $\lambda(A) \leq \sup\{.. \}$. $\lambda(A) \geq \sup\{.. \}$ holds by monotonicity.

When $\lambda(A) = 0$, $0 \leq \lambda(K)$ for all K so that $0 \leq \sup\{.. \}$. The opposite inequality holds by monotonicity. ■

Lemma 1.1. For each $k \in \mathbb{Z}$, define **dyadic cubes** in \mathbb{R}^d as set in the following form:

$$\prod_{i=1}^d [j_i 2^{-k}, (j_i + 1) 2^{-k}) \quad (38)$$

where $j_i \in \mathbb{Z}$ for every i . Let \mathcal{D} denote the collection of dyadic cubes.

Then, any open set $U \subseteq \mathbb{R}^d$ can be expressed as a countable union of some members of \mathcal{D} .

A dyadic cube of side length 2^{-k} has a unique parent of side length 2^{-k+1} and a unique grandparent with side length 2^{-k+2} .

Proof. Given open set U , let \mathcal{D}_U denote the set of all dyadic half open cubes D such that $D \subseteq U$ but the parent of U does not fully contain U .

Claim 1: $U = \bigcup_{D \in \mathcal{D}_U} D$. Obviously, $\bigcup_{D \in \mathcal{D}_U} D \subseteq U$. To show the converse, take any $x \in U$, since U is open, there exists $D \in \mathcal{D}_U$ such that $x \in D \subseteq U$.

Let D_0 be the earliest ancestor of D such that $x \in D_0 \subseteq U$. Obviously, $D_0 \in \mathcal{D}_U$. Therefore, $U \subseteq \bigcup_{D \in \mathcal{D}_U} D$.

Claim 2: Two dyadic cubes can overlap if and only if one is the ancestor of the other. By construction, dyadic cubes in \mathcal{D}_U are disjoint.

Claim 3: \mathcal{D}_U is countable because \mathcal{D} is itself countable. ■

Proposition 1.4. Lebesgue measure is the only measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ which assigns the *correct volume* to any d -dimensional intervals or even any d -dimensional dyadic cube.

Proof. Let λ denote the Lebesgue measure, let μ be another measure satisfying the desired property.

By lemma, for all open set U , $\mu(U) = \sum_{j=1}^{\infty} \mu(D_j) = \sum_{j=1}^{\infty} \lambda(D_j) = \lambda(U)$, where $\{D_j\}$ is a collection of disjoint dyadic cubes contains with union U . Therefore, $\lambda(A) = \mu(A)$ for all open Borel set A .

Let $A \in \mathcal{B}(\mathbb{R}^d)$, let open $U \supseteq A$, then $\mu(A) \leq \mu(U) = \lambda(U)$ for all U . Taking the infimum over all $U \supseteq A$, we conclude $\mu(A) \leq \lambda(A)$ for all Borel set A .

Next, take any bounded Borel set A , let V be a bounded open set containing A . Then,

$$\mu(V) = \mu(A) + \mu(V \setminus A) \quad (39)$$

$$\leq \lambda(A) + \lambda(V \setminus A) \quad (40)$$

$$= \lambda(V) \quad (41)$$

But we also know that $\mu(V) = \lambda(V)$ since V is open, the inequality holds as equality. Moreover, the previous conclusion implies $\mu(A) \leq \lambda(A)$ and $\mu(V \setminus A) \leq \lambda(V \setminus A)$, it must be $\mu(A) = \lambda(A)$ and $\mu(V \setminus A) = \lambda(V \setminus A)$. Therefore, $\mu(A) = \lambda(A)$ for all bounded Borel set A .

Lastly, any Borel set can be written as a countable disjoint union of bounded Borel set, therefore, $\mu(A) = \lambda(A)$ for all Borel set A . ■

Proposition 1.5. The Lebesgue outer measure on \mathbb{R}^d is translation invariant. In particular, Lebesgue measure is translation invariant and any translation of Lebesgue measurable set is Lebesgue measurable.

Proof. $\lambda^*(A+x) = \lambda^*(A)$ follows the definition of λ^* : translate all covering intervals by $+x$ and the volumes of these intervals stay the same. Since λ is simply the restriction of λ^* on Lebesgue measurable sets, λ is translation invariant as well.

Now take Lebesgue measurable B , for all $A \subseteq \mathbb{R}^d$:

$$\lambda^*(A) = \lambda^*(A \cap B) + \lambda^*(A \cap B^c) \quad (42)$$

$$\implies \lambda^*(A-x) = \lambda^*((A-x) \cap B) + \lambda^*((A-x) \cap B^c) \quad (43)$$

Note that

$$(A-x) + x = A \quad (44)$$

$$(A-x) \cap B + x = A \cap (B+x) \quad (45)$$

$$(A-x) \cap B^c + x = A \cap (B+x)^c \quad (46)$$

By translational invariance of λ^* ,

$$\lambda^*(A) = \lambda^*(A \cap (B+x)) + \lambda^*(A \cap (B+x)^c) \quad (47)$$

Therefore, $B+x$ is Lebesgue measurable as well. ■

Theorem 1.4. Let μ be a non-zero measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, which is finite on bounded Borel sets and translation invariant. Then, $\mu(A) = c\lambda(A)$ for all $A \in \mathcal{B}(\mathbb{R}^d)$, where λ is the Lebesgue measure.

Remark 1.1. Borel σ -algebra is closed under translation.

Proof. Let $c = \mu([0,1]^d) \in (0, \infty)$. Then $[0,1]^d$ is the disjoint union of 2^{dk} half-open dyadic intervals with side length 2^{-k} . All of these sub-intervals have the same μ since μ is translation invariant. Therefore, for every dyadic sub-interval with side length 2^{-k} , $\mu(D) = 2^{-dk}c$.

Let $\nu(A) = \frac{1}{c}\mu(A)$, then ν is a measure that is finite on bounded sets and agrees with λ on all half-open dyadic cubes. By the previous proposition, λ is the only measure assign correct volumes to dyadic cubes, therefore, $\nu = \lambda$. ■

Theorem 1.5. Under the axiom of choice, there exists a non-Lebesgue subset of \mathbb{R} .

Proof. Todo. ■

2 Functions

2.1 Measurable Functions

Definition 2.1. A function $f : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$ is **measurable** if $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}$.

In this course, we mainly consider functions with extended- \mathbb{R} as codomain: $Y = [-\infty, \infty]$, denoted as \mathbb{R}^* .

Definition 2.2. The σ -algebra on \mathbb{R}^* is defined to be the σ -algebra generated by $\mathcal{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}$.

Proposition 2.1.

$$\sigma(\mathcal{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}) = \mathcal{B}(\mathbb{R}) \cup \{B \cup \{\infty\} : B \in \mathcal{B}(\mathbb{R})\} \quad (48)$$

$$\cup \{B \cup \{-\infty\} : B \in \mathcal{B}(\mathbb{R})\} \quad (49)$$

$$\cup \{B \cup \{-\infty, \infty\} : B \in \mathcal{B}(\mathbb{R})\} \quad (50)$$

Proposition 2.2. Equivalently, f is measurable if for every $t \in \mathbb{R}$,

$$\{x \in X : f(x) \leq t\} \in \mathcal{A} \quad (51)$$

$$\{x \in X : f(x) < t\} \in \mathcal{A} \quad (52)$$

$$\{x \in X : f(x) \geq t\} \in \mathcal{A} \quad (53)$$

$$\{x \in X : f(x) > t\} \in \mathcal{A} \quad (54)$$

More generally, to determine the measurability of $f : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$, we only need to check whether $f^{-1}(C) \in \mathcal{A}$ for all C in a generating collection \mathcal{C} of \mathcal{B} . The converse holds true trivially.

Proof. Let $(X, \mathcal{A}), (Y, \mathcal{B})$ be two measurable spaces, let \mathcal{C} be a collection of subsets of Y generates \mathcal{B} .

(\implies) Let f be a measurable function, then for every $C \in \mathcal{C} \subseteq \mathcal{B}$. Obviously, $f^{-1}(C) \in \mathcal{A}$ by definition.

(\impliedby) Suppose $f^{-1}(C) \in \mathcal{A}$ for all $C \in \mathcal{C}$. Define

$$\mathcal{B}_0 := \{B \in \mathcal{B} : f^{-1}(B) \in \mathcal{A}\} \supseteq \mathcal{C} \quad (55)$$

It's easy to check \mathcal{B}_0 is in fact a σ -algebra : $f^{-1}(\emptyset) = \emptyset \in \mathcal{A}$, $f^{-1}(B^c) = (f^{-1}(B))^c$, and $f^{-1}(\bigcup B_i) = \bigcup f^{-1}(B_i)$. Therefore, $\mathcal{B} \subseteq \mathcal{B}_0$ and all $B \in \mathcal{B}$ satisfies $f^{-1}(B) \in \mathcal{A}$. ■

Example 2.1. $f(x) = \mathbb{1}\{x \in \mathbb{Q}\}$ is measurable.

2.2 Simple Functions

Definition 2.3. A function $f : (X, \mathcal{A}) \rightarrow (\mathbb{R}^*, \mathcal{B}(\mathbb{R}^*))$ is called **simple** if there exists finitely many disjoint sets A_1, \dots, A_n and real numbers a_1, \dots, a_n such that

$$f(x) = \begin{cases} a_i & \text{if } x \in A_i \\ 0 & \text{if } x \notin A_i \forall i \in [n] \end{cases} \quad (56)$$

Let \mathcal{S} denote the set of simple functions, and \mathcal{S}^+ denote the set of non-negative simple functions.

Proposition 2.3. All simple functions are measurable.

Proof. For any subset of \mathbb{R}^* , the pre-image is either X or a union of some (potentially none) A_i 's. ■

2.3 Properties of Measurable Functions

Example 2.2. Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$, then all of the following functions are measurable:

$$f(x, y) = x + y \quad (57)$$

$$f(x, y) = \max\{x, y\} \equiv x \vee y \quad (58)$$

$$f(x, y) = \min\{x, y\} \equiv x \wedge y \quad (59)$$

$$f(x, y) = x - y \quad (60)$$

$$f(x, y) = \alpha x \quad \alpha \in \mathbb{R} \quad (61)$$

Proposition 2.4 (Component-wise Measurable Functions). Let $f, g : (X, \mathcal{A}) \rightarrow (\mathbb{R}^*, \mathcal{B}(\mathbb{R}^*))$ be measurable, let $h(x) = (f(x), g(x)) \in \mathbb{R}^{*2}$, then f is measurable.

Proof.

$$h^{-1}([-\infty, t] \times [-\infty, s]) = f^{-1}([-\infty, t]) \cap g^{-1}([-\infty, s]) \in \mathcal{A} \quad (62)$$

And, $\mathcal{B}(\mathbb{R}^*)$ can be generated by sets with forms $[-\infty, t] \times [-\infty, s]$. ■

Proposition 2.5 (Composite of Measurable Functions). Let $(X, \mathcal{A}), (Y, \mathcal{B}), (Z, \mathcal{C})$ be measurable spaces, let $f : X \rightarrow Y$ and $g : Y \rightarrow Z$ be measurable functions. Then, the composite $g \circ f : X \rightarrow Z$ is measurable.

Corollary 2.1. Let $f, g : X \rightarrow \mathbb{R}$ be measurable functions, then $f + g, f - g, \max\{f, g\}$, and $\min\{f, g\}$ are all measurable.

Proof. $f + g$ and $f - g$ can be written as the composition of $h_1(x) = (f(x), g(x))$ and $h_2(x, y) = x \pm y$, which are all measurable.

$f \vee g$ and $f \wedge g$ are measurable as special cases of next proposition. ■

Proposition 2.6. Let f_1, f_2, \dots be a sequence of measurable maps from $(X, \mathcal{A}) \rightarrow (\mathbb{R}^*, \mathcal{B}(\mathbb{R}^*))$, then $\sup_n f_n$ and $\inf_n f_n$ are measurable.

Proof. Note $\{x \in X : \sup_n f_n \leq t\} = \bigcup_{n=1}^{\infty} \{x \in X : f_n \leq t\} \in \mathcal{A}$ for every t , therefore the supremum is measurable. ■

Corollary 2.2. $\limsup f_n$ and $\liminf f_n$ are measurable.

Proof. Let $g_k = \sup_{n \geq k} f_n$, g_k is measurable. $\limsup f_n = \inf_k g_k$ is measurable as well. Similar proof for the measurability of $\liminf f_n$. ■

Proposition 2.7. Let f and g be \mathbb{R}^* -valued measurable functions. Then sets

$$\{x \in A : f(x) < g(x)\}, \{x \in A : f(x) \leq g(x)\} \quad (63)$$

are measurable.

Proof.

$$\{x \in A : f(x) < g(x)\} = \bigcup_{r \in \mathbb{Q}} (\{x \in A : f(x) < r\} \cap \{x \in A : r < g(x)\}) \quad (64)$$

■

Corollary 2.3. Let $u, v : X \rightarrow \mathbb{R}^*$ be measurable functions, then $\{x \in X : u(x) = v(x)\}$ is measurable.

Proof. Note that $\{x \in X : u(x) = v(x)\} = \{x \in X : u(x) \leq v(x)\} \cap \{x \in X : u(x) \geq v(x)\}$. ■

Corollary 2.4. Let $\{f_n\}$ be a sequence of measurable functions from $(X, \mathcal{A}) \rightarrow (\mathbb{R}^*, \mathcal{B}(\mathbb{R}^*))$. Then,

$$\{x \in X : \lim_{n \rightarrow \infty} f_n(x) \text{ exists}\} \quad (65)$$

is measurable.

Proof. Note that $\{x \in X : \lim_{n \rightarrow \infty} f_n(x) \text{ exists}\} = \{x \in X : \liminf f_n(x) = \limsup f_n(x)\}$, the result follows from previous lemma. ■

Corollary 2.5. If $\{f_n\}$ is a sequence of measurable functions such that $\lim f_n(x)$ exists for all $x \in X$, then $\lim f_n$ is a measurable function on (X, \mathcal{A}) .

Proof. In this case, $\{x \in X : \lim_{n \rightarrow \infty} f_n(x) \text{ exists}\} = X$, and $\lim f_n = \liminf f_n$ on X . ■

Corollary 2.6. If $\{f_n\}$ is a sequence of measurable function from X to $[0, \infty]$, then $\sum_{n=1}^{\infty} f_n$ is measurable.

Proof. Follows the previous corollary directly: define $g_k = \sum_{n=1}^k f_n$ and $\lim_{k \rightarrow \infty} g_k = \sum_{n=1}^{\infty} f_n$. ■

3 Integrals

3.1 Integrating Simple Functions

Definition 3.1. Let $f \in \mathbb{S}^+$ with representation $\{(A_i, a_i)\}_{i=1}^n$. WLOG, $\bigcup_{i=1}^n A_i = X$. Then, define

$$\int_X f \, d\mu := \sum_{i=1}^n a_i \mu(A_i) \quad (66)$$

Proposition 3.1. The notion of integral on simple functions is well defined. Specifically, let $\{(A_i, a_i)\}_{i=1}^n$ and $\{(B_j, b_j)\}_{j=1}^m$ be any two representations of f , $\sum_{i=1}^n a_i \mu(A_i) = \sum_{j=1}^m b_j \mu(B_j)$.

Proof. First note that $\{A_i \cap B_j\}_{i,j}$ are disjoint sets with union X . Moreover, for any i, j , if $A_i \cap B_j \neq \emptyset$, take some $x \in A_i \cap B_j$, $f(x) = a_i = b_j$. Therefore, $a_i \mu(A_i \cap B_j) = b_j \mu(A_i \cap B_j)$ since either $a_i = b_j$ or $\mu(A_i \cap B_j) = \mu(\emptyset) = 0$.

$$\sum_{i=1}^n a_i \mu(A_i) = \sum_{i=1}^n a_i \sum_{j=1}^m \mu(A_i \cap B_j) \quad (67)$$

$$= \sum_{j=1}^m b_j \sum_{i=1}^n \mu(A_i \cap B_j) \quad (68)$$

$$= \sum_{j=1}^m b_j \mu(B_j) \quad (69)$$

■

3.2 Integrating Measurable Functions

Definition 3.2. For a non-negative measurable function $f : X \rightarrow [0, \infty]$, define its Lebesgue integral as

$$\int f \, d\mu = \sup \left\{ \int g \, d\mu : g \text{ is a non-negative simple function such that } g \leq f \right\} \quad (70)$$

For any measurable $f : X \rightarrow [-\infty, \infty]$, let

$$f^+(x) = \max\{f(x), 0\} \quad (71)$$

$$f^-(x) = -\min\{f(x), 0\} \quad (72)$$

So that $f = f^+ - f^-$, and f is measurable if and only if both f^+ and f^- are measurable.

If at least one of $\int f^+ \, d\mu$, $\int f^- \, d\mu$ is finite, the integral of f exists (well-defined) and is defined as

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu \quad (73)$$

If both $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ are finite, f is said to be **integrable**.

3.3 Properties of Integral of Non-negative Simple Functions

Proposition 3.2 (Linearity). If f, g are non-negative simple functions, then

$$\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu \quad (74)$$

Moreover, for any $\alpha \geq 0$,

$$\int \alpha f \, d\mu = \alpha \int f \, d\mu \quad (75)$$

Proof. Let f and g be simple functions represented by $\{(A_i, a_i)\}_{i=1}^n$ and $\{(B_j, b_j)\}_{j=1}^m$. WLOG, $\cup A_i = \cup B_j = X$. Then $f + g$ is a simple function with representation $\{(A_i \cap B_j, a_i + b_j)\}_{i,j}$, where $\cup_{i,j} A_i \cap B_j = X$. ■

Proposition 3.3. Let f, g be non-negative simple functions with $f \geq g$ everywhere. Then $\int f d\mu \geq \int g d\mu$.

Proof. Let f and g be simple functions represented by $\{(A_i, a_i)\}_{i=1}^n$ and $\{(B_j, b_j)\}_{j=1}^m$.

Claim: $a_i \mu(A_i \cap B_j) \geq b_j \mu(A_i \cap B_j)$ for every (i, j) . If $A_i \cap B_j \neq \emptyset$, then taking some $x \in A_i \cap B_j$ implies $a_i \geq b_j$. If $A_i \cap B_j = \emptyset$, the equality holds trivially.

Note that $\int f$ and $\int g$ can be written as $\sum_{i,j} a_i \mu(A_i \cap B_j)$ and $\sum_{i,j} b_j \mu(A_i \cap B_j)$ respectively, therefore $\int f \geq \int g$ by the previous claim. ■

Proposition 3.4 (Approximation using Simple Functions). Let $f : X \rightarrow [0, \infty]$ be a measurable function. Then there exists an increasing sequence of non-negative simple functions f_n such that $f_n \leq f_{n+1}$ and

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad (76)$$

for all x .

Proof. For each n and $1 \leq k \leq n2^n$, let

$$A_{n,k} = \left\{ x \in X : \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \right\} \quad (77)$$

Define

$$f_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{if } x \in A_{n,k} \\ n & \text{otherwise} \end{cases} \quad (78)$$

That is, for a $x \in X$, if $\frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}$ for some k , we take $f_n(x) = \frac{k-1}{2^n}$; if $f(x) \geq n$, we define $f_n(x) = n$. Clearly, each f_n is a simple function.

Claim 1: $f_n \leq f_{n+1}$. Easy to verify.

Claim 2: $\lim_{n \rightarrow \infty} f_n(x) = f(x)$. Let $x \in X$, (i) if $f(x) = \infty$, then $f_n(x) = n$ for all $n \in \mathbb{N}$ and $\lim_{n \rightarrow \infty} f_n(x) = \infty = f(x)$.

(ii) if $f(x) < \infty$, then $\exists n_0$ such that $f(x) < n_0$. For every $n \geq n_0$, $x \in A_{n,k}$ for some k such that $f_n(x) = \frac{k-1}{2^n}$ and $\frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}$. Therefore, for all $n \geq n_0$, $|f_n(x) - f(x)| < \frac{1}{2^n}$, which implies $\lim_{n \rightarrow \infty} f_n(x) = f(x)$. ■

Proposition 3.5 (Monotone Convergence 1: $\mathbb{S}_+ \uparrow \mathbb{S}_+$). Let f_n be a sequence of non-negative simple functions that increase to another non-negative simple function f at each point, then

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu \quad (79)$$

Proof. By monotonicity, $f_n \leq f$ for all n and $\int f \, d\mu \geq \lim \int f_n \, d\mu$.

Fix $0 < \varepsilon < 1$ and define $g = (1 - \varepsilon)f$. Suppose f is represented by (A_i, a_i) . Then for every n, i , define

$$A_{n,i} = \{x \in A_i : f_n(x) \geq (1 - \varepsilon)a_i\} \quad (80)$$

Define

$$g_n(x) = \begin{cases} (1 - \varepsilon)a_i & \text{if } x \in A_{n,i} \\ 0 & \text{otherwise} \end{cases} \quad (81)$$

In order to show $\int f \, d\mu \leq \lim \int f_n \, d\mu$, we are constructing this g_n satisfying

$$(1 - \varepsilon) \int f \, d\mu \leq \lim \int g_n \, d\mu \leq \lim \int f_n \, d\mu \leq \int f \, d\mu \quad (82)$$

where the last equality has been shown above. The equality can then be shown by taking $\varepsilon \rightarrow 0$ and using Squeeze theorem. Note that $(1 - \varepsilon) \int f \, d\mu \not\leq \int g_n \, d\mu$, only the limit does.

By construction, $g_n \leq f_n$ and $\int g_n \, d\mu \leq \int f_n \, d\mu$ as a result.

$$\lim_n \int f_n \, d\mu \geq \lim_n \int g_n \, d\mu \quad (83)$$

$$= \lim_n \sum_{i=1}^K (1 - \varepsilon)a_i \mu(A_{n,i}) \quad (84)$$

$$= \sum_{i=1}^K (1 - \varepsilon)a_i \lim_n \mu(A_{n,i}) \quad (85)$$

$$= \sum_{i=1}^K (1 - \varepsilon)a_i \mu(A_i) \text{ Since for all } i, A_{n,i} \uparrow A_i \text{ as } n \rightarrow \infty. \quad (86)$$

$$= (1 - \varepsilon) \int f \, d\mu \quad (87)$$

Taking $\varepsilon \rightarrow 0$ completes the proof. ■

Proposition 3.6 (Monotone Convergence 2: $\mathbb{S}_+ \uparrow$ Measurable). Let $f : X \rightarrow [0, \infty]$ be a measurable function. Let f_n be a sequence of non-negative simple functions such that $f_n \uparrow f$ point-wise. Then

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int f_n \, d\mu \quad (88)$$

Proof. The proof follows the previous proposition and the definition of $\int f \, d\mu$. Since $f_n \uparrow f$, $f_n \leq f$ and $\int f_n \leq \int f$ for all n . $\int f_n$ is a bounded monotone sequence, therefore $\lim \int f_n$ exists and $\leq \int f$.

To show the other equality, it suffices to prove $\lim \int f_n \geq \int g$ for any non-negative simple functions $g \leq f$.

Define $g_n = \min\{g, f_n\}$, easy to show that $g_n(x) \leq g_{n+1}(x)$.

$$\lim_{n \rightarrow \infty} g_n(x) = \lim_{n \rightarrow \infty} \min\{g, f_n\} \quad (89)$$

$$= \min\{g(x), f(x)\} \quad (90)$$

$$= g(x) \quad (91)$$

since $f_n \uparrow f$ and $g \leq f$.

By the previous proposition, $\int g \, d\mu = \lim \int g_n \, d\mu$ since g_n and g are non-negative simple functions. Since $g_n \leq f_n$ everywhere, so $\int g_n \, d\mu \leq \int f_n \, d\mu$. Taking limit on both sides implies $\int g \leq \lim \int f_n$. ■

Proposition 3.7 (Vector Space Properties for Non-negative Integrable Functions). Let $f, g : X \in [0, \infty]$ be integrable (of course, measurable as well) functions and $\alpha \geq 0$. Then

$$1. \int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu.$$

$$2. \int \alpha f \, d\mu = \alpha \int f \, d\mu.$$

$$3. \text{ If } f \geq g \text{ everywhere, then } \int f \, d\mu \geq \int g \, d\mu.$$

Proof. We know that there exists sequences of non-negative simple functions f_n and g_n such that $f_n \uparrow f$ and $g_n \uparrow g$. Note that $f_n + g_n$ is a sequence of simple functions increases to $f + g$. Therefore,

$$\int (f + g) d\mu = \lim_{n \rightarrow \infty} \int (f_n + g_n) \, d\mu \quad (92)$$

$$= \lim_{n \rightarrow \infty} \left(\int f_n \, d\mu + \int g_n \, d\mu \right) \quad (93)$$

$$= \lim_{n \rightarrow \infty} \int f_n \, d\mu + \lim_{n \rightarrow \infty} \int g_n \, d\mu \quad (94)$$

$$= \int f \, d\mu + \int g \, d\mu \quad (95)$$

Similarly, taking $\alpha f_n \uparrow \alpha f$ leads to the second result.

Finally, if $f \geq g$ everywhere, then

$$\{h \in \mathbb{S}_+ \text{ and } h \leq g\} \subseteq \{h \in \mathbb{S}_+ \text{ and } h \leq f\} \quad (96)$$

Therefore, the supremum of integrals of functions from a larger collection is larger. ■

3.4 Linearity of Lebesgue Integral for Arbitrary Integrable Functions

Theorem 3.1 (Vector Space Property of Integral Functions). Let (X, \mathcal{A}, μ) be a measure space, let $f, g : X \rightarrow \mathbb{R}^*$ be integrable functions, let $\alpha \in \mathbb{R}$. Then, $f + g$ and αf are integrable, and

$$\int f + g d\mu = \int f d\mu + \int g d\mu \quad (97)$$

$$\int \alpha f d\mu = \alpha \int f d\mu \quad (98)$$

Proof. It's easy to check that $(f + g)^+ \leq f^+ + g^+$ and $(f + g)^- \leq f^- + g^-$. By monotonicity, $\int (f + g)^+ d\mu, \int (f + g)^- d\mu < \infty$. Therefore, $f + g$ is integrable.

Moreover, $f + g = f^+ - f^- + g^+ - g^- \iff f + g + f^- + g^- = f^+ + g^+$. We can apply the linearity of non-negative integrable functions to derive the result.

When $\alpha \geq 0$, $(\alpha f)^+ = \alpha f^+$ and $(\alpha f)^- = \alpha f^-$. The proof for cases with $\alpha < 0$ is similar. ■

Corollary 3.1. Let f, g be integrable functions such that $f \geq g$, then $\int f d\mu \geq \int g d\mu$.

Proof. Let $h = f - g = f + (-1)g \geq 0$, which is integrable by the previous theorem. And $\int h d\mu \geq 0$ since it's the supremum of integrals for simple functions less than h , which includes the zero function (has zero integral). ■

Lemma 3.1. A function f is integrable if and only if $|f|$ is integrable.

Proof. Note that $|f| = f^+ + f^-$, and $\int f^+ + f^- d\mu < \infty$ by the integrability of f . Therefore, $|f|$ is integrable.

Moreover, $|f|^+ = f^+ + f^-$, therefore, the integrability of $|f|$ implies both $\int f^+ d\mu$ and $\int f^- d\mu$ are finite. ■

Proposition 3.8. All integrable function f satisfies the triangle inequality

$$\left| \int f d\mu \right| \leq \int |f| d\mu \quad (99)$$

Proof.

$$\left| \int f d\mu \right| = \left| \int f^+ - f^- d\mu \right| \quad (100)$$

$$= \left| \int f^+ d\mu - \int f^- d\mu \right| \quad (101)$$

$$\leq \left| \int f^+ d\mu \right| + \left| \int f^- d\mu \right| \quad (102)$$

$$= \int f^+ d\mu + \int f^- d\mu \quad (103)$$

$$= \int |f| d\mu \quad (104)$$

■

4 Limit Theorems (i.e., when we can exchange limits and integrals)

Theorem 4.1 (Monotone Convergence Theorem). Let (X, \mathcal{A}, μ) be a measure space, let $f_n : X \rightarrow [0, \infty]$ be a non-decreasing sequence of measurable functions converge to f . Then,

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int f_n \, d\mu \quad (105)$$

Proof. f measurable since $f = \lim_n f_n = \liminf_n f_n$. Moreover, $\int f_n \, d\mu$ is a non-decreasing sequence to the limit $\int f \, d\mu$, therefore $\int f \, d\mu \geq \lim_n \int f_n \, d\mu$.

For each $n \in \mathbb{N}$, there exists a non-decreasing sequence of non-negative simple functions $g_{n,k}$ converges to f_n . Define

$$h_n = \max\{g_{1,n}, g_{2,n}, \dots, g_{n,n}\} \quad (106)$$

Note that h_n is a non-decreasing sequence since

$$h_{n+1} = \max\{g_{1,n+1}, g_{2,n+1}, \dots, g_{n+1,n+1}\} \quad (107)$$

$$\geq \max\{g_{1,n+1}, g_{2,n+1}, \dots, g_{n,n+1}\} \quad (108)$$

$$\geq \max\{g_{1,n}, g_{2,n}, \dots, g_{n,n}\} = h_n \quad (109)$$

Moreover, for any $m \in \mathbb{N}$, for any $n \geq m$,

$$h_n(x) = \max\{g_{1,n}, g_{2,n}, \dots, g_{n,n}\} \geq g_{m,n} \quad (110)$$

Therefore, by taking the limit $n \rightarrow \infty$,

$$\lim_{n \rightarrow \infty} h_n(x) \geq \lim_{n \rightarrow \infty} g_{m,n} = f_m \quad (111)$$

Taking limit $m \rightarrow \infty$ on both sides

$$\lim_n h_n(x) = \lim_m \lim_n h_n(x) \geq \lim_m f_m = f \quad (112)$$

$$\implies \int \lim_n h_n(x) \, d\mu \geq \int f \, d\mu \quad (113)$$

Note that, by construction,

$$h_n = \max\{g_{1,n}, g_{2,n}, \dots, g_{n,n}\} \leq \max\{f_1, \dots, f_n\} = f_n \quad (114)$$

Therefore,

$$\int \lim_{n \rightarrow \infty} f_n(x) \, d\mu \geq \int f \, d\mu \quad (115)$$

■

Corollary 4.1. Let (f_n) be a sequence (not necessarily increasing) non-negative measurable functions, then

$$\int \sum_{n=1}^{\infty} f_n \, d\mu = \sum_{n=1}^{\infty} \int f_n \, d\mu \quad (116)$$

Theorem 4.2 (Fatou's Lemma). Let f_n be a sequence of non-negative measurable functions, then

$$\int \liminf_{n \rightarrow \infty} f_n \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n \, d\mu \quad (117)$$

Proof. Define $g_n = \inf_{k \geq n} f_k$, then g_n is an increasing sequence of non-negative functions. By construction, $\int g_n \, d\mu \leq \inf_{k \geq n} \int f_k \, d\mu$. By MCT,

$$\int \liminf_{n \rightarrow \infty} f \, d\mu = \int \lim_{n \rightarrow \infty} g_n \, d\mu \quad (118)$$

$$= \lim_{n \rightarrow \infty} \int g_n \, d\mu \quad (119)$$

$$\leq \lim_{n \rightarrow \infty} \inf_{k \geq n} \int f_k \, d\mu \quad (120)$$

$$= \liminf_{n \rightarrow \infty} \int f_n \, d\mu \quad (121)$$

■

Theorem 4.3 (Lebesgue's Dominated Convergence Theorem). Let (X, \mathcal{A}, μ) be a measure space, let f and f_n be \mathbb{R}^* -valued measurable functions on X such that $f_n \rightarrow f$ point-wise. If there exists a non-negative integrable function g such that $|f_n| \leq g$ for all n , then, all f and f_n are integrable, moreover,

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int f_n \, d\mu \quad (122)$$

Proof. Since $|f_n| \leq g$, all f_n are integrable. The limit f also satisfies $|f| \leq g$ and is integrable.

For now, assume f_n are \mathbb{R} -valued instead of \mathbb{R}^* -valued.

Note that $f + g = \lim_{n \rightarrow \infty} f_n + g$ is non-negative (because of the dominance) and integrable, by Fatou's lemma

$$\int f + g \, d\mu = \int \liminf_{n \rightarrow \infty} f + g \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n + g \, d\mu \quad (123)$$

$$= \liminf_{n \rightarrow \infty} \int f_n \, d\mu + \int g \, d\mu \quad (124)$$

$$\implies \int f \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n \, d\mu \quad (125)$$

Similarly, $g - f = \lim_{n \rightarrow \infty} g - f_n$ is non-negative and integrable as well, by Fatou's lemma

$$\int g - f \, d\mu = \int \liminf g - f_n \, d\mu \leq \liminf \int g - f_n \, d\mu \quad (126)$$

$$\implies - \int f \, d\mu \leq - \liminf \int f_n \, d\mu \quad (127)$$

$$\implies \int f \, d\mu \geq \limsup \int f_n \, d\mu \quad (128)$$

Also, $\liminf \int f_n \, d\mu \leq \limsup \int f_n \, d\mu$, therefore,

$$\liminf \int f_n \, d\mu \geq \int f \, d\mu \geq \limsup \int f_n \, d\mu \geq \liminf \int f_n \, d\mu \quad (129)$$

$$\implies \int f \, d\mu = \lim \int f_n \, d\mu \quad (130)$$

■

Proposition 4.1 (A Stronger Result). Given assumptions of the dominated convergence theorem, f_n L^1 -converges to f .

$$\lim_{n \rightarrow \infty} \int |f_n - f| \, d\mu = 0 \quad (131)$$

Proof. Note that $|f_n - f| \rightarrow 0$ point-wise, and $|f_n - f| \leq 2g$. The dominated convergence theorem suggests $\lim_{n \rightarrow \infty} \int |f_n - f| \, d\mu = \int 0 \, d\mu = 0$. ■

4.1 The Notion of Almost Everywhere

Definition 4.1. Let (X, \mathcal{A}, μ) be a measure space, a set $N \subseteq X$ (not necessarily measurable) is called "negligible w.r.t. μ " if $N \subseteq A$ for some $A \in \mathcal{A}$ with $\mu(A) = 0$.

A property is said to hold **almost everywhere** w.r.t. μ (denoted as μ -a.e.) if the set on which this property fails is negligible.

Proposition 4.2. Let $f : X \rightarrow [0, \infty]$ be an integrable function, then f is finite μ -a.e.

Proof. Let $A := f^{-1}(\infty)$, define $h_n(x) := n \mathbb{1}\{x \in A\}$. Clearly, h_n is a simple function $\leq f$ for every n , by monotonicity, $\int f \, d\mu \leq \int h_n \, d\mu = n\mu(A)$. Taking $n \rightarrow \infty$ leads to a contradiction. ■

Corollary 4.2. If $f : X \rightarrow \mathbb{R}^*$ is integrable w.r.t. μ , then $|f| < \infty$ μ -a.e.

Proof. f is integrable implies $\int f^+ \, d\mu, \int f^- \, d\mu < \infty$. Then, by the previous proposition, $f^+ < \infty$ except for a negligible set A , and $f^- < \infty$ except for a negligible set B . Therefore, $|f| = \infty$ on set $A \cup B$, which is negligible as well. ■

Proposition 4.3. Let $f : X \rightarrow [0, \infty]$ be measurable, then

$$\int f \, d\mu = 0 \iff f = 0 \, \mu - a.e. \quad (132)$$

Proof. (\implies) Suppose $f = 0$ a.e., for every simple function $g \leq f$, let (a_i, A_i) be the representation of g , then $\int g \, d\mu = 0$ by definition. Suppose $a_i > 0$ for some A_i , then $f(x) \geq a_i > 0$ for all $x \in A_i$, since $f = 0$ a.e., $\mu(A_i) = 0$. Therefore, $\int g \, d\mu = 0$, so is the integral of f .

(\impliedby), suppose $\int f \, d\mu = 0$, note that

$$\{x : f(x) > 0\} = \bigcup_{n=1}^{\infty} \{x : f(x) > 1/n\} \quad (133)$$

Define $A_n = \{x : f(x) > 1/n\}$, then A_n is an increasing sequence of sets, therefore, suppose there exists some A_n with $\mu(A_n) > 0$, define $g(x) = \frac{1}{n} \mathbb{1}\{x \in A_n\}$. $f \geq g$ by construction, so that $\int f \, d\mu \geq \int g \, d\mu = \frac{1}{n} \mu(A_n) > 0$. This leads to a contradiction, so all $\mu(A_n) = 0$, and $\mu(\{x : f(x) > 0\}) = \lim_n \mu(A_n) = 0$. \blacksquare

Corollary 4.3. Let $f : X \rightarrow \mathbb{R}^*$ be a measurable function,

$$f = 0 \, \mu - a.e. \implies \int f \, d\mu = 0 \quad (134)$$

Proof. $f = 0$ a.e. implies $f^+, f^- = 0$ a.e., apply the previous proposition, $\int f^+ \, d\mu = \int f^- \, d\mu = 0$, so is $\int f \, d\mu$.

Note the converse is not true, it is possible that $\int f^+ \, d\mu = \int f^- \, d\mu \neq 0$ such that $\int f \, d\mu = 0$. \blacksquare

Corollary 4.4. Let $f, g : X \rightarrow \mathbb{R}^*$ be integrable functions, then

$$f = 0 \, \mu - a.e. \implies \int f \, d\mu = \int g \, d\mu \quad (135)$$

Proof. Let $\tilde{f} = f(x) \mathbb{1}\{x \in \mathbb{R}\}$ and $\tilde{g} = g(x) \mathbb{1}\{x \in \mathbb{R}\}$, we are doing this to avoid subtracting infinity from infinity. $|\tilde{f}|$ and $|\tilde{g}|$ are bounded by $|f|$ and $|g|$ and are integrable. Moreover, $f = \tilde{f} = g = \tilde{g}$ a.e. by construction. Lastly, since $|\tilde{f}|, |\tilde{g}| < \infty$, we can write

$$\int \tilde{f} - \tilde{g} \, d\mu = \int \tilde{f} \, d\mu - \int \tilde{g} \, d\mu = 0 \quad (136)$$

$$\implies \int f \, d\mu = \int \tilde{f} \, d\mu = \int g \, d\mu = \int \tilde{g} \, d\mu \quad (137)$$

\blacksquare

Proposition 4.4. Monotone convergence theorem and dominated convergence theorem holds even if $f_n \rightarrow f$ a.e. In DCT, we can also have $|f_n| \leq g$ a.e.

Proof for MCT. Suppose $f_n \geq 0$ a.e.

$$A = \{x : f_n(x) \geq 0 \, \forall n \wedge \lim_{n \rightarrow \infty} f_n(x) = f(x)\} \quad (138)$$

Therefore, $A^c = \bigcup_n \{x : f_n(x) < 0\} \cup \{x : \lim_{n \rightarrow \infty} f_n(x) \neq f(x)\}$, which is a countable union of measure zero sets, hence $\mu(A^c) = 0$.

Define $\tilde{f}_n = \mathbb{1}_A f_n$ and $\tilde{f} = \mathbb{1}_A f$, apply the original version of MCT on \tilde{f}_n and \tilde{f} , then exert the fact that $\int \tilde{f}_n d\mu = \int f_n d\mu$ and $\int \tilde{f} d\mu = \int f d\mu$. ■

5 Integral of Complex-Valued Functions

Definition 5.1. A function $f : X \rightarrow \mathbb{C}$ is called **measurable** if both $\Re(f)$ and $\Im(f)$ (both are real-valued functions) are measurable. Similarly, f is **integrable** if both its real and imaginary parts are integrable.

Define

$$\int f d\mu = \int \Re(f) d\mu + i \int \Im(f) d\mu \quad (139)$$

Proposition 5.1. Let f, g be integrable complex-valued functions, then

1. $\int (f + g) d\mu = \int f d\mu + \int g d\mu$.
2. for all $\alpha \in \mathbb{C}$, $\int (\alpha f) d\mu = \alpha \int f d\mu$.

Proposition 5.2 (Triangle Inequality). Let $f : X \rightarrow \mathbb{C}$ be an integrable function, then

$$\left| \int f d\mu \right| \leq \int |f| d\mu \quad (140)$$

Proof. Note that there exists $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ such that

$$\left| \int f d\mu \right| = \alpha \int f d\mu \quad (141)$$

To see this point, let $z = re^{i\theta} \in \mathbb{C}$ so that $|z| = r$, let $\alpha = e^{-i\theta}$, which satisfies $\alpha z = r = |z|$.

Therefore,

$$\left| \int f d\mu \right| = \alpha \int f d\mu \quad (142)$$

$$= \int (\alpha f) d\mu \quad (143)$$

$$= \int \Re(\alpha f) d\mu + i \int \Im(\alpha f) d\mu \quad (144)$$

$$\implies \int \Im(\alpha f) d\mu = 0 \quad (145)$$

Therefore,

$$\left| \int f d\mu \right| = \int \Re(\alpha f) d\mu \leq \int |\alpha f| d\mu = \int |f| d\mu \quad (146)$$

where the last step holds because $|\alpha| = 1$. ■

6 Convergence of Measurable Functions

Definition 6.1. Let (X, \mathcal{A}, μ) be a measure space, let $\{f_n\}_n$ be a sequence of real-valued measurable functions on X , let $f : X \rightarrow \mathbb{R}$ be a measurable function.

Then $f_n \rightarrow f$ **in measure** if for every $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) = 0 \quad (147)$$

Remark 6.1. Convergence almost everywhere does not imply convergence in measure.

Counter-example. ■

Remark 6.2. Convergence in measure does not imply convergence almost everywhere (even if we are considering a finite measure).

Counter-example. ■

Proposition 6.1. Let μ be a finite measure, then convergence a.e. implies convergence in measure.

Proof. Suppose $f \rightarrow f_n$ a.e. Let $\varepsilon > 0$. Note that if there exists x such that $|f_n - f(x)| \geq \varepsilon$ for infinitely many n , then $f_n \not\rightarrow f$ at x . Therefore,

$$\mu(\{x : |f_n(x) - f(x)| > \varepsilon \text{ for infinitely many } n\}) \leq \mu(\{x : f_n(x) \not\rightarrow f(x)\}) = 0 \quad (148)$$

Further, note that

$$\{x : |f_n(x) - f(x)| > \varepsilon \text{ for infinitely many } n\} = \bigcap_{n=1}^{\infty} \underbrace{\bigcup_{k=n}^{\infty} \{x : |f_k(x) - f(x)| > \varepsilon\}}_{B_n} \quad (149)$$

Where $x \in B_n$ indicates there exists a $k \geq n$ such that $|f_k(x) - f(x)| > \varepsilon$. If we take the intersection of all B_n , it means for all $n \in \mathbb{N}$, there exists $k \geq n$ such that $|f_k(x) - f(x)| > \varepsilon$, which is equivalent to saying there are infinitely many k such that $|f_k(x) - f(x)| > \varepsilon$.

Clearly $B_1 \supseteq B_2 \supseteq \dots$, there must exist some B_i such that $\mu(B_i) < \infty$ since μ is a finite measure.

Therefore,

$$0 = \mu(\{x : |f_n(x) - f(x)| > \varepsilon \text{ for infinitely many } n\}) = \lim_{n \rightarrow \infty} \mu(B_n) \quad (150)$$

Hence, $\lim_{n \rightarrow \infty} \mu(B_n) = 0$. However, $B_n \supseteq \{x : |f_n(x) - f(x)| > \varepsilon\}$, therefore,

$$\lim_{n \rightarrow \infty} \mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) = 0 \quad (151)$$

■

Proposition 6.2. Let f_n be a sequence of measurable real-valued functions converge to a measurable f in measure, then there exists a subsequence of f_n converges to f a.e.

Proof. Suppose $f_n \rightarrow f$ in measure, take $\varepsilon = 1$, there exists infinitely many n_1 such that

$$\mu(\{x : |f_{n_1} - f(x)| > 1\}) < 2^{-1} \quad (152)$$

Then for every k , we can choose $n_k > n_{k-1}$ such that

$$\underbrace{\mu(\{x : |f_{n_k} - f(x)| > \frac{1}{k}\})}_{A_k} < 2^{-k} \quad (153)$$

Let $B = \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k$, define $B_j = \bigcup_{k=j}^{\infty} A_k$. Note that for all j , $B \subseteq B_j$, therefore,

$$\mu(B) \leq \mu(B_j) = \mu\left(\bigcup_{k=j}^{\infty} A_k\right) \leq \sum_{k=j}^{\infty} \mu(A_k) < \sum_{k=j}^{\infty} 2^{-j+1} \quad (154)$$

Take $j \rightarrow \infty$, $\mu(B) = 0$. If $x \notin B$, $x \in B^c = \bigcup_{j=1}^{\infty} \bigcap_{k=j}^{\infty} A_k^c$, which means $\exists j$ such that $x \in A_k^c$ for all $k \geq j$. That is

$$\exists j \text{ s.t. } \forall k \geq j \quad |f_{n_k} - f(x)| \leq \frac{1}{k} \quad (155)$$

Therefore, this subsequence n_k converges to $f(x)$ a.e. ■

Lemma 6.1 (Borel-Cantelli Lemma). If A_1, A_2, \dots , is a sequence of measurable sets such that

$$\sum_{k=1}^{\infty} \mu(A_k) < \infty \quad (156)$$

then

$$\mu(\{x : x \in \text{infinitely many } A_k\}) = 0 \quad (157)$$

Proof. Define

$$B = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \quad (158)$$

Easy to verify that $x \in B$ if and only if $x \in$ infinitely many A_k . For every j ,

$$B \subseteq \bigcup_{k=j}^{\infty} A_k \quad (159)$$

Hence

$$\mu(B) \leq \mu\left(\bigcup_{k=j}^{\infty} A_k\right) \leq \sum_{k=j}^{\infty} \mu(A_k) \rightarrow 0 \text{ as } j \rightarrow \infty \quad (160)$$

Therefore, $\mu(B) = 0$. ■

Theorem 6.1 (Egorov's Theorem). Let (X, \mathcal{A}, μ) be a measure space with $\mu(X) < \infty$. Let f_n be a sequence of measurable \mathbb{R} -valued functions converging a.e. to a \mathbb{R} -valued function f .

Then for all $\varepsilon > 0$, \exists a set $B \in \mathcal{A}$ such that

1. $\mu(B^c) < \varepsilon$,
2. and $f_n \rightarrow f$ uniformly on B .

Proof. Let $\varepsilon > 0$.

For all $n \in \mathbb{N}$, define

$$g_n(x) := \sup_{k \geq n} |f_k(x) - f(x)| \quad (161)$$

since $f_n \rightarrow f$ a.e., $g_n(x)$ is finite a.e. Moreover, $g_n(x) \rightarrow 0$ a.e. as $n \rightarrow \infty$ (both holds where $f_n \rightarrow f$).

Since $\mu(X) < \infty$, $g_n(x) \rightarrow 0$ in measure by previous results. Then, for every $k \in \mathbb{N}$, there exists n_k such that

$$\mu \left(\left\{ x : g_{n_k}(x) > \frac{1}{k} \right\} \right) < \frac{\varepsilon}{2^k} \quad (162)$$

Since there are infinitely many n_k to choose, we may choose an increasing sequence of n_k 's. Define

$$B^c = \left\{ x : g_{n_k}(x) > \frac{1}{k} \text{ for some } k \right\} \quad (163)$$

Then,

$$\mu(B^c) = \mu \left(\bigcup_{k=1}^{\infty} \left\{ x : g_{n_k}(x) > \frac{1}{k} \right\} \right) \quad (164)$$

$$\leq \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon \quad (165)$$

Lastly, we can show that $f_n \rightarrow f$ uniformly on B . Note that for every $\delta > 0$, take $k_\delta \geq \frac{1}{\delta}$, if $x \in B$, then $g_{n_{k_\delta}}(x) \leq \frac{1}{k_\delta} < \delta$. Therefore, $\sup_{n \geq n_{k_\delta}} |f_n(x) - f(x)| < \delta$.

Therefore, $\forall x \in B, n \geq n_{k_\delta}, |f_n(x) - f(x)| < \delta$ and $f_n \rightarrow f$ uniformly on B . ■

Definition 6.2. A sequence of measurable \mathbb{R} -valued functions f_n converges to a \mathbb{R} -valued measurable function f in L^1 if

$$\lim_{n \rightarrow \infty} \int |f_n - f| \, d\mu = 0 \quad (166)$$

Proposition 6.3 (Markov Inequality). If $g \geq 0$, then for all $t \geq 0$,

$$\mu(\{x : g(x) \geq t\}) \leq \frac{\int g \, d\mu}{t} \quad (167)$$

In probabilistic notations:

$$P(g \geq t) \leq \frac{\mathbb{E}[g]}{t} \quad (168)$$

Proof. Define $h(x) := t\mathbb{1}\{g \geq t\}$, obviously, $h \leq g$.

$$\int h \, d\mu = t\mu(\{x : g(x) \geq t\}) \leq \int g \, d\mu \quad (169)$$

The result follows. ■

Proposition 6.4. $f_n \xrightarrow{L^1} f \implies f_n \xrightarrow{\mu} f$.

Proof. Let $\varepsilon > 0$, apply Markov inequality on every $|f_n - f|$:

$$\mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\}) \leq \frac{\int |f_n - f| \, d\mu}{\varepsilon} \rightarrow 0 \text{ as } n \rightarrow \infty \quad (170)$$

Therefore, $f_n \xrightarrow{\mu} f$. ■

Remark 6.3.

1. $f_n \xrightarrow{a.e.} f \not\Rightarrow f_n \xrightarrow{L^1} f$.
2. $f_n \xrightarrow{L^1} f \not\Rightarrow f_n \xrightarrow{a.e.} f$.
3. $f_n \xrightarrow{\mu} f \not\Rightarrow f_n \xrightarrow{a.e.} f$.

Proposition 6.5 (Dominated Convergence Theorem II). Suppose $f_n \xrightarrow{\mu} f$, and \exists integrable g such that $|f_n| \leq g$ a.e. for all n . Then, $f_n \xrightarrow{L^1} f$, in particular, $\int f_n \, d\mu \rightarrow \int f \, d\mu$.

Proof. Suppose, for contradiction, $f_n \not\rightarrow f$ in L^1 . Equivalently, there exists ε and a subsequence f_{n_k} such that for all k :

$$\int |f_{n_k} - f| \, d\mu \geq \varepsilon \quad (\dagger) \quad (171)$$

But the convergence in measure implies $f_{n_k} \rightarrow f$ in measure as well. Then there exists a subsequence n_{k_ℓ} such that $f_{n_{k_\ell}} \rightarrow f$ almost everywhere.

By the previous dominated convergence theorem, $\lim_{\ell \rightarrow \infty} \int |f_{n_{k_\ell}} - f| \, d\mu = 0$, contradicts (\dagger) . ■

7 Normed Space

Definition 7.1. Let V be a vector space over \mathbb{R} (over \mathbb{C}), a **norm** on V is a function $\|\cdot\| : V \rightarrow \mathbb{R}$ such that

1. $\|x\| \geq 0 \ \forall x \in V$,
2. $\|x\| = 0 \iff x = 0$,
3. $\|ax\| = |a| \|x\|$ for all $a \in \mathbb{R}(\in \mathbb{C})$,
4. (Triangle Inequality) $\|x + y\| \leq \|x\| + \|y\| \ \forall x, y \in V$.

Example 7.1. For $V = \mathbb{R}^n$, for every $p \geq 1$, the L^p norm is defined as

$$\|x\|_{L^p} = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \quad (172)$$

Note: for $p < 1$, the triangle inequality fails.

Example 7.2. Let $C[a, b]$ denote the collection of continuous functions from $[a, b]$ to \mathbb{R} , where $[a, b]$ is a compact interval.

The sup norm is defined as

$$\|f\|_{\infty} = \sup_{x \in [a, b]} |f(x)| \quad (173)$$

The 1-norm is defined as

$$\|f\|_1 = \int_{[a, b]} |f| \, d\lambda \quad (174)$$

Definition 7.2. Let S be a set, a **metric** d on S is a function $d : S \times S \rightarrow \mathbb{R}$ such that for all $x, y, z \in S$:

1. $d(x, y) \geq 0$,
2. $d(x, y) = 0 \iff x = y$,
3. $d(x, y) = d(y, x)$,
4. $d(x, y) \leq d(x, z) + d(y, z)$.

Proposition 7.1. A norm induces a metric: $d(x, y) := \|x - y\|$.

Note: the converse is false, i.e., there are metrics not induced by any norm. For example, $d(x, y) := \mathbf{1}\{x = y\}$ is in general not induced by any norm.

Definition 7.3. Let S be a set with a metric d , a sequence of points x_n converges to $x \in S$ if

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0 \quad (175)$$

A sequence is **Cauchy** with respect to d if $\forall \varepsilon > 0, \exists n_0 \in \mathbb{N}$ s.t. $\forall m, n \geq n_0, d(x_m, x_n) < \varepsilon$.

Definition 7.4. A metric space w.r.t d is **complete** if every Cauchy sequence w.r.t. d converges to somewhere in the space.

Example 7.3. $C[a, b]$ with the supremum norm is complete.

Example 7.4. $C[a, b]$ with L^1 norm is not complete.

Proof. Using counter-example: for $[a, b] = [-1, 1]$,

$$f_n(x) = \begin{cases} 0 & \text{if } x \in [-1, 0] \\ nx & \text{if } x \in (0, 1/n) \\ 1 & \text{if } x \in [1/n, 1] \end{cases} \quad (176)$$

The sequence of f_n is Cauchy but converges to $f = \mathbf{1}_{\{x \geq 0\}} \notin C[a, b]$. ■