

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

Tianyu Du

November 10, 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.,  $\sigma$ -algebras.

### 1.2 Algebras and $\sigma$ -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  $\sigma$ -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  $\sigma$ -algebra on  $X$ ;  $\{\emptyset, X\}$  is a  $\sigma$ -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  $\sigma$ -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  $\sigma$ -algebra). The collection of subsets  $A \subseteq X$ , such that either  $A$  is countable or  $A^c$  is countable, forms a  $\sigma$ -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  $\sigma$ -algebras, then  $\bigcap_{i \in \mathcal{I}} \mathcal{A}_i$  is a  $\sigma$ -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  $\sigma$ -algebra  $\mathcal{A}$  containing  $\mathcal{P}$ . That is, for any  $\sigma$ -algebra  $\mathcal{B} \supseteq \mathcal{P}$ ,  $\mathcal{A} \subseteq \mathcal{B}$ .  $\mathcal{A}$  is defined as the  $\sigma$ -algebra **generated by**  $\mathcal{P}$ , denoted as  $\sigma(\mathcal{P})$ .

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

### 1.3 Borel $\sigma$ -algebra

**Definition 1.3.** The **Borel  $\sigma$ -algebra** of  $\mathbb{R}$ , denoted as  $\mathcal{B}(\mathbb{R})$ , is the  $\sigma$ -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  $\sigma$ -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  $\sigma$ -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  $\sigma$ -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  $\sigma$ -algebra  $\mathcal{A}$  of  $X$ , the pair  $(X, \mathcal{A})$  is called a **measurable space**.

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

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

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

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

**Example 1.7** (counting measure).

**Example 1.8** (point-mass measure).

**Proposition 1.2.** A measure  $\mu$  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  $\mu$  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  $\sigma$ -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.  $\lambda^*$  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  $\varepsilon$ .

(1.2) Let  $A \subseteq B$ ,  $\Omega_A$  and  $\Omega_B$  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  $b - a = \text{sum of lengths of } I'_i$ .

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  $\mu^*$ . A set  $B \subseteq X$  is  $\mu^*$ -**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  $\mu^*$  on it, let  $\mathcal{M}_{\mu^*}$  denote the set of all  $\mu^*$ -**measurable** sets. Then,  $\mathcal{M}_{\mu^*}$  is a  $\sigma$ -algebra and  $\mu^*|_{\mathcal{M}_{\mu^*}}$  ( $\mu^*$  restricted to  $\mathcal{M}_{\mu^*}$ ) is a measure.

*Proof.* To show  $B$  is  $\mu^*$ -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,  $\mathcal{M}_{\mu^*}$  is an algebra.

(1.3.2) Now show that  $\mathcal{M}_{\mu^*}$  is a  $\sigma$ -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  $\mathcal{M}_{\mu^*}$  is closed under countable disjoint unions.

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

(2) Let  $B_1, B_2, \dots$  be a sequence of disjoint sets from  $\mathcal{M}_{\mu^*}$ . 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,  $\mu^*$  is a measure on  $\mathcal{M}_{\mu^*}$ . ■

**Definition 1.10.** Let  $\lambda^*$  be the Lebesgue outer measure on  $\mathbb{R}$ , then the collection  $\mathcal{M}_{\lambda^*}$  of  $\lambda^*$ -measurable sets is called the **Lebesgue  $\sigma$ -algebra**. The restriction  $\lambda = \lambda^*|_{\mathcal{M}_{\lambda^*}}$ , which is a measure on  $\mathcal{M}_{\lambda^*}$ , is called the **Lebesgue measure**. Any set in  $\mathcal{M}_{\lambda^*}$  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  $\lambda^*$ -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  $\mu^*$ -measurable and  $\mathcal{M}_{\lambda^*}$  is a  $\sigma$ -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  $\lambda^*$  is an outer measure and  $\lambda^*(Q) = \prod_{i=1}^d (b_i - a_i)$ .
3.  $\mathcal{M}_{\lambda^*}$  is the Lebesgue  $\sigma$ -algebra on  $\mathbb{R}^d$ . Restricting  $\lambda^*$  on  $\mathcal{M}_{\lambda^*}$  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  $\mathcal{M}_{\lambda^*}$ .



## 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  $\lambda$  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  $\varepsilon$  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  $a \in A$ . 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  $\lambda$  denote the Lebesgue measure, let  $\mu$  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  $\lambda^*$ : translate all covering intervals by  $+x$  and the volumes of these intervals stay the same. Since  $\lambda$  is simply the restriction of  $\lambda^*$  on Lebesgue measurable sets,  $\lambda$  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  $\lambda^*$ ,

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

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

**Theorem 1.4.** Let  $\mu$  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  $\lambda$  is the Lebesgue measure.

**Remark 1.1.** Borel  $\sigma$ -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  $\mu$  since  $\mu$  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  $\nu$  is a measure that is finite on bounded sets and agrees with  $\lambda$  on all half-open dyadic cubes. By the previous proposition,  $\lambda$  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  $\sigma$ -algebra on  $\mathbb{R}^*$  is defined to be the  $\sigma$ -algebra generated by  $\mathcal{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}$ .

**Proposition 2.1.**

$$\sigma(\mathcal{B}(\mathbb{R}) \cup \{-\infty\} \cup \{\infty\}) = \mathcal{B}(\mathbb{R}) \cup \{B \cup \{\infty\} : B \in \mathcal{B}(\mathbb{R})\} \quad (1)$$

$$\cup \{B \cup \{-\infty\} : B \in \mathcal{B}(\mathbb{R})\} \quad (2)$$

$$\cup \{B \cup \{-\infty, \infty\} : B \in \mathcal{B}(\mathbb{R})\} \quad (3)$$

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

$$\{x \in X : f(x) \leq t\} \in \mathcal{A} \quad (4)$$

$$\{x \in X : f(x) < t\} \in \mathcal{A} \quad (5)$$

$$\{x \in X : f(x) \geq t\} \in \mathcal{A} \quad (6)$$

$$\{x \in X : f(x) > t\} \in \mathcal{A} \quad (7)$$

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

$$\mathcal{B}_0 := \{B \in \mathcal{B} : f^{-1}(B) \in \mathcal{A}\} \supseteq \mathcal{C} \quad (8)$$

It's easy to check  $\mathcal{B}_0$  is in fact a  $\sigma$ -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

$$f(x) = \begin{cases} a_i & \text{if } x \in A_i \\ 0 & \text{if } x \notin A_i \forall i \in [n] \end{cases} \quad (9)$$

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:

$$f(x, y) = x + y \quad (10)$$

$$f(x, y) = \max\{x, y\} \equiv x \vee y \quad (11)$$

$$f(x, y) = \min\{x, y\} \equiv x \wedge y \quad (12)$$

$$f(x, y) = x - y \quad (13)$$

$$f(x, y) = \alpha x \quad \alpha \in \mathbb{R} \quad (14)$$

**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.*

$$h^{-1}([-\infty, t] \times [-\infty, s]) = f^{-1}([-\infty, t]) \cap g^{-1}([-\infty, s]) \in \mathcal{A} \quad (15)$$

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  $\{x \in X : \sup_n f_n \leq t\} = \bigcap_{n=1}^{\infty} \{x \in X : f_n \leq t\} \in \mathcal{A}$  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

$$\{x \in A : f(x) < g(x)\}, \{x \in A : f(x) \leq g(x)\} \quad (16)$$

are measurable.

*Proof.*

$$\{x \in A : f(x) < g(x)\} = \bigcup_{r \in \mathbb{Q}} (\{x \in A : f(x) < r\} \cap \{x \in A : r < g(x)\}) \quad (17)$$

■

**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,

$$\{x \in X : \lim_{n \rightarrow \infty} f_n(x) \text{ exists}\} \quad (18)$$

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

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

**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  $\mu(A_i \cap B_j) = \mu(\emptyset) = 0$ .

$$\sum_{i=1}^n a_i \mu(A_i) = \sum_{i=1}^n a_i \sum_{j=1}^m \mu(A_i \cap B_j) \quad (2)$$

$$= \sum_{j=1}^m b_j \sum_{i=1}^n \mu(A_i \cap B_j) \quad (3)$$

$$= \sum_{j=1}^m b_j \mu(B_j) \quad (4)$$

■

### 3.2 Integrating Measurable Functions

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

$$\int f \, d\mu = \sup \left\{ \int g \, d\mu : g \text{ is a non-negative simple function such that } g \leq f \right\} \quad (5)$$

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

$$f^+(x) = \max\{f(x), 0\} \quad (6)$$

$$f^-(x) = -\min\{f(x), 0\} \quad (7)$$

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

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu \quad (8)$$

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

$$\int f + g \, d\mu = \int f \, d\mu + \int g \, d\mu \quad (9)$$

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

$$\int \alpha f \, d\mu = \alpha \int f \, d\mu \quad (10)$$

*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

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad (11)$$

for all  $x$ .

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

$$A_{n,k} = \left\{ x \in X : \frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n} \right\} \quad (12)$$

Define

$$f_n(x) = \begin{cases} \frac{k-1}{2^n} & \text{if } x \in A_{n,k} \\ n & \text{otherwise} \end{cases} \quad (13)$$

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

$$\int f d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu \quad (14)$$



*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

$$A_{n,i} = \{x \in A_i : f_n(x) \geq (1 - \varepsilon)a_i\} \quad (15)$$

Define

$$g_n(x) = \begin{cases} (1 - \varepsilon)a_i & \text{if } x \in A_{n,i} \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

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

$$(1 - \varepsilon) \int f \, d\mu \leq \lim \int g_n \, d\mu \leq \lim \int f_n \, d\mu \leq \int f \, d\mu \quad (17)$$

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.

$$\lim_n \int f_n \, d\mu \geq \lim_n \int g_n \, d\mu \quad (18)$$

$$= \lim_n \sum_{i=1}^K (1 - \varepsilon)a_i \mu(A_{n,i}) \quad (19)$$

$$= \sum_{i=1}^K (1 - \varepsilon)a_i \lim_n \mu(A_{n,i}) \quad (20)$$

$$= \sum_{i=1}^K (1 - \varepsilon)a_i \mu(A_i) \text{ Since for all } i, A_{n,i} \uparrow A_i \text{ as } n \rightarrow \infty. \quad (21)$$

$$= (1 - \varepsilon) \int f \, d\mu \quad (22)$$

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

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int f_n \, d\mu \quad (23)$$

*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  $g_n(x) \leq g_{n+1}(x)$ .

$$\lim_{n \rightarrow \infty} g_n(x) = \lim_{n \rightarrow \infty} \min\{g, f_n\} \quad (24)$$

$$= \min\{g(x), f(x)\} \quad (25)$$

$$= g(x) \quad (26)$$

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,

$$\int (f + g) d\mu = \lim_{n \rightarrow \infty} \int (f_n + g_n) \, d\mu \quad (27)$$

$$= \lim_{n \rightarrow \infty} \left( \int f_n \, d\mu + \int g_n \, d\mu \right) \quad (28)$$

$$= \lim_{n \rightarrow \infty} \int f_n \, d\mu + \lim_{n \rightarrow \infty} \int g_n \, d\mu \quad (29)$$

$$= \int f \, d\mu + \int g \, d\mu \quad (30)$$

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

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

$$\{h \in \mathbb{S}_+ \text{ and } h \leq g\} \subseteq \{h \in \mathbb{S}_+ \text{ and } h \leq f\} \quad (31)$$

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  $(X, \mathcal{A}, \mu)$  be a measure space, let  $f, g : X \rightarrow \mathbb{R}^*$  be integrable functions, let  $\alpha \in \mathbb{R}$ . Then,  $f + g$  and  $\alpha f$  are integrable, and

$$\int f + g d\mu = \int f d\mu + \int g d\mu \quad (32)$$

$$\int \alpha f d\mu = \alpha \int f d\mu \quad (33)$$

*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

$$\left| \int f d\mu \right| \leq \int |f| d\mu \quad (34)$$

*Proof.*

$$\left| \int f d\mu \right| = \left| \int f^+ - f^- d\mu \right| \quad (35)$$

$$= \left| \int f^+ d\mu - \int f^- d\mu \right| \quad (36)$$

$$\leq \left| \int f^+ d\mu \right| + \left| \int f^- d\mu \right| \quad (37)$$

$$= \int f^+ d\mu + \int f^- d\mu \quad (38)$$

$$= \int |f| d\mu \quad (39)$$

■

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

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

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int f_n \, d\mu \quad (1)$$

*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

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

Note that  $h_n$  is a non-decreasing sequence since

$$h_{n+1} = \max\{g_{1,n+1}, g_{2,n+1}, \dots, g_{n+1,n+1}\} \quad (3)$$

$$\geq \max\{g_{1,n+1}, g_{2,n+1}, \dots, g_{n,n+1}\} \quad (4)$$

$$\geq \max\{g_{1,n}, g_{2,n}, \dots, g_{n,n}\} = h_n \quad (5)$$

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

$$h_n(x) = \max\{g_{1,n}, g_{2,n}, \dots, g_{n,n}\} \geq g_{m,n} \quad (6)$$

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

$$\lim_{n \rightarrow \infty} h_n(x) \geq \lim_{n \rightarrow \infty} g_{m,n} = f_m \quad (7)$$

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

$$\lim_n h_n(x) = \lim_m \lim_n h_n(x) \geq \lim_m f_m = f \quad (8)$$

$$\implies \int \lim_n h_n(x) \, d\mu \geq \int f \, d\mu \quad (9)$$

Note that, by construction,

$$h_n = \max\{g_{1,n}, g_{2,n}, \dots, g_{n,n}\} \leq \max\{f_1, \dots, f_n\} = f_n \quad (10)$$

Therefore,

$$\int \lim_{n \rightarrow \infty} f_n(x) \, d\mu \geq \int f \, d\mu \quad (11)$$

■

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

$$\int \sum_{n=1}^{\infty} f_n \, d\mu = \sum_{n=1}^{\infty} \int f_n \, d\mu \quad (12)$$

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

$$\int \liminf_{n \rightarrow \infty} f_n \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n \, d\mu \quad (13)$$

*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,

$$\int \liminf_{n \rightarrow \infty} f_n \, d\mu = \int \lim_{n \rightarrow \infty} g_n \, d\mu \quad (14)$$

$$= \lim_{n \rightarrow \infty} \int g_n \, d\mu \quad (15)$$

$$\leq \lim_{n \rightarrow \infty} \inf_{k \geq n} \int f_k \, d\mu \quad (16)$$

$$= \liminf_{n \rightarrow \infty} \int f_n \, d\mu \quad (17)$$

■

**Theorem 4.3** (Lebesgue's Dominated Convergence Theorem). Let  $(X, \mathcal{A}, \mu)$  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,

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int f_n \, d\mu \quad (18)$$

*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

$$\int f + g \, d\mu = \int \liminf_{n \rightarrow \infty} f + g \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n + g \, d\mu \quad (19)$$

$$= \liminf_{n \rightarrow \infty} \int f_n \, d\mu + \int g \, d\mu \quad (20)$$

$$\implies \int f \, d\mu \leq \liminf_{n \rightarrow \infty} \int f_n \, d\mu \quad (21)$$

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

$$\int g - f \, d\mu = \int \liminf g - f_n \, d\mu \leq \liminf \int g - f_n \, d\mu \quad (22)$$

$$\implies - \int f \, d\mu \leq - \liminf \int f_n \, d\mu \quad (23)$$

$$\implies \int f \, d\mu \geq \limsup \int f_n \, d\mu \quad (24)$$

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

$$\liminf \int f_n \, d\mu \geq \int f \, d\mu \geq \limsup \int f_n \, d\mu \geq \liminf \int f_n \, d\mu \quad (25)$$

$$\implies \int f \, d\mu = \lim \int f_n \, d\mu \quad (26)$$

■

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

$$\lim_{n \rightarrow \infty} \int |f_n - f| \, d\mu = 0 \quad (27)$$

*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  $(X, \mathcal{A}, \mu)$  be a measure space, a set  $N \subseteq X$  (not necessarily measurable) is called **negligible w.r.t.**  $\mu$  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.  $\mu$  (denoted as  $\mu$ -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  $\mu$ -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,  $\int f \, d\mu \leq \int h_n \, d\mu = n\mu(A)$ . If  $\mu(A) > 0$ , taking  $n \rightarrow \infty$  leads to a contradiction. ■

*Alternative Proof.* Note: this intuitive proof is non-rigorous. Since  $f \geq 0$ , let  $A := f^{-1}(\infty)$ ,  $\int f \, d\mu \geq \int_A f \, d\mu = \infty\mu(A)$ ,  $\mu(A)$  must be zero. ■

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

*Proof.*  $f$  is integrable implies both  $\int f^+ \, d\mu, \int f^- \, d\mu < \infty$ . 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

$$\int f \, d\mu = 0 \iff f = 0 \, \mu - a.e. \quad (28)$$

*Proof.* ( $\Leftarrow$ ) Suppose  $f = 0$  a.e., for every simple function  $g \leq f$ , let  $(a_i, A_i)$  be the representation of  $g$ .

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,  $\int g \, d\mu = \sum_i a_i \mu(A_i) = 0$ , so is the integral of  $f$ .

( $\Rightarrow$ ) Suppose  $\int f \, d\mu = 0$ , note that

$$\{x : f(x) > 0\} = \bigcup_{n=1}^{\infty} \{x : f(x) > 1/n\} \quad (29)$$

Define  $A_n = \{x : f(x) > 1/n\}$ , then  $A_n$  is an increasing sequence of sets.

Suppose, for contradiction, there exists some  $A_n$  with  $\mu(A_n) > 0$ , define  $g(x) = \frac{1}{n} \mathbf{1}\{x \in A_n\}$ .  $f \geq g$  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  $\mu(\{x : f(x) > 0\}) = \lim_n \mu(A_n) = 0$ . ■

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

$$f = 0 \, a.e. \implies \int f \, d\mu = 0 \quad (30)$$

*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, obviously one may take  $f^+ = f^-$  so that  $\int f^+ \, d\mu = \int f^- \, d\mu \neq 0$  and  $\int f \, d\mu = 0$ . ■

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

$$f = 0 \, a.e. \implies \int f \, d\mu = \int g \, d\mu \quad (31)$$

*Proof.* Let  $\tilde{f} = f(x) \mathbf{1}\{x \in \mathbb{R}\}$  and  $\tilde{g} = g(x) \mathbf{1}\{x \in \mathbb{R}\}$ , we are doing this to avoid subtracting infinity from infinity.

Both  $|\tilde{f}|$  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  $|\tilde{f}|, |\tilde{g}| < \infty$ ,  $\tilde{f} - \tilde{g}$  is integrable and we can write

$$\int \tilde{f} - \tilde{g} \, d\mu = \int \tilde{f} \, d\mu - \int \tilde{g} \, d\mu = 0 \quad (32)$$

$$\implies \int f \, d\mu = \int \tilde{f} \, d\mu = \int g \, d\mu = \int \tilde{g} \, d\mu \quad (33)$$

■

**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.

$$A = \{x : f_n(x) \geq 0 \ \forall n \wedge \lim_{n \rightarrow \infty} f_n(x) = f(x)\} \quad (34)$$

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 use the fact that  $\int \tilde{f}_n d\mu = \int f_n d\mu$  and  $\int \tilde{f} d\mu = \int f d\mu$ . ■

*Proof for DCT.* The proof is similar, we can construct sets on which the desired properties holds denoted as  $A$ . Define  $\tilde{f}(x) := f(x)\mathbb{1}_{\{x \in A\}}$  and apply the original DCT. Lastly, use the fact that modifying  $f$  on a measure zero set  $A^c$  does not change the value of integral. ■



## 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  $\mathbb{R}$ -valued functions by construction of  $\mathbb{C}$ ) are measurable. Similarly,  $f$  is **integrable** if both its real and imaginary parts are integrable. Define

$$\int f \, d\mu = \int \Re(f) \, d\mu + i \int \Im(f) \, d\mu \in \mathbb{C} \quad (1)$$

**Proposition 5.1** (Linearity of Integral of Complex-Valued Functions). 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

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu \quad (2)$$

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

$$\left| \int f \, d\mu \right| = \alpha \int f \, d\mu \quad (3)$$

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,

$$\left| \int f \, d\mu \right| = \alpha \int f \, d\mu \quad (4)$$

$$= \int (\alpha f) \, d\mu \quad (5)$$

$$= \int \Re(\alpha f) \, d\mu + i \int \Im(\alpha f) \, d\mu \quad (6)$$

$$\implies \int \Im(\alpha f) \, d\mu = 0 \quad (7)$$

Therefore,

$$\left| \int f \, d\mu \right| = \int \Re(\alpha f) \, d\mu \leq \int |\alpha f| \, d\mu = \int |f| \, d\mu \quad (8)$$

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

## 6 Convergence of Measurable Functions

**Definition 6.1.** Let  $(X, \mathcal{A}, \mu)$  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  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) = 0 \quad (1)$$

*Note: this definition is a generalization of convergence in probability.*

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

*Counter-example.* Take  $\mu = \lambda$ , and define  $f_n(x) = \mathbb{1}\{x \in [n, \infty)\}$ , then  $f_n \rightarrow 0$  everywhere. However,  $\lambda(\{x : |f_n(x)| > 1/2\}) = \lambda([n, \infty)) = \infty \not\rightarrow 0$ . ■

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

*Counter-example.* Define

$$f_1(x) = 1 \quad (2)$$

$$f_2(x) = \mathbb{1}\{x \in [0, 1/2]\} \quad (3)$$

$$f_3(x) = \mathbb{1}\{x \in [1/2, 1]\} \quad (4)$$

$$f_4(x) = \mathbb{1}\{x \in [0, 1/4]\} \quad (5)$$

$$f_5(x) = \mathbb{1}\{x \in [1/4, 1/2]\} \quad (6)$$

$$f_6(x) = \mathbb{1}\{x \in [1/2, 3/4]\} \quad (7)$$

$$f_7(x) = \mathbb{1}\{x \in [3/4, 1]\} \quad (8)$$

$$f_8(x) = \mathbb{1}\{x \in [1/8, 1/4]\} \quad (9)$$

and so on. in general,  $\{x : f_n(x) = 1\}$  shrinks exponentially as  $n \rightarrow \infty$ , hence  $f_n \rightarrow 0$  in Lebesgue measure. However, for any fixed  $x \in [0, 1]$ , there are infinitely many  $n$  such that  $f_n(x) = 1$ , therefore,  $f_n$  does not converge to 0 pointwise. ■

**Proposition 6.1.** Let  $\mu$  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  $x$ . That is,

$$\{x : |f_n(x) - f(x)| > \varepsilon \text{ for infinitely many } n\} \subseteq \{x : f_n(x) \not\rightarrow f(x)\} \quad (10)$$

By monotonicity,

$$\mu(\{x : |f_n(x) - f(x)| > \varepsilon \text{ for infinitely many } n\}) \leq \mu(\{x : f_n(x) \not\rightarrow f(x)\}) = 0 \quad (11)$$

Further, note that

$$\{x : |f_n(x) - f(x)| > \varepsilon \text{ for infinitely many } n\} = \bigcap_{n=1}^{\infty} \underbrace{\bigcup_{k=n}^{\infty} \{x : |f_k(x) - f(x)| > \varepsilon\}}_{B_n} \quad (12)$$

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  $\mu(B_i) > 0$  since  $\mu$  is a finite measure. Therefore,

$$0 = \mu(\{x : |f_n(x) - f(x)| > \varepsilon \text{ for infinitely many } n\}) = \lim_{n \rightarrow \infty} \mu(B_n) \quad (13)$$

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

$$\lim_{n \rightarrow \infty} \mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) = 0 \quad (14)$$

■

**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

$$\mu(\{x : |f_{n_1} - f(x)| > 1\}) < 2^{-1} \quad (15)$$

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

$$\underbrace{\mu(\{x : |f_{n_k} - f(x)| > \frac{1}{k}\})}_{A_k} < 2^{-k} \quad (16)$$

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,

$$\mu(B) \leq \mu(B_j) = \mu\left(\bigcup_{k=j}^{\infty} A_k\right) \leq \sum_{k=j}^{\infty} \mu(A_k) < \sum_{k=j}^{\infty} 2^{-k+1} \quad (17)$$

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

$$\exists j \text{ s.t. } \forall k \geq j \quad |f_{n_k} - f(x)| \leq \frac{1}{k} \quad (18)$$

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

$$\sum_{k=1}^{\infty} \mu(A_k) < \infty \quad (19)$$

then

$$\mu(\{x : x \in \text{infinitely many } A_k\}) = 0 \quad (20)$$

*Proof.* Define

$$B = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \quad (21)$$

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

$$B \subseteq \bigcup_{k=j}^{\infty} A_k \quad (22)$$

Hence

$$\mu(B) \leq \mu\left(\bigcup_{k=j}^{\infty} A_k\right) \leq \sum_{k=j}^{\infty} \mu(A_k) \rightarrow 0 \text{ as } j \rightarrow \infty \quad (23)$$

Therefore,  $\mu(B) = 0$ . ■

**Theorem 6.1** (Egorov's Theorem). Let  $(X, \mathcal{A}, \mu)$  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  $B$ .

*On a finite measure space, convergence a.e. implies convergence uniformly on a slightly smaller set.*

*Proof.* Let  $\varepsilon > 0$ .

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

$$g_n(x) := \sup_{k \geq n} |f_k(x) - f(x)| \quad (24)$$

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

$$\mu\left(\left\{x : g_{n_k}(x) > \frac{1}{k}\right\}\right) < \frac{\varepsilon}{2^k} \quad (25)$$

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

$$B^c = \left\{x : g_{n_k}(x) > \frac{1}{k} \text{ for some } k\right\} \quad (26)$$

Then,

$$\mu(B^c) = \mu\left(\bigcup_{k=1}^{\infty} \left\{x : g_{n_k}(x) > \frac{1}{k}\right\}\right) \quad (27)$$

$$\leq \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = \varepsilon \quad (28)$$

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

$$\lim_{n \rightarrow \infty} \int |f_n - f| d\mu = 0 \quad (29)$$

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

$$\mu(\{x : g(x) \geq t\}) \leq \frac{\int g d\mu}{t} \quad (30)$$

In probabilistic notations:

$$P(g \geq t) \leq \frac{\mathbb{E}[g]}{t} \quad (31)$$

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

$$\int h d\mu = t\mu(\{x : g(x) \geq t\}) \leq \int g d\mu \quad (32)$$

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  $|f_n - f|$ :

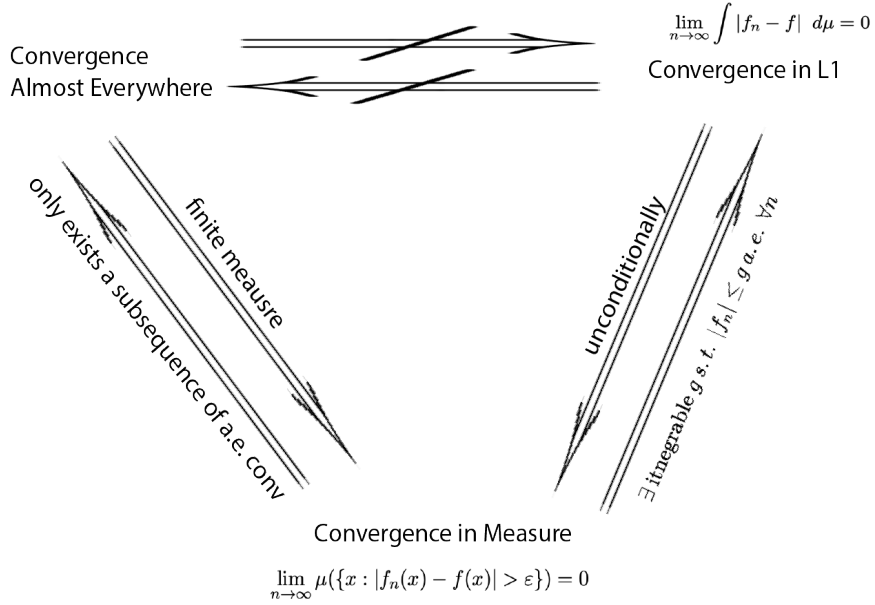
$$\mu(\{x : |f_n(x) - f(x)| \geq \varepsilon\}) \leq \frac{\int |f_n - f| d\mu}{\varepsilon} \rightarrow 0 \text{ as } n \rightarrow \infty \quad (33)$$

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.  $f_n \xrightarrow{\mu} f \not\Rightarrow f_n \xrightarrow{a.e.} f$ .

Figure 1: Modes of Convergences



**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,  $\int f_n d\mu \rightarrow \int f d\mu$ ).

*The convergence in measure version of the dominated convergence theorem.*

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

$$\int |f_{n_k} - f| d\mu \geq \varepsilon \quad (\dagger) \quad (34)$$

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

By the previous dominated convergence theorem,  $\lim_{\ell \rightarrow \infty} \int \left| f_{n_{k_\ell}} - f \right| d\mu = 0$ , 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 map  $\|\cdot\| : V \rightarrow \mathbb{R}$  satisfies the following properties:

1. (Non-negativity)  $\|x\| \geq 0 \ \forall x \in V$ ,
2.  $\|x\| = 0 \iff x = 0$ ,
3. (Linearity)  $\|ax\| = |a| \|x\|$  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  $\ell^p$  **norm** is defined as

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

*Note: we only define  $L^p$  norm for  $p \geq 1$ , since for  $p < 1$ , the triangle inequality fails.*

For  $p = \infty$ ,

$$\|x\|_{\ell^\infty} = \max_{1 \leq i \leq n} |x_i| \quad (2)$$

**Example 7.2.** Let  $C[a, b]$  denote the space of continuous functions map from  $[a, b]$  to  $\mathbb{R}$ , where  $[a, b]$  is a compact interval. The **sup-norm** is defined as

$$\|f\|_\infty = \sup_{x \in [a, b]} |f(x)| \quad (3)$$

This supremum is finite since  $|f|$  is continuous and  $|f|([a, b])$  is compact.

The **1-norm** is defined as

$$\|f\|_1 = \int_{[a, b]} |f| \, d\lambda \quad (4)$$

**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.  $d(x, y) \leq d(x, z) + d(y, z)$ .

**Definition 7.3.** A norm on a vector induces a metric, the **metric  $d$  induced by norm  $\|\cdot\|$**  is defined as

$$d(x, y) := \|x - y\| \quad (5)$$



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

**Definition 7.4.** Let  $S$  be a set with a metric  $d$ , a sequence of points  $\{x_n\}_{n=1}^{\infty}$  **converges** to  $x \in S$  if

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0 \quad (6)$$

A sequence is **Cauchy** with respect to  $d$  if

$$\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} \text{ s.t. } \forall m, n \geq n_0, d(x_m, x_n) < \varepsilon \quad (7)$$

**Definition 7.5.** A metric space w.r.t  $d$  is **complete** if every Cauchy sequence w.r.t.  $d$  converges to somewhere in the space.

**Remark 7.1.** In order to show the completeness of a metric space, take an arbitrary Cauchy sequence in this space, and show

1. construct the limit, in cases of functional spaces, we usually define the limit  $f$  as the point wise limit,
2. show this sequence converges to the proposed limit,
3. show the proposed limit is in the metric 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  $[a, b] = [-1, 1]$ ,

$$f_n(x) = \begin{cases} 0 & \text{if } x \in [-1, 0] \\ nx & \text{if } x \in (0, 1/n) \\ 1 & \text{if } x \in [1/n, 1] \end{cases} \quad (8)$$

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

**Proposition 7.1.**  $C[a, b]$  under sup-norm is complete.

*Proof.* Suppose  $f_n$  is a Cauchy sequence in  $C[a, b]$  under supremum norm. For all  $x \in [a, b]$ ,

$$f_n(x) - f_m(x) \leq \|f_n - f_m\|_{\infty} \rightarrow 0 \quad (9)$$

since  $f_n$  is Cauchy. Therefore,  $f_n(x)$  is a Cauchy sequence in  $\mathbb{R}$  and  $\lim_{n \rightarrow \infty} f_n(x)$  exists. Define  $f$  to be the point-wise limit of  $f_n$ .

Claim:  $f \in C[a, b]$  and  $f_n \rightarrow f$  in sup-norm.

For all  $\varepsilon > 0$ , there exists  $N$ , such that for all  $m, n \geq N$ ,

$$\|f_m - f_n\|_\infty < \varepsilon \quad (10)$$

Therefore, for all  $x \in [a, b]$ ,  $|f_n(x) - f_m(x)| < \|f_m - f_n\|_\infty < \varepsilon$ .

Fixing  $n$ , take  $m \rightarrow \infty$ , this shows for all  $n \geq N$ , for all  $x \in [a, b]$

$$|f(x) - f_n(x)| = \lim_{m \rightarrow \infty} |f_m(x) - f_n(x)| \leq \varepsilon \quad (11)$$

Therefore, for all  $n \geq N$ ,  $\|f - f_n\|_\infty \leq \varepsilon$ . Hence  $f \rightarrow f_n$  in sup-norm.

Now show the continuity of  $f$ : take  $x_0 \in [a, b]$ , given  $\varepsilon > 0$ , since  $f_n \rightarrow f$  in sup-norm, there exists  $N$  such that for all  $n \geq N$ ,

$$\|f - f_n\|_\infty \leq \frac{\varepsilon}{3} \quad (12)$$

In particular,  $\|f - f_N\|_\infty \leq \frac{\varepsilon}{3}$ .

Moreover, since  $f_N$  is continuous,  $\exists \delta > 0$  such that  $|x - x_0| < \delta \implies |f_N(x) - f_N(x_0)| < \varepsilon/3$  for every  $x$ . Take any  $x \in \mathcal{B}_\delta(x_0)$ , by triangle inequality,

$$|f(x) - f(x_0)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \quad (13)$$

$$\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \quad (14)$$

Hence,  $f \in C[a, b]$ . ■

## 8 Functional Analysis: $L^p$ Spaces

### 8.1 An Auxiliary Construction: the $\mathcal{L}^p$ Spaces

We will firstly define  $\mathcal{L}^p$  spaces, which is a little simpler than  $L^p$  spaces. The main difference is  $\mathcal{L}^p$  spaces are simply spaces of functions, while  $L^p$  does not distinguish functions that are equal almost everywhere. In fact,  $L^p$  spaces are spaces of equivalence classes of functions, an element  $f \in L^p$  actually denote the set of all functions that equal  $f$  almost everywhere.

**Definition 8.1.** Let  $(X, \mathcal{A}, \mu)$  be a measure space, for every  $1 \leq p < \infty$ , the  $\mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{R})$  space is the collection of all measurable functions  $f : X \rightarrow \mathbb{R}$  such that

$$\int |f|^p d\mu < \infty \quad (1)$$

Similarly,  $\mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{C})$  denotes the collection of all measurable functions  $f : X \rightarrow \mathbb{C}$  such that

$$\int |f|^p d\mu < \infty \quad (2)$$

Throughout this notes, we use  $\mathcal{L}^p$  to denote  $\mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{R})$  or  $\mathcal{L}^p(X, \mathcal{A}, \mu, \mathbb{C})$ , unless specified otherwise, statements about  $\mathcal{L}^p$  hold for both spaces.

**Proposition 8.1.**  $\mathcal{L}^p$  space is a vector space.

*Proof.*

1. Note that  $0 \in \mathcal{L}^p$ .
2. If  $f \in \mathcal{L}^p$  and  $\alpha \in \mathbb{R}$  or  $\mathbb{C}$ , then

$$\int |\alpha f|^p d\mu = |\alpha|^p \int |f|^p d\mu < \infty \quad (3)$$

Therefore,  $\alpha f \in \mathcal{L}^p$ .

3. For all  $x \in X$ ,

$$|f(x) + g(x)|^p \leq (|f(x)| + |g(x)|)^p \quad (4)$$

$$\leq (2 \max\{|f(x)|, |g(x)|\})^p \quad (5)$$

$$\leq 2^p \max\{|f(x)|^p, |g(x)|^p\} \quad (6)$$

$$\leq 2^p (|f(x)|^p + |g(x)|^p) \quad (7)$$

Thus,

$$\int |f + g|^p d\mu < \infty \quad (8)$$

$$\implies f + g \in \mathcal{L}^p \quad (9)$$

Hence,  $\mathcal{L}^p$  is a vector space. ■

**Definition 8.2.**  $\mathcal{L}^\infty(X, \mathcal{A}, \mu, \mathbb{R}/\mathbb{C})$  is defined to be the set of all bounded measurable  $f : X \rightarrow \mathbb{R}/\mathbb{C}$ .

**Definition 8.3.** For  $f \in \mathcal{L}^p$  with  $p < \infty$ , define

$$\|f\|_p = \left( \int |f|^p d\mu \right)^{\frac{1}{p}} \quad (10)$$

for  $p = \infty$ ,  $\|f\|_\infty$ 's definition is a little bit more complicated, for continuous functions, it collides with the sup-norm. However, it's not the same as sup-norm for discontinuous functions.

**Definition 8.4.** Given a measure space  $(X, \mathcal{A}, \mu)$ , a set  $B$  is called  **$\mu$ -null/negligible** if  $B \subseteq A$  for some  $A \in \mathcal{A}$  with  $\mu(A) = 0$  (note that  $B$  is not necessarily measurable).

A subset  $N \subseteq X$  is called **locally  $\mu$ -null** if  $\forall A \in \mathcal{A}$  with  $\mu(A) < \infty$ ,  $A \cap N$  is  $\mu$ -null. *A locally  $\mu$ -null set  $N$  shrinks any measurable set to  $\mu$ -null set by taking intersection.*

A property of elements of  $X$  is said to hold **locally a.e.** if the set on which it fails is locally  $\mu$ -null.

*We use this notion of locally null to circumvent non-sigma finite cases.*

**Definition 8.5.** For  $f \in \mathcal{L}^\infty$ , define

$$\|f\|_\infty = \inf \{M \geq 0 : \{x : |f(x)| > M\} \text{ is locally } \mu\text{-null.}\} \quad (11)$$

this is called the **essential supremum** of  $|f|$ . *Equivalently,  $\|f\|_\infty$  is the least (locally a.e.) upper bound of  $|f|$ .*

Note that  $\|f\|_\infty$  is only a semi-norm, we may modify a function on a measure-zero set without changing the value of  $\|f\|_\infty$ .

Our previous definitions of semi-norms on  $\mathcal{L}^p$  spaces satisfy

$$\|f\|_p = 0 \iff \int |f|^p d\mu = 0 \iff |f|^p = 0 \text{ a.e.} \iff f = 0 \text{ a.e.} \quad (12)$$

This definition of semi-norm on  $\mathcal{L}^\infty$  ensures  $\|f\|_\infty = 0 \iff f = 0 \text{ a.e.}$ .

**Example 8.1.** Take  $X = [0, 1]$  and  $\mu = \lambda$ ,

$$f(x) = \begin{cases} x & \text{if } x \neq \frac{1}{2} \\ 2 & \text{otherwise} \end{cases} \quad (13)$$

Then  $\|f\|_\infty = 1$  but  $\sup f = 2$ . To see this, note that  $\{x : |f(x)| > 1\} = \{1/2\}$  has zero measure. However, for any  $M < 1$ , the same has non-zero Lebesgue measure.

**Lemma 8.1.** Countable union of locally  $\mu$ -null sets is locally  $\mu$ -null.

*Proof.* Let  $B_1, B_2 \dots$  be  $\mu$ -null, then for any  $A \in \mathcal{A}$ ,

$$\mu \left( A \cap \bigcup_{i=1}^{\infty} B_i \right) = \mu \left( \bigcup_{i=1}^{\infty} A \cap B_i \right) \leq \sum_{i=1}^{\infty} \mu(A \cap B_i) = 0 \quad (14)$$

■

**Proposition 8.2.**

$$\mu(\{x : |f(x)| > \|f\|_{\infty}\}) \text{ is locally } \mu\text{-null.} \quad (15)$$

$$\mu(\{x : |f(x)| > c\}) \text{ is not locally } \mu\text{-null } \forall c < \|f\|_{\infty} \quad (16)$$

*Proof.* First, note that by definition of  $\|f\|_{\infty}$ , it follows that  $\{x : |f(x)| > c\}$  is not locally  $\mu$ -null for any  $c < \|f\|_{\infty}$ , which is the infimum. Moreover,

$$\{x : |f(x)| > \|f\|_{\infty}\} = \bigcup_{n=1}^{\infty} \{x : |f(x)| > \|f\|_{\infty} + 1/n\} \quad (17)$$

By the previous lemma, the result follows. ■

**Proposition 8.3.**  $\|f\|_p$  and  $\|f\|_{\infty}$  are semi-norms.

*Proof.*  $\|f\|_p$  or  $\|f\|_{\infty} = 0$  only implies  $f = 0$  almost everywhere but not everywhere, this fact makes them semi-norms.

Later in  $L^p$  spaces, we will define the zero vector to be the collection of functions that are zero almost everywhere, this modification guarantees  $\|\cdot\|$  to be a norm on  $L^p$ . ■

**Definition 8.6.** Given  $p \in (1, \infty)$ , the **conjugate exponent**  $q$  is defined as

$$\frac{1}{p} + \frac{1}{q} = 1 \quad (18)$$

That is,

$$q = \frac{p}{p-1} \quad (19)$$

For  $p = \infty$ ,  $q = 1$ .

**Lemma 8.2** (Young's Inequality). Take  $p \in (1, \infty)$ , let  $q$  be the conjugate exponent of  $p$ , then for all  $x, y \geq 0$ ,

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \quad (20)$$

*Proof.* ■

**Theorem 8.1** (Hölder's Inequality). Let  $(X, \mathcal{A}, \mu)$  be a measure space, take  $1 \leq p \leq \infty$ , and  $q$  be its conjugate exponent. Take  $f \in \mathcal{L}^p$ ,  $g \in \mathcal{L}^q$ , then the product

$$fg \in \mathcal{L}^1 \quad (21)$$

and

$$\|fg\|_1 \leq \|f\|_p \|g\|_q \quad (22)$$

*Proof.*  $p \in (1, \infty)$ . For all  $x$ , and for any function  $f$  and  $g$ , by Young's inequality,

$$|f(x)g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q} \quad (23)$$

Integrating both sides,

$$\|fg\|_1 \leq \frac{\|f\|_p^p}{p} + \frac{\|g\|_q^q}{q} \quad (24)$$

If  $\|f\|_p = \|g\|_q = 1$ , then

$$\|fg\|_1 \leq \frac{1}{p} + \frac{1}{q} = 1 \quad (25)$$

Now take arbitrary  $f \in \mathcal{L}^p$  and  $g \in \mathcal{L}^q$ , if  $\|f\|_p = 0$  or  $\|g\|_q = 0$ , then  $fg = 0$  a.e. and there is nothing to prove.

So assume  $\|f\|_p > 0$  and  $\|g\|_q > 0$ , let

$$\tilde{f} = \frac{f}{\|f\|_p} \quad \tilde{g} = \frac{g}{\|g\|_q} \quad (26)$$

By construction,  $\|\tilde{f}\|_p = 1 = \|\tilde{g}\|_q$ . By Equation (25),  $\|\tilde{f}\tilde{g}\|_1 \leq 1$ , but  $\|\tilde{f}\tilde{g}\|_1 = \frac{\|fg\|_1}{\|f\|_p \|g\|_q}$ . This proves the Hölder's inequality when  $p \in (1, \infty)$ . ■

*Proof.*  $p = 1$  and  $q = \infty$ . Let  $f \in \mathcal{L}^1$  and  $g \in \mathcal{L}^\infty$ . Claim:

$$\{x : |f(x)g(x)| > \|g\|_\infty |f(x)|\} \quad (27)$$

is  $\mu$ -null.

*Proof of the Claim.* Note that

$$\{x : |f(x)g(x)| > \|g\|_\infty |f(x)|\} = \bigcup_{n=1}^{\infty} (\{x : |f(x)| > 1/n\} \cap \{x : |g(x)| > \|g\|_\infty\}) \quad (28)$$

By Markov ineuqliaty,

$$\mu(\{x : |f(x)| > 1/n\}) \leq \frac{\int |f| d\mu}{1/n} < \infty \quad (29)$$

The intersection of a locally  $\mu$ -null set with a set of finite measure is  $\mu$ -null, moreover, the countable union of  $\mu$ -null sets is  $\mu$ -null. ■

By the claimed property,

$$\|fg\|_1 = \int |fg| d\mu \leq \int \|g\|_\infty |f| d\mu = \|g\|_\infty \|f\|_1 \quad (30)$$

This shows the Hölder's inequality. ■

**Example 8.2.** Take  $X = \{x_1, \dots, x_n\}$  and  $\mu$  to be the counting measure on  $X$ . Let  $p = q = 2$  and  $f, g \in \mathcal{L}^2$ . Define  $v = (f(x_1), \dots, f(x_n)) \in \mathbb{R}^n$  and  $u = (g(x_1), \dots, g(x_n)) \in \mathbb{R}^n$ .

$$\|fg\|_1 = \sum_{i=1}^n \mu(\{x_i\}) |f(x_i)g(x_i)| = \sum_{i=1}^n |f(x_i)g(x_i)| \quad (31)$$

Therefore,

$$|\langle v, u \rangle| = \left| \sum_{i=1}^n f(x_i)g(x_i) \right| \leq \|fg\|_1 \quad (32)$$

In this finite dimensional case with counting measure,

$$\|f\|_2 = \sqrt{\sum_{i=1}^n \mu(\{x_i\}) f(x_i)^2} = \sqrt{\sum_{i=1}^n f(x_i)^2} = \|v\|_2 \quad (33)$$

The same holds for  $g$ , in this case Hölder's inequality induces the Cauchy-Switchz inequality.

**Theorem 8.2** (Minkowski's Inequality). Let  $(X, \mathcal{A}, \mu)$  be a measure space. Take  $1 \leq p \leq \infty$ . If  $f, g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ , then  $f + g \in \mathcal{L}^p$  and

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \quad (34)$$

*Proof.* First, suppose that  $p \in (1, \infty)$ . Let  $q$  be the conjugate exponent of  $p$ . We have already shown that  $\mathcal{L}^p$  is a vector space, so  $f + g \in \mathcal{L}^p$ .

Note that

$$1/p + 1/q = 1 \implies (p+q)/(pq) = 1 \implies p+q = pq \implies p = (p-1)q \quad (35)$$

Therefore,

$$\int (|f+g|^{p-1})^q d\mu = \int |f+g|^p d\mu < \infty \quad (36)$$

Therefore,  $|f+g|^{p-1} \in \mathcal{L}^q$ . By Hölder's inequality,

$$\int |f+g|^p d\mu = \int |f+g| |f+g|^{p-1} d\mu \quad (37)$$

$$\leq \int |f| |f+g|^{p-1} d\mu + \int |g| |f+g|^{p-1} d\mu \quad (38)$$

$$\leq \|f\|_p \|f+g|^{p-1}\|_q + \|g\|_p \|f+g|^{p-1}\|_q \quad (39)$$

where

$$\|f+g|^{p-1}\|_q = \left( \int (|f+g|^{p-1})^q d\mu \right)^{1/q} = \left( \int |f+g|^p d\mu \right)^{1/q} \quad (40)$$

If  $\|f+g\|_p = 0$ , we are done. Suppose not, divide  $(\int |f+g|^p d\mu)^{1/q}$  on both sides,

$$\frac{\int |f+g|^p d\mu}{(\int |f+g|^p d\mu)^{1/q}} \leq \|f\|_p + \|g\|_p \quad (41)$$

$$\implies (\int |f+g|^p d\mu)^{1-1/q} = (\int |f+g|^p d\mu)^{1/p} = \|f+g\|_p \leq \|f\|_p + \|g\|_p \quad (42)$$

When  $p = 1$ ,

$$\|f+g\|_1 = \int |f+g| d\mu \leq \int (|f| + |g|) d\mu = \|f\|_1 + \|g\|_1 \quad (43)$$

When  $p = \infty$ , define

$$N_1 = \{x : |f(x)| > \|f\|_\infty\} \quad (44)$$

$$N_2 = \{x : |g(x)| > \|g\|_\infty\} \quad (45)$$

Then  $N_1$  and  $N_2$  are locally  $\mu$ -null, so is  $N_1 \cup N_2$ . For  $x \notin N_1 \cup N_2$ ,

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty \quad (46)$$

■

Thus, we have shown  $\|\cdot\|_p$  on  $\mathcal{L}^p$  satisfies

1. If  $f = 0$ , then  $\|f\|_p = 0$ ,



2.  $\|\alpha f\|_p = |\alpha| \|f\|_p$  for any scalar  $\alpha$ ,
3.  $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ .

Thus  $\|\cdot\|_p$  satisfies all conditions of being a norm except that  $\|f\|_p = 0 \not\Rightarrow f = 0$ , thus  $\|\cdot\|_p$  is a semi-norm on  $\mathcal{L}^p$ .

## 8.2 $L^p$ Spaces

Note that  $\|\cdot\|_p$  is a **semi-norm** on  $\mathcal{L}^p$ , to make it a norm, we introduce the  $L^p$  space.

**Definition 8.7.** For  $1 \leq p < \infty$ , define the class of zero vectors

$$\mathcal{N}^p := \{f \in \mathcal{L}^p : f \text{ is measurable and } f = 0 \text{ a.e.}\} \quad (47)$$

which is a subspace of  $\mathcal{L}^p$ . Define  $L^p$  to be the quotient space:

$$L^p(X, \mathcal{A}, \mu) := \mathcal{L}^p(X, \mathcal{A}, \mu) / \mathcal{N}^p(X, \mathcal{A}, \mu) \quad (48)$$

That is, an element  $[f] \in L^p$  (an equivalence class) is the collection of all  $g \in \mathcal{L}^p$  such that  $f - g = 0$  almost everywhere:

$$[f] := \{g \in \mathcal{L}^p : f - g \in \mathcal{N}^p\} \quad (49)$$

Then  $L^p$  is a vector space over  $\mathbb{R}$  or  $\mathbb{C}$ , and  $\|\cdot\|_p$  is well-defined: for any  $f$ , for all  $g \in [f]$ ,  $\|f\|_p = \|g\|_p$  since  $f = g$  almost everywhere so their integrals are the same. Most importantly,  $\|\cdot\|_p$  is a norm on  $L^p$ . For  $p = \infty$ , we define

$$\mathcal{N}^\infty := \{f : f \text{ is bounded, measure and } f = 0 \text{ a.e.}\} \quad (50)$$

Then  $L^\infty := \mathcal{L}^p / \mathcal{N}^p$ .

Note that  $L^p$  for  $1 \leq p \leq \infty$  is also a vector space with equivalence relations. In general, we treat  $L^p$  as a space of functions instead of a space of classes of functions.

**Proposition 8.4.** Convergence in  $L^p$  ( $1 \leq p < \infty$ ) implies convergence in measure.

*Proof.* By Markov's inequality,

$$\mu(\{x : |f_n(x) - f(x)| > \varepsilon\}) = \mu(\{x : |f_n(x) - f(x)|^p > \varepsilon^p\}) \quad (51)$$

$$\leq \frac{\int |f_n - f|^p d\mu}{\varepsilon^p} \rightarrow 0 \text{ as } n \rightarrow \infty \quad (52)$$

■

**Corollary 8.1.** Let  $f_n \rightarrow f$  in  $L^p$  with  $1 \leq p < \infty$ , then there exists a subsequence  $f_{n_k} \rightarrow f$  a.e.

*Proof.* As convergence in  $L^p$  implies convergence in measure, which further implies existence of a.e. converging subsequences. ■

**Theorem 8.3.** For any  $1 \leq p \leq \infty$ , the  $\|\cdot\|_p$  norm on  $L^p$  is complete.

*Proof.* For  $1 \leq p < \infty$ , let  $(f_n)$  be a Cauchy sequence in  $L^p$ .

*Step 1:* Find a subsequence  $(f_{n_k})$  such that  $\|f_{n_k} - f_{n_{k+1}}\|_p \leq 2^{-k}$  for all  $k$ . By Cauchy property, we may find  $n_1$  such that  $\|f_{n_1} - f_n\| \leq 2^{-1}$  for all  $n \geq n_1$ . Also, find a  $n_2 \geq n_1$  such that  $\|f_{n_2} - f_n\| \leq 2^{-2}$  for all  $n \geq n_2$ , etc.

*Step 2:* construct the limit Define

$$A_k := \{x : |f_{n_k}(x) - f_{n_{k+1}}(x)| > 2^{-k/2}\} \quad (53)$$

Then, by Markov's inequality,

$$\mu(A_k) \leq \frac{\int |f_{n_k} - f_{n_{k+1}}|^p d\mu}{(2^{-k/2})^p} \quad (54)$$

$$\leq \frac{2^{-kp}}{(2^{-k/2})^p} = 2^{-kp/2} \quad (55)$$

Thus,  $\sum_{k=1}^{\infty} \mu(A_k) < \infty$ . Define

$$B := \{x : x \in \text{infinitely many } A_k\} = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} A_j \quad (56)$$

By Borel-Cantelli lemma,  $\mu(B) = 0$ . Take any  $x \notin B$ , then for sufficiently large  $k$ ,

$$|f_{n_k}(x) - f_{n_{k+1}}(x)| \leq 2^{-k/2} \quad (57)$$

This shows for all  $x \notin B$ , the constructed  $(f_{n_k}(x))$  is a Cauchy sequence in  $\mathbb{R}$ , therefore, it's convergent.

Define the almost point-wise limit

$$f(x) := \begin{cases} \lim_{k \rightarrow \infty} f_{n_k}(x) & \text{if } x \notin B \\ 0 & \text{if } x \in B \end{cases} \quad (58)$$

*Step 3:* Show  $f_n \rightarrow f$  in  $L^p$ . Note that  $f_{n_k} \rightarrow f$  almost everywhere, so that  $|f|^p \rightarrow |f_{n_k}|^p$ . By Fatou's lemma,

$$\int |f|^p d\mu \leq \liminf_{k \rightarrow \infty} \int |f_{n_k}|^p d\mu \quad (59)$$

But the Cauchy property of  $f_n$  implies that  $\sup_n \|f_n\|_p < \infty$  (find  $n$  such that  $\|f_n - f_m\|_p \leq 1$  for all  $m \geq n$ . Thus,  $\forall m \geq n$ ,  $\|f_m\|_p \leq \|f_n - f_m\|_p + \|f_n\|_p \leq 1 + \|f_n\|_p$ . Therefore,  $\|f\|_p < \infty$ ).

For any  $\varepsilon > 0$ , we can find  $N$  so large that  $\|f_n - f_m\|_p < \varepsilon$  for all  $n, m \geq N$  since  $f_n$  is Cauchy.

By Fatou's lemma, for all  $n \geq N$ ,

$$\int |f_n - f|^p d\mu \leq \liminf_{n \rightarrow \infty} \int |f_n - f|^p d\mu \quad (60)$$

But when  $k$  is so large that  $n_K \geq N$ , we have

$$\int |f_n - f_{n_k}|^p d\mu = \|f_n - f_{n_k}\|_p^p \leq \varepsilon^p \quad (61)$$

Thus, for all  $n \geq N$ ,  $\|f - f_n\|_p \leq \varepsilon$ . ■

*Proof. for  $p = \infty$  case.* Let  $f_n$  be Cauchy in  $L^\infty$ , as before, find a subsequence  $f_{n_k}$  such that

$$\|f_{n_k} - f_{n_{k+1}}\|_\infty \leq 2^{-k} \quad \forall k \quad (62)$$

Then for all  $k$ , there exists a locally  $\mu$ -null set  $N_k$  such that for all  $x \notin N_k$ .

$$|f_{n_k}(x) - f_{n_{k+1}}(x)| \leq 2^{-k} \quad (63)$$

Let  $N = \bigcup_{k=1}^\infty N_k$ , so that  $N$  is locally  $\mu$ -null as well. Then for all  $x \notin N$ ,  $f_{n_k}(x)$  is a Cauchy sequence of real numbers, define  $f(x) = \lim_k f_{n_k}(x)$  outside  $N$  and  $f(x) = 0$  on  $N$ .

Claim:  $f_n \rightarrow f$  in  $L^\infty$ . Note that for all  $x \notin N$ , for all  $k$ ,

$$|f(x) - f_{n_k}(x)| \leq \sum_{j=k}^\infty |f_{n_j}(x) - f_{n_{j+1}}(x)| \leq \sum_{j=k}^\infty 2^{-j} = 2^{-k+1} \quad (64)$$

Thus,  $\|f - f_{n_k}\|_\infty \leq 2^{-k+1}$ .

Take any  $\varepsilon > 0$ , find  $N$  so large that  $\forall m, n \geq N$ ,  $\|f_m - f_n\|_\infty \leq \varepsilon$ . Then find  $k$  so large that  $n_k \geq N$  and  $2^{-k+1} \leq \varepsilon$ . Then for all  $n \geq N$ ,

$$\|f - f_n\|_\infty \leq \|f - f_{n_k}\|_\infty + \|f_{n_k} - f_n\|_\infty \leq 2\varepsilon \quad (65)$$

Taking  $\varepsilon' = \varepsilon/2$  concludes  $f_n \rightarrow f$  in  $L^\infty$ . ■

## 9 Signed and Complex Measures

**Definition 9.1.** Let  $(X, \mathcal{A})$  be a measurable space, let  $\mu : \mathcal{A} \rightarrow [-\infty, \infty]$  be a function. We say that  $\mu$  is a **signed measure** if

1.  $\mu(\emptyset) = 0$ ,
2. and  $\mu$  is countable additive: for all disjoint  $A_1, A_2, \dots \in \mathcal{A}$ ,  $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$ .

From now on, we use **measure** to denote the conventional notion of measure, that is,  $\mu : \mathcal{A} \rightarrow [0, \infty]$  with  $\mu(\emptyset) = 0$  and countable additivity. The term **signed measure** denotes functions  $\mu : \mathcal{A} \rightarrow [-\infty, \infty]$  with above properties.

**Remark 9.1.** Note that the countable additivity does not change if we permute  $A_i$ 's, thus, implies  $\sum_{i=1}^{\infty} \mu(A_i)$  should now change under any rearrangement of the terms. This implies that if  $\mu(\cup_{i=1}^{\infty} A_i)$  is finite,  $\sum_{i=1}^{\infty} |\mu(A_i)| < \infty$ .

**Proposition 9.1.** If  $\mu$  is a signed measure, then  $\mu$  cannot be both  $\infty$  and  $-\infty$ .

*Proof.* **Case 1:** if  $\mu(X) \in \mathbb{R}$ , then for any  $A$ ,  $\mu(X) = \mu(A) + \mu(A^c)$ , both of  $\mu(A)$  and  $\mu(A^c)$  must be finite.

**Case 2:** if  $\mu(X) = \infty$ , then  $\mu(A) + \mu(A^c) = \mu(X) = \infty$ , none of  $\mu(A)$  or  $\mu(A^c)$  can be  $-\infty$ .

**Case 3:** if  $\mu(X) = -\infty$ , then  $\mu(A) + \mu(A^c) = \mu(X) = -\infty$ , none of  $\mu(A)$  or  $\mu(A^c)$  can be  $\infty$ . ■

**Proposition 9.2** (Weak Monotonicity). If  $\mu(A)$  is finite (i.e., in  $\mathbb{R}$ ), then  $\mu(B)$  is finite for any  $B \subseteq A$ ,  $B \in \mathcal{A}$ .

*Proof.*  $\mu(A) = \mu(B) + \mu(A \setminus B) \in \mathbb{R}$ , both  $\mu(B)$  and  $\mu(A \setminus B)$  must be finite. ■

**Definition 9.2.** A signed measure is called **finite** if  $\mu(A)$  is finite for all  $A \in \mathcal{A}$ .

**Example 9.1** (Relationship between integrable function and measure). Let  $(X, \mathcal{A}, \mu)$  be a measure space, let  $f \in L^1$ , define  $\nu(A) = \int_A f d\mu$ , then  $\nu$  is a signed measure.

**Example 9.2** (Construction of signed measure). If  $\nu_1$  and  $\nu_2$  are measures and at least one of them is finite, then  $\nu_1 - \nu_2$  is a signed measure.

### 9.1 Hahn Decomposition Theorem

Let  $(X, \mathcal{A})$  be a measurable space and let  $\mu$  be a signed measure on  $(X, \mathcal{A})$ .

**Definition 9.3.** A set  $A \in \mathcal{A}$  is called a **positive set for  $\mu$**  if  $\mu(B) \geq 0$  for all  $B \subseteq A$ ,  $B \in \mathcal{A}$ . Similarly, a set  $A \in \mathcal{A}$  is called a **negative set for  $\mu$**  if  $\mu(B) \leq 0$  for all  $B \subseteq A$ ,  $B \in \mathcal{A}$ .

**Lemma 9.1.** If  $A \in \mathcal{A}$  satisfies  $-\infty < \mu(A) < 0$ , then there exists a negative set  $B \subseteq A$  such that  $\mu(B) \leq \mu(A)$ .

*Proof.* Let  $\delta_1 = \sup\{\mu(E) : E \in \mathcal{A}, E \subseteq A\}$ , note that  $\delta_1 \geq 0$  since  $\mu(\emptyset) = 0$ .

By the definition of  $\delta_1$  we can find  $A_1 \subseteq A$  such that  $\mu(A_1) \geq \delta_1/2$  if  $\delta_1 < \infty$ , or  $\mu(A_1) \geq 1$  if  $\delta_1 = \infty$ . Thus,  $\mu(A_1) \geq \min\{\delta_1/2, 1\}$ .

Let  $\delta_2 = \sup\{\mu(E) : E \in \mathcal{A}, E \subseteq A \setminus A_1\}$ , similarly, we can choose  $A_2 \subseteq A \setminus A_1$  and  $A_2 \in \mathcal{A}$  such that  $\mu(A_2) \geq \min\{\delta_2/2, 1\}$ .

Similarly, choose  $A_n \in \mathcal{A}$ ,  $A_n \subseteq A \setminus (A_1 \cup \dots \cup A_{n-1})$ , such that  $\mu(A_n) \geq \min\{\delta_n/2, 1\}$ . Then by definition,  $A_1, A_2, \dots$  are disjoint, they are all contained in  $A$ .

Let  $B = A \setminus (\bigcup_{i=1}^{\infty} A_i)$ .

Claim: this  $B$  is a negative set such that  $\mu(B) \leq \mu(A)$ .

Note that  $\mu(A) \in \mathbb{R} \implies \mu(B) \in \mathbb{R}$ . Thus,  $\mu(\bigcup_{i=1}^{\infty} A_i) = \mu(A) - \mu(B) \in \mathbb{R}$ .

But  $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$  since  $A_i$ 's are disjoint. Therefore,  $\mu(A_i) \rightarrow 0$  as  $i \rightarrow \infty$ .

However,  $\mu(A_i) \geq \min\{\delta_i/2, 1\} \geq 0$ . It must be  $\delta_i \rightarrow 0$  as  $i \rightarrow \infty$ .

Take any  $E \subseteq B$  such that  $E \in \mathcal{A}$ . Then  $E \subseteq B \subseteq A \setminus (A_1 \cup \dots \cup A_{n-1})$  for all  $n \in \mathbb{N}$ . So by definition of  $\delta_n$ , we have  $\mu(E) \leq \delta_n$ , thus  $\mu(E) \leq 0$  as we take  $n \rightarrow \infty$ . Hence  $B$  is a negative set.

Finally, since  $\mu(A_i) \rightarrow 0$ ,  $\mu(B) = \mu(A) - \sum_{i=1}^{\infty} \mu(A_i) \leq \mu(A)$ .

■

**Theorem 9.1** (Hahn Decomposition Theorem). Let  $(X, \mathcal{A})$  be a measurable space and  $\mu$  a signed measure on  $(X, \mathcal{A})$ . Then, there exists disjoint  $P \cup N$  in  $\mathcal{A}$  such that  $X = P \cup N$  such that  $P$  is a positive set for  $\mu$  and  $N$  is a negative set for  $\mu$ .

*Proof.* Since  $\mu$  is a signed measure, we know that it cannot take value at both  $\infty$  and  $-\infty$ . WLOG, suppose  $\mu$  never takes value  $-\infty$ . Let

$$L = \inf\{\mu(A) : A \in \mathcal{A} \text{ s.t. } A \text{ is negative}\} \quad (1)$$

Then there exists a sequence of negative sets  $A_n$  such that  $\mu(A_n) \rightarrow L$ . Define  $B = \bigcup_{n=1}^{\infty} A_n$ . For sure,  $B \in \mathcal{A}$ .

**Claim:**  $B$  is a negative set.

Take and  $E \subseteq B$  such that  $E \in \mathcal{A}$ , then

$$E = E \cap B = \bigcup_{i=1}^{\infty} E \cap A_i = \bigcup_{i=1}^{\infty} E \cap (A_i \setminus (A_1 \cup \dots \cup A_{i-1})) \quad (2)$$

where the last step holds because we only consider the net incremental at each step. Moreover,  $\{E \cap (A_i \setminus (A_1 \cup \dots \cup A_{i-1}))\}_i$  are disjoint.

Thus,

$$\mu(E) = \sum_{i=1}^{\infty} \underbrace{\mu(E \cap (A_i \setminus (A_1 \cup \dots \cup A_{i-1})))}_{\subseteq A_i} \quad (3)$$

Since  $A_i$ 's are all negative sets, we must have  $\mu(E) \leq 0$  and  $B$  is a negative set.

**Claim:**  $\mu(B) = L$ .

Since  $A_n \subseteq B$ ,

$$\mu(B) = \mu(A_n) + \mu(B \setminus A_n) \quad (4)$$

But  $B$  is a negative set, so  $\mu(B \setminus A_n) \leq 0$ . Thus,

$$\mu(B) \leq \mu(A_n) \quad \forall n \in \mathbb{N} \quad (5)$$

Thus,  $\mu(B) \leq \lim_n \mu(A_n) = L$ . But  $B$  itself is a negative set, and  $L$  is the infimum, so  $L \leq \mu(B)$ .

In particular, we've shown that  $L > -\infty$  since  $\mu$  never takes value at  $-\infty$ .

Let  $N = B$  and  $P = N^c$ . Since  $B \in \mathcal{A}$ , both  $N, P \in \mathcal{A}$ .

**Claim:**  $P$  is a positive set.

Suppose not, then  $\exists A \subseteq P$  such that  $A \in \mathcal{A}$  and  $-\infty < \mu(A) < 0$ .

By the lemma, there exists a negative set  $D \subseteq A$  and  $\mu(D) \leq \mu(A) < 0$ . Note that  $D \subseteq A \subseteq P$ , but then  $N \cup D$  is a negative set as a union of negative sets. Then,

$$\mu(N \cup D) = \mu(N) + \mu(D) = L + \mu(D) < L \quad (6)$$

which leads to a contradiction.

Consequently, this  $X = N \cup P$  is a Hahn decomposition. ■

**Theorem 9.2** (Jordan Decomposition Theorem). Every signed measure is the difference of two

measures, at least one of which is finite.

$$\mu = \mu^+ - \mu^- \quad (7)$$

*Proof.* Let  $\mu$  be a signed measure, let  $(N, P)$  be a Hahn decomposition of  $X$ .

For every  $A \in \mathcal{A}$ , define

$$\mu^+(A) = \mu(A \cap P) \quad (8)$$

$$\mu^-(A) = -\mu(A \cap N) \quad (9)$$

Since  $P$  is a positive set,  $\mu^+(A) \geq 0$ , similarly, since  $N$  is negative,  $\mu^-(A) \geq 0$  as well.

Let  $A_1, A_2, \dots$  be disjoint sets in  $\mathcal{A}$ , then

$$\mu^+(\cup_i A_i) = \mu(P \cap (\cup_i A_i)) \quad (10)$$

$$= \mu(\cup_i (P \cap A_i)) \quad (11)$$

$$= \sum_i \mu(P \cap A_i) \quad (12)$$

$$= \sum_i \mu^+(A_i) \quad (13)$$

So  $\mu^+$  is a measure. Similarly,  $\mu^-$  is a measure as well.

$$\mu^+(A) - \mu^-(A) = \mu(A \cap P) + \mu(A \cap N) = \mu(A) \quad (14)$$

Therefore,  $\mu = \mu^+ - \mu^-$ . Lastly, note that  $\mu(X) = \mu(P) + \mu(N) = \mu^+(P) - \mu^-(N)$ , we need at least one of them to be finite in order to avoid subtracting infinity from infinity. ■

**Proposition 9.3.** Let  $(\mu^+, \mu^-)$  be the decomposition of a signed measure from Hahn decomposition  $(P, N)$ , that is,  $\mu^+(A) = \mu(A \cap P)$  and  $\mu^-(A) = -\mu(A \cap N)$  for any  $A \in \mathcal{A}$ . Then,

$$\mu^+(A) = \sup\{\mu(B) : B \subseteq A, B \in \mathcal{A}\} \quad (15)$$

$$\mu^-(A) = \sup\{-\mu(B) : B \subseteq A, B \in \mathcal{A}\} \quad (16)$$

*Proof.* Take any  $A \in \mathcal{A}$ , take any  $B \subseteq A$  such that  $B \in \mathcal{A}$ . Then

$$\mu(B) = \mu^+(B) - \mu^-(B) \quad (17)$$

$$\leq \mu^+(B) \because \mu^-(B) \geq 0 \quad (18)$$

$$\leq \mu^+(A) \because \mu^+ \text{ is a measure} \quad (19)$$

Therefore,  $\mu^+(A) \geq \sup\{\mu(B) : B \subseteq A, B \in \mathcal{A}\}$ .

On the other hand,  $\mu^+(A) = \mu(A \cap P)$  by definition, take  $B = A \cap P \subseteq A$ , which satisfies  $A \cap P \in \mathcal{A}$ . Then  $\mu^+(A) \leq \sup\{\mu(B) : B \subseteq A, B \in \mathcal{A}\}$ .

The similar logic works for  $\mu^-$ . ■

**Definition 9.4.** The pair of  $(\mu^+, \mu^-)$  defined above is called the **Jordan decomposition** of the signed measure  $\mu$ , where  $\mu^+$  and  $\mu^-$  are called the **positive and negative parts of  $\mu$** .

**Definition 9.5.** The **variation** of  $\mu$  is defined to be the measure  $|\mu| = \mu^+ + \mu^-$ . The **total variation** of  $\mu$  is the number  $\|\mu\| = |\mu|(X)$ .

## 9.2 Complex Measures

**Definition 9.6.** Let  $(X, \mathcal{A})$  be a measurable space,  $\mu : \mathcal{A} \rightarrow \mathbb{C}$  is called a **complex measure** if for all disjoint  $A_1, A_2, \dots \in \mathcal{A}$ ,  $\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$  and  $\mu(\emptyset) = 0$ . In particular, this implies the sum is absolutely converged.

Any complex measure  $\mu$  can be written uniquely as

$$\mu = \mu' + i\mu'' \quad (20)$$

where

$$\mu'(A) = \Re(\mu(A)) \quad (21)$$

$$\mu''(A) = \Im(\mu(A)) \quad (22)$$

Let  $\mu' = \mu_1 - \mu_2$  and  $\mu'' = \mu_3 - \mu_4$  be Jordan compositions of  $\mu'$  and  $\mu''$  respectively. Then

$$\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4 \quad (23)$$

is called the **Jordan decomposition** of complex measure  $\mu$ .

**Definition 9.7.** The **variation** of a complex measure  $\mu$  is defined as

$$|\mu|(A) := \sup \left\{ \sum_{i=1}^n |\mu(A_i)| : A_1, \dots, A_n \in \mathcal{A} \text{ disjoint s.t. } \bigcup_{i=1}^n A_i = A \right\} \quad (24)$$

Note that the supremum is taken over all *finite partitions of  $A$* . It is easy to check that if  $\mu$  is a finite signed measure, this definition of variation is the same as the previous one.

**Lemma 9.2.** Suppose  $\mu : \mathcal{A} \rightarrow [0, \infty]$  is a function such that (i)  $\mu(\emptyset) = 0$  and (ii) is finite additivity (that is,  $\mu(A \cup B) = \mu(A) + \mu(B)$  for all disjoint  $A$  and  $B$ ). Moreover, if  $\lim_{n \rightarrow \infty} \mu(A_n) = 0$  for all  $A_n \searrow \emptyset$ , then  $\mu$  is a measure.

*Proof.* It suffices to check the countable additivity of  $\mu$ , let  $B_1, B_2, \dots$  be a disjoint sequence of sets in  $\mathcal{A}$ .

Let  $B = \bigcup_i B_i$  and define  $A_n := B \setminus \bigcup_{i=1}^{n-1} B_i$ . Easy to check  $A_n \searrow \emptyset$ . Therefore, by finite additivity of  $\mu$ :  $\mu(A_n) = \mu(B) - \sum_{i=1}^{n-1} \mu(B_i) \rightarrow 0$ . Taking  $n \rightarrow \infty$  implies  $\mu(B) = \sum_{i=1}^{\infty} \mu(B_i)$ . ■

**Proposition 9.4.** Let  $\mu$  be a complex measure, then  $|\mu|$  is a measure.

*Proof.* Obviously,  $|\mu|(\emptyset) = 0$ .



Take any disjoint  $A, B \in \mathcal{A}$ . Now show the finite additivity of  $|\mu|$ : let  $C_1, \dots, C_n$  be a measurable disjoint partition of  $A \cup B$ , so  $(C_i \cap A)$  and  $(C_i \cap B)$  are partitions of  $A$  and  $B$  respectively.

$$|\mu|(A) + |\mu|(B) \geq \sum |\mu(C_i \cap A)| + \sum |\mu(C_i \cap B)| \quad (25)$$

$$\geq \sum |\mu(C_i \cap A) + \mu(C_i \cap B)| \quad (26)$$

$$= \sum |\mu(C_i)| \because C_i \subseteq A \cup B \quad (27)$$

$$\geq |\mu|(A \cup B) \quad (28)$$

Conversely, let  $C_1, \dots, C_n$  be a partition of  $A$  and  $D_1, \dots, D_m$  be a partition of  $B$ , then  $C_1, \dots, C_n, D_1, \dots, D_m$  is a partition of  $A \cup B$ .

$$|\mu|(A \cup B) \geq \sum_{i=1}^n |\mu(C_i)| + \sum_{i=1}^m |\mu(D_i)| \quad (29)$$

Taking supremum of partitions  $(C_i)$  for  $A$  and  $(D_i)$  for  $B$ ,

$$|\mu|(A \cup B) \geq |\mu|(A) + |\mu|(B) \quad (30)$$

Therefore,  $|\mu|$  is finitely additive.

Now take a  $A_n \searrow \emptyset$  in  $\mathcal{A}$ , using the Jordan decomposition:  $\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$  where  $\mu_i$  are measures. By triangle inequality in  $\mathbb{C}$ ,

$$|\mu(A)| \leq \sum_{i=1}^4 \mu_i(A) \quad (31)$$

then for all measurable partitions  $A_1, \dots, A_n$  of  $A$ ,

$$\sum_{j=1}^n |\mu(A_j)| \leq \sum_{i=1}^4 \sum_{j=1}^n \mu_i(A_j) = \sum_{i=1}^4 \mu_i(A) \quad (32)$$

Taking supremum of all such partitions,

$$|\mu|(A) \leq \sum_{i=1}^4 \mu_i(A) \quad (33)$$

Since  $A_n \searrow \emptyset$  implies  $\mu_i(A_n) \rightarrow 0$  as  $\mu_i$ 's are finite measures (there is no  $\infty$  in  $\mathbb{C}$ ),  $|\mu|(A_n) \rightarrow 0$ . By Previous lemma,  $|\mu|$  is a measure. ■

**Proposition 9.5** (Completeness of Total Variation). The total variation is a norm on the space of finite signed/complex measures.

*Proof.* Obviously,  $\|\mu\|$  is a norm. Now show the completeness.

Let  $\{\mu_n\}$  be a Cauchy (in total variation) sequence of measures, for all  $A \in \mathcal{A}$ ,  $|\mu(A)| \leq |\mu|(A)$  since  $A$  is a trivial partition of  $A$ .

For any  $m, n \in \mathbb{N}$ ,  $A \in \mathcal{A}$ ,  $\mu_m - \mu_n$  is a signed measure,

$$|\mu_m(A) - \mu_n(A)| \leq |\mu_m - \mu_n|(A) \quad (34)$$

$$\leq \|\mu_m - \mu_n\| \quad (35)$$

Therefore,  $\{\mu_n(A)\}$  is a Cauchy sequence in  $\mathbb{R}$  for all  $A \in \mathcal{A}$ . Define  $\mu$  as the "set-wise" limit of  $\mu_n$ :

$$\mu(A) := \lim_{n \rightarrow \infty} \mu_n(A) \quad (36)$$

Now show  $\mu$  is a measure: observe that  $\mu_n \rightarrow \mu(A)$  uniformly over all  $A \in \mathcal{A}$  by Equation (35). The finite additivity of  $\mu$  follows its definition.

Fix arbitrary  $A_n \searrow \emptyset$ , show that  $\mu(A_n) \rightarrow 0$ . Take any  $\varepsilon > 0$ , find  $N$  so large that  $|\mu_N(A) - \mu(A)| < \varepsilon$  for all  $A$  by uniform convergence.

Find  $j_0$  so large such that for all  $j \geq j_0$ ,  $|\mu_N(A_j)| < \varepsilon/2$ . For all  $j \geq j_0$ ,

$$|\mu(A_j)| \leq |\mu(A_j) - \mu_N(A_j)| + |\mu_N(A_j)| < \varepsilon \quad (37)$$

Lastly, we show  $\|\mu_n - \mu\| \rightarrow 0$ . Take any partition  $A_1, \dots, A_k$  of  $X$ , take any  $\varepsilon > 0$ , the Cauchy property of  $\{\mu_n\}$  provides a  $N$  so large that for all  $m, n \geq N$ ,  $\|\mu_m - \mu_n\| < \varepsilon$ .

$$\sum_{j=1}^k |\mu_m(A_j) - \mu_n(A_j)| \leq \|\mu_m - \mu_n\| < \varepsilon \quad (38)$$

Take  $m \rightarrow \infty$ ,

$$\sum_{j=1}^k |\mu(A_j) - \mu_n(A_j)| \leq \varepsilon \quad (39)$$

Since above inequality holds for all partitions of  $X$ ,  $\|\mu - \mu_n\| < \varepsilon$ . ■

### 9.3 Integration w.r.t. Signed and Complex Measures

**Definition 9.8.** Let  $\mu = \mu^+ - \mu^-$  be a signed measure and its corresponding Jordan decomposition, define

$$\int f d\mu = \int f d(\mu^+ - \mu^-) = \int f d\mu^+ - \int f d\mu^- \quad (40)$$

Easy to check that  $f \mapsto \int f d\mu$  and  $\mu \mapsto \int f d\mu$  are linear maps.

When  $\mu$  is a complex measure:  $\mu = \mu' + i\mu''$ , define

$$\int f d\mu = \int f d\mu' + i \int f d\mu'' \quad (41)$$

## 10 Radon-Nikodym Theorem

**Definition 10.1.** Let  $(X, \mathcal{A})$  be a measurable space, let  $\mu, \nu$  be two measures on this space,  $\nu$  is **absolutely continuous** w.r.t.  $\mu$  if for every  $A \in \mathcal{A}$ :

$$\mu(A) = 0 \implies \nu(A) = 0 \quad (1)$$

denoted as  $\nu \ll \mu$ .

**Theorem 10.1** (Radon-Nikodym). Let  $(X, \mathcal{A})$  be a measurable space, let  $\mu, \nu$  be two  $\sigma$ -finite measures. Suppose  $\nu$  is absolutely continuous w.r.t.  $\mu$ , then there exists a measurable map  $g : X \rightarrow [0, \infty)$  such that

$$\nu(A) = \int_A g \, d\mu \quad (2)$$

for every  $A \in \mathcal{A}$ .

The map  $g$  is defined as the **Radon-Nikodym derivative**, denoted as  $\frac{d\nu}{d\mu}$ ,  $g$  is unique up to  $\mu$ -a.e. equivalence.

**Interpretations** Let  $\chi_A$  denote the indicator function of set  $A$ , recall that  $\int_A f \, d\mu \equiv \int f \chi_A \, d\mu$ . Then,  $\nu(A) = \int_A 1 \, d\nu = \int \chi_A \, d\nu = \int g \chi_A \, d\mu$  for all  $A$ . Moreover, for any integrable  $f$ ,

$$\int f \, d\nu = \int f g \, d\mu \quad (3)$$

This allows us to denote  $g$  as  $\frac{d\nu}{d\mu}$ .

**Example 10.1.** Suppose  $(X, \mathcal{A})$  is a metric space (take  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  here), suppose  $g$  is continuous w.r.t. the metric, let  $A = B(x, \varepsilon)$  be the  $\varepsilon$ -open ball around  $x \in X$ , then for sufficiently small  $\varepsilon$ :

$$\nu(A) = \nu(B(x, \varepsilon)) \quad (4)$$

$$\int_A g \, d\mu \approx g(x) \int_A d\mu = g(x) \mu(B(x, \varepsilon)) \quad (5)$$

That is,

$$\frac{d\nu}{d\mu} = g(x) \approx \frac{\nu(B(x, \varepsilon))}{\mu(B(x, \varepsilon))} \quad (6)$$

Actually,

$$g(x) = \lim_{\varepsilon \rightarrow 0} \frac{\nu(B(x, \varepsilon))}{\mu(B(x, \varepsilon))} \quad (7)$$

Therefore, the Radon-Nikodym derivative  $\frac{d\nu}{d\mu}$  captures the relative growth rate of  $\nu$  to  $\mu$  when we initially apply them on a small ball and expand the radius of this ball.

**Lemma 10.1.** Let  $(X, \mathcal{A})$  be a measurable space, let  $\nu$  be a measure on it, let  $\nu$  be a finite measure. Then,  $\nu \ll \mu$  if and only if

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ s.t. } \mu(A) < \delta \implies \nu(A) < \varepsilon \quad \forall A \in \mathcal{A} \quad (8)$$

Recall the definition of uniform continuity and  $\frac{df(x)}{dx}$ .

*Proof.* ( $\Leftarrow$ ) Suppose  $\mu(A) = 0, \nu(A) < \varepsilon$  for all  $\varepsilon > 0$ , it must be  $\nu(A) = 0$ .

( $\Rightarrow$ ) Suppose  $\nu \ll \mu$ , suppose the condition fails,  $\exists \varepsilon > 0$  such that  $\forall \delta > 0, \exists A$  with  $\mu(A) < \delta$  but  $\nu(A) \geq \varepsilon$ .

We can find a sequence  $A_1, A_2, \dots$  such that  $\mu(A_j) < \delta_j = 2^{-j}$  for all  $j$  and  $\nu(A_j) \geq \varepsilon$ . It follows  $\sum \mu(A_j) < \infty$ . By Borel-Cantelli lemma,

$$\mu \left( \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} A_k \right) = 0 \quad (9)$$

Define  $B_j = \bigcup_{k=j}^{\infty} A_k$  and  $B = \bigcap_{j=1}^{\infty} B_j$ . Since  $B_j \searrow B$  and  $\nu$  is a finite measure,  $\nu(B) = \lim_j \nu(B_j)$ . But for any  $j, \nu(B_j) \geq \nu(A_j) \geq \varepsilon$ . Therefore,  $\nu(B) \geq \varepsilon$ , which contradicts  $\nu \ll \mu$ . ■

*Proof of Radon-Nikodym Theorem.* Let  $\nu, \mu$  be finite measures, let

$$\mathcal{F} := \left\{ f : X \rightarrow [0, \infty] : f \text{ measurable and } \int_A f \, d\mu \leq \nu(A) \quad \forall A \in \mathcal{A} \right\} \quad (10)$$

We are choosing the largest  $g \in \mathcal{F}$  as  $\frac{d\nu}{d\mu}$ .

Claim:  $f, g \in \mathcal{F} \implies f \vee g \equiv \max\{f, g\} \in \mathcal{F}$ .

*Proof.* Let  $B := \{x : f(x) \geq g(x)\}$ , for any  $A \in \mathcal{A}$ ,

$$\int_A f \vee g \, d\mu = \int_{A \cap B} f \vee g \, d\mu + \int_{A \cap B^c} f \vee g \, d\mu \quad (11)$$

$$= \int_{A \cap B} f \, d\mu + \int_{A \cap B^c} g \, d\mu \leq \nu(A \cap B) + \nu(A \cap B^c) = \nu(A) \quad (12)$$

■

Let  $(f_n) \in \mathcal{F}$  be a sequence such that

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu = \sup \left\{ \int f \, d\mu : f \in \mathcal{F} \right\} \quad (13)$$

For every  $n \in \mathbb{N}$ , take  $g_n(x) = \max_{j \leq n} f_j(x)$ ,  $g_n \in \mathcal{F}$  by previous claim. Moreover,  $g_n(x) \uparrow$  for all  $x \in X$ .

$$\int f_n \, d\mu \leq \int g_n \, d\mu \leq \sup \left\{ \int f \, d\mu : f \in \mathcal{F} \right\} \quad (14)$$

By squeeze theorem,  $\lim_{n \rightarrow \infty} \int g_n \, d\mu = \sup\{\int f \, d\mu : f \in \mathcal{F}\}$ .

Define  $g(x) = \lim_{n \rightarrow \infty} g_n(x)$ , which always exists but is potentially infinity. By MCT,

$$\int g \, d\mu = \lim_{n \rightarrow \infty} \int g_n \, d\mu = \sup\{\int f \, d\mu : f \in \mathcal{F}\} \quad (15)$$

Note that  $\forall A \in \mathcal{A}$ ,

$$\int_A g \, d\mu = \lim_{n \rightarrow \infty} \int_A g_n \, d\mu \leq \nu(A) \quad (16)$$

So  $g \in \mathcal{F}$  and attains the supremum, in terms of integral, over  $\mathcal{F}$ .

Claim:  $\forall A \in \mathcal{A}$ ,  $\int_A g \, d\mu = \nu(A)$ .

*Proof.* Define  $\nu_0(A) = \nu(A) - \int_A g \, d\mu$ . Since  $\nu$  is a measure and  $A \mapsto \int_A g \, d\mu$  is also a finite measure. Therefore,  $\nu_0$  is a finite signed measure. Moreover, since  $g \in \mathcal{F}$ ,  $\nu_0(A) \geq 0$  for all  $A \in \mathcal{A}$ .

Suppose, for contradiction,  $\nu_0(A) > 0$  for some  $A \in \mathcal{A}$ . It must be  $\nu_0(X) > 0$ . But  $\mu(X) < \infty$ , there exists  $\varepsilon > 0$  such that  $\nu_0(X) > \varepsilon\mu(X)$ . Note that  $\nu_0 - \varepsilon\mu$  is a finite signed measure, let  $(P, N)$  be the Hahn decomposition of  $\nu_0 - \varepsilon\mu$ . Then for any  $A \in \mathcal{A}$ ,

$$\nu(A) = \int_A g \, d\mu + \nu_0(A) \quad (17)$$

$$\geq \int_A g \, d\mu + \nu_0(A \cap P) \quad (18)$$

$$= \int_A g \, d\mu + \underbrace{\nu_0(A \cap P) - \varepsilon\mu(A \cap P)}_{\geq 0} + \varepsilon\mu(A \cap P) \quad (19)$$

$$\geq \int_A g \, d\mu + \varepsilon\mu(A \cap P) \quad (20)$$

$$= \int_A g + \varepsilon\chi_{A \cap P} \, d\mu \quad (21)$$

Therefore,  $g + \varepsilon\chi_{A \cap P} \in \mathcal{F}$ . Take  $A = X$ :

$$\int g + \varepsilon\chi_{A \cap P} \, d\mu = \int g \, d\mu + \varepsilon\mu(P \cap A) \geq \int g \, d\mu \quad (22)$$

Suppose, for contradiction,  $\mu(P) \leq 0$ , it must be  $\mu(P) = 0$ , by absolute continuity,  $\nu \ll \mu$ ,  $\nu(P) = 0$  as well. Then, since  $\int_P g \, d\mu$  is bounded on a measure zero set, it must be zero,

$$\nu_0(P) = \nu(P) - \int_P g \, d\mu = 0 \quad (23)$$

Thus

$$(\nu_0 - \varepsilon\mu)(P) = 0 \quad (24)$$

$$\implies (\nu_0 - \varepsilon\mu)(X) = (\nu_0 - \varepsilon\mu)(P) + (\nu_0 - \varepsilon\mu)(N) \leq 0 \quad (25)$$

Contradicts  $\nu_0(X) \geq \varepsilon\mu(X)$ , therefore,  $\mu(P) > 0$ .

This leads to a contradiction since  $g + \varepsilon\chi_{A \cap P}$  has strictly larger integral than  $g$ . Therefore,  $\nu_0 = 0$ . ■

Suppose  $\mu$  and  $\nu$  are  $\sigma$ -finite. Let  $B_1, B_2, \dots \in \mathcal{A}$  be a partition of  $X$  such that  $\mu(B_n), \nu(B_n)$  are finite. Moreover, define  $\mu_n(A) := \mu(A \cap B_n)$  and  $\nu_n(A) := \nu(A \cap B_n)$ , both of  $\mu_n$  and  $\nu_n$  are finite on  $X$  (in particular, on  $B_n$ ) and  $\nu_n \ll \mu_n$ .

For every  $n \in \mathbb{N}$ , there exists measurable  $g_n : X \rightarrow [0, \infty]$  such that

$$\nu_n(A) = \int_A g_n \, d\mu \quad (26)$$

Therefore,

$$\nu(A \cap B_n) = \int g_n \chi_{A \cap B_n} d\mu \quad (27)$$

$$= \int g_n \chi_{B_n} \chi_A d\mu \quad (28)$$

$$= \int_A g_n \chi_{B_n} d\mu \quad (29)$$

Let  $g = \sum_{n=1}^{\infty} g_n \chi_{B_n}$ , then

$$\nu(A) = \sum_{n=1}^{\infty} \nu(A \cap B_n) \quad (30)$$

$$= \sum_{n=1}^{\infty} \int g_n \chi_{B_n} \chi_A d\mu \quad (31)$$

$$= \sum_{n=1}^{\infty} \chi_A \int g_n \chi_{B_n} d\mu \quad (32)$$

$$= \int \chi_A \sum_{n=1}^{\infty} g_n \chi_{B_n} d\mu \quad (33)$$

$$= \int_A g d\mu \quad (34)$$

$$(35)$$

Since  $g_n < \infty$  everywhere for all  $n$ , so is  $g$ . ■

**Remark 10.1** (Uniqueness of Radon-Nikodym Derivative). Let  $g$  and  $h$  be two Radon-Nikodym derivatives of  $\nu$  w.r.t.  $\mu$ .

**Case 1:** suppose  $\nu(X) < \infty$ , then for all  $A \in \mathcal{A}$ , by definitoin,

$$\int_A g d\mu = \nu(A) = \int_A h d\mu \quad (36)$$

Let  $B := \{x, g(x) > h(x)\}$ ,  $(g - h)\chi_A \geq 0$  and  $(g - h)\chi_A = 0$  a.e. on  $A$ . Similarly,  $(h - g)\chi_{A^c} \geq 0$  and  $(h - g)\chi_{A^c} = 0$  a.e. on  $A^c$ . Add them together,  $g - h = 0$  a.e. on  $X$ .

**Case 2:** suppose  $\nu$  is  $\sigma$ -finite, let  $B_1, B_2, \dots$  be disjoint measurable sets such that  $\nu(B_n) < \infty$  and  $\cup_n B_n = X$ . Since  $g = h$  a.e. on every  $B_n$  as shown before,  $g = h$  a.e. on  $X$ .

**Theorem 10.2** (Radon-Nikodym Theorem for Finite Signed and Complex Measures). Let  $(X, \mathcal{A})$  be a measurable space, let  $\mu$  be a  $\sigma$ -finite measure on  $X$ . Let  $\nu$  be a finite signed or complex measure on  $X$ . Suppose that  $|\nu| \ll \mu$  (in this case, we simply say  $\nu \ll \mu$ ). Then there exists  $g \in \mathcal{L}^1(X, \mathcal{A}, \mu, \mathbb{R})$  or  $\mathcal{L}^1(X, \mathcal{A}, \mu, \mathbb{C})$  such that

$$\nu(A) = \int_A g d\mu \quad \forall A \in \mathcal{A} \quad (37)$$

Moreover,  $g$  is unique up to  $\mu$ -a.e. equivalence.

*Proof.*  $\nu = \nu_1 - \nu_2 + i\nu_3 - i\nu_4$  where  $\nu_1, \nu_2, \nu_3, \nu_4$  are finite.  $|\nu| \ll \mu \implies \nu_i \ll \mu$  for  $i = 1, 2, 3, 4$ .

Let  $g_i = \frac{d\nu_i}{d\mu}$ , then  $g = g_1 - g_2 + ig_3 - ig_4$  is the Radon-Nikodym derivative of  $\nu$  w.r.t.  $\mu$ . ■

Check



## 11 Lebesgue Decomposition Theorem

**Definition 11.1.** Let  $(X, \mathcal{A})$  be a measurable space, let  $\mu$  be a measure on  $(X, \mathcal{A})$ , then  $\mu$  is **concentrated** on a set  $E \in \mathcal{A}$  if  $\mu(E^c) = 0$ .

We say that a signed or complex measure  $\mu$  is concentrated on  $E$  if the measure  $|\mu|$  is concentrated on  $E$ .

**Definition 11.2.** Two measures / signed measures / complex measures  $\mu$  and  $\nu$  on measurable space  $(X, \mathcal{A})$  are **mutually singular** if  $\exists E \in \mathcal{A}$  such that  $\mu$  is concentrated on  $E$  and  $\nu$  is concentrated on  $E^c$ .

$$\mu \perp \nu \tag{1}$$

**Example 11.1.** Any measure on  $\mathbb{R}$  that is concentrated on  $\mathbb{Z}$  is mutually singular to the Lebesgue measure, which is concentrated on  $\mathbb{Z}^c$ .

**Theorem 11.1** (Lebesgue Decomposition). Let  $(X, \mathcal{A})$  be a measurable space, let  $\mu$  be a measure (the reference measure) on it. Let  $\nu$  be a finite signed, complex measure, or  $\sigma$ -finite measure on  $(X, \mathcal{A})$ , then there is a unique decomposition

$$\nu = \nu_a + \nu_s \tag{2}$$

such that

$$\nu_a \ll \mu \tag{3}$$

$$\nu_s \perp \mu \tag{4}$$

*Proof.* **Case 1: suppose  $\nu$  is a finite measure.** Define  $\mathcal{N}$  to be the collection of  $\mu$ -negligible sets:

$$\mathcal{N} := \{B \in \mathcal{A} : \mu(B) = 0\} \tag{5}$$

Let

$$S := \sup\{\nu(B) : B \in \mathcal{N}\} < \infty \text{ since } \nu \text{ is finite.} \tag{6}$$

Then there exists a sequence of sets  $B_n \in \mathcal{N}$  such that  $S = \lim_{n \rightarrow \infty} \nu(B_n)$ , define

$$N = \bigcup_{n=1}^{\infty} B_n \tag{7}$$

Easy to verify that  $\mu(N) \leq \sum_{n=1}^{\infty} \mu(B_n) = 0$ , so  $N \in \mathcal{N}$ . Obviously,  $\nu(N) \leq S$  since  $N \in \mathcal{N}$ . Moreover, since  $\nu(N) \geq \nu(B_n)$  for every  $n \in \mathbb{N}$ ,  $\nu(N) \geq \lim_n \nu(B_n) = S$ . Thus,  $\nu(N) = S$ , so that  $N$  is the  $\nu$ -maximal set in  $\mathcal{N}$ .

For every  $A \in \mathcal{A}$ , define

$$\nu_a(A) = \nu(A \cap N^c) \quad (8)$$

$$\nu_s(A) = \nu(A \cap N) \quad (9)$$

So that  $\nu = \nu_a + \nu_s$ .

**Claim:**  $\nu_s \perp \mu$ .

Easy to verify that  $\mu(N) = 0$  and  $\nu_s(N^c) = \nu(N^c \cap N) = 0$ .

**Claim:**  $\nu_a \ll \mu$ .

Take any  $B \in \mathcal{A}$  such that  $\mu(B) > 0$ . Suppose, for contradiction,  $\nu_a(B) \neq 0$ , that is,  $\nu(B \cap N^c) \neq 0$ . Since we assumed  $\nu$  is a finite measure (not signed),  $\nu(B \cap N^c) > 0$ . Thus,

$$\nu(N \cup B) = \nu(N) + \nu(B \cap N^c) > \nu(N) = S \quad (10)$$

but  $N \cup B \in \mathcal{N}$ , this leads to a contradiction.

**Case 2: suppose  $\nu$  is a finite signed or complex measure**, we can find  $N$  as above for  $|\nu|$  and define  $\nu_a(A) = \nu(A \cap N^c)$  and  $\nu_s(A) = \nu(A \cap N)$ .

**Case 3: if  $\nu$  is a  $\sigma$ -finite measure**, we can firstly express  $X$  as a disjoint union  $D_1, D_2, \dots$  with finite  $\nu$  measure, and then find  $N_i \subseteq D_i$  as the  $\nu$ -maximal element among all  $\mu$ -zero subsets of  $D_i$ . Lastly, define  $N = \bigcup_{i=1}^{\infty} N_i$  and follow the construction before.

**Uniqueness:** suppose

$$\nu = \nu_a + \nu_s = \nu'_a + \nu'_s \quad (11)$$

Assume  $\nu$  is a finite / finite signed / complex measure, then

$$\nu_a - \nu'_a = \nu'_s - \nu_s \quad (12)$$

The left hand side is absolutely continuous and the right hand side is singular to  $\mu$  by the following lemma, hence, they must be both zero.

**Lemma 11.1.** The notion of absolute continuity and singularity are closed under linear combinations.

*Proof.* **TODO** ■

**Lemma 11.2.** If a measure is both absolutely continuous and singular with respect to  $\mu$ , then it must be zero.

*Proof.* **TODO** ■



## 12 Product Measure and Fubini's Theorem

### 12.1 Dynkin's $\pi$ - $\lambda$ System

We firstly construct a relatively weaker notion than  $\sigma$ -algebras, namely the  $\pi$ -system and  $\lambda$ -system.

**Definition 12.1.** Let  $X$  be a set, a collection  $\mathcal{P}$  of subsets of  $X$  is called a  **$\pi$ -system** if it's closed under intersection:

$$A, B \in \mathcal{P} \implies A \cap B \in \mathcal{P} \quad (1)$$

**Definition 12.2.** Let  $X$  be a set, a collection  $\mathcal{D}$  of subsets of  $X$  is called a  **$\lambda$ -system** (or Dynkin class or a d-system) if

1.  $X \in \mathcal{D}$ ;
2. (closure under set difference)  $A, B \in \mathcal{D}, A \subseteq B \implies B \setminus A \in \mathcal{D}$ ;
3. (closure under ascending union) if  $A_1 \subseteq A_2 \subseteq \dots \in \mathcal{D}$ , then  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$ .

**Remark 12.1** (An equivalent definiton). The third requirement of  $\lambda$ -system may be replaced by closure under countable disjoint union.

*Proof.* **TODO** ■

**Remark 12.2.** A  $\sigma$ -algebra is always a  $\lambda$ -system but not converse.

**Example 12.1.** Take any two probability measures  $\mu$  and  $\nu$  on  $\mathbb{R}$ , let

$$\mathcal{D} = \{A \in \mathcal{B}(\mathbb{R}) : \mu(A) = \nu(A)\} \quad (2)$$

Then  $\mathcal{D}$  is always a  $\lambda$ -system but not necessarily a  $\sigma$ -algebra:

*Proof.* Let  $\mu$  and  $\nu$  be two probability measures, since  $\mu(X) = \nu(X) = 1$ , so  $X \in \mathcal{D}$ .

Let  $A, B \in \mathcal{D}$  such that  $A \subseteq B$ , since probability measures are finite,  $\mu(B \setminus A) = \mu(B) - \mu(A) = \nu(B) - \nu(A) = \nu(B \setminus A)$ .

Let  $A_n$  be an ascending sequence of sets in  $\mathcal{D}$ ,  $\mu(\bigcup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mu(A_n) \leq 1$ . Since  $\nu(A_n) = \mu(A_n)$ , given convergence, the limit must be the same. ■

*Counter-Example.* Consider  $X = \{1, 2, 3, 4\}$ , define probability measures

$$\mu(\{1\}) = \mu(\{2\}) = \mu(\{3\}) = \mu(\{4\}) = \frac{1}{4} \quad (3)$$

$$\nu(\{1\}) = \frac{1}{2}, \nu(\{2\}) = 0, \nu(\{3\}) = \frac{1}{2}, \nu(\{4\}) = 0 \quad (4)$$

Take  $A = \{1, 2\}, B = \{2, 3\}$ , both in  $\mathcal{D}$ , however,  $A \cap B \notin \mathcal{D}$ , thus  $\mathcal{D}$  fails to be a  $\sigma$ -algebra. ■

**Theorem 12.1** (Dynkin's  $\pi$ - $\lambda$  theorem). Let  $X$  be a set, let  $\mathcal{P}$  be a  $\pi$ -system on  $X$  and  $\mathcal{D}$  be a  $\lambda$ -system on  $X$ . Then

$$\mathcal{P} \subseteq \mathcal{D} \implies \sigma(\mathcal{P}) \subseteq \mathcal{D} \quad (5)$$

*Usage:* suppose we wish to check some property on  $\mathcal{A}$ , and we find some  $\pi$ -system  $\mathcal{P}$  that generates  $\mathcal{A}$ , it suffices to check this property on any  $\lambda$ -system covers  $\mathcal{P}$ .

*Proof.* Note that an arbitrary intersection of  $\lambda$ -system is a  $\lambda$ -system.

Let  $\mathcal{D}$  be the smallest (i.e., the intersection)  $\lambda$ -system contains  $\mathcal{P}$ . Suppose  $\mathcal{P} \subseteq \mathcal{D}$ , define:

$$\mathcal{D}_1 = \{A \in \mathcal{D} : A \cap B \in \mathcal{D} \quad \forall B \in \mathcal{P}\} \quad (6)$$

Since  $\mathcal{P}$  is a  $\pi$ -system, take any  $A \in \mathcal{P}$ ,  $A \cap B \in \mathcal{P}$  for any  $B \in \mathcal{P}$ ,  $A \in \mathcal{D}_1$ , therefore,  $\mathcal{P} \subseteq \mathcal{D}_1$ .

Note that

1. Note that  $X \in \mathcal{D}$ . And,  $\forall B \in \mathcal{P}$ ,  $X \cap B = B \in \mathcal{P} \subseteq \mathcal{D}$ , therefore,  $X \in \mathcal{D}_1$ .
2. Let  $A, B \in \mathcal{D}_1$ , such that  $A \subseteq B$ ,  $\forall C \in \mathcal{P}$ ,  $A \cap C, B \cap C \in \mathcal{D}$ . But  $\mathcal{D}$  is a  $\lambda$ -system,

$$(B \cap C) \setminus (A \cap C) = (B \setminus A) \cap C \in \mathcal{D} \quad (7)$$

Hence,  $B \setminus A \in \mathcal{D}_1$ .

3. If  $A_1 \subseteq A_2 \subseteq \dots \in \mathcal{D}_1$ , then for all  $B \in \mathcal{P}$ ,  $A_i \cap B \in \mathcal{D}$  and

$$\left( \bigcup_{i=1}^{\infty} A_i \right) \cap B = \bigcup_{i=1}^{\infty} (A_i \cap B) \in \mathcal{D} \quad (8)$$

Therefore,  $\bigcup A_i \in \mathcal{D}_1$ , so  $\mathcal{D}_1$  is a  $\lambda$ -system.

Since  $\mathcal{D}_1$  is a  $\lambda$ -system contains  $\mathcal{P}$ , hence  $\mathcal{D} \subseteq \mathcal{D}_1$ . Therefore,  $\mathcal{D}_1 = \mathcal{D}$ .

This shows  $\forall A \in \mathcal{D}, \forall B \in \mathcal{P}, A \cap B \in \mathcal{D}$ . (†)

Define

$$\mathcal{D}_2 = \{A \in \mathcal{D} : A \cap B \in \mathcal{D} \quad \forall B \in \mathcal{P}\} \quad (9)$$

By (†),  $\forall A \in \mathcal{P} \subseteq \mathcal{D}, \forall B \in \mathcal{D}, A \cap B \in \mathcal{D}$ , therefore,  $\mathcal{P} \subseteq \mathcal{D}_2$ . Also,  $\mathcal{D}_2$  is a  $\lambda$ -system:

1.  $X \in \mathcal{D}_2$ .
2. Let  $A \subseteq B$  both in  $\mathcal{D}_2$ , take any  $C \in \mathcal{P}$ ,

$$(B \setminus A) \cap C = (B \cap C) \setminus (A \cap C) \in \mathcal{D} \quad (10)$$

3. Same as in Equation (8).

Why is it sufficient to show this holds on the smallest  $\mathcal{D}$ ?

Is this proof overshoot?

Therefore,  $\mathcal{D}_2$  is also a  $\lambda$ -system containing  $\mathcal{P}$ , this implies  $\mathcal{D}_2 = \mathcal{D}$ .

Moreover, for all  $A, B \in \mathcal{D}$ ,  $A \cap B \in \mathcal{D}$ , so that  $\mathcal{D}$  is also a  $\pi$ -system.

**Lemma 12.1** (Another Definition of  $\sigma$ -algebra). A collection of sets is both  $\pi$  and  $\lambda$  if and only if its a  $\sigma$ -algebra.

*Proof.* For a  $\lambda$ -system  $\mathcal{D}$ , it contains  $X^c = \emptyset$  and is closed under complement (take one of two sets to be  $X$ ).

To show closure under countable union, let  $A_1, A_2, \dots \in \mathcal{D}$ , we may define  $B_n = A_1 \cup \dots \cup A_n$ , so that  $\bigcup A_n = \bigcup B_n$  and  $B_n$  is an increasing sequence. In particular, since  $\mathcal{D}$  is closed under complement (as a  $\lambda$ -system) and finite intersection (as a  $\pi$ -system),  $\mathcal{D}$  is closed under finite union, each  $B_n \in \mathcal{D}$  as well. By definition of  $\lambda$ -system,  $\bigcup A_n \in \mathcal{D}$ .

The converse is trivial, every  $\sigma$ -algebra is both  $\lambda$  and  $\pi$ . ■

Therefore,  $\mathcal{D}$  is a  $\sigma$ -algebra containing  $\mathcal{P}$ , it follows  $\sigma(\mathcal{P}) \subseteq \mathcal{D}$ . ■

**Corollary 12.1.** Let  $\mu$  and  $\nu$  be  $\sigma$ -finite measures on a measurable space  $(X, \mathcal{A})$ . If  $\mu$  and  $\nu$  agree on a  $\pi$ -system  $\mathcal{P}$  that generate  $\mathcal{A}$ , then  $\mu = \nu$  on  $\mathcal{A}$ .

*Proof.* We know that  $\mathcal{D} = \{A \in \mathcal{A} : \mu(A) = \nu(A)\}$  is a  $\lambda$ -system,  $\mathcal{P} \subseteq \mathcal{D}$  implies  $\sigma(\mathcal{P}) = \mathcal{A} \subseteq \mathcal{D}$ . ■

**Corollary 12.2.** Let  $\mu$  and  $\nu$  be measures on a measurable space  $(X, \mathcal{A})$ . Let  $\mathcal{P}$  be a  $\pi$ -system on  $X$  such that

1.  $\sigma(\mathcal{P}) = \mathcal{A}$ ,
2.  $\forall A \in \mathcal{P}, \mu(A) = \nu(A) < \infty$ ,
3.  $\exists$  a sequence  $A_1 \subseteq A_2 \subseteq \dots \in \mathcal{P}$  such that  $\bigcup A_i = X$ .

Then  $\mu = \nu$ .

*Intuition:* for a  $\pi$ -system that approximates the entire space  $X$  via an ascending sequence and generates  $\mathcal{A}$ , then it suffices to show  $\mu = \nu$  on the  $\pi$ -system in order to show  $\mu = \nu$ .

*Proof.* Case 1: finite measures. Suppose  $\mu$  and  $\nu$  are finite measures, define

$$\mathcal{D} = \{A \in \mathcal{A} : \mu(A) = \nu(A)\} \tag{11}$$

Clearly,  $\mathcal{P} \subseteq \mathcal{D}$ . We are going to show  $\mathcal{D} = \mathcal{A}$ .

Firstly, we show  $\mathcal{D}$  is a  $\lambda$ -system.

(1) Using property (3), we may construct a sequence in  $\mathcal{P}$  increasing to  $X$ , taking the limit shows  $\mu(X) = \nu(X)$  and  $X \in \mathcal{D}$  as a result.

(2) Let  $A, B \in \mathcal{D}$  such that  $A \subseteq B$ , since  $\mu$  and  $\nu$  are finite on  $\mathcal{P}$ ,

$$\mu(B \setminus A) = \mu(B) - \mu(A) \tag{12}$$

$$= \nu(B) - \nu(A) \tag{13}$$

$$= \nu(B \setminus A) \tag{14}$$

Thus  $B \setminus A \in \mathcal{D}$ .

(3) If  $A_1 \subseteq A_2 \subseteq \dots \in \mathcal{D}$ , then

$$\mu(\cup A_i) = \lim \mu(A_i) = \lim \nu(A_i) = \nu(\cup A_i) \quad (15)$$

Therefore  $\mathcal{D}$  is a  $\lambda$ -system. Since  $\mathcal{P} \subseteq \mathcal{D}$ , the  $\pi$ - $\lambda$  theorem implies  $\sigma(\mathcal{P}) \subseteq \mathcal{D}$ . Thus,  $\mathcal{A} = \mathcal{D}$ . ■

*Proof. The general case.* There exists  $A_1 \subseteq A_2 \subseteq \dots \in \mathcal{P}$  such that  $\bigcup A_i = X$ . Moreover,  $\mu(A_i) = \nu(A_i) < \infty$  for every  $i$ . Define

$$\mathcal{D}_i = \{B \in \mathcal{A} : \mu(B \cap A_i) = \nu(B \cap A_i)\} \quad (16)$$

$\mathcal{D}_i$  is a  $\lambda$ -system containing  $\mathcal{P}$ , so that  $\mathcal{A} = \sigma(\mathcal{P}) \subseteq \mathcal{D}_i$ . Hence,  $\mathcal{D}_i = \mathcal{A}$ .

For every  $B \in \mathcal{A}$ ,  $\mu(B \cap A_i) = \nu(B \cap A_i)$  for all  $i$ . But

$$\mu(B) = \lim \mu(B \cap A_i) = \lim \nu(B \cap A_i) = \nu(B) \quad (17)$$

Thus,  $\mu = \nu$ . ■

## 12.2 Product Measures

**Definition 12.3.** Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be two measure spaces, suppose  $\mu$  and  $\nu$  are  $\sigma$ -finite. Let  $X \times Y$  be the Cartesian product of  $X$  and  $Y$

$$X \times Y := \{(x, y) : x \in X, y \in Y\} \quad (18)$$

The **product  $\sigma$ -algebra**, denoted as  $\mathcal{A} \times \mathcal{B}$ , is the  $\sigma$ -algebra generated by the following collection of sets:

$$\{A \times B : A \in \mathcal{A} \text{ and } B \in \mathcal{B}\} \quad (19)$$

sets in this collection are called **rectangles**.

**Theorem 12.2** (Product Measure). There exists a unique measure  $\mu \times \nu$  on  $(X \times Y, \mathcal{A} \times \mathcal{B})$  that satisfies  $\forall A \in \mathcal{A}, B \in \mathcal{B}$ ,

$$\mu \times \nu(A \times B) = \mu(A)\nu(B) \quad (20)$$

Here, we only require the product measure to be well behave on rectangles but not other sets in  $\mathcal{A}$ .

*Proof. Uniqueness.* Observe that the set of all rectangles  $\{A \times B : A \in \mathcal{A}, B \in \mathcal{B}\}$  is a  $\pi$ -system:

$$(A \times B) \cap (A' \times B') = \{(a, b) : a \in A, b \in B, a \in A', b \in B'\} \quad (21)$$

$$= \{(a, b) : a \in A \cap A', b \in B \cap B'\} \quad (22)$$

$$= \underbrace{(A \cap A')}_{\in \mathcal{A}} \times \underbrace{(B \cap B')}_{\in \mathcal{B}} \quad (23)$$

Since  $\mu$  and  $\nu$  are  $\sigma$ -finite, there exists  $A_1, A_2, \dots \in \mathcal{A}$  such that  $\bigcup A_i = X$  and  $\mu(A_i) < \infty$  for all  $i \in \mathbb{N}$ . Similarly, there exists  $B_1, B_2, \dots$  such that  $\bigcup B_i = Y$  and  $\nu(B_i) < \infty$  for all  $i \in \mathbb{N}$ . Combining these two sequences,

$$A_1 \times B_1 \subseteq A_2 \times B_2 \subseteq \dots \quad (24)$$

Such that  $\mu \times \nu(A_i \times B_i) = \mu(A_i)\nu(B_i) < \infty$  for all  $i$  and  $\bigcup (A_i \times B_i) = X \times Y$ .

If  $\gamma_1$  and  $\gamma_2$  are two candidates for  $\mu \times \nu$ . We have shown that there exists sequence of rectangles with the following properties:

1.  $R_i = A_i \times B_i$ ,
2.  $R_1 \subseteq R_2 \subseteq \dots$ ,
3.  $\bigcup R_i = X \times Y$ ,
4.  $\gamma_1(R_i) = \gamma_2(R_i)$  for all  $i$ .

By the previous corollary,  $\gamma_1 = \gamma_2$  on a  $\pi$ -system that generates  $\mathcal{A}$ , thus  $\gamma_1 = \gamma_2$  on  $\mathcal{A}$ . ■

*Proof. Existence.*  $\forall E \in \mathcal{A} \times \mathcal{B}$  and  $\forall x \in X, y \in Y$ , define

$$E_x = \{y \in Y : (x, y) \in E\} \quad (25)$$

$$E_y = \{x \in X : (x, y) \in E\} \quad (26)$$

Similarly, for any measurable  $f : X \times Y \rightarrow \mathbb{R}^*$ , define

$$f_x : Y \rightarrow \mathbb{R}^* \quad f_x(y) = f(x, y) \quad (27)$$

$$f_y : X \rightarrow \mathbb{R}^* \quad f_y(x) = f(x, y) \quad (28)$$



**Lemma 12.2.** The projection of a measurable set is measurable. That is,  $\forall E \in \mathcal{A} \times \mathcal{B}$ ,  $\forall x \in X$ ,  $E_x \in \mathcal{B}$ ;  $\forall y \in Y$ ,  $E_y \in \mathcal{A}$ .

*Proof.* Take any  $x \in X$ , let

$$\mathcal{F} = \{E \in \mathcal{A} \times \mathcal{B} : E_x \in \mathcal{B}\} \quad (29)$$

We show that  $\mathcal{F} = \mathcal{A} \times \mathcal{B}$ .

Note that  $\forall x \in X$ , for every rectangle,  $A \in \mathcal{A}$ ,  $B \in \mathcal{B}$ ,  $(A \times B)_x = B \in \mathcal{B}$ . Thus  $\mathcal{F}$  contains all rectangles.

(i)  $\emptyset \in \mathcal{F}$ .

(ii) Let  $E \in \mathcal{F}$ , then  $(E^c)_x = (E_x)^c \in \mathcal{B}$ , therefore,  $E^c \in \mathcal{F}$ .

(iii) Let  $E_1, E_2, \dots \in \mathcal{F}$ , then  $(\cup E_i)_x = \cup \underbrace{(E_i)_x}_{\in \mathcal{B}} \in \mathcal{B}$ .

Therefore,  $\mathcal{F}$  is a  $\sigma$ -algebra containing all rectangles, thus  $\mathcal{F} \supseteq \sigma(\text{Rectangles}) = \mathcal{A} \times \mathcal{B}$ . Hence,  $\mathcal{F} = \mathcal{A} \times \mathcal{B}$ .

The same proof works for  $E_y$ . ■

**Lemma 12.3.** The projection of measurable function is measurable.

*Proof.* Take any measurable  $f : X \times Y \rightarrow \mathbb{R}^*$ , for all  $B \in \mathcal{B}(\mathbb{R}^*)$ , for all  $x \in X$ ,

$$f_x^{-1}(B) = \{y : f_x(y) \in B\} \quad (30)$$

$$= \{y : f(x, y) \in B\} \quad (31)$$

$$= \{y : (x, y) \in f^{-1}(B)\} \quad (32)$$

$$= \underbrace{(f^{-1}(B))_x}_{\in \mathcal{A} \times \mathcal{B}} \in \mathcal{B} \quad (33)$$

This shows  $f_x$  is measurable, a similar argument works for  $f_y$ . ■

**Proposition 12.1.** Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be  $\sigma$ -finite measure spaces.  $\forall E \in \mathcal{A} \times \mathcal{B}$ ,

$$x \mapsto \nu(E_x) \text{ is measurable with respect to } \mathcal{A} \quad (34)$$

$$y \mapsto \mu(E_y) \text{ is measurable with respect to } \mathcal{B} \quad (35)$$

*Intuitively*,  $x \mapsto \nu(E_x)$  computes the side length at a particular level of  $x$ .

*Proof.* First, suppose  $\mu$  and  $\nu$  are finite measures.

$$\mathcal{D} = \{E \in \mathcal{A} \times \mathcal{B} : x \mapsto \nu(E_x) \text{ is } \mathcal{A} \text{ measurable}\} \quad (36)$$

Note that for a rectangle  $E = A \times B$  for some  $A \in \mathcal{A}$ ,  $B \in \mathcal{B}$ , then

$$\nu(E_x) = \begin{cases} \nu(B) & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \quad (37)$$

So  $\nu(E_x) = \nu(B)\chi_A(x)$ . This is a measurable function on  $(X, \mathcal{A})$  since it's a constant multiplied by an indicator function. Thus  $\mathcal{D}$  contains all rectangles.

Let  $E_1, E_2 \in \mathcal{D}$  such that  $E_1 \supseteq E_2$ , then

$$\nu((E_1 \setminus E_2)_x) = \nu((E_1)_x \setminus (E_2)_x) \quad (38)$$

$$= \nu((E_1)_x) - \nu((E_2)_x) \text{ since } \nu \text{ is finite} \quad (39)$$

The map  $x \mapsto \nu((E_1)_x)$  and  $x \mapsto \nu((E_2)_x)$  are both measurable, thus  $E_1 \setminus E_2 \in \mathcal{D}$ .

Finally, take an increasing sequence  $E_1 \subseteq E_2 \subseteq \dots \in \mathcal{D}$ , then

$$\left(\bigcup E_i\right)_x = \bigcup (E_i)_x \quad (40)$$

Moreover,  $(E_1)_x \subseteq (E_2)_x \subseteq \dots$ , so

$$\nu\left(\left(\bigcup E_i\right)_x\right) = \nu\left(\bigcup (E_i)_x\right) \quad (41)$$

$$= \lim_{i \rightarrow \infty} \nu((E_i)_x) \quad (42)$$

Thus  $x \mapsto \nu((\bigcup E_i)_x)$  is the limit of a sequence of measurable maps  $x \mapsto \nu((E_i)_x)$ , thus it is measurable and  $\bigcup E_i \in \mathcal{D}$ .

Therefore,  $\mathcal{D}$  is a  $\lambda$ -system containing all rectangles, thus  $\mathcal{D} = \mathcal{A} \times \mathcal{B}$ . ■

*Proof. Suppose  $\nu$  is  $\sigma$ -finite.* Then there exists a sequence of disjoint sets  $D_1, D_2, \dots \in \mathcal{B}$  such that  $\bigcup D_i = Y$  and  $\nu(D_i) < \infty$  for all  $i \in \mathbb{N}$ . It's easy to show that  $x \mapsto \nu(D_i \cap E_x)$  is measurable: simply define  $\nu_i(B) = \nu(B \cap D_i)$ , which is a finite measure, then apply our previous reasoning,  $x \mapsto \nu_i(B)$  is measurable.

But,

$$\nu(E_x) = \sum_{i=1}^{\infty} \nu(E_x \cap D_i) \quad (43)$$

Being a series of measurable functions (as the limit of measurable partial sums),  $x \mapsto \nu(D_i \cap E_x)$  is measurable. ■

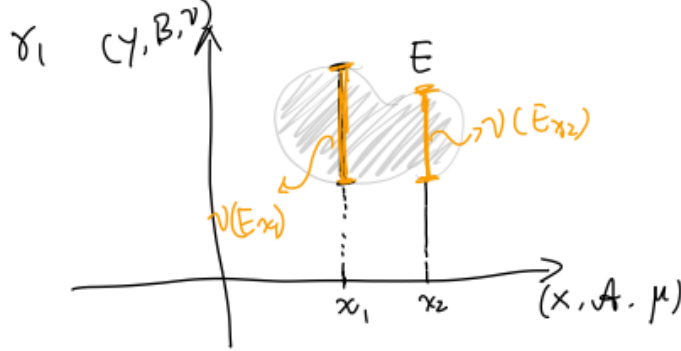
For every  $E \in \mathcal{A} \times \mathcal{B}$ , define

$$\gamma_1(E) = \int_X \nu(E_x) \, d\mu(x) \quad (44)$$

$$\gamma_2(E) = \int_Y \nu(E_y) \, d\nu(y) \quad (45)$$

Intuitively, for  $\gamma_1$ , at each  $x \in X$ , we compute the height of vertical section  $E_x$ , then we sum all these lengths across different locations in  $X$ .

Figure 2: Intuition for  $\gamma_1$



(1) Obviously,  $\gamma_1(\emptyset) = 0$  since  $(\emptyset)_x = 0$ . The same holds for  $\gamma_2$ .

(2) Let  $E_1, E_2 \dots \in \mathcal{A} \times \mathcal{B}$  be a sequence of disjoint sets,  $\{(E_i)_x\}_{i=1}^\infty$  are disjoint for any  $x \in X$ . Since  $x \mapsto \nu((E_i)_x)$  is a non-negative measurable function, by (corollary of) the monotone convergence theorem,

$$\sum_{i=1}^{\infty} \gamma_1(E_i) = \sum_{i=1}^{\infty} \int_X \nu((E_i)_x) d\mu(x) \quad (46)$$

$$= \int_X \sum_{i=1}^{\infty} \nu((E_i)_x) d\mu(x) \quad (47)$$

$$= \int_X \nu(\cup (E_i)_x) d\mu(x) \quad (48)$$

$$= \gamma_1((\cup E_i)_x) \quad (49)$$

The same applies to  $\gamma_2$ , both  $\gamma_1$  and  $\gamma_2$  are measures.

Now, for any rectangle  $A \times B$ ,  $(A \times B)_x = B$  if  $x \in A$  and is  $\emptyset$  otherwise.

$$\gamma_1(A \times B) = \int_X \nu((A \times B)_x) d\mu(x) \quad (50)$$

$$= \int_A \nu(B) d\mu(x) + \int_{A^c} \nu(\emptyset) d\mu(x) \quad (51)$$

$$= \nu(B)\mu(A) \quad (52)$$

Similarly, we can show that  $\gamma_2(A \times B) = \mu(A)\nu(B)$ . By the uniqueness of product measure,  $\gamma_1 = \gamma_2$ . ■

**Definition 12.4.** We define the **product measure** as  $\mu \times \nu = \gamma_1 = \gamma_2$ ,

$$\gamma_1(E) = \int_X \nu(E_x) d\mu(x) \quad (53)$$

$$\gamma_2(E) = \int_Y \nu(E_y) d\nu(y) \quad (54)$$

### 12.3 Fubini's Theorem

**Theorem 12.3** (Tonelli's Theorem). Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be  $\sigma$ -finite measure spaces, let  $f : X \times Y \rightarrow [0, \infty]$  be a measurable function (not necessarily  $(\mu \times \nu)$ -integrable). Then,

1. The map  $x \mapsto \int_Y f_x d\nu$  is  $\mathcal{A}$ -measurable, and  $y \mapsto \int_X f_y d\mu$  is  $\mathcal{B}$ -measurable.
2. the following iterated formula holds:

$$\int_{X \times Y} f d(\mu \times \nu) = \int_X \left( \int_Y f_x d\nu \right) d\mu \quad (55)$$

$$= \int_Y \left( \int_X f_y d\mu \right) d\nu \quad (56)$$

*Proof.* If  $f = \chi_E$  for some set  $E \in \mathcal{A}$ , then we have the first conclusion since

$$\int_Y f_x d\nu = \int_Y (\chi_E)_x d\nu = \nu(E_x) \quad (57)$$

The same holds for  $f_y$ , we have shown this kind of maps are measurable.

The second part follows from the construction of product measure:

$$\int_{X \times Y} \chi_E d(\mu \times \nu) = (\mu \times \nu)(E) \quad (58)$$

$$= \int_X \nu(E_x) d\mu \text{ by definition of product measure} \quad (59)$$

$$= \int_X \int_Y f_x d\nu d\mu \text{ by Equation (57)} \quad (60)$$

By linearity the theorem holds for any non-negative simple function  $f$ . For any non-negative measurable function  $f$ , there exists an increasing sequence of simple functions  $f_n \rightarrow f$ , each  $f_n$  has above properties. By monotone convergence theorem, the limit function  $f$  also satisfies these properties. ■

**Theorem 12.4** (Fubini's Theorem). Let  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be two  $\sigma$ -finite measure spaces, let  $f : X \times Y \rightarrow [-\infty, \infty]$  be a measurable mapping and  $(\mu \times \nu)$ -integrable, then

1. For  $\mu$  almost every  $x$ ,  $f_x$  is  $\nu$ -integrable,
2. for  $\nu$  almost every  $y$ ,  $f_y$  is  $\mu$ -integrable.

3. Define

$$I_f(x) := \begin{cases} \int_Y f_x \, d\nu & \text{if } f_x \text{ is } \nu\text{-integrable} \\ 0 & \text{otherwise} \end{cases} \quad (61)$$

$$J_f(y) := \begin{cases} \int_X f_y \, d\mu & \text{if } f_y \text{ is } \mu\text{-integrable} \\ 0 & \text{otherwise} \end{cases} \quad (62)$$

The following iterated formula holds:

$$\int_{X \times Y} f \, d(\mu \times \nu) = \int_X I_f \, d\mu = \int_Y J_f \, d\nu \quad (63)$$

*Proof.* Let  $f = f^+ - f^-$  and  $f$  is  $(\mu \times \nu)$ -integrable, both  $f^+$  and  $f^-$  are measurable. **Why?** By Tonelli's theorem,

$$\int_X \int_Y |f_x(y)| \, d\nu(y) \, d\mu(x) = \int_{X \times Y} |f| \, d(\mu \times \nu) < \infty \quad (64)$$

Thus

$$\int_X \int_Y f_x^+ \, d\nu(y) \, d\mu(x) < \infty \quad (65)$$

But we know that  $x \mapsto \int_Y f_x^+(y) \, d\nu(y)$  is measurable by Tonelli's theorem. So this finiteness in Equation (64) shows

$$\int_Y f_x^+ \, d\nu(y) < \infty \, \mu\text{-a.e.} \quad (66)$$

The same for  $f_x^-$ . So  $\mu$ -a.e.  $x$ ,  $f_x$  is integrable w.r.t.  $\nu$ . On the set where  $f_x^+$  and  $f_x^-$  are both integrable, we have

$$I_f(x) = \int_Y f_x \, d\nu = \int_Y f_x^+ \, d\nu - \int_Y f_x^- \, d\nu \quad (67)$$

Outside of this set,

$$\int_X I_f(x) \, d\mu(x) = 0 \quad (68)$$

Thus,

$$\int_X I_f(x) \, d\mu(x) = \int_X \left[ \int_Y f_x^+ \, d\nu - \int_Y f_x^- \, d\nu \right] \, d\mu \quad (69)$$

$$= \int_X \int_Y f_x^+ \, d\nu \, d\mu - \int_X \int_Y f_x^- \, d\nu \, d\mu \quad (70)$$

$$= \int_{X \times Y} f^+ \, d(\mu \times \nu) - \int_{X \times Y} f^- \, d(\mu \times \nu) \quad (71)$$

$$= \int_{X \times Y} f \, d(\mu \times \nu) \quad (72)$$

■

## 13 Riesz Representation Theorem

This section needs to be revised.

### 13.1 Locally Compact Hausdorff Spaces

**Definition 13.1.** A **Hausdorff space** is a topological space where for any two distinct points in it, there exists neighbourhoods of each which are disjoint.

*Distinct points are separated.*

**Definition 13.2.** A Hausdorff space is called **locally compact** if every point has an open neighbourhood whose closure is compact.

**Lemma 13.1.** Let  $X$  be a Hausdorff space, let  $K$  and  $L$  be disjoint compact subsets of  $X$ . Then, there exists disjoint open sets  $U, V$  such that  $K \subseteq U$  and  $L \subseteq V$ .

*Disjoint compact sets are separated by open sets.*

*Proof.* **This proof needs revising!** WLOG, assume  $K, L \neq \emptyset$ , suppose  $K$  consists of single a single point  $x$ .

$\forall y \in L$ ,  $\exists$  disjoint open sets,  $U_x \ni x$  and  $V_y \ni y$  since  $K \cap L = \emptyset$  and  $X$  is Hausdorff. Then  $\{V_y\}_{y \in L}$  is an open cover for  $L$ . By compactness of  $L$ , there exists a finite sub-cover  $V_{y_1}, V_{y_2}, \dots, V_{y_n}$ .

Let

$$U = \bigcap_{i=1}^n U_{y_i} \quad (1)$$

$$V = \bigcup_{i=1}^n V_{y_i} \quad (2)$$

Then  $U$  and  $V$  are open disjoint and  $x \in U$ ,  $L \subseteq V$ . Let  $K$  be an arbitrary compact set.  $\forall y \in L$ ,  $\exists$  disjoint open sets  $U_y \supseteq K$  and  $V_y \ni y$ .

Again,  $\{V_y\}_{y \in K}$  is an open cover for  $L$ , there exists a finite sub-cover, and take  $U = \bigcap_{i=1}^n U_{y_i}$  and  $V = \bigcup_{i=1}^n V_{y_i}$ . ■

**Lemma 13.2.** Let  $X$  be a locally compact Hausdorff space, take  $x \in X$  and an open neighbourhood  $U$  of  $x$ . Then, there exists open set  $V$  such that  $x \in V \subseteq \overline{V} \subseteq U$ , and  $\overline{V}$  is compact.

*$U$  is locally compact as well.*

*Proof.* The local compactness implies there exists open  $W \ni x$  such that  $\overline{W}$  is compact. Let

$$W_1 = W \cap U \quad (3)$$

then  $W_1$  is open and  $x \in W_1$ . Also  $\overline{W_1}$  is a closed subset of compact set is also compact thus  $\overline{W_1}$  is compact.

Let  $K = \overline{W_1} \setminus W_1 = \overline{W_1} \cap W_1^c$ .  $K$  (the boundary) is a closed set contained in compact  $\overline{W_1}$ . So  $K$  is compact. So by Lemma 1, there exists disjoint open sets  $V_1, V_2$  such that  $K \subseteq V_1$  and  $x \in V_2$ .

Let  $V = V_2 \cap W$ , note that

1.  $x \in V$ ,
2.  $V$  is open,
3.  $V \subseteq U$ ,
4.  $\overline{V}$  is a closed subset of the compact set  $\overline{W_1}$ .
5.  $\overline{V} \subseteq U$ :  $V \subseteq W$  and  $V$  is separated from the boundary of  $W_1$  by an open set. From this, it is not hard to see that  $\overline{V} \subseteq W_1$  thus  $\overline{V} \subseteq U$ .

■

**Lemma 13.3.** Let  $X$  be a locally compact Hausdorff space, let  $K$  be a compact subset of  $X$ , suppose there exists an open  $U$  such that  $K \subseteq U$ . Then, there exists open  $V$  such that  $K \subseteq V \subseteq \overline{V} \subseteq U$ , moreover,  $\overline{V}$  is compact.

*Proof.* For each  $x \in K$ , find an open set  $V_x$  such that

$$x \in V_x \subseteq \overline{V_x} \subseteq U \quad (4)$$

and  $\overline{V_x}$  is compact.  $\{V_x\}_{x \in K}$  is an open cover for  $K$ , thus take a finite sub-cover of it:  $V_{x_1}, V_{x_2}, \dots, V_{x_n}$ . Let  $V = \bigcup_{i=1}^n V_{x_i}$ , then  $V$  is open, contains  $K$  and  $\overline{V} \subseteq U$ .  $\overline{V}$  being a closed subset of compact set  $\bigcup \overline{V_x}$  is also compact. ■

**Definition 13.3.** A topological space is called **normal** if it is Hausdorff and any pair of disjoint closed sets can be separated by disjoint open sets.

**Lemma 13.4.** Any compact Hausdorff space is normal.

**Theorem 13.1** (Urysohn's Lemma). Let  $X$  be a normal topological space, let  $E$  and  $F$  be disjoint closed subsets of  $X$ . Then,  $\exists$  a continuous function  $f : X \rightarrow [0, 1]$  such that  $f = 0$  on  $E$  and  $f = 1$  on  $F$ .

*Proof.* Let  $D$  be the set of Dyadic rationals in  $(0, 1)$ , i.e., all numbers of the form  $\frac{k}{2^n}$ . We will individually construct a family of open sets  $\{U_r\}_{r \in D}$ . First note that  $E$  and  $F$  being closed sets and  $X$  normal, then there exists disjoint open sets  $U \supseteq E$  and  $V \supseteq F$  such that

$$U \subseteq V^c \quad (5)$$

where  $V^c$  is closed,  $E \cap U^c = \emptyset$ , and  $\overline{U} \cap F = \emptyset$ . Moreover,  $U \subseteq \overline{U} \subseteq V^c$ .

Let  $U_{1/2} = U$ , applying the same argument on  $(E, U^c)$  and get  $U_{1/4} \subseteq U = U_{1/2}$ . Same for  $(\overline{U}, F)$ , get  $U_{3/4}$  such that

$$U = U_{1/2} \subseteq U_{3/4} \subseteq \overline{U_{3/4}} \subseteq V^c \quad (6)$$

Continuous by induction, we find  $\{U_r\}_{r \in D}$  such that



1.  $E \subseteq U_r, \overline{U_r} \subseteq F^c$  for all  $r \in D$ .
2. For all  $r < s, \overline{U_r} \subseteq U_s$ .

Define

$$f(x) = \begin{cases} 1 & \text{if } x \notin \bigcup_{r \in D} U_r \\ \inf\{r : x \in U_r\} & \text{otherwise} \end{cases} \quad (7)$$

To show the continuity, since  $f$  is real-valued, it suffices to show that  $f^{-1}((r, s))$  is open for any Dyadic rational  $(r, s)$  since all intervals of this form generates the Euclidean topology on the real line.

First, suppose the  $0 < r < s < 1, x \in f^{-1}((r, s))$  if and only if  $r < f(x) < s$ , then

1.  $x \notin \overline{U_q}$  for some  $q > r$ :  $f(x) > r$  if and only if  $f(x) > q' > q > r$  for some  $q', q \in D$  implies  $f(x) \notin U_{q'}$ , but  $\overline{U_q} \subseteq U_{q'}$ , so  $f \notin \overline{U_q}$ .
2.  $x \in U_p$  for some  $p < s$ .

if and only if

$$x \in \left( \bigcup_{q > r} \overline{U_q^c} \right) \cap \left( \bigcup_{p < s} U_p \right) \quad (8)$$

So  $f^{-1}((r, s))$  is open for all  $r, s \in (0, 1)$ . Similar arguments work for  $r \leq 0 < s < 1, 0 < r < 1 \leq s$  and other cases. ■