

MATH597  
Analysis II

Pingbang Hu

April 24, 2022

## Abstract

Notice that since in this course, the cross-referencing between theorems, lemmas, and propositions are quite complex and hard to keep track of, hence in this note, whenever you see a **!** over  $=$ , like  $\stackrel{!}{=}$ , then that **!** is *clickable*! It will direct you to the corresponding theorem, lemma, or proposition we're using to deduce that particular equality.

Notice that there are some proofs is **intended** left as assignments, and for completeness, I put them in [Appendix A](#), use it in your **own risks**! You'll lose the chance to practice and really understand the materials.

Additionally, we'll use Folland[[FF99](#)] as our main text, while using Tao[[Tao13](#)] and Axler[[Ax19](#)] as supplementary references.

# Contents

<b>1</b>	<b>Measure</b>	<b>3</b>
1.1	$\sigma$ -algebras	4
1.2	Measures	7
1.3	Outer Measures	11
1.4	Hahn-Kolmogorov Theorem	17
1.5	Borel Measures on $\mathbb{R}$	21
1.6	Lebesgue-Stieltjes Measure on $\mathbb{R}$	25
1.7	Regularity Properties of Lebesgue-Stieltjes Measures	28
<b>2</b>	<b>Integration</b>	<b>32</b>
2.1	Measurable Function	32
2.2	Integration of Nonnegative Functions	36
2.3	Integration of Complex Functions	41
2.4	$L^1$ Space	43
2.5	Riemann Integrability	46
2.6	Modes of Convergence	47
<b>3</b>	<b>Product Measure</b>	<b>52</b>
3.1	Product $\sigma$ -algebra	52
3.2	Product Measures	55
3.3	Monotone Class Lemma	56
3.4	Fubini-Tonelli Theorem	58
3.5	Lebesgue Measure on $\mathbb{R}^d$	59
<b>4</b>	<b>Differentiation on Euclidean Space</b>	<b>62</b>
4.1	Hardy-Littlewood Maximal Function	62
4.2	Lebesgue Differentiation Theorem	64
<b>5</b>	<b>Normed Vector Space</b>	<b>67</b>
5.1	Metric Spaces and Normed Spaces	67
5.2	$L^p$ Space	68
5.3	Embedding Properties of $L^p$ Spaces	72
5.4	Banach Spaces	74
5.5	Bounded Linear Transformations	77
5.6	Dual of $L^p$ Spaces	79
<b>6</b>	<b>Signed and Complex Measures</b>	<b>82</b>
6.1	Signed Measures	82
6.2	Absolutely Continuous Measures	87
6.3	Lebesgue Differentiation Theorem for Regular Borel Measures	91
6.4	Monotone Differentiation Theorem	94
6.5	Functions of Bounded Variation	97
<b>7</b>	<b>Hilbert Spaces</b>	<b>104</b>
7.1	Inner Product Spaces	104
7.2	Orthonormal Bases	107

---

<b>8</b>	<b>Introduction to Fourier Analysis</b>	<b>112</b>
8.1	Fourier Series . . . . .	112
<b>A</b>	<b>Additional Proofs</b>	<b>119</b>
A.1	Measure . . . . .	119
A.2	Integration . . . . .	122
A.3	Produce Measure . . . . .	122
A.4	Differentiation on Euclidean Space . . . . .	122
A.5	Normed-Vector-Space . . . . .	122
A.6	Signed and Complex Measures . . . . .	122
A.7	Hilbert Spaces . . . . .	123
A.8	Introduction to Fourier Analysis . . . . .	123

# Chapter 1

## Measure

### Lecture 1: $\sigma$ -algebra

Before we start, we first see some examples.

05 Jan. 11:00

**Example (Finite power set).** Let  $X = \{a, b, c\}$ . Then

$$\mathcal{P}(X) := \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\},$$

which is the *power set* of  $X$ . We see that

$$\#X = n \Rightarrow \#\mathcal{P}(X) = 2^n$$

for  $n < \infty$ .

**Example (Infinite power set).** If  $n = \infty$ , say  $X = \mathbb{N}$ , then

$$\mathcal{P}(\mathbb{N})$$

is an uncountable set while  $\mathbb{N}$  is a countable set. We can see this as follows. Consider

$$\phi: \mathcal{P}(\mathbb{N}) \rightarrow [0, 1], \quad A \mapsto 0.a_1a_2a_3 \dots \text{ (base 2),}$$

where

$$a_i = \begin{cases} 1, & \text{if } i \in A \\ 0, & \text{if } i \notin A, \end{cases}$$

and for example,  $A$  can be  $A = \{2, 3, 6, \dots\} \subseteq \mathbb{N}$ . Note that  $\phi$  is surjective, hence we have

$$\#\mathcal{P}(\mathbb{N}) \geq \#[0, 1].$$

But since  $[0, 1]$  is uncountable, so is  $\mathcal{P}(\mathbb{N})$ .

We like to *measure* the *size* of subsets of  $X$ . Hence, we are intriguing to define a map  $\mu$  such that

$$\mu: \mathcal{P}(X) \rightarrow [0, \infty].$$

**Example.** Let  $X = \{0, 1, 2\}$ . Then we want to define  $\mu: \mathcal{P}(X) \rightarrow [0, \infty]$ , we can have

- $\mu(A) = \#A$ . Then we have
  - $\mu(\{0, 1\}) = 2$
  - $\mu(\{0\}) = 1$

- $\mu(A) = \sum_{i \in A} 2^i$ . Then we have  

$$- \mu(\{0, 1\}) = 2^0 + 2^1 = 3$$

**Example.** Let  $X = \{0\} \cup \mathbb{N}$ . Then we want to define  $\mu: \mathcal{P}(\mathbb{N}) \rightarrow [0, \infty]$ , we can have

- $\mu(A) = \#A$ . Then we have  

$$- \mu(\{2, 3, 4, 5, \dots\}) = \infty = \mu(\{\text{even numbers}\})$$
- $\mu(A) = e^{-1} \sum_{i \in A} \frac{1}{i!}$ . Then we have  

$$- \mu(\{0, 2, 4, 6, \dots\}) = e^{-1} \left(1 + \frac{1}{2!} + \frac{1}{3!} + \dots\right)$$
- $\mu(A) = \sum_{i \in A} a_i$

**Example.** Let  $X = \mathbb{R}$ . Then we want to define  $\mu: \mathcal{P}(\mathbb{R}) \rightarrow [0, \infty]$ , we can have

- $\mu(A) = \#A$
- $\mu((a, b)) = b - a$ .

**Problem.** Can we extend this map to all of  $\mathcal{P}(\mathbb{R})$ ?

**Answer.** No! ⊗

- $\mu((a, b)) = e^b - e^a$ .

**Problem.** Can we extend this map to all of  $\mathcal{P}(\mathbb{R})$ ?

**Answer.** No! ⊗

We immediately see the problems. To extend our native measure method into  $\mathbb{R}$  is hard and will cause something counter-intuitive!<sup>1</sup> Hence, rather than define measurement on *all* subsets in the power set of  $X$ , we only focus on *some* subsets. In other words, we want to define

$$\mu: \mathcal{P}(\mathbb{R}) \supset \mathcal{A} \rightarrow [0, \infty].$$

## 1.1 $\sigma$ -algebras

We start from the definition of the most fundamental element in measure theory.

**Definition 1.1.1 ( $\sigma$ -algebra).** Let  $X$  be a set. A collection  $\mathcal{A}$  of subsets of  $X$ , i.e.,  $\mathcal{A} \subset \mathcal{P}(X)$  is called a  $\sigma$ -algebra on  $X$  if

- $\emptyset \in \mathcal{A}$ .
- $\mathcal{A}$  is closed under complements. i.e., if  $A \in \mathcal{A}$ ,  $A^c = X \setminus A \in \mathcal{A}$ .
- $\mathcal{A}$  is closed under countable unions. i.e., if  $A_i \in \mathcal{A}$ , then  $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ .

**Remark.** There are some easy properties we can immediately derive.

- $X \in \mathcal{A}$  from  $X = X \setminus \underbrace{\emptyset}_{\in \mathcal{A}}$  and  $\mathcal{A}$  is closed under complement.

<sup>1</sup>[https://en.wikipedia.org/wiki/Banach-Tarski\\_paradox](https://en.wikipedia.org/wiki/Banach-Tarski_paradox)

- $\bigcap_{i=1}^{\infty} A_i = \left( \bigcup_{i=1}^{\infty} A_i^c \right)^c$ , namely  $\mathcal{A}$  is closed under countable intersections.
- $A_1 \cup A_2 \cup \dots \cup A_n = A_1 \cup A_2 \cup \dots \cup A_n \cup \emptyset \cup \emptyset \cup \dots$ , hence  $\mathcal{A}$  is closed under finite unions and intersections.

**Note.** The [definition of  \$\sigma\$ -algebra](#) should remind us the definition of topological basis, and this is indeed the case. We can consider a topological space and put some structure on the  $\sigma$ -algebra  $\mathcal{A}$ , which gives us the following.

**Definition 1.1.2 (Borel set).** Given a topological space  $X$ , a *Borel set* is any set in  $X$  that can be formed from open sets through the operations of countable union, countable intersection and relative complement.

## Lecture 2: Measure

**Example.** We first see some examples.

- (1) Let  $\mathcal{A} = \mathcal{P}(X)$ , which is the power  $\sigma$ -algebra.
- (2) Let  $\mathcal{A} = \{\emptyset, X\}$ , which is a trivial  $\sigma$ -algebra.
- (3) Let  $B \subset X$ ,  $B \neq \emptyset$ ,  $B \neq X$ . Then we see that  $\mathcal{A} = \{\emptyset, B, B^c, X\}$  is a  $\sigma$ -algebra.

07 Jan. 11:00

**Lemma 1.1.1.** Let  $\mathcal{A}_\alpha$ ,  $\alpha \in I$ , be a family of  $\sigma$ -algebra on  $X$ . Then

$$\bigcap_{\alpha \in I} \mathcal{A}_\alpha$$

is a  $\sigma$ -algebra on  $X$ .

**Proof.** A simple proof can be made as follows. Firstly,  $\emptyset \in \mathcal{A}_\alpha$  for every  $\alpha$  clearly. Moreover, closure under complement and countable unions for every  $\mathcal{A}_\alpha$  implies the same must be true for  $\bigcap_{\alpha \in I} \mathcal{A}_\alpha$ . Hence,  $\bigcap_{\alpha \in I} \mathcal{A}_\alpha$  is a  $\sigma$ -algebra.

**Remark.** Notice that  $I$  may be an uncountable intersection. ■

The above allows us to give the following definition.

**Definition 1.1.3 (Generation of  $\sigma$ -algebra).** Given  $\mathcal{E} \subset \mathcal{P}(X)$ , where  $\mathcal{E}$  is not necessarily a  $\sigma$ -algebra. Let  $\langle \mathcal{E} \rangle$  be the intersection of all  $\sigma$ -algebras on  $X$  containing  $\mathcal{E}$ , then we call  $\langle \mathcal{E} \rangle$  the  $\sigma$ -algebra generated by  $\mathcal{E}$ .

**Remark.** Clearly,  $\langle \mathcal{E} \rangle$  is the smallest  $\sigma$ -algebra containing  $\mathcal{E}$ , and it is unique. To check the uniqueness, we suppose there are two different  $\langle \mathcal{E} \rangle_1$  and  $\langle \mathcal{E} \rangle_2$  generated from  $\mathcal{E}$ . It's easy to show

$$\langle \mathcal{E} \rangle_1 \subseteq \langle \mathcal{E} \rangle_2,$$

and by symmetry, they are equal.

**Example.** We see that  $\{\emptyset, B, B^c, X\} = \langle \{B\} \rangle = \langle \{B^c\} \rangle$ .

**Lemma 1.1.2.** We have

- (1) Given  $\mathcal{A}$  a  $\sigma$ -algebra,  $\mathcal{E} \subset \mathcal{A} \subset \mathcal{P}(X) \Rightarrow \langle \mathcal{E} \rangle \subset \mathcal{A}$
- (2)  $\mathcal{E} \subset \mathcal{F} \subset \mathcal{P}(X) \Rightarrow \langle \mathcal{E} \rangle \subset \langle \mathcal{F} \rangle$

**Proof.** We'll see that after proving the first claim, the second follows smoothly.

- (1) The first claim is trivial, since we know that  $\langle \mathcal{E} \rangle$  is the smallest  $\sigma$ -algebra containing  $\mathcal{E}$ , then if  $\mathcal{E} \subset \mathcal{A}$ , we clearly have  $\langle \mathcal{E} \rangle \subset \mathcal{A}$  by the definition.
- (2) The second claim is also easy. From the first claim and the definition, we have

$$\mathcal{E} \subset \mathcal{F} \subset \langle \mathcal{F} \rangle \Rightarrow \langle \mathcal{E} \rangle \subset \langle \mathcal{F} \rangle.$$

■

At this point, we haven't put any specific structure on  $X$ . Now we try to describe those spaces with good structure, which will give the space some nice properties.

**Definition 1.1.4 (Borel  $\sigma$ -algebra).** For a topological space  $X$ , the *Borel  $\sigma$ -algebra on  $X$* , denoted as  $\mathcal{B}(X)$ , is the  $\sigma$ -algebra generated by the collection of all open sets in  $X$ .

**Example.** We see that  $\mathcal{B}(\mathbb{R})$  contains

- $\mathcal{E}_1 = \{(a, b) \mid a < b; a, b \in \mathbb{R}\}$ .
- $\mathcal{E}_2 = \{[a, b] \mid a < b; a, b \in \mathbb{R}\}$  since  $[a, b] = \bigcap_{n=1}^{\infty} (a - \frac{1}{n}, b + \frac{1}{n})$ .
- $\mathcal{E}_3 = \{(a, b] \mid a < b; a, b \in \mathbb{R}\}$  since  $(a, b] = \bigcap_{n=1}^{\infty} (a, b + \frac{1}{n})$ .
- $\mathcal{E}_4 = \{[a, b) \mid a < b; a, b \in \mathbb{R}\}$  since  $[a, b) = \bigcap_{n=1}^{\infty} (a - \frac{1}{n}, b)$ .
- $\mathcal{E}_5 = \{(a, \infty) \mid a \in \mathbb{R}\}$  since  $(a, \infty) = \bigcup_{n=1}^{\infty} (a, a + n)$ .
- $\mathcal{E}_6 = \{[a, \infty) \mid a \in \mathbb{R}\}$  since  $[a, \infty) = \bigcup_{n=1}^{\infty} [a, a + n)$ .
- $\mathcal{E}_7 = \{(-\infty, b) \mid b \in \mathbb{R}\}$  since  $(-\infty, b) = \bigcup_{n=1}^{\infty} (b - n, b)$ .
- $\mathcal{E}_8 = \{(-\infty, b] \mid b \in \mathbb{R}\}$  since  $(-\infty, b] = \bigcup_{n=1}^{\infty} (b - n, b]$ .

**Proposition 1.1.1.**  $\mathcal{B}(\mathbb{R}) = \langle \mathcal{E}_i \rangle$  for each  $i = 1, \dots, 8$  in the above example.

**Proof.** Firstly, we see that  $\mathcal{E}_i \subset \mathcal{B}(\mathbb{R}) \Rightarrow \langle \mathcal{E}_i \rangle \subset \mathcal{B}(\mathbb{R})$  by Lemma 1.1.2. Secondly, by definition,  $\mathcal{B}(\mathbb{R}) = \langle \mathcal{E} \rangle$  where

$$\mathcal{E} = \{O \subseteq \mathbb{R} \mid O \text{ is open in } \mathbb{R}\}.$$

It's enough to show  $\mathcal{E} \subset \langle \mathcal{E}_i \rangle$  since if so,  $\langle \mathcal{E} \rangle \subseteq \langle \mathcal{E}_i \rangle$ , and clearly  $\langle \mathcal{E} \rangle \supseteq \langle \mathcal{E}_i \rangle = \mathcal{B}(\mathbb{R})$ , then we will have  $\langle \mathcal{E} \rangle = \langle \mathcal{E}_i \rangle$ . Let  $O \subset \mathbb{R}$  be an open set, i.e.,  $O \in \mathcal{E}$ . We claim that every open set in  $\mathbb{R}$  is a countable union of disjoint open intervals.<sup>a</sup>



Thus,

$$O = \bigcup_{j=1}^{\infty} I_j,$$

where  $I_j$  open interval with the form of  $(a, b), (-\infty, b), (a, \infty), (-\infty, \infty)$ .

For example,  $\mathcal{E}_1$  is trivially true, and

$$(a, b) = \bigcup_{n=1}^{\infty} \underbrace{\left[ a + \frac{1}{n}, b - \frac{1}{n} \right]}_{\in \mathcal{E}_2} \underbrace{\hspace{10em}}_{\in \langle \mathcal{E}_2 \rangle}$$

shows the case for  $\mathcal{E}_2$  and

$$(a, \infty) = \bigcup_{k=1}^{\infty} (a, a + k)$$

shows the case for  $\mathcal{E}_5$ . It's now straightforward to check open intervals are in  $\langle \mathcal{E}_i \rangle$  for every  $i$ . ■

<sup>a</sup><https://math.stackexchange.com/questions/318299/any-open-subset-of-bbb-r-is-a-countable-union-of-disjoint-open-intervals>

Now, to put a structure on a space, we define the following.

**Definition.** Given a space  $X$ , we have the following.

**Definition 1.1.5 (Measurable space).** A *measurable space* is a tuple of  $X$  and a  $\sigma$ -algebra  $\mathcal{A}$  on  $X$ , denoted by  $(X, \mathcal{A})$ .

In particular, if the  $\sigma$ -algebra is the **Borel  $\sigma$ -algebra** of  $X$ , then we give it a special name.

**Definition 1.1.6 (Borel space).** A *Borel space* is a tuple of  $X$  and  $\mathcal{B}(X)$ , denoted by  $(X, \mathcal{B}(X))$ .

**Remark.** This means that  $X$  implicitly has some topological structure.

**Definition 1.1.7 ( $\mathcal{A}$ -measurable set).** Given a **measurable space**  $(X, \mathcal{A})$ , every  $E \in \mathcal{A}$  is a so-called  *$\mathcal{A}$ -measurable set*.

## 1.2 Measures

With the definition of **measurable space**, we now can refine our **measure** function  $\mu$  as follows.

**Definition 1.2.1 (Measure).** Given a **measurable space** on  $(X, \mathcal{A})$ , a *measure* is a function

$$\mu: \mathcal{A} \rightarrow [0, \infty]$$

such that

- (null empty set)  $\mu(\emptyset) = 0$ .
- (countable additivity)  $\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$  if  $A_1, A_2, \dots \in \mathcal{A}$  are disjoint.

**Definition 1.2.2 (Measure space).** We denote  $(X, \mathcal{A}, \mu)$  as so-called a *measure space* given  $\mu$  is the **measure** on  $(X, \mathcal{A})$ .

**Notation.** We denote  $[0, \infty] := [0, \infty) \cup \{\infty\}$ .

**Remark.** We only want **countable additivity** but not uncountable additivity.

**Proof.** Consider the most intuitive **measure** on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

Since we have

$$(0, 1] = (1/2, 1] \cup (1/4, 1/2] \cup (1/8, 1/4] \cup \dots$$

and also

$$(0, 1] = \bigcup_{x \in (0, 1]} \{x\}.$$

Specifically, in the first case, we are claiming that

$$1 = \underbrace{\frac{1}{2}}_{\mu((\frac{1}{2}, 1])} + \underbrace{\frac{1}{4}}_{\mu((\frac{1}{4}, \frac{1}{2}])} + \underbrace{\frac{1}{8}}_{\mu((\frac{1}{8}, \frac{1}{4}])} + \dots;$$

while in the second case, we are claiming that

$$1 = \sum_{x \in (0, 1]} 0$$

since  $\mu(x) = 0$  for  $x \in \mathbb{R}$ , which is clearly not what we want. \*

**Example (Counting measure).** For any  $(X, \mathcal{A})$ , we let  $\mu(A) := \#A$ . This is the so-called *counting measure*.

**Example (Dirac-Delta measure).** Let  $x_0 \in X$ . For any  $(X, \mathcal{A})$ , the *Dirac-Delta measure* at  $x_0$  is

$$\mu(A) = \begin{cases} 1, & \text{if } x_0 \in A; \\ 0, & \text{if } x_0 \notin A \end{cases}$$

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

**Example.** For  $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ , given  $A \in \mathcal{P}(\mathbb{N})$ ,

$$\mu(A) = \sum_{i \in A} a_i$$

where  $a_1, a_2, \dots \in [0, \infty)$ .

## Lecture 3: Construct a Measure

After seeing examples of **measures**, we now want to construct it from ground up.

10 Jan. 11:00

**Note.** If  $A, B \in \mathcal{A}$  and  $A \subset B$ , then

$$\mu(B \setminus A) + \mu(A) = \mu(B) \Rightarrow \mu(B \setminus A) = \mu(B) - \mu(A) \text{ if } \mu(A) < \infty.$$

**Theorem 1.2.1.** Given  $(X, \mathcal{A}, \mu)$  be a **measure space**. Then the following hold.

(1) Monotonicity.

$$A, B \in \mathcal{A}, A \subset B \Rightarrow \mu(A) \leq \mu(B).$$

(2) Countable subadditivity.

$$A_1, A_2, \dots \in \mathcal{A} \Rightarrow \mu\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

(3) Continuity from below.

$$\begin{cases} A_1, A_2, \dots \in \mathcal{A} \\ A_1 \subset A_2 \subset A_3 \subset \dots \end{cases} \Rightarrow \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

(4) Continuity from above.

$$\begin{cases} A_1, A_2, \dots \in \mathcal{A} \\ A_1 \supset A_2 \supset A_3 \supset \dots \\ \mu(A_1) < \infty \end{cases} \Rightarrow \mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

**Proof.** We prove this theorem one by one.

(1) Since  $A \subset B$ , hence we have

$$\mu(B) = \mu\left(\underbrace{(B \setminus A) \cup A}_{\text{disjoint}}\right) \stackrel{!}{=} \underbrace{\mu(B \setminus A)}_{\geq 0} + \mu(A) \geq \mu(A).$$

(2) This should be trivial from [countable additivity](#) with the fact that  $\mu(A) \geq 0$  for all  $A$ .

DIY!

(3) Let  $B_1 = A_1$ ,  $B_i = A_i \setminus A_{i-1}$  for  $i \geq 2$ , then

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$$

are a disjoint union and  $B_i \in \mathcal{A}$ , hence we see that

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \mu(B_i).$$

With  $\mu\left(\bigcup_{i=1}^n B_i\right) = \mu(A_n)$ , we have

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \mu(B_i) = \lim_{n \rightarrow \infty} \mu\left(\bigcup_{i=1}^n B_i\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

(4) Let  $E_i = A_1 \setminus A_i \Rightarrow E_i \in \mathcal{A}$ ,  $E_1 \subset E_2 \subset \dots$ . We then have

$$\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} (A_1 \setminus A_i) = A_1 \setminus \left(\bigcap_{i=1}^{\infty} A_i\right),$$

which implies

$$\bigcap_{i=1}^{\infty} A_i = A_1 \setminus \left(\bigcup_{i=1}^{\infty} E_i\right) \Rightarrow \mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \mu(A_1) - \mu\left(\bigcup_{i=1}^{\infty} E_i\right)$$

since  $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \mu(A_1) < \infty$ . Then from [continuity from below](#), we further have

$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \mu(A_1) - \lim_{n \rightarrow \infty} \mu(E_n) = \mu(A_1) - \lim_{n \rightarrow \infty} (\mu(A_1) - \mu(A_n)).$$

From [monotonicity](#), we see that  $\mu(A_n) \leq \mu(A_1) < \infty$ , hence we can split the limit and further get

$$\mu\left(\bigcap_{i=1}^{\infty} A_i\right) = \mu(A_1) - \mu(A_1) + \lim_{n \rightarrow \infty} \mu(A_n) = \lim_{n \rightarrow \infty} \mu(A_n).$$

■

**Note.** Sometimes we also call [continuity from below](#) property as monotone convergence theorem for sets. We'll later see the important [monotone convergence theorem](#) for integral, which is in different content.

**Remark** (Condition of continuity from above). The condition  $\mu(A_1) < \infty$  in [continuity from above](#) is necessary.

**Proof.** Given  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$  where  $\mu$  is the [counting measure](#). Then we see

- $A_n = \{n, n+1, n+2, \dots\} \Rightarrow \mu(A_n) = \infty$
- $A_1 \supset A_2 \supset A_3 \supset \dots$
- $\bigcap_{i=1}^{\infty} A_i = \emptyset \Rightarrow \mu\left(\bigcap_{i=1}^{\infty} A_i\right) = 0$

We see that in this case, since  $\mu(A_1) \not< \infty$ , hence [continuity from above](#) doesn't hold. ⊛

We now try to characterize some properties of a [measure space](#).

**Definition.** Given  $(X, \mathcal{A}, \mu)$ , we have the following.

**Definition 1.2.3** ( $\mu$ -null set).  $A \subset X$  is a  $\mu$ -null set if  $A \in \mathcal{A}$  and  $\mu(A) = 0$ .

**Definition 1.2.4** ( $\mu$ -subnull set).  $A \subset X$  is a  $\mu$ -subnull set if there exists a  $\mu$ -null set  $B$  such that  $A \subset B$ .

**Definition 1.2.5** (Complete measure space).  $(X, \mathcal{A}, \mu)$  is a complete [measure space](#) if every  $\mu$ -subnull set is  $\mathcal{A}$ -measurable.

**Note.** We see that for a  $\mu$ -subnull set, it's not necessary  $\mathcal{A}$ -measurable if the [measure space](#) is not complete.

**Remark.** From the property of [measure](#), the condition for a [measure space](#)  $(X, \mathcal{A}, \mu)$  being complete is equivalent to saying that every  $\mu$ -subnull set is a  $\mu$ -null set.

**Proof.** This follows from the [monotonicity](#) of a [measure](#) and the fact that a [measure](#) is always non-negative. Finally, a  $\mu$ -subnull set is always in  $\mathcal{A}$ . ⊛

There are some useful terminologies we'll use later relating to  $\mu$ -null.

**Definition 1.2.6** (Almost everywhere). Given  $(X, \mathcal{A}, \mu)$ , a statement  $P(x)$ ,  $x \in X$  holds  $\mu$ -almost everywhere (a.e.) if the set

$$\{x \in X : P(x) \text{ does not hold}\}$$

is  $\mu$ -null.

It's always pleasurable working with finite rather than infinite, hence we give the following definition.

**Definition.** Given  $(X, \mathcal{A}, \mu)$ , we have the following.

**Definition 1.2.7** (Finite measure).  $\mu$  is a *finite measure* if  $\mu(X) < \infty$ .

**Definition 1.2.8** ( $\sigma$ -finite measure).  $\mu$  is a  $\sigma$ -finite measure if  $X = \bigcup_{n=1}^{\infty} X_n$ ,  $X_n \in \mathcal{A}$ ,  $\mu(X_n) < \infty$ .

**Exercise.** Every *measure space* can be *completed*. Namely, we can always find a bigger  $\sigma$ -algebra to *complete* the space.

### 1.3 Outer Measures

As we said before, we're now going to construct a *measure*. And a modern way to do this is to start with something called *outer measure*.

**Definition 1.3.1** (Outer measure). An *outer measure* on  $X$  is a function

$$\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$$

such that

- (null empty set)  $\mu^*(\emptyset) = 0$ .
- (monotonicity)  $\mu^*(A) \leq \mu^*(B)$  if  $A \subset B$ .
- (countable subadditivity)  $\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$  for every  $A_i \subset X$ .

**Example.** For  $A \subset \mathbb{R}$ ,

$$\mu^*(A) = \inf \left\{ \sum_{i=1}^{\infty} (b_i - a_i) : \bigcup_{i=1}^{\infty} (a_i, b_i) \supset A \right\}$$

is an *outer measure*.

**Proof.** This follows directly from the [Proposition 1.3.1](#) we're going to show. ⊛

**Remark.** We see that an *outer measure* need not be a *measure*.

**Proposition 1.3.1.** Let  $\mathcal{E} \subset \mathcal{P}(X)$  such that  $\emptyset, X \in \mathcal{E}$ . Let

$$\rho : \mathcal{E} \rightarrow [0, \infty]$$

such that  $\rho(\emptyset) = 0$ . Then

$$\mu^*(A) := \inf \left\{ \sum_{i=1}^{\infty} \rho(E_i) : \forall_{i \in \mathbb{N}} E_i \in \mathcal{E}, \bigcup_{i=1}^{\infty} E_i \supset A \right\}$$

is an **outer measure** on  $X$ .

**Theorem 1.3.1** (Tonelli's Theorem for series). Recall the Tonelli's Theorem<sup>a</sup> for series, i.e., if  $a_{ij} \in [0, \infty]$ ,  $\forall i, j \in \mathbb{N}$ , then

$$\sum_{(i,j) \in \mathbb{N}^2} a_{ij} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} a_{ij} = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} a_{ij}.$$

<sup>a</sup>[https://en.wikipedia.org/wiki/Fubini%27s\\_theorem](https://en.wikipedia.org/wiki/Fubini%27s_theorem)

**Proof.** Read Tao[Tao13] Theorem 0.0.2. ■

## Lecture 4: Carathéodory Extension Theorem

**As previously seen.** Last time we skip the proof of **Proposition 1.3.1**, which is a quite important theorem for building a **measure**. To see this, we note that from **Proposition 1.3.1**, given a positive function  $\rho$  defined on a subset of the power set of  $X$  with  $\rho(\emptyset) = 0$ , we can induce an **outer measure** from  $\rho$ .

12 Jan. 11:00

**Note.** We'll see later that how can we further induce a natural **measure** from the induced **outer measure**.

We now prove **Proposition 1.3.1**.

**Proof of Proposition 1.3.1.** We need to prove the following.

**Claim.**  $\mu^*$  is well-defined, i.e., inf is taken over a non-empty set.

**Proof.** This is trivial since  $X \in \mathcal{E}$  and  $X \supset A$  for any  $A \in \mathcal{E}$ . ⊛

**Claim.** Null empty set holds, i.e.,  $\mu^*(\emptyset) = 0$ .

**Proof.** Since  $\emptyset \in \mathcal{E}$  and

$$\mu^*(\emptyset) = \inf \left\{ \sum_{i=1}^{\infty} \rho(E_i) : \forall_{i \in \mathbb{N}} E_i \in \mathcal{E}, \bigcup_{i=1}^{\infty} E_i \supset \emptyset \right\} = 0$$

since  $\rho(\emptyset) = 0$  for all  $i$  and further, by Squeeze Theorem, we see that  $\lim_{n \rightarrow \infty} \sum_{i=1}^n \rho(\emptyset) = 0$ . ⊛

**Claim.** Monotonicity holds, i.e.,  $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$ .

**Proof.** We show this by contradiction. Suppose  $A \subset B$  and  $\mu^*(A) > \mu^*(B)$ , then by definition of  $\mu^*$ , we have

$$\begin{aligned}\mu^*(A) &= \inf \left\{ \sum_{i=1}^{\infty} \rho(E_i) : \forall_{i \in \mathbb{N}} E_i \in \mathcal{E}, \bigcup_{i=1}^{\infty} E_i \supset A \right\} \\ &> \inf \left\{ \sum_{i=1}^{\infty} \rho(E_i) : \forall_{i \in \mathbb{N}} E_i \in \mathcal{E}, \bigcup_{i=1}^{\infty} E_i \supset B \right\} = \mu^*(B).\end{aligned}$$

Now, let  $B =: (B \setminus A) \cup A$ , then we have

$$\begin{aligned}\mu^*(A) &= \inf \left\{ \sum_{i=1}^{\infty} \rho(E_i) : \forall_{i \in \mathbb{N}} E_i \in \mathcal{E}, \bigcup_{i=1}^{\infty} E_i \supset A \right\} \\ &> \inf \left\{ \sum_{i=1}^{\infty} \rho(E_i) : \forall_{i \in \mathbb{N}} E_i \in \mathcal{E}, \bigcup_{i=1}^{\infty} E_i \supset (B \setminus A) \cup A \right\} = \mu^*(B).\end{aligned}$$

Now, since  $B \setminus A \supseteq \emptyset$ , then this inequality can't hold, hence a contradiction  $\nexists$ .  $\otimes$

**Claim.** Countable subadditivity holds, i.e.,  $\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$  for every  $A_i \subset X$ .

**Proof.** Let  $A_1, A_2, \dots \in X$ . If one of  $\mu^*(A_n) = \infty$ , then result holds. So we may assume  $\mu^*(A_n) < \infty$  for all  $n \in \mathbb{N}$ . Now, fix any  $\epsilon > 0$ , we will show that

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n) + \epsilon.$$

For each  $n \in \mathbb{N}$ ,  $\exists E_{n,1}, E_{n,2}, \dots \in \mathcal{E}$  such that  $\bigcup_{k=1}^{\infty} E_{n,k} \supset A_n$  and  $\mu^*(A_n) + \epsilon/2^n \geq \sum_{k=1}^{\infty} \rho(E_{n,k})$ .

**Remark.** This is an important trick! We often set the error term as  $\epsilon/2^n$  instead of  $\epsilon$  as in above to accommodate the summation over a countable set.

Then we see that

$$\bigcup_{k=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} E_{k,n} = \bigcup_{(n,k) \in \mathbb{N}^2} E_{k,n},$$

which implies

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{(n,k) \in \mathbb{N}^2} \rho(E_{k,n}) \stackrel{!}{=} \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \rho(E_{k,n}) \leq \sum_{n=1}^{\infty} \left( \mu^*(A_n) + \frac{\epsilon}{2^n} \right)$$

from the inequality just derived. Now, since the last term is just

$$\sum_{n=1}^{\infty} \left( \mu^*(A_n) + \frac{\epsilon}{2^n} \right) = \sum_{n=1}^{\infty} \mu^*(A_n) + \epsilon,$$

hence we finally have

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n) + \epsilon$$

for arbitrarily small fixed  $\epsilon > 0$ , hence the subadditivity is proved.  $\otimes$

■

**Definition 1.3.2** (Carathéodory measurable). Let  $\mu^*$  be an **outer measure** on  $X$ . We say  $A \subset X$  is *Carathéodory measurable with respect to  $\mu^*$*  if

$$\forall E \subset X, \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A).$$

**Note.** We sometimes write *C-measurable* instead of **Carathéodory measurable** for convenience.

**Lemma 1.3.1.** Let  $\mu^*$  be an **outer measure** on  $X$ . Suppose  $B_1, \dots, B_N$  are disjoint **C-measurable** sets. Then,

$$\forall E \subset X, \mu^*\left(E \cap \left(\bigcup_{i=1}^N B_i\right)\right) = \sum_{i=1}^N \mu^*(E \cap B_i).$$

**Proof.** Since we have

$$\begin{aligned} \mu^*\left(E \cap \left(\bigcup_{i=1}^N B_i\right)\right) &= \mu^*(E' \cap B_1) + \mu^*(E' \setminus B_1) \\ &= \mu^*\left(E \cap \left(\bigcup_{i=1}^N B_i \cap B_1\right)\right) + \mu^*\left(E \cap \left(\bigcup_{i=1}^N B_i\right) \cap B_1^c\right) \\ &= \mu^*(E \cap B_1) + \mu^*\left(E \cap \left(\bigcup_{i=2}^N B_i\right)\right) \end{aligned}$$

where the equality comes from the fact that  $B_1$  is **C-measurable** and disjoint from  $B_i, i \neq 1$ . Then, we simply iterate this argument and have the result. Note that in the first inequality, we define  $E' := E \cap \left(\bigcup_{i=1}^N B_i\right)$  for the simplicity of notation. ■

**Remark.** This implies that if we restrict an **outer measure** on a **C-measurable** set, then it becomes finite additive.

**Theorem 1.3.2** (Carathéodory extension Theorem). Let  $\mu^*$  be an **outer measure** on  $X$ . Let  $\mathcal{A}$  be the collection of **C-measurable** sets (with respect to  $\mu^*$ ). Then,

- (1)  $\mathcal{A}$  is a  **$\sigma$ -algebra** on  $X$ .
- (2)  $\mu = \mu^*|_{\mathcal{A}}$  is a **measure** on  $(X, \mathcal{A})$ .
- (3)  $(X, \mathcal{A}, \mu)$  is a **complete measure space**.

**Proof.** We divide the proof in several steps.

- (1) We show  $\mathcal{A}$  is a  **$\sigma$ -algebra** by showing

- We first show  $\emptyset \in \mathcal{A}$ .

**Claim.**  $\emptyset \in \mathcal{A}$ .

**Proof.** To show this, we simply check that  $\emptyset$  is **C-measurable**. We see that

$$\forall_{E \subset X} \mu^*(E) = \mu^*(E \cap \emptyset) + \mu^*(E \setminus \emptyset) = \mu^*(E),$$

which just shows  $\emptyset \in \mathcal{A}$ . ⊛

- Then we show  $\mathcal{A}$  is closed under complements.

**Claim.**  $\mathcal{A}$  closed under complements.



**Proof.** This is equivalent to say that if  $A$  is **C-measurable**, so is  $A^c$ . We see that if  $A$  is **C-measurable**, then for every  $E \subset X$ ,

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A).$$

Observing that  $E \cap A = E \setminus A^c$  and  $E \setminus A = E \cap A^c$ , hence

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

We immediately see that above implies  $A^c \in \mathcal{A}$ . ⊗

- We now show  $\mathcal{A}$  is closed under countable unions.

**Note.** To show  $\mathcal{A}$  closed under countable unions, we show that  $\mathcal{A}$  is closed under:

$$\text{finite unions} \xrightarrow{\text{then}} \text{countable } \underline{\text{disjoint}} \text{ unions} \xrightarrow{\text{then}} \text{countable unions}.$$

Hence, we first show  $\mathcal{A}$  is closed under finite unions.

**Claim.**  $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$ .

**Proof.** Fix  $E \subset X$  arbitrary. We need to show that

$$\mu^*(E) = \mu^*(E \cap (A \cup B)) + \mu^*(E \setminus (A \cup B)),$$

i.e.,  $\mu^*(1 \cup 2 \cup 3 \cup 4) = \mu^*(1 \cup 2 \cup 3) + \mu^*(4)$  given  $A, B \in \mathcal{A}$  and the following figure.



- Since  $A$  is **C-measurable**,
  - \*  $\mu^*(1 \cup 2 \cup 3 \cup 4) = \mu^*(1 \cup 2) + \mu^*(3 \cup 4)$
  - \*  $\mu^*(1 \cup 2 \cup 3) = \mu^*(1 \cup 2) + \mu^*(3)$
- Since  $B$  is **C-measurable**,
  - \*  $\mu^*(3 \cup 4) = \mu^*(3) + \mu^*(4)$

Hence, we have

$$\begin{aligned} \mu^*(1 \cup 2 \cup 3 \cup 4) &= \mu^*(1 \cup 2) + \mu^*(3 \cup 4) \\ &= \mu^*(1 \cup 2) + \mu^*(3) + \mu^*(4) = \mu^*(1 \cup 2 \cup 3) + \mu^*(4). \end{aligned}$$

⊗

We now show  $\mathcal{A}$  is closed under countable disjoint unions.

**Claim.**  $\mathcal{A}$  is closed under countable disjoint unions.

**Proof.** Let  $A_1, A_2, \dots \in \mathcal{A}$  and disjoint. Fix  $E \subset X$  arbitrary. Since  $\mu^*$  is countably subadditive,

$$\mu^*(E) \leq \mu^*\left(E \cap \bigcup_{i=1}^{\infty} A_i\right) + \mu^*\left(E \setminus \bigcup_{i=1}^{\infty} A_i\right),$$

hence we only need to show another way around.

Fix  $N \in \mathbb{N}$ , we have  $\bigcup_{n=1}^N A_n \in \mathcal{A}$  since  $N$  is finite, and

$$\begin{aligned} \mu^*(E) &= \mu^*\left(E \cap \left(\bigcup_{n=1}^N A_n\right)\right) + \mu^*\left(E \setminus \left(\bigcup_{n=1}^N A_n\right)\right) \\ &\geq \underbrace{\sum_{n=1}^N \mu^*(E \cap A_n)}_{\stackrel{!}{=} \mu^*\left(E \cap \left(\bigcup_{n=1}^N A_n\right)\right)} + \underbrace{\mu^*\left(E \setminus \bigcup_{n=1}^{\infty} A_n\right)}_{\leq \mu^*\left(E \setminus \left(\bigcup_{n=1}^N A_n\right)\right)}. \end{aligned}$$

Now, take  $N \rightarrow \infty$  then we are done. ⊗

We can then extend this to the case of countable unions.

**Exercise.** Show  $\mathcal{A}$  is closed under countable unions.

**Answer.** \_\_\_\_\_ ⊗

DIY

Above shows that  $\mathcal{A}$  is a  $\sigma$ -algebra.

The proof will be continued...

## Lecture 5: Hahn-Kolmogorov Theorem

Firstly, we see a stronger version of Lemma 1.3.1 we have seen before.

14 Jan. 11:00

**Lemma 1.3.2.** Let  $\mu^*$  be an outer measure on  $X$ . Suppose  $B_1, B_2, \dots$  are disjoint C-measurable sets. Then,

$$\forall E \subset X, \mu^*\left(E \cap \left(\bigcup_{i=1}^{\infty} B_i\right)\right) = \sum_{i=1}^{\infty} \mu^*(E \cap B_i).$$

**Proof.**

$$\sum_{n=1}^{\infty} \mu^*(E \cap B_i) \geq \mu^*\left(E \cap \bigcup_{n=1}^{\infty} B_n\right) \geq \mu^*\left(E \cap \left(\bigcup_{n=1}^N B_n\right)\right) \stackrel{!}{=} \sum_{n=1}^N \mu^*(E \cap B_n).$$

Now, we just take  $N \rightarrow \infty$  and since  $N \in \mathbb{N}$  is arbitrary, we then get the result according to Squeeze Theorem. ■

Let's continue the proof of Theorem 1.3.2.

**Proof of Theorem 1.3.2 (cont.)** The 1. is proved, now we prove 2. and 3.

2. Since from Definition 1.2.1, to show  $\mu$  is a measure, we need to show the following.

- Null empty set property.

**Claim.**  $\mu(\emptyset) = 0$

**Proof.** This means that we need to show  $\mu^*|_{\mathcal{A}}(\emptyset) = 0$ . Since  $\emptyset \in \mathcal{A}$  and  $\mu^*$  is an outer measure, hence from the property of outer measure, it clearly holds. ⊗

- **Countable additivity** property.

**Claim.**  $\mu^*$  on  $\mathcal{A}$  has **Countable additivity** property.

**Proof.** It follows from [Lemma 1.3.2](#) with  $E = X$  ⊗

3. The proof is given in [Theorem A.1.1](#). ■

## 1.4 Hahn-Kolmogorov Theorem

We see that we can start with any collection of open sets  $\mathcal{E}$  and any  $\rho$  such that it assigns **measure** on  $\mathcal{E}$ , then it induces an **outer measure** by [Proposition 1.3.1](#), finally **complete** the **outer measure** by [Theorem 1.3.2](#).

Specifically, we have

$$(\mathcal{E}, \rho) \xrightarrow{\text{Proposition 1.3.1}} (\mathcal{P}(X), \mu^*) \xrightarrow{\text{Theorem 1.3.2}} (\mathcal{A}, \mu)$$

To introduce this concept, we see that we can start with a more general definition compared to  **$\sigma$ -algebra** we are working on till now.

**Definition 1.4.1 (Algebra).** Let  $X$  be a set. A collection  $\mathcal{A}$  of subsets of  $X$ , i.e.,  $\mathcal{A} \subset \mathcal{P}(X)$  is called an *algebra on  $X$*  if

- $\emptyset \in \mathcal{A}$ .
- $\mathcal{A}$  is closed under complements. i.e., if  $A \in \mathcal{A}$ ,  $A^c = X \setminus A \in \mathcal{A}$ .
- $\mathcal{A}$  is closed under **finite** unions. i.e., if  $A_i \in \mathcal{A}$ , then  $\bigcup_{i=1}^n A_i \in \mathcal{A}$  for  $n < \infty$ .

**Remark.** The only difference between an **algebra** and a  **$\sigma$ -algebra** is whether they closed under **countable** unions in the definition.

Now, we can look at a more general setup compared to an **outer measure**.

**Definition 1.4.2 (Pre-measure).** Let  $\mathcal{A}_0$  be an **algebra** on  $X$ . A *pre-measure* on  $X$  with respect to  $\mathcal{A}_0$  is a function

$$\mu_0: \mathcal{A}_0 \rightarrow [0, \infty]$$

such that

- (null empty set)  $\mu_0(\emptyset) = 0$
- (finite additivity)  $\mu_0\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mu_0(A_i)$  if  $A_1, \dots, A_n \in \mathcal{A}_0$  are disjoint.
- (countable additivity within the **algebra**) If  $A \in \mathcal{A}_0$  and  $A = \bigcup_{n=1}^{\infty} A_n$ ,  $A_n \in \mathcal{A}_0$ , disjoint, then

$$\mu_0(A) = \sum_{n=1}^{\infty} \mu_0(A_n).$$

**Lemma 1.4.1.** The **null empty set** property and **countable additivity within the algebra** implies **finite additivity** in [Definition 1.4.2](#).

**Proof.** It's easy to see that since  $\mu_0$  is monotone. ■

**Theorem 1.4.1** (Hahn-Kolmogorov Theorem). Let  $\mu_0$  be a pre-measure on algebra  $\mathcal{A}_0$  on  $X$ . Let  $\mu^*$  be the outer measure induced by  $(\mathcal{A}_0, \mu_0)$  in Proposition 1.3.1. Let  $\mathcal{A}$  and  $\mu$  be the Carathéodory  $\sigma$ -algebra and measure for  $\mu^*$ , then  $(\mathcal{A}, \mu)$  extends  $(\mathcal{A}_0, \mu_0)$ , i.e.,

$$\mathcal{A} \supset \mathcal{A}_0, \quad \mu|_{\mathcal{A}_0} = \mu_0.$$

**Proof.** We prove this theorem in two parts. We first show that  $\mathcal{A} \supset \mathcal{A}_0$ .

**Claim.**  $\mathcal{A} \supset \mathcal{A}_0$ .

**Proof.** Let  $A \in \mathcal{A}_0$ , we want to show  $A \in \mathcal{A}$ , i.e.,  $A$  is  $\mathcal{C}$ -measurable, i.e.,

$$\forall E \subset X \quad \mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

We first fix an  $E \subset X$ . From countable subadditivity of  $\mu^*$ , we have

$$\mu^*(E) \leq \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Hence, we only need to show another direction. If  $\mu^*(E) = \infty$ , then  $\mu^*(E) = \infty \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$  clearly. So, assume  $\mu^*(E) < \infty$ .

Fix  $\epsilon > 0$ . By the Proposition 1.3.1 of  $\mu^*$ ,  $\exists B_1, B_2, \dots \in \mathcal{A}_0$ ,  $\bigcup_{n=1}^{\infty} B_n \supset E$  such that

$$\mu^*(E) + \epsilon \geq \sum_{n=1}^{\infty} \mu_0(B_n) = \sum_{n=1}^{\infty} \left( \underbrace{\mu_0(B_n \cap A)}_{\in \mathcal{A}_0} + \underbrace{\mu_0(B_n \cap A^c)}_{\in \mathcal{A}_0} \right)$$

by the finite additivity of  $\mu_0$ . Note that

$$\left\{ \begin{array}{l} \bigcup_{n=1}^{\infty} (B_n \cap A) \supset E \cap A \\ \bigcup_{n=1}^{\infty} (B_n \cap A^c) \subset E \cap A^c \end{array} \right. \Rightarrow \mu^*(E) + \epsilon \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

since

$$\mu^*(E \cap A) \leq \mu^* \left( \bigcup_{n=1}^{\infty} (B_n \cap A) \right) \leq \sum_{n=1}^{\infty} \mu^*(B_n \cap A)$$

and

$$\mu^*(E \cap A^c) \leq \mu^* \left( \bigcup_{n=1}^{\infty} (B_n \cap A^c) \right) \leq \sum_{n=1}^{\infty} \mu^*(B_n \cap A^c).$$

We then see that for any  $\epsilon > 0$ , the inequality

$$\mu^*(E) + \epsilon \geq \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

holds, hence so does

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c),$$

which implies  $\mathcal{A} \supset \mathcal{A}_0$ . ⊗

The proof will be continued...

## Lecture 6: Hahn-Kolmogorov Theorem and Extension.

Let's continue the proof of Theorem 1.4.1.

18 Jan. 11:00

**Proof of Theorem 1.4.1 (cont.)** We proved the first part already, now we prove the part left.

**Claim.**  $\mu|_{\mathcal{A}_0} = \mu_0$ .

**Proof.** Let  $A \in \mathcal{A}_0$ , we want to show that

$$\mu(A) = \mu_0(A).$$

- Firstly, let

$$B_i = \begin{cases} A, & \text{if } i = 1 \\ \emptyset, & \text{if } i \geq 2 \end{cases} \in \mathcal{A}_0,$$

hence  $\bigcup_{i=1}^{\infty} B_i = A$ , then we see that

$$\mu^*(A) \leq \sum_{i=1}^{\infty} \mu_0(B_i) = \mu_0(A)$$

from the definition of  $\mu^*$  and countable additivity within the algebra of  $\mu_0$ .

- Secondly, let  $B_i \in \mathcal{A}_0$ ,  $\bigcup_{i=1}^{\infty} B_i \supset A$  be arbitrary. Let  $C_1 = A \cap B_1 \in \mathcal{A}_0$ ,  $C_i = A \cap B_i \setminus \left( \bigcup_{j=1}^{i-1} B_j \right) \in \mathcal{A}_0$  for  $i \geq 2$  since the operations are finite. Then we see

$$A = \bigcup_{i=1}^{\infty} C_i \in \mathcal{A}_0$$

are disjoint countable unions, by countable additivity within the algebra, we therefore have

$$\mu_0(A) = \sum_{i=1}^{\infty} \mu_0(C_i) \leq \sum_{i=1}^{\infty} \mu_0(B_i) \Rightarrow \mu_0(A) \leq \mu^*(A)$$

by taking the infimum from the definition of  $\mu^*$ .

Combine these two inequality, we see that

$$\mu^*(A) = \mu_0(A),$$

for every  $A \in \mathcal{A}_0$ , which implies  $\mu(A) = \mu_0(A)$  for every  $A \in \mathcal{A}_0$  from Theorem 1.3.2, where we extend  $\mu^*$  to  $\mu$  respect to  $\mathcal{A}_0$ .  $\otimes$

■

**Definition 1.4.3** (Hahn-Kolmogorov extension).  $(\mathcal{A}, \mu)$  obtained from Theorem 1.4.1 is the *Hahn-Kolmogorov extensions* of  $(\mathcal{A}_0, \mu_0)$ .

**Note.** We sometimes say *HK extension* instead of *Hahn-Kolmogorov extensions* for simplicity.

We can show the uniqueness of *HK extension*.

**Theorem 1.4.2** (Uniqueness of HK extension). Let  $\mathcal{A}_0$  be an algebra on  $X$ ,  $\mu_0$  be a pre-measure on  $\mathcal{A}_0$ . Let  $(\mathcal{A}, \mu)$  be the HK extension of  $(\mathcal{A}_0, \mu_0)$ . Let  $(\mathcal{A}', \mu')$  be another extension of  $(\mathcal{A}_0, \mu_0)$ . Then if  $\mu_0$  is  $\sigma$ -finite,  $\mu = \mu'$  on  $\mathcal{A} \cap \mathcal{A}'$ .

**Proof.** First, we note the following.

**Note.** Notice that  $\mathcal{A}_0 \subset \mathcal{A}, \mathcal{A}'$  since they both extend  $\mathcal{A}_0$ .

Let  $A \in \mathcal{A} \cap \mathcal{A}'$ , we need to show

$$\underbrace{\mu(A)}_{\mu^*(A)} = \mu'(A).$$

We'll show this by showing two inequalities. We first show that  $\mu^*(A) \geq \mu'(A)$ .

**Claim.**  $\mu^*(A) \geq \mu'(A)$ .

**Proof.** It's easy to show that  $\mu^*(A) \geq \mu'(A)$  by choosing the arbitrary cover of  $A$  and using the [definition](#) of  $\mu^*$ .  $\otimes$

Secondly, we will show that  $\mu(A) \leq \mu'(A)$ .

**Claim.**  $\mu(A) \leq \mu'(A)$ .

**Proof.** We split this into two cases.

- Assume  $\mu(A) < \infty$ , and fix  $\epsilon > 0$ . Then there exists  $B_i \in \mathcal{A}_0$  with  $B := \bigcup_{i=1}^{\infty} B_i \supset A$  such that

$$\mu(A) + \epsilon = \mu^*(A) + \epsilon \geq \sum_{i=1}^{\infty} \mu_0(B_i) \stackrel{!}{=} \sum_{i=1}^{\infty} \mu(B_i) \geq \mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \mu(B).$$

This implies that

$$\mu(B \setminus A) = \mu(B) - \mu(A) \leq \epsilon,$$

where the equality comes from  $A \subset B$  and  $\mu(A) < \infty$ . On the other hand,

$$\mu(B) = \lim_{N \rightarrow \infty} \mu\left(\bigcup_{i=1}^N B_i\right) = \lim_{N \rightarrow \infty} \mu'\left(\bigcup_{i=1}^N B_i\right) = \mu'(B)$$

where the middle equality follows from  $\mu = \mu'$  on  $\mathcal{A}_0$ , hence,

$$\mu(A) \leq \mu(B) = \mu'(B) = \mu'(A) + \mu'(B \setminus A) \leq \mu'(A) + \mu(B \setminus A) \leq \mu'(A) + \epsilon$$

for arbitrary  $\epsilon$ , so we conclude  $\mu(A) \leq \mu'(A)$ .

- Assume  $\mu(A) = \infty$ . Since  $\mu_0$  is [σ-finite](#), so we know  $X = \bigcup_{n=1}^{\infty} X_n$  for some  $X_n \in \mathcal{A}_0$  such that  $\mu_0(X_n) < \infty$ . Replacing  $X_n$  by  $X_1 \cup \dots \cup X_n \in \mathcal{A}_0$ , we may assume that  $X_1 \subset X_2 \subset \dots$ . Then,

$$\forall_{n \in \mathbb{N}} \mu(A \cap X_n) < \infty \stackrel{!}{\Rightarrow} \mu(A \cap X_n) \leq \mu'(A \cap X_n).$$

By [continuity from above](#), we have

$$\mu(A) = \lim_{n \rightarrow \infty} \mu(A \cap X_n) \leq \lim_{n \rightarrow \infty} \mu'(A \cap X_n) = \mu'(A).$$

Combine above two inequalities, the result follows.  $\otimes$

**Corollary 1.4.1.** Let  $\mu_0$  be a [pre-measure](#) on [algebra](#)  $\mathcal{A}_0$  on  $X$ . Suppose  $\mu_0$  is [σ-finite](#), then there exists a unique [measure](#)  $\mu$  on  $\langle \mathcal{A}_0 \rangle$  that extends  $\mathcal{A}_0$ .

Furthermore,

(1) The **completion** of  $(X, \langle \mathcal{A}_0 \rangle, \mu)$  is the **HK extension** of  $(\mathcal{A}_0, \mu_0)$ .<sup>a</sup>

(2)  $\mu(A) = \inf \left\{ \sum_{i=1}^{\infty} \mu_0(B_i) \mid \forall_{i \in \mathbb{N}} B_i \in \mathcal{A}_0, \bigcup_{i=1}^{\infty} B_i \supset A \right\}$  for all  $A \in \overline{\langle \mathcal{A}_0 \rangle}$ .

<sup>a</sup>This really means the pair of the  $\sigma$ -algebra and the **measure** in the **completion** of  $(X, \langle \mathcal{A}_0 \rangle, \mu)$  is the **HK extension** of  $(\mathcal{A}_0, \mu_0)$ , not the whole tuple of the **measurable-space**.

## Lecture 7: Borel Measures

### 1.5 Borel Measures on $\mathbb{R}$

21 Jan. 11:00

**As previously seen.** Recall that when we say something is **Borel**, we assume some sort of topological structure on the underlying space implicitly. In our context, we're considering the usual topology on  $\mathbb{R}^n$  specifically. We'll focus on one dimensional case for now.

**Definition 1.5.1** (Distribution function). An increasing<sup>a</sup> function

$$F: \mathbb{R} \rightarrow \mathbb{R}$$

and right-continuous.  $F$  is then a *distribution function*.

<sup>a</sup>Here, increasing means  $F(x) \leq F(y)$  for  $x < y$ .

**Example.** Here are some examples of right-continuous functions.

(1)  $F(x) = x$ .

(2)  $F(x) = e^x$ .

(3)  $F(x) = \begin{cases} 1, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$

(4) Let  $\mathbb{Q} := \{r_1, r_2, \dots\}$ . Define  $F_n(x) = \begin{cases} 1, & \text{if } x \geq r_n; \\ 0, & \text{if } x < r_n, \end{cases}$  and  $F(x) := \sum_{n=1}^{\infty} \frac{F_n(x)}{2^n}$ .

Then  $F$  is a distribution function (hence right-continuous). This is shown in [Lemma A.1.1](#).

**Note.** If  $F$  is increasing, and

$$F(\infty) := \lim_{x \nearrow \infty} F(x), \quad F(-\infty) := \lim_{x \searrow -\infty} F(x)$$

exist in  $[-\infty, \infty]$ .

In probability theory, cumulative distribution function (CDF) is a distribution function with  $F(\infty) = 1, F(-\infty) = 0$ .<sup>a</sup>

<sup>a</sup>There are distributions [FF99] Ch9., but these are different from distribution functions.

Now, we see a new definition which is essential to our discussion.

**Definition 1.5.2** (Borel measure). A *Borel measure* is any **measure**  $\mu$  defined on the  $\sigma$ -algebra of **Borel sets**.

Since we're now considering a topological space, hence it's reasonable to define the following because we have the concept of compact set now.

**Definition 1.5.3** (Locally finite). Let  $X$  be a Hausdorff topological space,  $\mu$  on  $(X, \mathcal{B}(X))$  is called *locally finite* if  $\mu(K) < \infty$  for every compact set  $K \subset X$ .

**Note.** Some authors will require a **Borel measure** equipped with the **locally finite** property. But formally, this is not so common.

**Lemma 1.5.1.** Let  $\mu$  be a **locally finite Borel measure** on  $\mathbb{R}$ , then

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

is a **distribution function**.

**Proof.** We need to show two things.

**Claim.**  $F_\mu$  is increasing.

**Proof.** To show  $F_\mu$  is increasing, consider  $x < y$  such that

$$F_\mu(x) \leq F_\mu(y)$$

by considering

- $x > 0$ : Then  $F_\mu(x) = \mu((0, x])$  and

$$F_\mu(y) = \mu((0, y]) = \mu((0, x] \cup (x, y]) \geq \mu((0, x]) = F_\mu(x).$$

- $x = 0$ : Then  $F_\mu(x) = 0$  and

$$F_\mu(y) = \mu((0, y]) \geq 0 = F_\mu(0)$$

since  $y > 0$ .

- $x < 0$ : Follows the same argument with  $x > 0$ .

⊗

We now show  $F_\mu$  is right-continuous.

**Claim.**  $F_\mu$  is right-continuous.

**Proof.** Firstly, assume that  $x \geq 0$ , then we see that

$$F_\mu(x) = \mu((0, x]) = \mu((0, x^+])$$

from the fact that a measure is right-continuous.<sup>a</sup> Now, if  $x \leq 0$ , the same argument follows since multiplying  $-1$  will not change the fact that a **measure** is continuous. ⊗

<sup>a</sup>Actually, a measure is always continuous.

■

**Definition 1.5.4 (Half intervals).** We call  $\emptyset, (a, b], (a, \infty), (-\infty, b]$ , and  $(-\infty, \infty)$  *half-intervals*.

**Lemma 1.5.2.** Let  $\mathcal{H}$  be the collection of finite disjoint unions of **half-intervals**. Then,  $\mathcal{H}$  is an **algebra** on  $\mathbb{R}$ .

**Proof.** We observe that  $\emptyset \in \mathcal{H}$  and  $\mathcal{H}$  is closed under finite unions is obvious, hence we only need to show that  $\mathcal{H}$  is closed under complements.



**Claim.**  $\mathcal{H}$  is closed under complements.

**Proof.** We have

- $\emptyset^c = \mathbb{R} = (-\infty, \infty) \in \mathcal{H}$ .
- $(a, b]^c = (-\infty, a] \cup (a, \infty) \in \mathcal{H}$  since it's a two disjoint union of half intervals.
- $(a, \infty)^c = (-\infty, a] \in \mathcal{H}$ .
- $(-\infty, b]^c = (b, \infty) \in \mathcal{H}$ .
- $(-\infty, \infty)^c = \emptyset \in \mathcal{H}$ .

⊗

■

**Proposition 1.5.1** (Distribution function defines a pre-measure). Let  $F: \mathbb{R} \rightarrow \mathbb{R}$  be a **distribution function**. For a **half interval**  $I$ , define

$$\ell(I) := \ell_F(I) = \begin{cases} 0, & \text{if } I = \emptyset; \\ F(b) - F(a), & \text{if } I = (a, b]; \\ F(\infty) - F(a), & \text{if } I = (a, \infty]; \\ F(b) - F(-\infty), & \text{if } I = (-\infty, b]; \\ F(\infty) - F(-\infty), & \text{if } I = (-\infty, \infty). \end{cases}$$

Define  $\mu_0 := \mu_{0,F}: \mathcal{H} \rightarrow [0, \infty]$  by

$$\mu_0(A) = \sum_{k=1}^N \ell(I_k) \text{ if } A = \bigcup_{k=1}^N I_k,$$

where  $A$  is a finite disjoint union of **half intervals**  $I_1, \dots, I_N$ . Then,  $\mu_0$  is a **pre-measure** on  $\mathcal{H}$ .

**Proof.** Firstly, we note that  $\mu_0$  is well-defined. And also,  $\mu_0$  satisfies **null empty set** and **Finite additivity** properties clearly. The only nontrivial part needs a proof is the **Countable additivity within  $\mathcal{H}$**  properties. To show that **Countable additivity within  $\mathcal{H}$**  holds, we proceed as follows.

Suppose  $A \in \mathcal{H}$  where  $A = \bigcup_{i=1}^{\infty} A_i$  is a countable disjoint union. It is enough to consider the case that  $A = I$ ,  $A_k = I_k$  are all half-intervals.

**Remark.** Since  $\mathcal{H}$  is only a collection of *finite* disjoint **half intervals**, hence after considering  $A = I$ , we can apply the same argument iteratively and stop in finite steps. Formally, we can consider  $H \in \mathcal{H}$ ,  $H = \bigcup_{i=1}^{\infty} A^i$ , where  $A^i$  being a **half interval**. Then by the above argument, we have  $A^i = I^i$  and so on.

Focus on the case  $I = (a, b]$ . Let  $(a, b] = \bigcup_{n=1}^{\infty} (a_n, b_n]$ , which is a disjoint union. Then we only need to check  $F(b) - F(a) = \sum_{n=1}^{\infty} (F(b_n) - F(a_n))$ .

**Claim.**  $F(b) - F(a) \geq \sum_{n=1}^{\infty} (F(b_n) - F(a_n))$ .

**Proof.** Since  $(a, b] \supset \bigcup_{n=1}^N (a_n, b_n]$  for any fixed  $N \in \mathbb{N}$ , hence

$$\forall_{N \in \mathbb{N}} F(b) - F(a) \geq \sum_{n=1}^N (F(b_n) - F(a_n)).$$

By letting  $N \rightarrow \infty$ , we have  $F(b) - F(a) \geq \sum_{n=1}^{\infty} (F(b_n) - F(a_n))$ . ⊗

**Claim.**  $F(b) - F(a) \leq \sum_{n=1}^{\infty} (F(b_n) - F(a_n))$ .

**Proof.** Fix  $\epsilon > 0$ . Since  $F$  is right-continuous,  $\exists a' > a$  such that  $F(a') - F(a) < \epsilon$ . For each  $n \in \mathbb{N}$ ,  $\exists b'_n > b_n$  such that

$$F(b'_n) - F(b_n) < \frac{\epsilon}{2^n}.$$

Then, we have  $[a', b] \subset \bigcup_{n=1}^{\infty} (a_n, b'_n)$ , hence  $\exists_{N \in \mathbb{N}} [a', b] \subset \bigcup_{n=1}^N (a_n, b'_n)$ , which is only finitely many unions now.

**Remark.** This essentially follows from the fact that open sets are closed under countable unions, hence the equality will not hold, even after taking the limit.

In this case, we have

$$F(b) - F(a') \leq \sum_{n=1}^N F(b'_n) - F(a_n).$$

Finally, we see that

$$\begin{aligned} F(b) - F(a) &\leq F(b) - F(a') + \epsilon \\ &\leq \sum_{n=1}^{\infty} (F(b'_n) - F(a_n)) + \epsilon \\ &\leq \sum_{n=1}^{\infty} \left( F(b_n) - F(a_n) + \frac{\epsilon}{2^n} \right) + \epsilon = \sum_{n=1}^{\infty} (F(b_n) - F(a_n)) + 2\epsilon \end{aligned}$$

for any fixed  $\epsilon > 0$ , hence

$$F(b) - F(a) \leq \sum_{n=1}^{\infty} (F(b_n) - F(a_n)).$$

**Remark.** It's again the  $\epsilon/2^n$  trick we saw before! ⊗

Combine these two inequalities, we have

$$F(b) - F(a) = \sum_{n=1}^{\infty} (F(b_n) - F(a_n))$$

as we desired. ■

## Lecture 8: Lebesgue-Stieltjes Measure on $\mathbb{R}$

To classify all [measures](#), we now see this last theorem to complete the task.

24 Jan. 11:00

**Theorem 1.5.1** (Locally finite Borel measures on  $\mathbb{R}$ ). We have

- (1)  $F: \mathbb{R} \rightarrow \mathbb{R}$  a **distribution function**, then there exists a **unique locally finite Borel measure**  $\mu_F$  on  $\mathbb{R}$  satisfying

$$\mu_F((a, b]) = F(b) - F(a)$$

for every  $a < b$ .

- (2) Suppose  $F, G: \mathbb{R} \rightarrow \mathbb{R}$  are **distribution functions**. Then,

$$\mu_F = \mu_G$$

on  $\mathcal{B}(\mathbb{R})$  if and only if  $F - G$  is a constant function.

**Proof.** ■

HW.

**Remark.** Theorem 1.5.1 simply states that given a **distribution function**, if we restrict our attention on **locally finite measures** on  $\mathbb{R}$  following our usual convention, then it defines the **measure** on  $\mathcal{B}(\mathbb{R})$  uniquely up to a *constant shift*.

## 1.6 Lebesgue-Stieltjes Measure on $\mathbb{R}$

We see that

$$F \text{ distribution function} \xrightarrow{!} \mu_F \text{ on Carathéodory } \sigma\text{-algebra } \mathcal{A}_{\mu_F} \supset \mathcal{B}(\mathbb{R}).$$

Furthermore, we have

$$(\mathcal{A}_{\mu_F}, \mu_F) = \overline{(\mathcal{B}(\mathbb{R}), \mu_F)}.$$

**Definition 1.6.1** (Lebesgue-Stieltjes measure). Given a **distribution function**  $F$ , we say  $\mu_F$  on  $\mathcal{A}_{\mu_F}$  is called the *Lebesgue-Stieltjes measure* corresponding to  $F$ .

**Definition.** From Definition 1.6.1, if  $F(x) = x$ , then the induced  $(\mathcal{A}_{\mu_F}, \mu_F)$  is denoted as  $(\mathcal{L}, m)$ .

**Definition 1.6.2** (Lebesgue measure).  $m$  is called *Lebesgue measure*.

**Definition 1.6.3** (Lebesgue  $\sigma$ -algebra).  $\mathcal{L}$  is called *Lebesgue  $\sigma$ -algebra*.

**Remark.** Recall that  $\mathcal{L}$  is induced by Theorem 1.3.2, namely given  $m$ , for all  $A \subset \mathbb{R}$ , we have

$$\mathcal{L} := \left\{ A \subset \mathbb{R} \mid \forall_{E \subset \mathbb{R}} m(A) = m(A \cap E) + m(A \setminus E) \right\}.$$

**Note.** We see that since  $F$  is right-continuous and increasing, hence

$$F(x^-) \leq F(x) = F(x^+).$$

Some text will use  $x-$  and  $x+$  instead of  $x^-$  and  $x^+$ , respectively.

We now see some examples.

**Example** (Discrete measure).  $\mu_F((a, b]) = F(b) - F(a)$ . Then

- $\mu_F(\{a\}) = F(a) - F(a^-)$

- $\mu_F([a, b]) = F(b) - F(a^-)$
- $\mu_F((a, b)) = F(b^-) - F(a)$

This is so-called *discrete measure*.

**Example (Dirac measure).** We define

$$F(x) = \begin{cases} 1, & \text{if } x \geq 0; \\ 0, & \text{if } x < 0. \end{cases}$$

Then

- $\mu_F(\{0\}) = 1$
- $\mu_F(\mathbb{R}) = 1$
- $\mu_F(\mathbb{R} \setminus \{0\}) = 0$ . This is easy to see since  $\mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$ , hence

$$\mu_F(\mathbb{R} \setminus \{0\}) = \mu_F((-\infty, 0) \cup (0, \infty)) = \underbrace{\mu_F((-\infty, 0))}_{0-0} + \underbrace{\mu_F((0, \infty))}_{1-1} = 0,$$

where  $\mu_F((-\infty, 0)) = 0$  follows from  $F(0^-) - F(-\infty) = 0 - 0 = 0$ , while  $\mu_F((0, \infty)) = 0$  follows from  $F(\infty) - F(0) = 1 - 1 = 0$ .

We call  $\mu_F$  the *Dirac measure* at 0.

**Example.** Denote  $\mathbb{Q} = \{r_1, r_2, \dots\}$ , and we define

$$F(x) = \sum_{n=1}^{\infty} \frac{F_n(x)}{2^n} \text{ where } F_n(x) = \begin{cases} 1, & \text{if } x \geq r_n; \\ 0, & \text{if } x < r_n. \end{cases}$$

Then

- $\mu_F(\{r_i\}) > 0$  for all  $r_i \in \mathbb{Q}$ .
- $\mu_F(\mathbb{R} \setminus \mathbb{Q}) = 0$ .

This is shown in [Lemma A.1.2](#).

**Example.** If  $F$  is continuous at  $a$ , then  $\mu_F(\{a\}) = 0$ .

**Example (Lebesgue measure).**  $F(x) = x$ , then recall that we denote  $\mu_F := m$ , and we have

- $m((a, b]) = m((a, b)) = m([a, b]) = b - a$ .

**Example.**  $F(x) = e^x$

- $\mu_F((a, b]) = \mu_F((a, b)) = e^b - e^a$ .

**Example** (Middle thirds Cantor set). Let  $C := \bigcap_{n=1}^{\infty} K_n$ , where we have

$$\begin{aligned} K_0 &:= [0, 1] \\ K_1 &:= K_0 \setminus \left( \frac{1}{3}, \frac{2}{3} \right) \\ K_2 &:= K_1 \setminus \left( \frac{1}{9}, \frac{2}{9} \right) \cup \left( \frac{7}{9}, \frac{8}{9} \right) \\ &\vdots \\ K_n &:= K_{n-1} \setminus \bigcup_{k=1}^{3^{n-1}} \left( \frac{3k+1}{3^{n+1}}, \frac{3k+2}{3^{n+1}} \right). \end{aligned}$$

We see that  $C$  is uncountable and with  $m(C) = 0$ . And observe that  $x \in C$  if and only if  $x = \sum_{n=1}^{\infty} \frac{a_n}{3^n}$  for some  $a_n \in \{0, 2\}$ . Hence, we can instead formulate  $K_n$  by

$$K_n = \bigcup_{\substack{a_i \in \{0, 2\} \\ 1 \leq i \leq n}} \left[ \sum_{i=1}^{\infty} \frac{a_i}{3^i}, \sum_{i=1}^{\infty} \frac{a_i}{3^i} + \frac{1}{3^n} \right].$$

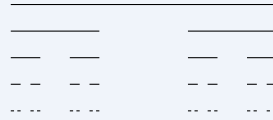


Figure 1.1: The top line corresponds to  $K_0$ , and then  $K_1$ , etc.

The proof of  $m(C) = 0$  is given in [Lemma A.1.3](#).

### Cantor Function

Consider  $F$  as follows. We define a function  $F$  to be 0 to the left of 0, and 1 to the right of 1. Then, define  $F$  to be  $\frac{1}{2}$  on  $(\frac{1}{3}, \frac{2}{3})$ ,  $\frac{1}{4}$  on  $(\frac{1}{9}, \frac{2}{9})$ ,  $\frac{3}{4}$  on  $(\frac{7}{9}, \frac{8}{9})$  and so on. This is so-called *Cantor Function*. We can show  $F$  is continuous and increasing, which makes  $F$  a distribution function. Also, we see that the measure this  $F$  induced is called *Cantor measure*.



Figure 1.2: Cantor Function (Devil's Staircase).

We see that  $F$  is *continuous* and increasing. Furthermore,

Cantor Measure $\mu_F$	Lebesgue Measure $m$
$\mu_F(\mathbb{R} \setminus C) = 0$	$m(\mathbb{R} \setminus C) = \infty > 0$
$\mu_F(C) = 1$	$m(C) = 0$
$\mu_F(\{a\}) = 0$	$m(\{a\}) = 0$

**Remark.**  $\mu_F$  and  $m$  are said to be **singular** to each other.

## 1.7 Regularity Properties of Lebesgue-Stieltjes Measures

We first see a lemma.

**Lemma 1.7.1.** Let  $\mu$  be **Lebesgue-Stieltjes measure** on  $\mathbb{R}$ . Then we have

$$\mu(A) \stackrel{!}{=} \inf \left\{ \sum_{i=1}^{\infty} \mu((a_i, b_i]) \mid \bigcup_{i=1}^{\infty} (a_i, b_i] \supset A \right\} = \inf \left\{ \sum_{i=1}^{\infty} \mu((a_i, b_i)) \mid \bigcup_{i=1}^{\infty} (a_i, b_i) \supset A \right\}$$

for every  $A \in \mathcal{A}_\mu$

**Proof.** The second equality follows from the **continuity of the measure**. ■

**Remark.** This is similar to

$$(a, b) = \bigcup_{n=1}^{\infty} (a, b - 1/n], \quad (a, b] = \bigcap_{n=1}^{\infty} (a, b + 1/n].$$

## Lecture 9: Properties of Lebesgue-Stieltjes measure

**As previously seen.** Let  $X \subset [0, \infty]$ . Recall that

- Finite supremum.

$$\alpha = \sup X < \infty \Leftrightarrow \begin{cases} \forall_{x \in X} \alpha \geq x \\ \forall_{\epsilon > 0} \exists_{x \in X} x + \epsilon \geq \alpha. \end{cases}$$

- Infinite supremum.

$$\alpha = \sup X = \infty \Leftrightarrow \forall_{L > 0} \exists_{x \in X} x \geq L.$$

This should be useful latter on.

26 Jan. 11:00

**Theorem 1.7.1 (Regularity).** Let  $\mu$  be **Lebesgue-Stieltjes measure**. Then, for all  $A \in \mathcal{A}_\mu$ ,

- (1) (outer regularity)  $\mu(A) = \inf\{\mu(O) \mid O \supset A, O \text{ is open}\}$
- (2) (inner regularity)  $\mu(A) = \sup\{\mu(K) \mid K \subset A, K \text{ is compact}\}$

**Proof.** We check them separately.

(1)

- (2) Let  $s := \sup\{\mu(K) \mid K \subset A, K \text{ is compact}\}$ , then by **monotonicity**, we have  $\mu(A) \geq s$ . To show the other direction, we consider

**Claim.** **Inner regularity** holds if  $A$  is a bounded set.

DIY

**Proof.** Then  $\bar{A} \in \mathcal{B}(\mathbb{R}) \subset \mathcal{A}_\mu$ ,  $\bar{A}$  is also bounded  $\Rightarrow \mu(\bar{A}) < \infty$ . Fix  $\epsilon > 0$ , then by **outer regularity**, there exists an open  $O \supset \bar{A} \setminus A$ , and  $\mu(O) - \mu(\bar{A} \setminus A) = \mu(O \setminus (\bar{A} \setminus A)) \leq \epsilon$ . Let  $K := \underbrace{A \setminus O}_{K \subset A} = \underbrace{\bar{A} \setminus O}_{\text{compact}}$ , we show that

$$\mu(K) \geq \mu(A) - \epsilon.$$

\*

DIY

**Claim.** **Inner regularity** holds if  $A$  is an unbounded set with  $\mu(A) < \infty$ .

**Proof.** Let  $A = \bigcup_{n=1}^{\infty} A_n$ ,  $A_n = A \cap [-n, n]$  where  $A_1 \subset A_2 \subset \dots$ , then

$$\lim_{n \rightarrow \infty} \mu(A_n) = \mu(A) < \infty.$$

\*

**Claim.** **Inner regularity** holds if  $A$  is an unbounded set with  $\mu(A) = \infty$ .

**Proof.** We can show that

$$\lim_{n \rightarrow \infty} \mu(A_n) = \mu(A) = \infty.$$

Fix  $L > 0$ , then  $\exists N$  such that  $\mu(A_N) \geq L$ .

\*

■

**Definition.** Let  $X$  be a topological space. Then

**Definition 1.7.1** ( $G_\delta$ -set). A  $G_\delta$ -set is  $G = \bigcap_{i=1}^{\infty} O_i$ ,  $O_i$  open.

**Definition 1.7.2** ( $F_\sigma$ -set). A  $F_\sigma$ -set is  $F = \bigcup_{i=1}^{\infty} F_i$ ,  $F_i$  closed.

**Theorem 1.7.2.** Let  $\mu$  be a **Lebesgue-Stieltjes measure**. Then **TFAE**<sup>a</sup>:

- (1)  $A \in \mathcal{A}_\mu$
- (2)  $A = G \setminus M$ ,  $G$  is a  $G_\delta$ -set,  $M$  is a  $\mu$ -null set.
- (3)  $A = F \setminus N$ ,  $F$  is a  $F_\sigma$ -set,  $N$  is a  $\mu$ -null set.

<sup>a</sup> TFAE: The following are equivalent.

**Proof.** We see that (2)  $\Rightarrow$  (1) and (3)  $\Rightarrow$  (1) are clear.

**Claim.** (1)  $\Rightarrow$  (3).

**Proof.** We consider two cases.

- Assume  $\mu(A) < \infty$ . From the [inner regularity](#), we have

$$\forall n \in \mathbb{N} \exists \text{compact } K_n \subset A \text{ such that } \mu(K_n) + \frac{1}{n} \geq \mu(A).$$

Let  $F = \bigcup_{n=1}^{\infty} K_n$ , then  $N = A \setminus F$  is [μ-null](#).

Check!

- Assume  $\mu(A) = \infty$ . Let  $A = \bigcup_{k \in \mathbb{Z}} A_k$ ,  $A_k = A \cap (k, k+1]$ . From what we have just shown above,

$$\forall k \in \mathbb{Z} \ A_k = F_k \cup N_k, \quad A = \underbrace{\left( \bigcup_k F_k \right)}_{F_\sigma\text{-set}} \cup \underbrace{\left( \bigcup_k N_k \right)}_{\mu\text{-null}}.$$

⊗

**Claim.** (1)  $\Rightarrow$  (2).

**Proof.** We see that

$$A^c = F \cup N, \quad A = F^c \cap N^c = F^c \setminus N.$$

⊗

■

**Proposition 1.7.1.** Let  $\mu$  be a [Lebesgue-Stieltjes measure](#), and  $A \in \mathcal{A}_\mu$ ,  $\mu(A) < \infty$ . Then we have

$$\forall \epsilon > 0 \exists I = \bigcup_{i=1}^{N(\epsilon)} I_i$$

disjoint open intervals such that  $\mu(A \triangle I) \leq \epsilon$ .

**Proof.** Using [outer regularity](#) and the fact that every open set is  $\bigcup_{i=1}^{\infty} I_i$ , where  $I_i$  are disjoint open intervals.

■

DIY

We now see some properties of [Lebesgue measure](#).

**Theorem 1.7.3.** Let  $A \in \mathcal{L}$ , then we have  $A + s \in \mathcal{L}$ ,  $rA \in \mathcal{L}$  for all  $r, s \in \mathbb{R}$ . i.e.,

$$m(A + s) = m(A), \quad m(rA) = |r| \cdot m(A).$$

**Proof.**

■

DIY

**Example.** We now see some examples.

- (1) Let  $\mathbb{Q} =: \{r_i\}_{i=1}^{\infty}$  which is dense in  $\mathbb{R}$ . Let  $\epsilon > 0$ , and

$$O = \bigcup_{i=1}^{\infty} \left( r_i - \frac{\epsilon}{2^i}, r_i + \frac{\epsilon}{2^i} \right).$$



---

We see that  $O$  is open and dense<sup>a</sup> in  $\mathbb{R}$ . But we see

$$m(O) \leq \sum_{i=1}^{\infty} \frac{2\epsilon}{2^i} = 2\epsilon.$$

Furthermore,  $\partial O = \overline{O} \setminus O$ ,  $m(\partial O) = \infty$

- (2) There exists uncountable set  $A$  with  $m(A) = 0$ .
- (3) There exists  $A$  with  $m(A) > 0$  but  $A$  contains no non-empty open intervals.
- (4) There exists  $A \notin \mathcal{L}$ . e.g. Vitali set.<sup>b</sup>
- (5) There exists  $A \in \mathcal{L} \setminus \mathcal{B}(\mathbb{R})$ .

---

<sup>a</sup>[https://en.wikipedia.org/wiki/Dense\\_set](https://en.wikipedia.org/wiki/Dense_set)

<sup>b</sup>[https://en.wikipedia.org/wiki/Vitali\\_set](https://en.wikipedia.org/wiki/Vitali_set)

# Chapter 2

## Integration

### Lecture 10: Integration

#### 2.1 Measurable Function

26 Jan. 11:00

We start with a definition.

**Definition 2.1.1** (Measurable function). Suppose  $(X, \mathcal{A}), (Y, \mathcal{B})$  are measurable spaces. Then we say  $f: X \rightarrow Y$  is  $(\mathcal{A}, \mathcal{B})$ -measurable if

$$\forall_{B \in \mathcal{B}} f^{-1}(B) \in \mathcal{A}.$$

**Remark.** If  $\mathcal{A}$  and  $\mathcal{B}$  are given, we'll sometimes say  $f$  is measurable if it'll not cause any confusions.

**Lemma 2.1.1.** Given two measurable spaces  $(X, \mathcal{A})$  and  $(Y, \mathcal{B})$ , and suppose  $\mathcal{B} = \langle \mathcal{E} \rangle$  for some  $\mathcal{E} \subset Y$ . Then,

$$f: X \rightarrow Y \text{ is } (\mathcal{A}, \mathcal{B})\text{-measurable} \Leftrightarrow \forall_{E \in \mathcal{E}} f^{-1}(E) \in \mathcal{A}.$$

**Proof.** We see that the *only if* part ( $\Rightarrow$ ) is clear. On the other direction, we consider the following. Let  $\mathcal{D} = \{E \subset Y \mid f^{-1}(E) \in \mathcal{A}\}$ , then

- $\mathcal{E} \subset \mathcal{D}$  by assumption
- $\mathcal{D}$  is a  $\sigma$ -algebra

hence, we see that  $\langle \mathcal{E} \rangle = \mathcal{B} \subset \mathcal{D}$  from Lemma 1.1.2. The result then follows from the definition of  $(\mathcal{A}, \mathcal{B})$ -measurable. ■

Check!

**Note.** Recall that

- $f^{-1}(E^c) = f^{-1}(E)^c$
- $f^{-1}\left(\bigcup_{\alpha} E_{\alpha}\right) = \bigcup_{\alpha} f^{-1}(E_{\alpha})$

**Definition 2.1.2** ( $\mathcal{A}$ -measurable). Let  $(X, \mathcal{A})$  be a measurable space. Then,

$$\left. \begin{array}{l} f: X \rightarrow \mathbb{R} \\ f: X \rightarrow \overline{\mathbb{R}} \\ f: X \rightarrow \mathbb{C} \end{array} \right\} \text{ is } \mathcal{A}\text{-measurable if } \left\{ \begin{array}{l} f \text{ is } (\mathcal{A}, \mathcal{B}(\mathbb{R}))\text{-measurable} \\ f \text{ is } (\mathcal{A}, \mathcal{B}(\overline{\mathbb{R}}))\text{-measurable} \\ \operatorname{Re} f, \operatorname{Im} f: X \rightarrow \mathbb{R} \text{ are } \mathcal{A}\text{-measurable.} \end{array} \right.$$

**Notation.** Notice that

- $\overline{\mathbb{R}} = [-\infty, \infty]$
- $\mathcal{B}(\overline{\mathbb{R}}) = \{E \subset \overline{\mathbb{R}} \mid E \cap \mathbb{R} \in \mathcal{B}(\mathbb{R})\}$ .
- $\operatorname{Re} f$  is the real part of  $f$ , while  $\operatorname{Im} f$  is the imaginary part of  $f$ .

**Example.** We see that

- $\mathcal{A} = \mathcal{P}(X) \Rightarrow$  Every function is  $\mathcal{A}$ -measurable.
- $\mathcal{A} = \{\emptyset, X\} \Rightarrow$  The only  $\mathcal{A}$ -measurable functions are constant functions.

There are two very common kinds of measurable functions are worth mentioning.

**Definition.** Given a measurable function  $f$ , we have the following.

**Definition 2.1.3** (Lebesgue measurable function).  $f$  is a Lebesgue measurable function if  $f: (\mathbb{R}, \mathcal{L}) \rightarrow (\mathbb{C}, \mathcal{B}(\mathbb{C}))$ .

**Definition 2.1.4** (Borel measurable function).  $f$  is a Borel measurable function if  $f: (\mathbb{R}, \mathcal{B}(\mathbb{R})) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

**Lemma 2.1.2.** Given  $f: X \rightarrow \mathbb{R}$ , TFAE.

- (1)  $f$  is  $\mathcal{A}$ -measurable
- (2)  $\forall a \in \mathbb{R}, f^{-1}((a, \infty)) \in \mathcal{A}$
- (3)  $\forall a \in \mathbb{R}, f^{-1}([a, \infty)) \in \mathcal{A}$
- (4)  $\forall a \in \mathbb{R}, f^{-1}((-\infty, a)) \in \mathcal{A}$
- (5)  $\forall a \in \mathbb{R}, f^{-1}((-\infty, a]) \in \mathcal{A}$

**Proof.** The result follows from Lemma 2.1.1 we just saw. ■

**Remark** (Operations preserve  $\mathcal{A}$ -measurability). Given  $f, g: X \rightarrow \mathbb{R}$  and is  $\mathcal{A}$ -measurable, then

- (1)  $\phi: \mathbb{R} \rightarrow \mathbb{R}$ ,  $\mathcal{A}$ -measurable, then

$$\phi \circ f: X \rightarrow \mathbb{R}$$

is  $\mathcal{A}$ -measurable.

- (2)  $-f, 3f, f^2, |f|$  are all  $\mathcal{A}$ -measurable, and  $\frac{1}{f}$  is  $\mathcal{A}$ -measurable if  $f(x) \neq 0, \forall x \in X$ .
- (3)  $f + g$  is  $\mathcal{A}$ -measurable. We see this from

$$(f + g)^{-1}((a, \infty)) = \bigcup_{r \in \mathbb{Q}} (f^{-1}((r, \infty)) \cap g^{-1}((a - r, \infty)))$$

with Lemma 2.1.2.

- (4)  $f \cdot g$  is  $\mathcal{A}$ -measurable. We see this from

$$f(x)g(x) = \frac{1}{2} ((f(x) + g(x))^2 - f(x)^2 - g(x)^2).$$

(5) We see that

$$(f \vee g)(x) := \max\{f(x), g(x)\} \text{ and } (f \wedge g)(x) := \min\{f(x), g(x)\}$$

are  $\mathcal{A}$ -measurable.

(6) Let  $f_n: X \rightarrow \overline{\mathbb{R}}$  be  $\mathcal{A}$ -measurable. Then

$$\sup_{n \in \mathbb{N}} f_n, \inf_{n \in \mathbb{N}} f_n, \limsup_{n \rightarrow \infty} f_n, \liminf_{n \rightarrow \infty} f_n$$

are  $\mathcal{A}$ -measurable.

**Proof.** Consider  $\sup_{n \in \mathbb{N}} f_n =: g$ , then

$$g^{-1}((a, \infty]) = \bigcup_{n \in \mathbb{N}} f_n^{-1}((a, \infty])$$

for  $\sup_n f_n(x) = g(x) > a$ . A similar argument can prove the case of  $\inf_{n \in \mathbb{N}} f_n$ .

And notice that  $\limsup_{n \rightarrow \infty} f_n = \inf_{k \in \mathbb{N}} \sup_{n \geq k} f_n$ , then the similar argument also proves this case. ⊗

check

(7) If  $\lim_{n \rightarrow \infty} f_n(x)$  converges for every  $x \in X$ , then  $f$  is  $\mathcal{A}$ -measurable.

(8) If  $f: \mathbb{R} \rightarrow \mathbb{R}$  is continuous  $\Rightarrow f$  is Borel measurable  $\Rightarrow f$  is Lebesgue measurable since the preimage of an open set of a continuous function is open, then we consider  $f^{-1}((a, \infty))$ .

**Definition 2.1.5 (Support).** The *support* of function  $f: X \rightarrow \overline{\mathbb{R}}$  is

$$\text{supp } f := \{x \in X \mid f(x) \neq 0\}.$$

**Definition.** For  $f: X \rightarrow \overline{\mathbb{R}}$ , let  $f^+ := f \vee 0$  and  $f^- := (-f) \vee 0$ , i.e.,  $f^+(x) = \max\{f(x), 0\}$ ,  $f^-(x) = \max\{-f(x), 0\}$ . Then we have the following.

**Definition 2.1.6 (Positive part).**  $f^+$  is the *positive part* of  $f$ .

**Definition 2.1.7 (Negative part).**  $f^-$  is the *negative part* of  $f$ .

**Remark.** If  $\text{supp } f^+ \cap \text{supp } f^- = \emptyset$  and  $f(x) = f^+(x) - f^-(x)$ , then

$$f \text{ is } \mathcal{A}\text{-measurable} \Leftrightarrow f^+, f^- \text{ are } \mathcal{A}\text{-measurable}.$$

**Definition 2.1.8 (Characteristic (Indicator) function).** For  $E \subset X$ , the *characteristic (indicator) function* of  $E$  is

$$\mathcal{X}_E(x) = \mathbb{1}_E(x) = \begin{cases} 1, & \text{if } x \in E; \\ 0, & \text{if } x \in E^c. \end{cases}$$

**Remark.** We see that  $\mathbb{1}_E$  is  $\mathcal{A}$ -measurable  $\Leftrightarrow E \in \mathcal{A}$ .

**Definition 2.1.9 (Simple function).** Let  $(X, \mathcal{A})$  be a measurable space. Then a *simple function*  $\phi: X \rightarrow \mathbb{C}$  that is  $\mathcal{A}$ -measurable and takes only finitely many values.

**Remark.** We see that if  $\phi(X) = \{c_1, \dots, c_N\}$ , then

$$E_i = \phi^{-1}(\{c_i\}) \in \mathcal{A} \Rightarrow \phi = \sum_{i=1}^N \underbrace{c_i}_{\neq \pm\infty} \underbrace{\mathbb{1}_{E_i}}_{\in \mathcal{A}}.$$

## Lecture 11: Integration of nonnegative functions

As previously seen. For a **simple function**  $\phi$ ,  $c_i$  can actually be in  $\mathbb{C}$ .

31 Jan. 11:00

**Theorem 2.1.1.** Given a **measurable space**  $(X, \mathcal{A})$  and let  $f: X \rightarrow [0, \infty]$ , the following are equivalent.

- (1)  $f$  is a  **$\mathcal{A}$ -measurable** function.
- (2) There exists **simple functions**  $0 \leq \phi_1(x) \leq \phi_2(x) \leq \dots \leq f(x)$  such that

$$\forall_{x \in X} \lim_{n \rightarrow \infty} \phi_n(x) = f(x)$$

i.e.,  $f$  is a **pointwise** upward limit of **simple functions**.

**Proof.** We'll prove both directions.

**Claim.** (2)  $\Rightarrow$  (1).

**Proof.** It's clear from the fact that  $f(x) = \sup_n \phi_n(x)$  and **the remark**. ⊗

**Claim.** (1)  $\Rightarrow$  (2).

**Proof.** Assume  $f$  is  **$\mathcal{A}$ -measurable**, and fix  $n \in \mathbb{N}$ .

Let  $F_n = f^{-1}([2^n, \infty]) \in \mathcal{A}$ . Also, for  $0 \leq k \leq 2^{2n} - 1$ ,  $E_{n,k} = f^{-1}([\frac{k}{2^n}, \frac{k+1}{2^n}]) \in \mathcal{A}$ .

Then, define  $\phi_n$  be

$$\phi_n = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n} \mathbb{1}_{E_{n,k}} + 2^n \mathbb{1}_{F_n},$$

we have

- $0 \leq \phi_1(x) \leq \phi_2(x) \leq \dots \leq f(x)$  for every  $x \in X$
- $\forall x \in X \setminus F_n$ , we have  $0 \leq f(x) - \phi_n(x) \leq \frac{1}{2^n}$

Furthermore, we see that

$$F_1 \supset F_2 \supset \dots, \quad \bigcap_{n=1}^{\infty} F_n = f^{-1}(\{\infty\}),$$

then

- $x \in f^{-1}([0, \infty]) = X \setminus \bigcap_{n=1}^{\infty} F_n \Rightarrow \lim_{n \rightarrow \infty} \phi_n(x) = f(x)$
- $x \in f^{-1}(\{\infty\}) = \bigcap_{n=1}^{\infty} F_n \Rightarrow f_n(x) \geq 2^n \Rightarrow \lim_{n \rightarrow \infty} \phi_n(x) = \infty = f(x)$

⊗

■

**Corollary 2.1.1.** If  $f$  is bounded on a set  $A \subset \mathbb{R}$ , i.e.,  $\exists L > 0$  such that

$$\forall_{x \in A} |f(x)| \leq L,$$

then there exists a sequence of **simple functions**  $\{\phi_n\}$  such that  $\phi_n \rightarrow f$  **uniformly** on  $A$ .

**Proof.** ■

DIY

**Corollary 2.1.2.** If  $f: X \rightarrow \mathbb{C}$  is a **measurable function** if and only if there exists **simple functions**  $\phi_n: X \rightarrow \mathbb{C}$  such that

$$0 \leq |\phi_1(x)| \leq |\phi_2(x)| \leq \dots \leq |f(x)|$$

with

$$\forall_{x \in X} \lim_{n \rightarrow \infty} \phi_n(x) = f(x).$$

**Proof.** ■

DIY

## 2.2 Integration of Nonnegative Functions

We start with our first definition about integral.

**Definition 2.2.1** (Integration of nonnegative function). Let  $(X, \mathcal{A}, \mu)$  be a **measure space**, and  $\phi: X \rightarrow [0, \infty]$  such that

$$\phi = \sum_{i=1}^N c_i \mathbb{1}_{E_i}$$

be a **simple function**. Define

$$\int \phi = \int \phi \, d\mu = \int_X \phi \, d\mu = \sum_{i=1}^N c_i \mu(E_i).$$

Furthermore, for  $A \in \mathcal{A}$ ,

$$\int_A \phi = \int_A \phi \, d\mu = \int \phi \mathbb{1}_A \, d\mu.$$

**Note.** Note that

- In the expression  $\sum_{i=1}^N c_i \mu(E_i)$ , we're using the convention  $0 \cdot \infty = 0$ .
- The function  $\phi \mathbb{1}_A$  is also a **simple function** since both  $\phi$  and  $\mathbb{1}_A$  are **simple function**.

**Proposition 2.2.1.** Suppose we have  $\phi, \psi \geq 0$  be two **simple functions**. Then,

- (1) **Definition 2.2.1** is well-defined.
- (2)  $\int c\phi = c \int \phi$  for  $c \in [0, \infty)$ .
- (3)  $\int \phi + \psi = \int \phi + \int \psi$ .
- (4)  $\phi(x) \geq \psi(x)$  for all  $x \Rightarrow \int \phi \geq \int \psi$ .
- (5)  $\nu(A) = \int_A \phi \, d\mu$  is a **measure** on  $(X, \mathcal{A})$ .

**Proof.** ■

DIY

**Definition 2.2.2** (Generalization of Integration of nonnegative function). Given  $(X, \mathcal{A}, \mu)$  with  $f: X \rightarrow [0, \infty]$  be  $\mathcal{A}$ -measurable. Define

$$\int f = \int f d\mu = \sup \left\{ \int \phi : 0 \leq \phi \leq f \text{ such that } \phi \text{ is simple} \right\}.$$

**Note.** Note that

- If  $f$  is a simple function, the Definition 2.2.1 and Definition 2.2.2 of  $\int f$  are the same.
- $\int cf = c \int f$  for  $c \in [0, \infty)$ .
- If  $f \geq g \geq 0 \Rightarrow \int f \geq \int g$ .
- But  $\int f + g = \int f + \int g$  is not trivial.

**Theorem 2.2.1** (Monotone Convergence Theorem). Given  $(X, \mathcal{A}, \mu)$  be a measure space. Then if

- $f_n: X \rightarrow [0, \infty]$  be  $\mathcal{A}$ -measurable for every  $n \in \mathbb{N}$ ;
- $0 \leq f_1(x) \leq f_2(x) \leq \dots$  for every  $x \in X$ ;
- $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  for every  $x \in X$ ,

we have

$$\lim_{n \rightarrow \infty} \int f_n = \int f.$$

**Proof.** Note that if  $\lim_{n \rightarrow \infty} \int f_n$  exists, then it's equal to  $\sup_n \int f_n$ .

Then

- $f_n \leq f \Rightarrow \int f_n \leq \int f \Rightarrow \lim_{n \rightarrow \infty} \int f_n \leq \int f$ .
- Fix a simple function  $0 \leq \phi \leq f$ , then it's enough to show  $\lim_{n \rightarrow \infty} \int f_n \geq \int \phi$ .

We first fix  $\alpha = (0, 1)$ , then it's also enough to show

$$\lim_{n \rightarrow \infty} \int f_n \geq \alpha \int \phi.$$

Let  $A_n := \{x \in X \mid f_n(x) \geq \alpha \phi(x)\}$ , then since  $f_n$  is measurable,

- $A_n \in \mathcal{A}$
- $A_1 \subset A_2 \subset A_3 \subset \dots$
- $\bigcup_{n=1}^{\infty} A_n = X$

We then have

$$\int f_n \geq \int f_n \mathbb{1}_{A_n} \geq \int \alpha \phi \mathbb{1}_{A_n} = \alpha \int_{A_n} \phi = \alpha \nu(A_n)$$

where  $\nu(A) = \int_A \phi$  is a measure. This implies

$$\lim_{n \rightarrow \infty} \int f_n \geq \alpha \lim_{n \rightarrow \infty} \nu(A_n) \stackrel{!}{=} \alpha \nu(X) = \alpha \int \phi.$$

■

Check!

**Corollary 2.2.1** (Linearity of nonnegative integral). Let  $f, g \geq 0$  be measurable, then

$$\int f + g = \int f + \int g.$$

**Proof.** There exists simple functions  $\phi_n$  and  $\psi_n$  such that

- $0 \leq \phi_1 \leq \phi_2 \leq \dots$  and  $\phi_n \rightarrow f$  pointwise
- $0 \leq \psi_1 \leq \psi_2 \leq \dots$  and  $\psi_n \rightarrow g$  pointwise

Then,

$$\int (f + g) \stackrel{!}{=} \lim_{n \rightarrow \infty} \int (\phi_n + \psi_n) = \lim_{n \rightarrow \infty} \int \phi_n + \int \psi_n \stackrel{!}{=} \int f + \int g.$$

■

## Lecture 12: Fatou's Lemma

We start with a useful corollary.

2 Feb. 11:00

**Corollary 2.2.2** (Tonelli's theorem for nonnegative series and integrals). Given  $g_n \geq 0$  for every  $n \in \mathbb{N}$  and let  $g_n$  be measurable, then

$$\int \sum_{n=1}^{\infty} g_n = \sum_{n=1}^{\infty} \int g_n.$$

**Proof.** Let  $f_N := \sum_{n=1}^N g_n$  such that  $\lim_{N \rightarrow \infty} f_N = \sum_{n=1}^{\infty} g_n =: f$ , then since  $g_n \geq 0$ , we have  $0 \leq f_1 \leq f_2 \leq \dots$  with

$$\lim_{N \rightarrow \infty} f_N(x) = \sum_{n=1}^{\infty} g_n(x).$$

By Theorem 2.2.1, we have

$$\lim_{N \rightarrow \infty} \underbrace{\int \sum_{n=1}^N g_n}_{f_N} = \underbrace{\int \sum_{n=1}^{\infty} g_n}_f.$$

Now, since the terms in the limit on the left-hand side is just a finite sum, by Corollary 2.2.1, we have

$$\underbrace{\lim_{N \rightarrow \infty} \sum_{n=1}^N \int g_n}_{\sum_{n=1}^{\infty} \int g_n} = \lim_{N \rightarrow \infty} \int \sum_{n=1}^N g_n = \int \sum_{n=1}^{\infty} g_n,$$

hence

$$\int \sum_{n=1}^{\infty} g_n = \sum_{n=1}^{\infty} \int g_n.$$

■

**Remark.** Recall that we have seen two series case before. We'll later see two integrals cases.

**Theorem 2.2.2** (Fatou's Lemma). Suppose  $f_n \geq 0$  and measurable, then

$$\int \liminf_{n \rightarrow \infty} f_n \leq \liminf_{n \rightarrow \infty} \int f_n.$$

**Proof.** Before we start we note the following.



**Remark.** Recall that

$$\liminf_{n \rightarrow \infty} f_n := \lim_{k \rightarrow \infty} \inf_{n \geq k} f_n = \sup_{k \in \mathbb{N}} \inf_{n \geq k} f_n$$

and

$$\exists \lim_{n \rightarrow \infty} a_n \Leftrightarrow \limsup_{n \rightarrow \infty} a_n = \liminf_{n \rightarrow \infty} a_n.$$

Let  $g_k = \inf_{n \geq k} f_n$ , then  $g_k$  is measurable and  $0 \leq g_1 \leq g_2 \leq \dots$ . Now, from Theorem 2.2.1, we have

$$\int \lim_{k \rightarrow \infty} g_k = \lim_{k \rightarrow \infty} \int g_k.$$

Notice that the left-hand side is just  $\int \liminf_{n \rightarrow \infty} f_n$ , while the right-hand side is just  $\lim_{k \rightarrow \infty} \int \inf_{n \geq k} f_n$ , i.e.,

$$\int \liminf_{n \rightarrow \infty} f_n = \lim_{k \rightarrow \infty} \int \inf_{n \geq k} f_n.$$

We see that we want to take the inf outside the integral on the right-hand side. Observe that

$$\forall_{m \geq k} \inf_{n \geq k} f_n \leq f_m \Rightarrow \forall_{m \geq k} \int \inf_{n \geq k} f_n \leq \int f_m \Rightarrow \int \inf_{n \geq k} f_n \leq \inf_{m \geq k} \int f_m.$$

Then, we have

$$\int \liminf_{n \rightarrow \infty} f_n = \lim_{k \rightarrow \infty} \int \inf_{n \geq k} f_n \leq \lim_{k \rightarrow \infty} \inf_{m \geq k} \int f_m = \liminf_{m \rightarrow \infty} \int f_m.$$

■

**Example** (Escape to horizontal infinity). Given  $(\mathbb{R}, \mathcal{L}, m)$ , let  $f_n := \mathbb{1}_{(n, n+1)}$ . We immediately see that

- $f_n \rightarrow 0$  pointwise
- $\int f_n = 1$  for every  $n$
- $\int f = 0$

From Theorem 2.2.2, we have a strict inequality

$$0 = \int \liminf_{n \rightarrow \infty} f_n, \liminf_{n \rightarrow \infty} \int f_n = 1.$$

**Example** (Escape to width infinity). Given  $(\mathbb{R}, \mathcal{L}, m)$ , let  $f_n := \frac{1}{n} \mathbb{1}_{(0, n)}$ .

**Example** (Escape to vertical infinity). Given  $(\mathbb{R}, \mathcal{L}, m)$ , let  $f_n := n \mathbb{1}_{(0, \frac{1}{n})}$ .

**Lemma 2.2.1** (Markov's inequality). Let  $f \geq 0$  be measurable. Then

$$\forall_{c \in (0, \infty)} \mu(\{x \mid f(x) \geq c\}) \leq \frac{1}{c} \int f.$$

**Proof.** Denote  $\{x \mid f(x) \geq c\} =: E$ , then

$$f(x) \geq c \mathbb{1}_E(x) \Rightarrow \int f \geq c \int \mathbb{1}_E = c \cdot \mu(E).$$

■

**Remark.** Notice that  $E = f^{-1}([c, \infty))$ , hence  $E$  is measurable.

**Proposition 2.2.2.** Let  $f \geq 0$  be measurable. Then,

$$\int f = 0 \Leftrightarrow f = 0 \text{ a.e.}$$

i.e.,

$$\int f \, d\mu = 0 \Leftrightarrow \mu(A) = 0$$

where  $A = \{x \mid f(x) > 0\} = f^{-1}((0, \infty))$ .

**Proof.** Firstly, assume that  $f = \phi$  is a simple function. We may write

$$\phi = \sum_{i=1}^N c_i \mathbb{1}_{E_i}$$

where  $E_i$  are disjoint and  $c_i \in (0, \infty)$ . Then,

$$\int \phi = \sum_{i=1}^N c_i \mu(E_i) = 0 \Leftrightarrow \mu(E_1) = \dots = \mu(E_N) = 0 \Leftrightarrow \mu(A) = 0, \quad A = \bigcup_{i=1}^N E_i.$$

Now, assume that  $f$  is a general function where  $f \geq 0$  is the only constraint, and we consider two cases.

- Assume  $\mu(A) = 0$  (i.e.,  $f = 0$  a.e.). Let  $0 \leq \phi \leq f$ , where  $\phi$  is simple. Then

$$\forall_{x \in A^c} \phi(x) = 0$$

since  $f(x) = 0, \forall x \in A^c$ . This implies that  $\phi = 0$  a.e. since  $\mu(A) = 0$ , so  $\int \phi = 0$ . We then have

$$\int f = 0$$

from Definition 2.2.2.

- Assume  $\int f = 0$ . Let  $A_n = f^{-1}([\frac{1}{n}, \infty))$ . Then we see that
  - $A_1 \subset A_2 \subset \dots$
  - $\bigcup_{n=1}^{\infty} A_n = f^{-1}\left(\bigcup_{n=1}^{\infty} [\frac{1}{n}, \infty)\right) = f^{-1}((0, \infty)) = A$ .

We then have

$$\mu(A_n) = \mu\left(\left\{x \mid f(x) \geq \frac{1}{n}\right\}\right) \stackrel{!}{\leq} n \int f = 0,$$

which further implies

$$\mu(A) = \lim_{n \rightarrow \infty} \mu(A_n) = 0$$

from the continuity of measure from below. ■

**Corollary 2.2.3.** If  $f, g \geq 0$  are both measurable and  $f = g$  a.e., then

$$\int f = \int g.$$

**Proof.** Let  $A = \{x \mid f(x) \neq g(x)\}^a$ . Then by assumption,  $\mu(A) = 0$ , hence

$$f \mathbb{1}_A = 0 \text{ a.e.}, \quad g \mathbb{1}_A = 0 \text{ a.e.}$$

This further implies that

$$\begin{aligned} \int f &= \int f(\mathbb{1}_A + \mathbb{1}_{A^c}) \stackrel{!}{=} \int f \mathbb{1}_A + \int f \mathbb{1}_{A^c} \\ &= \int f \mathbb{1}_{A^c} = \int g \mathbb{1}_{A^c} = \int g \mathbb{1}_{A^c} + \int g \mathbb{1}_A = \int g. \end{aligned}$$

■

<sup>a</sup> $A$  is measurable indeed.

**Corollary 2.2.4.** Let  $f_n \geq 0$  be measurable. Then

- $$\begin{aligned} (1) \quad & \left. \begin{array}{l} 0 \leq f_1 \leq f_2 \leq \dots \leq f \text{ a.e.} \\ \lim_{n \rightarrow \infty} f_n = f \text{ a.e.} \end{array} \right\} \Rightarrow \lim_{n \rightarrow \infty} \int f_n = \int f. \\ (2) \quad & \lim_{n \rightarrow \infty} f_n = f \text{ a.e.} \Rightarrow \int f \leq \liminf_{n \rightarrow \infty} \int f_n. \end{aligned}$$

**Proof.**

■

DIY

**Remark.** Almost all the theorems we've proved can be replaced by theorems dealing with **almost everywhere** condition.

## Lecture 13: Integration of Complex Functions

### 2.3 Integration of Complex Functions

4 Feb. 11:00

As usual, we start with a definition.

**Definition 2.3.1 (Integrable).** Let  $(X, \mathcal{A}, \mu)$  be a **measure space** and let  $f: X \rightarrow \overline{\mathbb{R}}$  and  $g: X \rightarrow \mathbb{C}$  be **measurable**.<sup>a</sup>

Then  $f, g$  are called *integrable* if  $\int |f| < \infty$  and  $\int |g| < \infty$ , and we define

$$\int f = \int f^+ - \int f^-, \quad \int g = \int \operatorname{Re} g + i \int \operatorname{Im} g.$$

Furthermore, for  $f: X \rightarrow \overline{\mathbb{R}}$ , we define

$$\int f = \begin{cases} \infty, & \text{if } \int f^+ = \infty, \int f^- < \infty; \\ -\infty, & \text{if } \int f^+ < \infty, \int f^- = \infty. \end{cases}$$

<sup>a</sup>Recall that for a complex-valued function like  $g$ , this means that both  $\operatorname{Re} g$  and  $\operatorname{Im} g$  are measurable.

We now see a lemma.

**Lemma 2.3.1.** Let  $f, g: X \rightarrow \overline{\mathbb{R}}$  or  $\mathbb{C}$  **integrable**. Assume that  $f(x) + g(x)$  is well-defined for all  $x \in X$ .<sup>a</sup> Then we have

- (1)  $f + g, cf$  for all  $c \in \mathbb{C}$  are **integrable**.
- (2)  $\int f + g = \int f + \int g$ . This is not trivial since  $(f + g)^+ \neq f^+ + g^+$ .

$$(3) \left| \int f \right| \leq \int |f|.$$

<sup>a</sup>That is, we never see  $\infty + (-\infty)$  or  $(-\infty) + \infty$ .

**Proof.** Check [FF99] page 53. ■

**Lemma 2.3.2.** Let  $(X, \mathcal{A}, \mu)$  be a **measure space** and let  $f$  be an **integrable** function on  $X$ . Then

- (1)  $f$  is finite **a.e.** i.e.,  $\{x \in X \mid |f(x)| = \infty\}$  is a **null set**.
- (2) The set  $\{x \in X \mid f(x) \neq 0\}$  is  **$\sigma$ -finite**.

**Proof.** ■

**Proposition 2.3.1.** Let  $(X, \mathcal{A}, \mu)$  be a **measure space**, then

- (1) If  $h$  is **integrable** on  $X$ , then

$$\forall E \in \mathcal{A} \quad \int_E h = 0 \Leftrightarrow \int |h| = 0 \Leftrightarrow h = 0 \text{ a.e.}$$

- (2) If  $f, g$  are **integrable** on  $X$ , then

$$\forall E \in \mathcal{A} \quad \int_E f = \int_E g \Leftrightarrow f = g \text{ a.e.}$$

**Proof.** We prove this one by one.

- (1) We see that the second equivalence is done in **Proposition 2.2.2**, hence we prove the first equivalence only. Since we have

$$\int |h| = 0 \Rightarrow \left| \int_E h \right| \leq \int_E |h| \leq \int |h| = 0,$$

which shows one implication. Now assume that  $\int_E h = 0$  for all  $E \in \mathcal{A}$ , then we can write  $h$  as

$$h = u + iv = (u^+ - u^-) + i(v^+ - v^-).$$

Let  $B := \{x \in X \mid u^+(x) > 0\}$ , then by assumption, we have

$$0 = \int_B h = \operatorname{Re} \int_B h = \int_B u = \int_B u^+ = \int_B u^+ + \int_{B^c} u^+ = \int u^+,$$

hence  $u^+ = 0$  **almost everywhere**. Similarly, we have  $u^-, v^+, v^-$  are all zero **almost everywhere**. This gives us that  $h$  is zero **almost everywhere** as desired.

- (2) ■

HW 5  
Q8 by  
Lemma 2.2.1

DIY

**Theorem 2.3.1** (Dominated Convergence Theorem). Let  $(X, \mathcal{A}, \mu)$  be a **measure space**, and

- Let  $f_n$  be **integrable** on  $X$ .
- $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  **almost everywhere**.
- There is a  $g: X \rightarrow [0, \infty]$  such that  $g$  is **integrable** and

$$\forall n \in \mathbb{N} \quad |f_n(x)| \leq g(x) \text{ a.e.}$$

Then we have

$$\lim_{n \rightarrow \infty} \int f_n = \int f = \int \lim_{n \rightarrow \infty} f_n.$$

**Proof.** Let  $F$  be the countable union of [null set](#) on which the three conditions may fail. Then we see that after modifying the definition of  $f_n$ ,  $f$  and  $g$  on  $F$ , we may assume that all three conditions hold everywhere since modifying on a [null set](#) does not change the integral.

We now consider the  $\mathbb{R}$ -valued case only. Note that the second and the third conditions imply that  $f$  is [integrable](#) since  $|f| \leq g(x)$ . We then see that  $g + f_n \geq 0$  and  $g - f_n \geq 0$  because  $-g \leq f_n \leq g$ . From [Theorem 2.2.2](#), we have

$$\int g + f \leq \liminf_{n \rightarrow \infty} \int g + f_n, \quad \int g - f \leq \liminf_{n \rightarrow \infty} \int g - f_n.$$

From the [linearity of integral](#), we have

$$\int g + \int f \leq \int g + \liminf_{n \rightarrow \infty} \int f_n, \quad \int g - \int f \leq \int g - \liminf_{n \rightarrow \infty} \int f_n.$$

Now, since  $\int g < \infty$ , we can cancel it, which gives

$$\int f \leq \liminf_{n \rightarrow \infty} \int f_n, \quad -\int f \leq \liminf_{n \rightarrow \infty} \int -f_n = -\limsup_{n \rightarrow \infty} \int f_n,$$

which implies

$$\int f \leq \liminf_{n \rightarrow \infty} \int f_n \leq \limsup_{n \rightarrow \infty} \int f_n \leq \int f.$$

This shows that the limit exists, and the desired result indeed holds. ■

**Corollary 2.3.1** (Tonelli's theorem for series and integrals). Suppose  $f_n$  are [integrable](#) functions such that

$$\sum_{n=1}^{\infty} \int |f_n| < \infty,$$

then we have

$$\int \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} \int f_n.$$

**Proof.** Take  $G(x)$  to be

$$G(x) := \sum_{n=1}^{\infty} |f_n(x)|,$$

then we see  $G(x) \geq |F_N(x)|$  where  $F_N(x) := \sum_{n=1}^N f_n(x)$ . By [Corollary 2.2.2](#), we have

$$\int G(x) = \sum_{n=1}^{\infty} \int |f_n(x)| < \infty.$$

Lastly, from [Theorem 2.3.1](#), the result follows. ■

**Remark.** Compare to [Corollary 2.2.2](#), we see that we further generalize the result!

## Lecture 14: $L^1$ Space

### 2.4 $L^1$ Space

7 Feb. 11:00

We now introduce another space called  $L^p$  spaces, which are function spaces defined using a natural generalization of the [p-norm](#) for finite-dimensional vector spaces. We sometimes call it Lebesgue spaces

also.

Before we start, we need to define a *norm*.

**Definition 2.4.1 (Norm).** Let  $V$  be a vector space over field  $\mathbb{R}$  or  $\mathbb{C}$ . A *norm* is a *seminorm*, defined as

**Definition 2.4.2 (Seminorm).** A *seminorm* on  $V$  is

$$\|\cdot\| : V \rightarrow [0, \infty)$$

such that

- $\|cv\| = |c| \|v\|$  for every  $v \in V$  and every scalar  $c$ .
- $\|v + w\| \leq \|v\| + \|w\|$  for every  $v, w \in V$ .

with an additional condition

- $\|v\| = 0 \Leftrightarrow v = 0$ .

**Lemma 2.4.1.** A *normed* vector space is a metric space with metric

$$\rho(v, w) = \|v - w\|.$$

**Proof.** ■

DIY

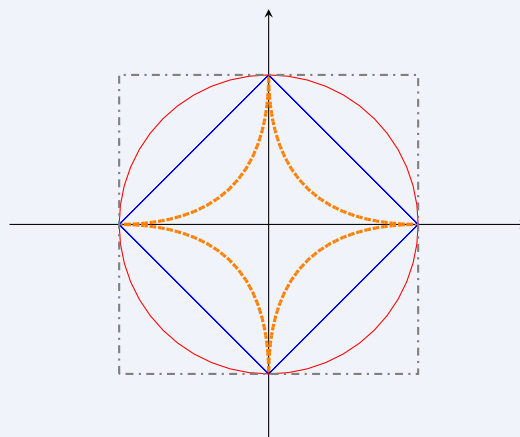
**Example ( $p$ -norm).**  $V = \mathbb{R}^d$  with

$$\|x\|_p = \begin{cases} \left( \sum_{i=1}^d |x_i|^p \right)^{1/p}, & \text{if } p \in [0, \infty); \\ \max_{1 \leq i \leq d} |x_i|, & \text{if } p = \infty \end{cases}$$

is a *normed* vector space. The unit ball

$$\{x \in \mathbb{R}^d \mid \|x\|_p \leq 1\}$$

for different  $p$  has the following figures.



**Remark.** All  $p$ -norms induce the same topology. i.e., if  $U$  is open in  $p$ -norm, it is open in  $p'$ -norm as well.

**Note.** Recall that we say  $f$  is **integrable** means

$$\int |f| < \infty,$$

and if  $f = g$  **a.e.**, then

$$\int f = \int g$$

**Definition 2.4.3** ( $L^1$  Space). Given  $(X, \mathcal{A}, \mu)$ ,

$$f \in L^1(X, \mathcal{A}, \mu) (= L^1(X, \mu) = L^1(X) = L^1(\mu))$$

means that  $f$  is an **integrable** function on  $X$ .

**Lemma 2.4.2.**  $L^1(X, \mathcal{A}, \mu)$  is a vector space with **seminorm**

$$\|f\|_1 = \int |f|.$$

**Proof.** ■

Check this is indeed a **seminorm**.

**Definition 2.4.4** ( $L^1$  Space with equivalence class). Define  $f \sim g$  if  $f = g$  **a.e.**, then

$$L^1(X, \mathcal{A}, \mu) / \sim = L^1(X, \mathcal{A}, \mu),$$

i.e., we simply denote the collection of equivalence classes by itself.<sup>a</sup>

<sup>a</sup>By some abusing of notation of  $L^1$ .

**Remark.** We have

- With **Definition 2.4.4**,  $L^1(X, \mathcal{A}, \mu)$  is a **normed** vector space.
- We say that the  $L^1$ -metric  $\rho(f, g)$  is simply

$$\rho(f, g) = \int |f - g|.$$

## Dense Subsets of $L^1$

**Note.** Recall the definition of a **dense set**<sup>a</sup>.

<sup>a</sup>[https://en.wikipedia.org/wiki/Dense\\_set](https://en.wikipedia.org/wiki/Dense_set)

**Definition 2.4.5** (Step function). A *step function* on  $\mathbb{R}$  is

$$\psi = \sum_{i=1}^N c_i \mathbb{1}_{I_i},$$

where  $I_i$  is an interval.

**Notation.** We denote the collection of continuous functions with compact support by  $C_c(\mathbb{R})$ .

**Theorem 2.4.1.** We have the following.

- (1)  $\{\text{integrable simple functions}\}$  is dense in  $L^1(X, \mathcal{A}, \mu)$  (with respect to  $L^1$ -metric).
- (2)  $(X, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{A}_\mu, \mu)$ , where  $\mu$  is a **Lebesgue-Stieltjes-measure**. Then the set of **integrable simple** functions is dense in  $L^1(\mathbb{R}, \mathcal{A}_\mu, \mu)$ .
- (3)  $C_c(\mathbb{R})$  is dense in  $L^1(\mathbb{R}, \mathcal{L}, m)$ .

**Proof.** We prove this one by one.

- (1) Since there exists **simple functions**  $0 \leq |\phi_1| \leq |\phi_2| \leq \dots \leq |f|$ , where  $\phi_n \rightarrow f$  **pointwise**. Then by **Theorem 2.3.1**, we have

$$\lim_{n \rightarrow \infty} \int \underbrace{|f_n - f|}_{\leq |\phi_n| + |f| \leq 2|f|} = 0$$

where  $2|f|$  is in  $L^1$ .

- (2) Let  $\mathbb{1}_E$  **approximate** by  $\sum_{i=1}^{\infty} c_i \mathbb{1}_{I_i}$ . From **Theorem 1.7.1** for **Lebesgue-Stieltjes-measure**,

$$\forall \epsilon' > 0 \exists I = \bigcup_{i=1}^N I_i \text{ such that } \mu(E \Delta I) \leq \epsilon'.$$

- (3) To approximate  $\mathbb{1}_{(a,b)}$ , we simply consider function  $g \in C_c(\mathbb{R})$  such that

$$\int |\mathbb{1}_{(a,b)} - g| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

■

## Lecture 15: Riemann Integral

### 2.5 Riemann Integrability

9 Feb. 11:00

We are now working in  $(\mathbb{R}, \mathcal{L}, m)$ . Let's first revisit the definition of Riemann Integral. Let  $P$  be a partition of  $[a, b]$  as

$$P = \{a = t_0 < t_1 < \dots < t_k = b\}.$$

Then the *lower Riemann sum* of  $f$  using  $P$  is equal to  $L_P$ , which is defined as

$$L_P = \sum_{i=1}^K \left( \inf_{[t_{i-1}, t_i]} f \right) (t_i - t_{i-1}),$$

and the *upper Riemann sum* of  $f$  using  $P$  is equal to  $U_P$ , which is defined as

$$U_P = \sum_{i=1}^K \left( \sup_{[t_{i-1}, t_i]} f \right) (t_i - t_{i-1}).$$

Then we call

- *Lower Riemann integral* of  $f = \underline{I} = \sup_P L_P$
- *Upper Riemann integral* of  $f = \overline{I} = \inf_P U_P$

**Definition 2.5.1** (Riemann (Darboux) integrable). A **bounded** function  $f: [a, b] \rightarrow \mathbb{R}$  is called *Rie-*



mann (Darboux) integrable if  $\underline{I} = \bar{I}$ . If so, then we write

$$\underline{I} = \bar{I} = \int_a^b f(x) dx.$$

**Note.** We see that

- If  $P \subset P'$ , then

$$L_P \leq L_{P'}, \quad U_{P'} \leq U_P.$$

- Recall that continuous functions on  $[a, b]$  are [Riemann integrable](#) on  $[a, b]$ .

**Theorem 2.5.1.** Let  $f: [a, b] \rightarrow \mathbb{R}$  be a [bounded](#) function. Then

- (1) If  $f$  is [Riemann integrable](#), then  $f$  is [Lebesgue measurable](#), thus [Lebesgue integrable](#). Further,

$$\int_a^b f(x) dx = \int_{[a,b]} f dm.$$

- (2) If  $f$  is [Riemann integrable](#)  $\Leftrightarrow f$  is continuous [Lebesgue a.e.](#)<sup>a</sup>

<sup>a</sup>Here, we mean that the set where  $f$  is discontinuous is a [null set](#) under [Lebesgue measure](#).

**Proof.** There exists  $P_1 \subset P_2 \subset \dots$  such that  $L_{P_n} \nearrow \underline{I}$  and  $U_{P_n} \searrow \bar{I}$ .

**Note.** Here, we took refinements of  $P_n$  if needed.

Now, define [simple \(step\) functions](#)

$$\begin{aligned} \bullet \phi_n &= \sum_{i=1}^K \left( \inf_{[t_{i-1}, t_i]} f \right) \mathbb{1}_{(t_{i-1}, t_i]} \\ \bullet \psi_n &= \sum_{i=1}^K \left( \sup_{[t_{i-1}, t_i]} f \right) \mathbb{1}_{(t_{i-1}, t_i]} \end{aligned}$$

if  $P_n = \{a = t_0 < t_1 < \dots < t_K\}$ . Let  $\phi := \sup_n \phi_n$  and  $\psi := \inf_n \psi_n$ . We then see that  $\phi, \psi$  are [Lebesgue \(Borel\) measurable function](#).

**Note.** Note that

- $\exists M > 0$  such that  $\forall_{n \in \mathbb{N}} |\phi_n|, |\psi_n| \leq M \mathbb{1}_{[a,b]}$
- $\int \phi_n dm = L_{P_n}, \int \psi_n dm = U_{P_n}$

By [Theorem 2.3.1](#) and the fact that  $M \mathbb{1}_{[a,b]} \in L^1(\mathbb{R}, \mathcal{L}, m)$ , we have

$$\underline{I} = \lim_{n \rightarrow \infty} \int \phi_n dm = \int \phi dm, \quad \bar{I} = \lim_{n \rightarrow \infty} \int \psi_n dm = \int \psi dm.$$

Thus,

$$f \text{ is } \text{Riemann integrable} \Leftrightarrow \int \phi = \int \psi \Leftrightarrow \int (\psi - \phi) = 0 \Leftrightarrow \psi = \phi \text{ Lebesgue a.e.}$$

■

## 2.6 Modes of Convergence

As we should already see, there are different *modes* of convergence. Let's formalize them.

**Definition.** Let  $f_n, f: X \rightarrow \mathbb{C}$ , and  $S \subset X$ . Then we have the following definitions.

**Definition 2.6.1** (Pointwise convergence).  $f_n \rightarrow f$  *pointwise* on  $S$  if

$$\forall_{x \in S} \forall_{\epsilon > 0} \exists_{N \in \mathbb{N}} \forall_{n \geq N} |f_n(x) - f(x)| < \epsilon.$$

**Definition 2.6.2** (Uniformly convergence).  $f_n \rightarrow f$  *uniformly* on  $S$  if

$$\forall_{\epsilon > 0} \exists_{N \in \mathbb{N}} \forall_{x \in S} \forall_{n \geq N} |f_n(x) - f(x)| < \epsilon.$$

**Remark.** We see that we can replace  $\forall \epsilon > 0$  by  $\forall k \in \mathbb{N}$  with  $\epsilon$  changing to  $\frac{1}{k}$ .

**Lemma 2.6.1.** Let  $B_{n,k}$  be

$$B_{n,k} := \left\{ x \in X \mid |f_n(x) - f(x)| < \frac{1}{k} \right\}.$$

Then

(1)  $f_n \rightarrow f$  *pointwise* on  $S$  if and only if

$$S \subset \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} B_{n,k}.$$

(2)  $f_n \rightarrow f$  *uniformly* on  $S$  if and only if  $\exists N_1, N_2, \dots \in \mathbb{N}$  such that

$$S \subset \bigcap_{k=1}^{\infty} \bigcap_{n=N_k}^{\infty} B_{n,k}.$$

**Proof.** This essentially follows from [Definition 2.6.1](#). ■

**Definition.** Let  $(X, \mathcal{A}, \mu)$  be a *measure space*. Assuming that  $f_n, f$  are *measurable functions*, then we have the following.

**Definition 2.6.3** (Converge almost everywhere).  $f_n \rightarrow f$  *almost everywhere* means

$$\exists \text{ null set } E \text{ such that } f_n \rightarrow f \text{ pointwise on } E^c.$$

**Definition 2.6.4** (Converge in  $L^1$ ).  $f_n \rightarrow f$  *in  $L^1$*  means

$$\lim_{n \rightarrow \infty} \|f_n - f\| = 0.$$

**Example.** Given  $(\mathbb{R}, \mathcal{L}, m)$  and let  $f = 0$ . Consider the following functions.

- (1)  $f_n = \mathbb{1}_{(n, n+1)}$
- (2)  $f_n = \frac{1}{n} \mathbb{1}_{(0, n)}$
- (3)  $f_n = n \mathbb{1}_{(0, \frac{1}{n})}$
- (4) **Typewriter functions.**



**Proof.** We see that different function sequences converge in different senses.

**Exercise.** Classify in what senses do (1), (2), (3) and the **type write** function converge.

⊛

## Lecture 16: Modes of Convergence

Let's start with a proposition.

11 Feb. 11:00

**Proposition 2.6.1** (Fast  $L^1$  convergence leads to a.e. convergence). Let  $(X, \mathcal{A}, \mu)$  be a **measure space**, and  $f_n, f$  are all **measurable** functions on  $X$ . Then

$$\sum_{n=1}^{\infty} \|f_n - f\|_1 < \infty \Rightarrow f_n \rightarrow f \text{ a.e.}$$

**Proof.** Let

$$E := \bigcup_{k=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} B_{n,k}^c = \{x \in X \mid f_n(x) \not\rightarrow f(x)\}.$$

By **Lemma 2.2.1**, we see that

$$\forall_k \forall_N \mu(B_{n,k}^c) \leq k \int |f_n - f| \Rightarrow \forall_k \mu\left(\bigcup_{n=N}^{\infty} B_{n,k}^c\right) \leq \sum_{n=N}^{\infty} k \|f_n - f\|_1 \rightarrow 0$$

as  $N \rightarrow \infty$ . Now, by [continuity of measure from above](#),

$$\forall_k \mu \left( \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} B_{n,k}^c \right) = \lim_{N \rightarrow \infty} \mu \left( \bigcup_{n=N}^{\infty} B_{n,k}^c \right) = 0 \Rightarrow \mu(E) = 0$$

since  $f_n \rightarrow f$  [pointwise](#) on  $E^c$ . ■

**Corollary 2.6.1.** Given  $\{f_n\}_n$  such that  $f_n \rightarrow f$  in  $L^1$ , there exists a subsequence  $\{f_{n_j}\}_{n_j}$  where  $f_{n_j} \rightarrow f$  a.e.

**Proof.** Since

$$\forall_{j \in \mathbb{N}} \forall_{n_j \in \mathbb{N}} \|f_{n_j} - f\|_1 \leq \frac{1}{j^2}.$$

Then,

$$\sum_{j=1}^{\infty} \|f_{n_j} - f\|_1 < \infty.$$

From [Proposition 2.6.1](#), we have the desired result. ■

**Definition 2.6.5** (Converge in measure). Let  $f_n, f$  be [measurable functions](#) on  $(X, \mathcal{A}, \mu)$ . Then  $f_n \rightarrow f$  in measure if

$$\forall_{\epsilon > 0} \lim_{n \rightarrow \infty} \mu(\{x \in X \mid |f_n(x) - f(x)| \geq \epsilon\}) = 0.$$

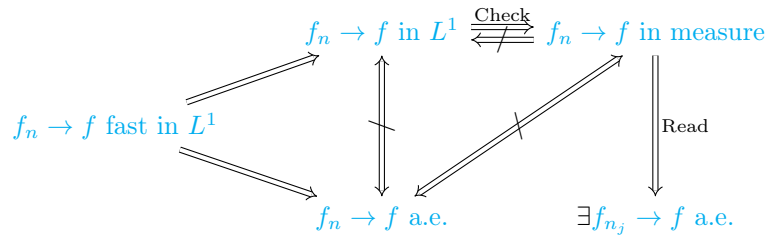
**Example.** Let  $f_n = n \mathbb{1}_{(0, \frac{1}{n})}$  and  $f = 0$  on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ , then  $f_n \rightarrow f$  in measure.

**Proof.** We see that

$$\forall \epsilon > 0 \quad \{x \in X \mid |f_n(x) - f(x)| > \epsilon\} = \left(0, \frac{1}{n}\right),$$

$f_n \rightarrow 0$  in measure. (Recall that  $f_n \not\rightarrow 0$  in  $L^1$ ) ⊛

**Remark.** We see that



Finally, we have the following.

**Definition.** Let  $f_n, f$  be [measurable functions](#) on  $(X, \mathcal{A}, \mu)$ .

**Definition 2.6.6** (Uniformly almost everywhere).  $f_n \rightarrow f$  uniformly almost everywhere if  $\exists$  null set  $F$  such that  $f_n \rightarrow f$  uniformly on  $F^c$ .

**Definition 2.6.7** (Almost uniformly).  $f_n \rightarrow f$  almost uniformly if  $\forall \epsilon > 0 \exists F \in \mathcal{A}$  such that  $\mu(F) < \epsilon$ ,  $f_n \rightarrow f$  uniformly on  $F^c$ .

**Lemma 2.6.2.** We have

$$f_n \rightarrow f \text{ uniformly on } S \Leftrightarrow \exists N_1, N_2, \dots \in \mathbb{N} \ S \subset \bigcap_{k=1}^{\infty} \bigcap_{n=N_k}^{\infty} B_{n,k}.$$

**Theorem 2.6.1** (Egorov's Theorem). Let  $f_n, f$  be measurable functions on  $(X, \mathcal{A}, \mu)$ . Suppose  $\mu(X) < \infty$ , then

$$f_n \rightarrow f \text{ a.e.} \Leftrightarrow f_n \rightarrow f \text{ almost uniformly.}$$

**Proof.** We prove two directions.

( $\Leftarrow$ )

( $\Rightarrow$ ) Fix  $\epsilon > 0$ . We see that

$$f_n \rightarrow f \text{ a.e.} \Rightarrow \mu \left( \bigcup_{k=1}^{\infty} \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} B_{n,k}^c \right) = 0 \Rightarrow \forall_k \mu \left( \bigcap_{N=1}^{\infty} \bigcup_{n=N}^{\infty} B_{n,k}^c \right) = 0.$$

From continuity of measure from above and  $\mu(X) < \infty$ , we further have

$$\forall_k \lim_{N \rightarrow \infty} \mu \left( \bigcup_{n=N}^{\infty} B_{n,k}^c \right) = 0 \Rightarrow \forall_k \exists_{N_k \in \mathbb{N}} \mu \left( \bigcup_{n=N_k}^{\infty} B_{n,k}^c \right) < \frac{\epsilon}{2^k}.$$

Now, let

$$F := \bigcup_{k=1}^{\infty} \bigcup_{n=N_k}^{\infty} B_{n,k}^c,$$

we see that  $\mu(F) < \epsilon$ , hence  $f_n \rightarrow f$  uniformly. ■

DIY

# Chapter 3

## Product Measure

### 3.1 Product $\sigma$ -algebra

Before we start, we see the setup.

- Product space.

$$X = \prod_{\alpha \in I} X_{\alpha}$$

where  $x = (x_{\alpha})_{\alpha \in I} \in X$ .

- Coordinate map.

$$\pi_{\alpha}: X \rightarrow X_{\alpha}.$$

Now we see the formal definition.

**Definition 3.1.1** (Product  $\sigma$ -algebra). Let  $(X_{\alpha}, \mathcal{A}_{\alpha})$  be a measurable space for all  $\alpha \in I$ . Then a product  $\sigma$ -algebra on  $X = \prod_{\alpha \in I} X_{\alpha}$  is

$$\bigotimes_{\alpha \in I} \mathcal{A}_{\alpha} = \left\langle \bigcup_{\alpha \in I} \pi_{\alpha}^{-1}(\mathcal{A}_{\alpha}) \right\rangle,$$

where  $\pi_{\alpha}^{-1}(\mathcal{A}_{\alpha}) = \{\pi_{\alpha}^{-1}(E) \mid E \in \mathcal{A}_{\alpha}\}$ .

**Notation.** We denote  $I = \{1, \dots, d\} \Rightarrow X = \prod_{i=1}^d X_i, x = (x_1, \dots, x_d)$ . Also,

$$\bigotimes_{i=1}^d \mathcal{A}_i = \mathcal{A}_1 \otimes \dots \otimes \mathcal{A}_d.$$

**Lemma 3.1.1.** If  $I$  is countable, then

$$\bigotimes_{i=1}^{\infty} \mathcal{A}_i = \left\langle \left\{ \prod_{i=1}^{\infty} E_i \mid \forall_i E_i \in \mathcal{A}_i \right\} \right\rangle.$$

**Proof.** If  $E_i \in \mathcal{A}_i$ , then  $\pi_i^{-1}(E_i) = \prod_{j=1}^{\infty} E_j$ , where  $E_j = X_j$  for  $j \neq i$ . On the other hand, since

$$\prod_{i=1}^{\infty} E_i = \bigcap_{i=1}^{\infty} \pi_i^{-1}(E_i),$$

from Lemma 1.1.2, the result follows. ■

## Lecture 17: Product Measure

We now see a lemma.

14 Feb. 11:00

**Lemma 3.1.2.** Suppose  $\mathcal{A}_\alpha = \langle \mathcal{E}_\alpha \rangle$  for every  $\alpha \in I$ . Then

- (1)  $\pi_\alpha^{-1}(\mathcal{A}_\alpha) = \langle \pi_\alpha^{-1}(\mathcal{E}_\alpha) \rangle$
- (2)  $\bigotimes_\alpha \mathcal{A}_\alpha = \langle \bigcup_\alpha \pi_\alpha^{-1}(\mathcal{E}_\alpha) \rangle$
- (3) If  $I$  is countable, then

$$\bigotimes_{i=1}^{\infty} \mathcal{A}_i = \left\langle \left\{ \prod_{i=1}^{\infty} E_i \mid \forall_i E_i \in \mathcal{E}_i \right\} \right\rangle$$

**Proof.** We prove this one by one.

- (1) Note that for  $f: Y \rightarrow Z$ , and  $\mathcal{B}$  be a  $\sigma$ -algebra on  $Z$ , then  $f^{-1}(\mathcal{B})$  is also a  $\sigma$ -algebra since  $f^{-1}(\emptyset) = \emptyset$ ,  $f^{-1}(B)^c = f^{-1}(B^c)$ , and  $\bigcup_n f^{-1}(B_n) = f^{-1}(\bigcup_n B_n)$ . Hence,  $\pi_\alpha^{-1}$  is a  $\sigma$ -algebra on  $X$ , i.e.,

$$\pi_\alpha^{-1}(\mathcal{E}_\alpha) \subset \pi_\alpha^{-1}(\mathcal{A}_\alpha) \stackrel{!}{\Rightarrow} \langle \pi_\alpha^{-1}(\mathcal{E}_\alpha) \rangle \subset \pi_\alpha^{-1}(\mathcal{A}_\alpha).$$

To show the other direction, let  $\mathcal{M}$  being

$$\mathcal{M} = \{B \subset X_\alpha \mid \pi_\alpha^{-1}(B) \in \langle \pi_\alpha^{-1}(\mathcal{E}_\alpha) \rangle\}.$$

We now check

- $\mathcal{M}$  is a  $\sigma$ -algebra.
- $\mathcal{E}_\alpha \subset \mathcal{M}$ . This is true by definition of  $\mathcal{M}$ .

Thus,  $\langle \mathcal{E}_\alpha \rangle = \mathcal{A}_\alpha \subset \mathcal{M}$ . Hence, if  $E \in \mathcal{A}_\alpha$ ,  $E \in \mathcal{M}$ , implying

$$\pi_\alpha^{-1}(E) \in \langle \pi_\alpha^{-1}(\mathcal{E}_\alpha) \rangle,$$

i.e.,  $\mathcal{A}_\alpha \subset \langle \pi_\alpha^{-1}(\mathcal{E}_\alpha) \rangle$ .

(2)

(3)

Check  
(Easy)!

DIY

DIY

■

**Theorem 3.1.1.** Suppose  $X_1, \dots, X_d$  are metric spaces. Let  $X = \prod_{i=1}^d X_i$  with product metric defined as

$$\rho(x, y) = \sum_{i=1}^d \rho_i(x_i, y_i).$$

Then,

$$(1) \bigotimes_{i=1}^d \mathcal{B}(X_i) \subset \mathcal{B}(X)$$

(2) If in addition, each  $X_i$  has a countable dense subset,

$$\bigoplus_{i=1}^d \mathcal{B}(X_i) = \mathcal{B}(X).$$

**Proof.** ■

DIY

**Remark.** We see that

- $\mathcal{B}(\mathbb{R}^d) = \mathcal{B}(\mathbb{R}) \otimes \dots \otimes \mathcal{B}(\mathbb{R})$
- let  $f = u + iv: X \rightarrow \mathbb{C}$ , and  $\mathcal{A}$  be a  $\sigma$ -algebra on  $X$ . Then

$$\forall_{E \in \mathcal{B}(\mathbb{R})} u^{-1}(E), v^{-1}(E) \in \mathcal{A} \Leftrightarrow f^{-1}(F) \in \mathcal{A}, \forall F \in \mathcal{B}(\mathbb{C})$$

with  $\mathcal{B}(\mathbb{C}) = \mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ .

We first focus on 2 dimensional case. Specifically, we think of our coordinate is  $x$  and  $y$  on  $\mathbb{R}^2$ .

**Definition.** Let  $X, Y$  be two sets, then we have the following.

**Definition 3.1.2** (*x-section, y-section for set*). For  $E \subset X \times Y$ ,

$$E_x = \{y \in Y \mid (x, y) \in E\}, \quad E^y = \{x \in X \mid (x, y) \in E\},$$

where  $E_x$  is called the *x-section* of  $E$ , while  $E_y$  is called the *y-section* of  $E$ .

**Definition 3.1.3** (*x-section, y-section for function*). For  $f: X \times Y \rightarrow \mathbb{C}$ , define  $f_x: Y \rightarrow \mathbb{C}$ ,  $f^y: X \rightarrow \mathbb{C}$  by

$$f_x(y) = f(x, y) = f^y(x),$$

where  $f_x(y)$  is called the *x-section* of  $f$ , while  $f_y(x)$  is called the *y-section* of  $f$ .

**Example.** We see that

$$(\mathbb{1}_E)_x = \mathbb{1}_{E_x}$$

and

$$(\mathbb{1}_E)^y = \mathbb{1}_{E^y}.$$

**Proposition 3.1.1.** Given two measurable spaces  $(X, \mathcal{A})$  and  $(Y, \mathcal{B})$ , then

(1) If  $E \in \mathcal{A} \otimes \mathcal{B}$ , then

$$\forall_{x \in X} \forall_{y \in Y} E_x \in \mathcal{B}, E^y \in \mathcal{A}.$$

(2) If  $f: X \times Y \rightarrow \mathbb{C}$  is  $\mathcal{A} \otimes \mathcal{B}$ -measurable, then

$$\forall_{x \in X} \forall_{y \in Y} f_x \text{ is } \mathcal{B}\text{-measurable, } f^y \text{ is } \mathcal{A}\text{-measurable.}$$

**Proof.** We prove this one by one.

(1) Let  $\mathcal{F} := \left\{ E \subset X \times Y \mid \forall_{x \in X} \forall_{y \in Y} E_x \in \mathcal{B}, E^y \in \mathcal{A} \right\}$ , then



- $\mathcal{F}$  is a  $\sigma$ -algebra.
  - $\emptyset \in \mathcal{F}$ .
  - $(E^c)_x = E_x^c$ .
  - $\left(\bigcup_{j=1}^{\infty} E_j\right)_x = \bigcup_{j=1}^{\infty} (E_j)_x$ .

And the same is true for  $y$ .

- Let  $\mathcal{R}_0 := \{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\} \subset \mathcal{F}$ , which is again easy to show from definition.

Hence, we see that  $\langle \mathcal{R}_0 \rangle = \mathcal{A} \otimes \mathcal{B} \subset \mathcal{F}$ .

(2) Since

$$(f_x)^{-1}(B) = (f^{-1}(B))_x$$

and

$$(f^y)^{-1}(B) = (f^{-1}(B))^y,$$

the result follows from (1). ■

## 3.2 Product Measures

We start with the definition.

**Definition 3.2.1 (Rectangle).** Given two measurable spaces  $(X, \mathcal{A}), (Y, \mathcal{B})$ , a (measurable) rectangle is  $R = A \times B$  where  $A \in \mathcal{A}$  and  $B \in \mathcal{B}$ . Furthermore, we let

$$\mathcal{R}_0 := \{R = A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\},$$

and

$$\mathcal{R} := \left\{ \bigcup_{i=1}^N R_i \mid N \in \mathbb{N}, R_1, \dots, R_N \text{ disjoint rectangles} \right\}.$$

**Note.** Whenever we're talking about rectangle, they're always measurable.

**Lemma 3.2.1.**  $\mathcal{R}$  is an algebra, and

$$\langle \mathcal{R}_0 \rangle = \langle \mathcal{R} \rangle = \mathcal{A} \otimes \mathcal{B}.$$

**Proof.** Simply observe that

$$(A \times B)^c = (A^c \times Y) \cup (X \times B^c).$$
■

DIY

## Lecture 18: Monotone Class

Let's start with a theorem.

16 Feb. 11:00

**Theorem 3.2.1.** Let  $(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$  be measure spaces. Then

(1) There is a measure  $\mu \times \nu$  on  $\mathcal{A} \otimes \mathcal{B}$  satisfying

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

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

(2) If  $\mu, \nu$  are  $\sigma$ -finite, then  $\mu \times \nu$  is unique.

**Proof.** We prove this one by one.

(1) Define  $\pi: \mathcal{R} \rightarrow [0, \infty]$  by  $\pi(A \times B) = \mu(A)\nu(B)$ , and extending linearly, we have

$$\pi(A \times B) = \mu(A)\nu(B),$$

hence

$$\pi\left(\prod_{i=1}^N A_i \times B_i\right) = \sum_{i=1}^n \pi(A_i \times B_i).$$

We claim that  $\pi$  is a **pre-measure**. To show this, it's enough to check that  $\pi(A \times B) = \sum_{n=1}^{\infty} \pi(A_n \times B_n)$  if  $A \times B = \prod_n A_n \times B_n$ . Since  $A_n \times B_n$  are disjoint, so

$$\mathbb{1}_{A \times B}(x, y) = \sum_{n=1}^{\infty} \mathbb{1}_{A_n \times B_n}(x, y).$$

Thus,

$$\mathbb{1}_A(x) \mathbb{1}_B(y) = \sum_{n=1}^{\infty} \mathbb{1}_{A_n}(x) \mathbb{1}_{B_n}(y).$$

Integrating with respect to  $x$ , and applying [Theorem 1.3.1](#), we have

$$\int_X \mathbb{1}_A(x) \mathbb{1}_B(y) d\mu(x) = \sum_{n=1}^{\infty} \int_X \mathbb{1}_{A_n}(x) \mathbb{1}_{B_n}(y) d\mu(x),$$

which implies

$$\mu(A) \mathbb{1}_B(y) = \sum_{n=1}^{\infty} \mu(A_n) \mathbb{1}_{B_n}(y)$$

for every  $y$ . We can then integrate again with respect to  $y$  and apply [Theorem 1.3.1](#), we have

$$\int_Y \mu(A) \mathbb{1}_B(y) d\nu(y) = \sum_{n=1}^{\infty} \int_Y \mu(A_n) \mathbb{1}_{B_n}(y) d\nu(y),$$

which gives us

$$\mu(A)\nu(B) = \sum_{n=1}^{\infty} \mu(A_n)\nu(B_n).$$

Hence, we see that  $\mu$  is indeed a **pre-measure**, so [Theorem 1.4.1](#) gives  $\mu \times \nu$  on  $\langle \mathcal{R} \rangle = \mathcal{A} \otimes \mathcal{B}$  extending  $\pi$  on  $\mathcal{R}$ .

(2) If  $\mu, \nu$  are  $\sigma$ -finite, then  $\pi$  is  $\sigma$ -finite on  $\mathcal{R}$ , then [Theorem 1.4.2](#) applies. Moreover, we have that

$$(\mu \times \nu)(E) = \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i) \nu(B_i) \mid E \subset \bigcup_{i=1}^{\infty} A_i \times B_i, A_i \in \mathcal{A}, B_i \in \mathcal{B} \right\}.$$

■

### 3.3 Monotone Class Lemma

Let's start with a definition.

**Definition 3.3.1 (Monotone Class).** If  $X$  is a set, and  $C \subset \mathcal{P}(X)$ , we say that  $C$  is a *monotone class* on  $X$  if

- $C$  is closed under countable increasing unions.
- $C$  is closed under countable decreasing intersections.

**Example.** Every  $\sigma$ -algebra is a monotone class.

**Example.** If  $C_\alpha$  are (arbitrarily many) monotone classes on a set  $X$ , then  $\bigcap_{\alpha} C_\alpha$  is a monotone class. Furthermore, if  $\mathcal{E} \subset \mathcal{P}(X)$ , there is a unique smallest monotone class containing  $\mathcal{E}$ , denoted by  $\langle \mathcal{E} \rangle$ , which follows the same idea as in Definition 1.1.3.

**Theorem 3.3.1 (Monotone Class Lemma).** Suppose  $\mathcal{A}_0$  is an algebra on  $X$ . Then  $\langle \mathcal{A}_0 \rangle^a$  is the monotone class generated by  $\mathcal{A}_0$ .

$^a\langle \mathcal{A}_0 \rangle$  is the  $\sigma$ -algebra generated by  $\mathcal{A}_0$  by Definition 1.1.3.

**Proof.** Let  $\mathcal{A} = \langle \mathcal{A}_0 \rangle$  and let  $\mathcal{C}$  be the monotone class generated by  $\mathcal{A}_0$ . Since  $\mathcal{A}$  is a  $\sigma$ -algebra, it's a monotone class. Note that it contains  $\mathcal{A}_0$ , hence  $\mathcal{A} \supset \mathcal{C}$ .

To show  $\mathcal{C} \supset \mathcal{A}$ , it's enough to show that  $\mathcal{C}$  is a  $\sigma$ -algebra. We check that

1.  $\emptyset \in \mathcal{A}_0 \subseteq \mathcal{C}$ .
2. Let  $\mathcal{C}' = \{E \subset X \mid E^c \in \mathcal{C}\}$ .
  - $\mathcal{C}'$  is a monotone class.
  - $\mathcal{A}_0 \subset \mathcal{C}'$  because if  $E \in \mathcal{A}_0$ , then  $E^c \in \mathcal{A}_0$ , so  $E^c \in \mathcal{C}$ , thus  $E \in \mathcal{C}'$ .

We see that  $\mathcal{C}' \subset \mathcal{C}'$ , so  $\mathcal{C}$  is closed under complements.

3. For  $E \subset X$ , let  $\mathcal{D}(E) = \{F \in \mathcal{C} \mid E \cup F \in \mathcal{C}\}$ .
  - $\mathcal{D}(E) \subset \mathcal{C}$ .
  - $\mathcal{D}(E)$  is a monotone class.
  - If  $E \in \mathcal{A}_0$ , then  $\mathcal{A}_0 \subset \mathcal{D}(E)$ . We see this by picking  $F \in \mathcal{A}_0$ , then  $E \cup F \in \mathcal{A}_0 \subset \mathcal{C}$ .

Hence,  $\mathcal{C} = \mathcal{D}(E)$  if  $E \in \mathcal{A}_0$ .

4. Let  $\mathcal{D} = \{E \in \mathcal{C} \mid \mathcal{D}(E) = \mathcal{C}\}$ . That is  $\mathcal{D} = \{E \in \mathcal{C} \mid E \cup F \in \mathcal{C}, \forall F \in \mathcal{C}\}$ . Then we have
  - $\mathcal{A}_0 \subset \mathcal{D}$  by 3.
  - $\mathcal{D}$  is a monotone class.
  - $\mathcal{D} \subset \mathcal{C}$  by definition.

Thus,  $\mathcal{D} = \mathcal{C}$ , so if  $E, F \in \mathcal{C}$ , then  $E \cup F \in \mathcal{C}$ . This implies that  $\mathcal{C}$  is closed under finite unions.

5. Now to show that  $\mathcal{C}$  is closed under countable unions, let  $E_1, E_2, \dots \in \mathcal{C}$ . We may then define

$$F_N = \bigcup_{n=1}^N E_n \in \mathcal{C}.$$

Then we see that  $F_1 \subset F_2 \subset \dots$ , hence  $\bigcup_N F_N \in \mathcal{C}$ . But this simply implies

$$\bigcup_N F_N = \bigcup_n E_n,$$

so we're done. ■

## Lecture 19: Fubini-Tonelli's Theorem

**As previously seen.** If  $E \in \mathcal{A} \otimes \mathcal{B} \Rightarrow E_x \in \mathcal{B}, E^y \in \mathcal{A} \forall x \in X, \forall y \in Y$ . Note that the reverse is not true.

18 Feb. 11:00

### 3.4 Fubini-Tonelli Theorem

We start with a theorem.

**Theorem 3.4.1** (Tonelli's theorem for characteristic functions). Given  $(X, \mathcal{A}, \mu)$  and  $(Y, \mathcal{B}, \nu)$  be  $\sigma$ -finite measure space. Suppose  $E \in \mathcal{A} \otimes \mathcal{B}$ , then

- (1)  $\alpha(x) := \nu(E_x): X \rightarrow [0, \infty]$  is a  $\mathcal{A}$ -measurable function.
- (2)  $\beta(y) := \mu(E^y): Y \rightarrow [0, \infty]$  is a  $\mathcal{B}$ -measurable function.
- (3)  $(\mu \times \nu)(E) = \int_X \nu(E_x) d\mu(x) = \int_Y \mu(E^y) d\nu(y)$ .

**Proof.** We prove this in two cases.

1. Assume that  $\mu, \nu$  are finite measures. Let

$$C := \{E \in \mathcal{A} \otimes \mathcal{B} \mid \text{Conditions (1), (2), (3) hold}\}.$$

It's enough to prove that  $\langle \mathcal{R} \rangle = \mathcal{A} \otimes \mathcal{B} \subset C$ . We further observe that from the Theorem 3.3.1 and the fact that  $\mathcal{R}$  is an algebra, it's also enough to show that

- $\mathcal{R} \subset C$ .
- $C$  is a monotone class.

From condition (1),

$$\alpha(x) = \nu((A \times B)_x) = \begin{cases} \nu(B), & \text{if } x \in A; \\ 0, & \text{if } x \notin A \end{cases} = \nu(B) \mathbb{1}_A.$$

And from condition (2),

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

and

$$\int_X \nu((A \times B)_x) d\mu(x) = \nu(B)\mu(A).$$

Let  $E_n \in C, E_1 \subset E_2 \subset \dots$ . We need to show  $E = \bigcup_{n=1}^{\infty} E_n \in C$ . We now see that

$$E_x = \bigcup_{n=1}^{\infty} (E_n)_x, (E_1)_x \subset (E_2)_x \subset \dots \Rightarrow \alpha(x) = \nu(E_x) \stackrel{!}{=} \lim_{n \rightarrow \infty} \nu((E_n)_x) \quad \forall x \in X.$$

This implies that (1) is proved.

For (3), we see that

$$(\mu \times \nu)(E) \stackrel{!}{=} \lim_{n \rightarrow \infty} (\mu \times \nu)(E_n) = \lim_{n \rightarrow \infty} \int_X \nu((E_n)_x) d\mu(x) \stackrel{!}{=} \int_X \nu(E_x) d\mu(x).$$

Now let  $F_n \in C, F_1 \supset F_2 \supset \dots$ . We need to show that  $F = \bigcap_{n=1}^{\infty} F_n \in C$ . Instead of using Theorem 2.2.1, we now want to use Theorem 2.3.1, which is applicable since  $\mu(X), \nu(Y) < \infty$  by assumption.

2. Assume  $\mu$  and  $\nu$  are  $\sigma$ -finite measures. We then have a sequence  $\{X_n \times Y_n\}$  of rectangles of with only finite measure. Now, just consider if  $E \in \mathcal{A} \otimes \mathcal{B}$ , 1. applies to  $E \cap (X_n \times Y_n)$  for each  $n$ , with

$$X \times Y = \bigcup_{n=1}^{\infty} (X_n \times Y_n), \begin{cases} X_1 \subset X_2 \subset \dots, & \mu(X_k) < \infty \\ Y_1 \subset Y_2 \subset \dots, & \nu(Y_k) < \infty, \end{cases}$$

we have

$$\mu \times \nu(E \cap (X_n \times Y_n)) = \int \mathbb{1}_{X_n}(x) \cdot \nu(E_x \cap Y_n) d\mu(x) = \int \mathbb{1}_{Y_n}(y) \mu(E^y \cap X_n) d\nu(y).$$

By applying Theorem 2.2.1, the result follows. ■

**Theorem 3.4.2** (Fubini-Tonelli's Theorem). Given two  $\sigma$ -finite measure space  $(X, \mathcal{A}, \mu), (Y, \mathcal{B}, \nu)$ , we have the following.

**Theorem** (Tonelli's Theorem). If  $f: X \times Y \rightarrow [0, \infty]$  is  $\mathcal{A} \otimes \mathcal{B}$ -measurable, then

- (1)  $g(x) := \int_Y f(x, y) d\nu(y)$ ,  $X \rightarrow [0, \infty]$  is a  $\mathcal{A}$ -measurable function.
- (2)  $h(x) := \int_X f(x, y) d\mu(x)$ ,  $Y \rightarrow [0, \infty]$  is a  $\mathcal{B}$ -measurable function.
- (3) We have

$$\int_{X \times Y} f d(\mu \times \nu) = \int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left( \int_X f(x, y) d\mu(x) \right) d\nu(y).$$

**Theorem** (Fubini's Theorem). If  $f \in L^1(X \times Y, \mu \times \nu)$ , then

- (1)  $f_x \in L^1(Y, \nu)$  for  $\mu$ -a.e.  $x$ , and  $g(x) \in L^1(X, \mu)$  defined  $\mu$ -a.e.
- (2)  $f^y \in L^1(X, \mu)$  for  $\nu$ -a.e.  $y$ , and  $h(y) \in L^1(Y, \nu)$  defined  $\nu$ -a.e.
- (3) The iterated integral formulas hold. Namely, we have

$$\int_{X \times Y} f d(\mu \times \nu) = \int_X \left( \int_Y f(x, y) d\nu(y) \right) d\mu(x) = \int_Y \left( \int_X f(x, y) d\mu(x) \right) d\nu(y).$$

**Proof.** Read [FF99]. ■

**Remark.** The Fubini and Tonelli's theorem are frequently used in tandem. Say that if one want to reverse the order of integration in a double integral  $\iint f d\mu d\nu$ . We first verify that  $\int |f| d(\mu \times \nu) < \infty$  by using Tonelli's theorem to evaluate this integral as an iterated integral. Then, we apply Fubini theorem to conclude that

$$\iint f d\mu d\nu = \iint f d\nu d\mu.$$

## Lecture 20: Lebesgue Measure on $\mathbb{R}^d$

### 3.5 Lebesgue Measure on $\mathbb{R}^d$

21 Feb. 11:00

**Example.**  $(\mathbb{R}^2, \mathcal{L} \otimes \mathcal{L}, m \times m)$  is not complete.

**Proof.** We see that

- Let  $A \in \mathcal{L}$ ,  $A \neq \emptyset$ ,  $m(A) = 0$ .
- Let  $B \subset [0, 1]$ ,  $B \notin \mathcal{L}$  (Vital set for example).
- Let  $E = A \times B$ ,  $F = A \times [0, 1]$ .

We see that  $E \subset F$ ,  $F \in \mathcal{L} \otimes \mathcal{L}$ ,  $(m \times m)(F) = m(A)m([0, 1]) = 0$ , i.e.,  $F$  is a **null** set. But  $E$  is **not**  $\mathcal{L} \otimes \mathcal{L}$ -measurable since otherwise, its sections are all **measurable**.  $\circledast$

**Definition 3.5.1.** Let  $(\mathbb{R}^d, \mathcal{L}^d, m^d)$  be the *completion* of

$$(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), m \times \dots \times m),$$

which is same as the *completion* of

$$(\mathbb{R}^d, \mathcal{L} \otimes \dots \otimes \mathcal{L}, m \times \dots \times m).$$

**Remark.** We see that

$$\mathcal{L}^d \supsetneq \mathcal{L} \otimes \dots \otimes \mathcal{L} = \left\langle \left\{ \prod_{i=1}^d E_i \mid E_i \in \mathcal{L} \right\} \right\rangle.$$

**Definition 3.5.2** (Rectangle in  $\mathbb{R}^d$ ). A *rectangle* in  $\mathbb{R}^d$  is  $R = \prod_{i=1}^d E_i$  where  $E_i \in \mathcal{B}(\mathbb{R})$ .

**Definition 3.5.3** (Lebesgue measure in  $\mathbb{R}^d$ ). We let the *Lebesgue measure* in  $\mathbb{R}^d$ , denoted as  $m^d$ , defined as

$$m^d(E) := \inf \left\{ \sum_{k=1}^{\infty} m^d(R_k) \mid E \subset \bigcup_{k=1}^{\infty} R_k, R_k \text{ is rectangles} \right\}.$$

**Theorem 3.5.1.** Let  $E \subset \mathcal{L}^d$ . Then

- (1)  $m^d(E) = \inf \{m^d(O) \mid \text{open } O \supset E\} = \sup \{m^d(K) \mid \text{compact } K \subset E\}$ .
- (2)  $E = A_1 \cup N_1 = A_2 \setminus N_2$ , where  $A_1$  is  $F_\sigma$ ,  $A_2$  is  $G_\delta$ , and  $N_i$  are **null**.
- (3) If  $m^d(E) < \infty$ ,  $\forall \epsilon > 0$ ,  $\exists R_1, \dots, R_m$  **rectangles** whose sides are intervals such that

$$m^d \left( E \triangle \left( \bigcup_{i=1}^m R_i \right) \right) < \epsilon.$$

**Proof.** Similar to  $d = 1$  case.  $\blacksquare$

**Theorem 3.5.2.** **Integrable step functions** and  $C_c(\mathbb{R}^d)$ , the collection of continuous functions, are dense in  $L^1(\mathbb{R}^d, \mathcal{L}^d, m^d)$

**Proof.** See [FF99].  $\blacksquare$

**Theorem 3.5.3.** **Lebesgue measure** in  $\mathbb{R}^d$  is translation-invariant.

**Proof.** See [FF99].  $\blacksquare$

**Theorem 3.5.4** (Effect of linear transformation on Lebesgue measure). If  $T \in \text{GL}(\mathbb{R}^d)$ ,  $e \in \mathcal{L}^d$ , then  $T(E)$  is **measurable** and

$$m(T(E)) = |\det T| \cdot m(E).$$

**Proof.** See [FF99].



## Chapter 4

# Differentiation on Euclidean Space

**As previously seen.** Given  $f: [a, b] \rightarrow \mathbb{R}$ , there are two versions of **fundamental theorem of calculus**:

(1)

$$\int_a^b f'(x) dx = f(b) - f(a).$$

(2)

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

which follows from

$$\lim_{r \rightarrow 0^+} \frac{1}{r} \int_x^{x+r} f(t) dt = f(x) = \lim_{r \rightarrow 0^+} \frac{1}{r} \int_{x-r}^x f(t) dt.$$

**Remark.** We see that

$$\lim_{r \rightarrow 0^+} \frac{1}{r} \int_x^{x+r} (f(t) - f(x)) dt = 0 = \lim_{r \rightarrow 0^+} \frac{1}{r} \int_{x-r}^x (f(t) - f(x)) dt,$$

where we have

$$f(x) = \frac{1}{r} \int_x^{x+r} f(t) dt.$$

This generalized to  $f: \mathbb{R}^d \rightarrow \mathbb{R}$ , namely

$$\lim_{r \rightarrow 0^+} \frac{1}{\text{vol}(B(x, r))} \int_{B(x, r)} (f(t) - f(x)) \underbrace{dt}_{\mathbb{R}^d} \stackrel{?}{=} 0.$$

### 4.1 Hardy-Littlewood Maximal Function

We first see our notation.

**Notation.** Given a(n) (open) ball in  $\mathbb{R}^d$ ,  $B = B(a, r)$ , denote  $cB = B(a, cr)$  for  $c > 0$ .

**Lemma 4.1.1 (Vitali-type covering lemma).** Let  $B_1, \dots, B_k$  be a finite collection of open balls in  $\mathbb{R}^d$ . Then there exists a sub-collection  $B'_1, \dots, B'_m$  of disjoint open balls such that

$$\bigcup_{i=1}^m (3B'_i) \supset \bigcup_{i=1}^k B_i.$$



**Proof.** Greedy Algorithm. ■

## Lecture 21: Hardy-Littlewood Maximal Function and Inequality

**Notation.** We let

$$\int_E f \, dm = \int_E f(x) \, dx.$$

The problem we're working on is

$$\frac{1}{m(B(w, r))} \int_{B(x, r)} f(y) \, dy \xrightarrow[?]{r \rightarrow 0} f(x).$$

**Definition 4.1.1 (Locally integrable).** Given  $f: \mathbb{R}^d \rightarrow \mathbb{C}$  be [Lebesgue measurable](#) function. Then we say  $f$  is *locally integrable* if for every compact  $K \subset \mathbb{R}^d$ ,

$$\int_K |f| \, dm < \infty.$$

In this case, we write  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ .

**Definition 4.1.2 (Hardy-Littlewood maximal function).** Given  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ , the *Hardy-Littlewood maximal function* for  $f$  is defined as

$$Hf(x) := \sup \{A_r(x) \mid r > 0\},$$

where

$$A_r(x) := \frac{1}{m(B(x, r))} \int_{B(x, r)} |f(y)| \, dy.$$

**Note.** We note that  $A_r(\cdot)$  means *averaging function* over an open ball with radius being  $r$ .

**Lemma 4.1.2.** Let  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ , then

- (1)  $A_r(x)$  is jointly continuous for  $(x, r) \in \mathbb{R}^d \times (0, \infty)$ .
- (2)  $Hf(x)$  is [Borel measurable](#).

**Proof.** We outline the proof.

- (1) Let  $(x, r) \rightarrow (x^*, r^*) \Rightarrow A_r(x) \rightarrow A_{r^*}(x^*)$ . Let  $(x_n, r_n)$  be any sequence which converges to  $x^*, r^*$ , then we consider  $\lim_{n \rightarrow \infty} A_{r_n}(x_n)$  and we can calculate

$$\int \underbrace{|f(y)| \mathbb{1}_{B(x_n, r_n)}(y)}_{:= h_n(y)},$$

then we apply [Theorem 2.3.1](#) to  $h_n$ .

- (2) Observe that

$$(Hf)^{-1}(\underbrace{(a, \infty)}_{\text{open}}) = \bigcup_{r>0} A_r^{-1}((a, \infty))$$

is open, since  $A_r^{-1}((a, \infty))$  is open from the 1. Note that the equality comes from the fact that  $Hf = \sup_r A_r$ .

25 Feb. 11:00

**Theorem 4.1.1** (Hardy-Littlewood maximal inequality). There exists  $C_d > 0$  such that for every  $f \in L^1(\mathbb{R}^d)$ ,

$$\forall_{\alpha > 0} m(\{x \in \mathbb{R}^d \mid Hf(x) > \alpha\}) \leq \frac{C_d}{\alpha} \int |f(x)| \, dx.$$

**Proof.** We first fix  $f \in L^1$  and  $\alpha > 0$ . We define

$$E := \{x \mid Hf(x) > \alpha\},$$

which is a Borel measurable set by Lemma 4.1.2. Then

$$x \in E \Rightarrow \exists_{r_x > 0} A_{r_x}(x) > \alpha \Rightarrow m(B(x, r_x)) < \frac{1}{\alpha} \int_{B(x, r_x)} |f(y)| \, dy.$$

From inner regularity, we have

$$m(E) = \sup \{m(K) \mid \text{compact } K \subset E\}.$$

Let  $K \subset E$  be compact, then

$$K \subset \bigcup_{x \in K} B(x, r_x) \stackrel{K \text{ compact}}{\Rightarrow} K \subset \bigcup_{i=1}^N B_i \stackrel{!}{\Rightarrow} K \subset \bigcup_{i=1}^m \{3B'_j\}.$$

From here, we further have

$$m(K) \leq \sum_{i=1}^m m(3B'_j) = 3^d \sum_{j=1}^m m(B'_j) \leq \frac{3^d}{\alpha} \sum_{j=1}^m \int_{B'_j} |f(y)| \, dy.$$

Now, since  $B'_1, \dots, B'_m$  are disjoint, hence we finally have

$$m(K) \leq \frac{3^d}{\alpha} \int_{\mathbb{R}^d} |f(y)| \, dy.$$

## Lecture 22: Lebesgue Differentiation Theorem

We should compare the Hardy-Littlewood maximal inequality to Markov's inequality. Namely, there exists  $C_d > 0$  (can take  $3^d$ ) such that for all  $f \in L^1(\mathbb{R}^d)$ ,  $\alpha > 0$ , we have

$$\begin{cases} m(\{x \mid Hf(x) > \alpha\}) \leq \frac{C_d}{\alpha} \int |f|; \\ m(\{x \mid |f(x)| > \alpha\}) \leq \frac{1}{\alpha} \int |f|. \end{cases}$$

## 4.2 Lebesgue Differentiation Theorem

We start with a theorem!

**Theorem 4.2.1** (Lebesgue Differentiation Theorem). Let  $f \in L^1$ , then

$$\lim_{r \rightarrow 0} \frac{1}{m(B(x, r))} \int_{B(x, r)} |f(y) - f(x)| \, dy = 0$$

for a.e.  $x$ .

**Proof.** The result holds for  $f \in C_c(\mathbb{R}^d)$ , namely for those continuous functions with **compact support**. This is because for any  $\epsilon > 0$ , if  $r$  is small and  $|f(y) - f(x)| < \epsilon$ , then

$$\frac{1}{m(B(x, r))} \int_{B(x, r)} |f(y) - f(x)| \, dy < \epsilon.$$

Now, let  $f \in L^1(\mathbb{R}^d)$  and fix  $\epsilon > 0$ . By density, there exists  $g \in C_c(\mathbb{R}^d)$  with  $\|f - g\|_1 < \epsilon$ . We then have

$$\int_{B_r(x)} |f(y) - f(x)| \, dy \leq \int_{B_r(x)} |f(y) - g(y)| \, dy + \int_{B_r(x)} |g(y) - g(x)| \, dy + \int_{B_r(x)} |g(x) - f(x)| \, dy.$$

**Note.** We use  $B_r(x)$  above to denote  $B(x, r)$  for spacing reason only. Nothing tricky here.

Divide all of these by  $m(B(x, r))$ , and take  $\limsup_{r \rightarrow \infty}$ , we need to understand the error terms, namely

$$\frac{1}{m(B(x, r))} \int_{B(x, r)} |f(x) - g(x)| \, dy = |g(x) - f(x)|$$

and

$$\frac{1}{m(B(x, r))} \int_{B(x, r)} |f(y) - g(y)| \, dy \leq (H(f - g))(x).$$

We define

$$Q(x) := \limsup_{r \rightarrow \infty} \frac{1}{m(B(x, r))} \int_{B(x, r)} |f(y) - f(x)| \, dy.$$

We want to show  $m(\{x \in X \mid Q(x) > 0\}) = 0$ . Let  $E_\alpha = \{x \in X \mid Q(x) > \alpha\}$ . It is enough to show  $m(E_\alpha) = 0$  for all  $\alpha > 0$  because  $\{x \in X \mid Q(x) > 0\} = \bigcup_n E_{\frac{1}{n}}$ . We know by the above that

$$Q(x) \leq (H(f - g))(x) + 0 + |g(x) - f(x)|.$$

Therefore,

$$E_\alpha \subset \{x \in X \mid (H(f - g))(x) > \alpha/2\} \cup \{x \in X \mid |g(x) - f(x)| > \alpha/2\}.$$

By the [Hardy-Littlewood maximal inequality](#) and [Markov's inequality](#), we have

$$\begin{cases} m(\{x \mid (H(f - g))(x) > \alpha/2\}) \leq \frac{2C_d}{\alpha} \int |f - g|; \\ m(\{x \mid |g(x) - f(x)| > \alpha/2\}) \leq \frac{2}{\alpha} \int |f - g|. \end{cases}$$

Thus,

$$0 \leq m(E_\alpha) \leq \frac{2C_d}{\alpha} \|f - g\|_1 + \frac{2}{\alpha} \|f - g\|_1 \leq \frac{2(C_d + 1)}{\alpha} \epsilon.$$

Taking  $\epsilon \rightarrow 0$ ,  $m(E_\alpha)$  does not depend on  $\epsilon$  and  $g$ , hence  $m(E_\alpha) = 0$ . ■

**Corollary 4.2.1.** [Theorem 4.2.1](#) also holds for  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ .

**Proof.** Using the fact that  $m^d$  is  $\sigma$ -finite, and apply [Theorem 4.2.1](#). Specifically, partition  $\mathbb{R}^d$  into countably many compact sets  $K_i$  and apply [Theorem 4.2.1](#) to  $f \mathbb{1}_{K_i}$  for all  $i$ . ■

**Corollary 4.2.2.** For  $f \in L^1_{\text{loc}}$ , we have

$$\lim_{r \rightarrow 0} \frac{1}{m(B(x, r))} \int_{B(x, r)} f(y) \, dy = f(x)$$

for a.e.  $x$ .

**Proof.** Use that

$$f(x) = \frac{1}{m(B(x, r))} \int_{B(x, r)} f(y) \, dy$$

and the triangle inequality. ■

DIY

**Definition 4.2.1** (Lebesgue point). Let  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ , the point  $x \in \mathbb{R}^d$  is called a *Lebesgue point* of  $f$  if

$$\lim_{r \rightarrow 0} \frac{1}{m(B(x, r))} \int_{B(x, r)} |f(y) - f(x)| \, dy = 0.$$

**Remark.** Corollary 4.2.1 tells us that almost all points in  $\mathbb{R}^d$  are Lebesgue points for  $f$ .

**Definition 4.2.2** (Shrink nicely). We say that  $\{E_r\}_{r>0}$  *shrinks nicely* to  $x$  as  $r \rightarrow 0$  if  $E_r \subset B(x, r)$  and

$$\exists_{c>0} c \cdot m(B(x, r)) \leq m(E_r).$$

**Corollary 4.2.3.** Suppose  $E_r$  shrink nicely to 0, and  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ , and  $x$  is a Lebesgue point. Then

$$\begin{cases} \lim_{r \rightarrow 0} \frac{1}{m(E_r)} \int_{E_r+x} |f(y) - f(x)| \, dy = 0; \\ \lim_{r \rightarrow 0} \frac{1}{m(E_r)} \int_{E_r+x} |f(y)| \, dy = f(x). \end{cases}$$

**Corollary 4.2.4.** If  $f \in L^1_{\text{loc}}(\mathbb{R})$ , then  $F(x) = \int_0^x f(y) \, dy$  is differentiable and  $F'(x) = f(x)$  almost everywhere.

# Chapter 5

## Normed Vector Space

### Lecture 23: Metric, normed and $L^p$ Spaces

#### 5.1 Metric Spaces and Normed Spaces

09 Mar. 11:00

We have seen the definition of a [norm](#) before, now we formally introduce the concept of *metric*.

**Definition 5.1.1 (Metric).** Let  $Y$  be a set, a function  $\rho: Y \times Y \rightarrow [0, \infty)$  is a *metric* on  $Y$  if

- $\rho(x, y) = \rho(y, x)$  for all  $x, y \in Y$ .
- $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$  for all  $x, y, z \in Y$ .
- $\rho(x, y) = 0$  if and only if  $x = y$ .

**Note.** The following make sense in a [metric](#) space.

- (1) Open/closed balls.
- (2) Open/closed sets.
- (3) Convergence sequences ( $x_n \rightarrow x$  with respect to  $\rho$  if and only if  $\lim_{n \rightarrow \infty} \rho(x_n, x) = 0$ ).
- (4) Continuous functions.

**Example.** We have the following [metric](#) spaces.

- (1)  $\mathbb{Q}$  with  $\rho(x, y) = |x - y|$ .
- (2)  $\mathbb{R}$  with  $\rho(x, y) = |x - y|$ .
- (3)  $\mathbb{R}_+$  with  $\rho(x, y) = |\ln(y/x)|$ .
- (4)  $\mathbb{R}^d$  with

$$\rho_p(x, y) = \left( \sum_{i=1}^d |x_i - y_i|^p \right)^{1/p}$$

and

$$\rho_\infty(x, y) = \max_{1 \leq i \leq d} |x_i - y_i|.$$

These all give the same open sets, hence they are topologically equivalent.

- (5)  $C([0, 1])$  with

$$\rho_p(f, g) = \left( \int_0^1 |f - g|^p \right)^{1/p}$$

and

$$\rho_\infty(f, g) = \max_{x \in [0, 1]} |f(x) - g(x)|.$$

- (6) Let  $(X, \mathcal{A}, \mu)$  be a **measure space** with  $\mu(X) < \infty$ . Let  $Y$  be the set of **measurable functions** on  $X$ , then

$$\rho(f, g) = \int \min\{|f(x) - g(x)|, 1\} \, d\mu(x)$$

is a **metric** and  $f_n \rightarrow f$  in  $\rho$  if and only if  $f_n \rightarrow f$  in **measure**.

Let  $V$  be a vector space over scalar field  $K = \mathbb{R}$  or  $K = \mathbb{C}$ .

**As previously seen** (Metric induced by a norm). Recall the definition of **seminorm** and **norm**. We see that a **norm** induces a metric

$$\rho(v, w) := \|v - w\|,$$

and we have

$$v_n \rightarrow v \Leftrightarrow \lim_{n \rightarrow \infty} \|v_n - v\| = 0.$$

**Example.** We first see some common examples of **normed** vector space.

- (1)  $L^1(X, \mathcal{A}, \mu)$  with  $\|f\|_1 := \int |f| \, d\mu$ .
- (2)  $C([0, 1])$  with  $\|f\|_1 := \int_0^1 |f(x)| \, dx$ ,  $\|f\|_\infty := \max_{0 \leq x \leq 1} |f(x)|$ .
- (3) For  $\mathbb{R}^d$  and  $0 < p < \infty$ , we have

$$\|x\|_p := \left( \sum_{i=1}^d |x_i|^p \right)^{1/p}, \quad \|x\|_\infty := \max_{1 \leq i \leq d} |x_i|.$$

## 5.2 $L^p$ Space

It turns out that we can generalize  $L^1$  into  $L^p$ .

**Definition 5.2.1** ( $L^p$  space). Given a **measure space**  $(X, \mathcal{A}, \mu)$  and a **measurable function**  $f$  and  $p$  such that  $0 < p < \infty$ , we define a **seminorm**  $\|\cdot\|_p$  such that

$$\|f\|_p := \left( \int_X |f|^p \, d\mu \right)^{1/p},$$

which induces the so-called  $L^p$  space  $L^p(X, \mathcal{A}, \mu)$ , where

$$L^p(X, \mathcal{A}, \mu) := \left\{ f \mid \|f\|_p < \infty \right\}.$$

**Remark.** Note that  $\|\cdot\|_p$  is only a **seminorm**. But if we identify functions which are equal **almost everywhere**, then it's indeed a **norm**.

**Example.**  $(\mathbb{R}, \mathcal{L}, m)$  has  $f(x) = x^{-\alpha} \mathbb{1}_{(1, \infty)}(x) \in L^p$  if and only if  $\alpha p > 1$ . In contrast,  $g(x) = x^{-\beta} \mathbb{1}_{(0, 1)}(x) \in L^p$  if and only if  $\beta p < 1$ .

Similar to **Definition 5.2.1**, we have the following.

**Definition 5.2.2** ( $\ell^p$  space). If  $(X, \mathcal{P}(X), \nu)$  is equipped with the **counting measure**, then we say it's

an  $\ell^p$  space such that

$$\ell^p(X) := L^p(X, \mathcal{P}(X), \nu).$$

**Remark.** We are interested in  $\ell^p(\mathbb{N})$  in particular. We have

$$\ell^p := \ell^p(\mathbb{N}) = \left\{ a = (a_1, a_2, \dots) \mid \|a\|_p = \left( \sum_{i=1}^{\infty} |a_i|^p \right)^{1/p} < \infty \right\}.$$

**Lemma 5.2.1.**  $L^p(X, \mathcal{A}, \nu)$  is a vector space for all  $p \in (0, \infty)$ .

**Proof.** We verify the following.

- $c \cdot f \in L^p(X, \mathcal{A}, \mu)$  for  $c \in \mathbb{R}$ . Indeed, since

$$\|cf\|_p = \left( \int |cf|^p d\mu \right)^{1/p} = |c| \|f\|_p < \infty \Leftrightarrow \|f\|_p < \infty,$$

which implies  $c \cdot f \in L^p(X, \mathcal{A}, \mu)$ .

- $f + g \in L^p(X, \mathcal{A}, \mu)$ . Indeed, since for any real numbers  $\alpha, \beta$ , we have

$$(\alpha + \beta)^p \leq (2 \cdot \max\{|\alpha|, |\beta|\})^p = 2^p \cdot \max\{|\alpha|^p, |\beta|^p\} \leq 2^p (|\alpha|^p + |\beta|^p),$$

which implies that for  $f, g \in L^p(X, \mathcal{A}, \mu)$ , we have

$$\|f + g\|_p < \infty \Leftrightarrow \|f + g\|_p^p = \int |f + g|^p d\mu \leq 2^p \int (|f|^p + |g|^p) < \infty.$$

This further implies

$$\|f + g\|_p < \infty \Leftrightarrow \|f\|_p, \|g\|_p < \infty,$$

which is what we want. ■

We see that in the above derivation, it doesn't give us the triangle inequality, namely

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p,$$

hence we need some new results.

**Theorem 5.2.1 (Hölder's inequality).** Let  $1 < p < \infty$ , and let  $q := p/(p-1)$  so that  $1/p + 1/q = 1$ . Then we have

$$\|f \cdot g\|_1 \leq \|f\|_p \|g\|_q.$$

**Proof.** We prove this in steps.

**Claim.** We have

$$t \leq \frac{t^p}{p} + 1 - \frac{1}{p} = \frac{t^p}{p} + \frac{1}{q}$$

for all  $t \geq 0$ .

**Proof.** By taking  $F(t) := t - t^p/p$  and  $t \geq 0$ , we see that the maximum of  $F$  implies the above inequality. ⊗

**Claim (Young's Inequality).** We have

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}$$

for  $\alpha, \beta > 0$ .<sup>a</sup>

<sup>a</sup>[https://en.wikipedia.org/wiki/Young's\\_inequality\\_for\\_products](https://en.wikipedia.org/wiki/Young's_inequality_for_products)

**Proof.** This follows by taking  $t := \alpha/\beta^{q-1}$  in the first inequality we obtained.  $\otimes$

Then, without loss of generality, we can assume that  $0 < \|f\|_p, \|g\|_q < \infty$ . Now, consider  $F(x) = f(x)/\|f\|_p$ ,  $G(x) = g(x)/\|g\|_q$ . We know that  $\|F\|_p = 1 = \|G\|_q$ . Then by Young's Inequality, we have

$$\int |F(x)G(x)| \, d\mu \leq \int \frac{|F(x)|^p}{p} + \frac{|G(x)|^q}{q} \Rightarrow \frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq \frac{1}{p} + \frac{1}{q} = 1,$$

which implies our desired result.  $\blacksquare$

**Example.** For  $p = q = 2$ ,  $X = \{1, \dots, d\}$  with  $\mu$  being the **counting measure**, then for any  $x, y \in \mathbb{R}^d$ , we have

$$\sum_{i=1}^d |x_i y_i| \leq \sqrt{\sum_{i=1}^d x_i^2} \sqrt{\sum_{i=1}^d y_i^2}$$

We now see how we can obtain the desired triangle inequality.

**Theorem 5.2.2 (Minkowski's Inequality).** Let  $1 \leq p < \infty$ , then for  $f, g \in L^p$ ,

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

**Proof.** For  $p = 1$ , it's easy since it's just triangle inequality. Now, we assume that  $1 < p < \infty$ , and we may assume also that  $\|f + g\| \neq 0$  without loss of generality. Then

$$\begin{aligned} \int |f(x) + g(x)|^p &\leq \int |f(x) + g(x)|^{p-1} (|f(x)| + |g(x)|) \\ &\leq \left( \int |f + g|^{(p-1)q} \right)^{1/q} \left[ \left( \int |f|^p \right)^{1/p} + \left( \int |g|^p \right)^{1/p} \right] \\ &\leq \left( \int |f + g|^p \right)^{1/q} (\|f\|_p + \|g\|_p). \end{aligned}$$

We then see that

$$\underbrace{(|f(x) + g(x)|^p)^{1-1/q}}_{(|f(x)+g(x)|^p)^{1/p}} \leq \|f\|_p + \|g\|_p,$$

which is just  $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ .  $\blacksquare$

## Lecture 24: Embedding $L^p$ Space

**Definition 5.2.3 (Essential supremum).** For a **measurable function**  $f$  on  $(X, \mathcal{A}, \mu)$ , we define

$$S := \{\alpha \geq 0 \mid \mu(\{x \mid |f(x)| > \alpha\}) = 0\} = \{\alpha \geq 0 \mid |f(x)| \leq \alpha \text{ a.e.}\}.$$

11 Mar. 11:00



Then, we say that the *essential supremum* of  $f$ , denoted as  $\|f\|_\infty$ , is defined as

$$\|f\|_\infty := \begin{cases} \inf S, & \text{if } S \neq \emptyset; \\ \infty, & \text{if } S = \emptyset. \end{cases}$$

**Definition 5.2.4** ( $L^\infty$  space). Let  $L^\infty(X, \mathcal{A}, \mu)$  be

$$L^\infty(X, \mathcal{A}, \mu) = \{f \mid \|f\|_\infty < \infty\}.$$

**Definition 5.2.5** ( $\ell^\infty$  space). We let  $\ell^\infty$  be defined as

$$\ell^\infty = L^\infty(\mathcal{N}, \mathcal{P}(\mathcal{N}), \nu),$$

where  $\nu$  is the [counting measure](#).

**Example.** Consider  $(\mathbb{R}, \mathcal{L}, m)$ . Then

$$\begin{aligned} f(x) &= \frac{1}{x} \mathbb{1}_{(0, \infty)}(x) \notin L^\infty; \\ g(x) &= x \mathbb{1}_{\mathbb{Q}}(x) + \frac{1}{1+x^2} \in L^\infty. \end{aligned}$$

If  $f$  is continuous on  $(\mathbb{R}, \mathcal{L}, m)$ , then  $\|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$ . For  $a \in \ell^\infty$ , we have  $\|a\|_\infty = \sup_{i \in \mathbb{N}} |a_i|$ , and sequences in  $\ell^\infty$  are exactly the bounded sequences.

**Lemma 5.2.2.** We have the following.

(1) Suppose  $f \in L^\infty(X, \mathcal{A}, \mu)$ . Then,

$$\begin{cases} \mu(\{x \mid |f(x)| > \alpha\}) = 0, & \text{if } \alpha \geq \|f\|_\infty; \\ \mu(\{x \mid |f(x)| > \alpha\}) > 0, & \text{if } \alpha < \|f\|_\infty. \end{cases}$$

(2)  $|f(x)| \leq \|f\|_\infty$  [almost everywhere](#).

(3)  $f \in L^\infty$  if and only if there exists a bounded [measurable function](#)  $g$  such that  $f = g$  [almost everywhere](#).

**Proof.** ■

DIY

**Theorem 5.2.3.** We have the following.

(1)  $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$ .

(2)  $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$ .

(3)  $f_n \rightarrow f$  in  $L^\infty$  if and only if  $f_n \rightarrow f$  [uniformly almost everywhere](#).

**Proof.** We'll do one implication in (3). Let  $A_n = \{x \mid |f_n(x) - f(x)| > \|f_n - f\|_\infty\}$ . Then  $\mu(A_n) = 0$ . Let  $A = \bigcup_n A_n$ , we see that  $\mu(A) = 0$  as well.

For  $x \in A^c$  and for every  $n$ , we have

$$|f_n(x) - f(x)| \leq \|f_n - f\|_\infty.$$

Given  $\epsilon > 0$ , there is an  $N$  so that  $\|f_n - f\|_\infty < \epsilon$  for all  $n \geq N$ . But then for all  $x \in A^c$ ,  $|f_n(x) - f(x)| < \epsilon$  as well. ■

DIY (1)  
and (2)

**Remark.** The motivation for 1. is that

$$\frac{1}{1} + \frac{1}{\infty} = 1,$$

and we want to have the similar result as in [Theorem 5.2.1](#).

**Proposition 5.2.1.** We have the following.

- (1) For  $1 \leq p < \infty$ , the collection of [simple functions](#) with finite measure [support](#) is dense in  $L^p(X, \mathcal{A}, \mu)$ .
- (2) For  $1 \leq p < \infty$ , the collection of [step functions](#) with finite measure [support](#) is dense in  $L^p(\mathbb{R}, \mathcal{L}, m)$ , so is  $C_c(\mathbb{R})$ .
- (3) For  $p = \infty$ , the collection of [simple functions](#) is dense in  $L^\infty(X, \mathcal{A}, \mu)$ .

**Proof.** ■

DIY

**Remark.** Note that  $C_c(\mathbb{R})$  is **not** dense in  $L^\infty(\mathbb{R}, \mathcal{L}, m)$ .

### 5.3 Embedding Properties of $L^p$ Spaces

**Definition 5.3.1** (Equivalent norm). Two [norms](#)  $\|\cdot\|, \|\cdot\|'$  on  $V$  are *equivalent* if there exists  $c_1, c_2 > 0$ , such that

$$c_1 \|v\| \leq \|v\|' \leq c_2 \|v\|$$

for all  $v \in V$ .

**Note.** We see that

- (1) These [norms](#) gives the same topological properties (open sets, closed sets, convergence, etc.).
- (2) [Definition 5.3.1](#) is an equivalence relation on [norms](#).

**Example.** For  $\mathbb{R}^d$  we have the [norms](#)  $\|\cdot\|_p$  for  $1 \leq p \leq \infty$ . All of these are equivalent. We see that for  $1 \leq p < \infty$ ,

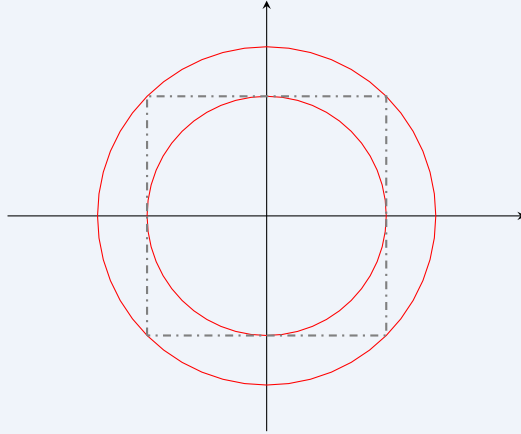
$$\|x\|_p = \left( \sum_{i=1}^d |x_i|^p \right)^{1/p} \leq (d \|x\|_\infty^p)^{1/p} = d^{1/p} \|x\|_\infty.$$

Also,

$$\|x\|_p = \left( \sum_{i=1}^d |x_i|^p \right)^{1/p} \geq (\|x\|_\infty^p)^{1/p} = \|x\|_\infty.$$

Thus,  $\|\cdot\|_p$  is equivalent to  $\|\cdot\|_\infty$  for every  $1 \leq p < \infty$ , and transitivity gives that they are all equivalent.

Another way of thinking of this, by assuming  $v \neq 0$ , and scaling by some  $t$ , we may assume  $v$  lies on the unit circle in one of the [norms](#). Then we are squeezing a unit circle in  $\|\cdot\|'$  between two circles of radius  $c_1, c_2$  in  $\|\cdot\|$ . In picture, we have to show that  $\|\cdot\|_2$  and  $\|\cdot\|_\infty$  are equivalent, we have



since the circles in  $\|\cdot\|_\infty$  are squares.

**Example.** For  $1 \leq p, q \leq \infty$ , we have  $L^p(\mathbb{R}, m)$ -norm and  $L^q(\mathbb{R}, m)$ -norm are not equivalent, even worse, we have that

$$L^p(\mathbb{R}, m) \not\subseteq L^1(\mathbb{R}, m), \quad L^p(\mathbb{R}, m) \not\supseteq L^1(\mathbb{R}, m).$$

## Lecture 25: Banach Spaces

**Proposition 5.3.1.** Suppose  $\mu(X) < \infty$ , then for every  $0 < p < q \leq \infty$ ,  $L^q \subseteq L^p$ .

14 Mar. 11:00

**Proof.** Suppose  $q < \infty$ , then

$$\int |f|^p \leq \left( \int (|f|^p)^{q/p} \right)^{p/q} \left( \int 1^{q/(q-p)} \right)^{1-p/q} = \left( \int |f|^q \right)^{p/q} \mu(X)^{1-p/q} < \infty$$

where we split  $\int |f|^p$  into  $\int |f|^p \cdot 1$ . From Hölder's inequality with  $q/p > 1$ , we have

$$\|f\|_p \leq \|f\|_q \mu(X)^{1/p-1/q} < \infty.$$

The case that  $q = \infty$  is left as an exercise. ■

DIY

**Proposition 5.3.2.** If  $0 < p < q \leq \infty$ , then  $\ell^p \subseteq \ell^q$ .

**Proof.** We consider two cases.

- When  $q = \infty$ , we have

$$\|a\|_\infty^p = \left( \sup_i |a_i| \right)^p = \sup_i |a_i|^p \leq \sum_{i=1}^{\infty} |a_i|^p.$$

Thus,  $\|a\|_\infty \leq \|a\|_p$ .

- When  $q < \infty$ , we see that

$$\sum_{i=1}^{\infty} |a_i|^q = \sum_{i=1}^{\infty} |a_i|^p \cdot |a_i|^{q-p} \leq \|a\|_\infty^{q-p} \sum_{i=1}^{\infty} |a_i|^p \leq \|a\|_\infty^{q-p} \cdot \|a\|_p^p = \|a\|_p^q.$$

Therefore,

$$\|a\|_q \leq \|a\|_p. \quad \blacksquare$$

**Proposition 5.3.3.** For all  $0 < p < q < r \leq \infty$ ,  $L^p \cap L^r \subseteq L^q$ .

**Proof.** ■

DIY

## 5.4 Banach Spaces

Let's start with a definition.

**Definition 5.4.1 (Cauchy sequence).** Let  $Y, \rho$  be a [metric](#) space. We call  $x_n$  a *Cauchy sequence* if for every  $\epsilon > 0$ , there exists an  $N \in \mathbb{N}$  such that for all  $n, m \geq N$ ,  $\rho(x_n, x_m) < \epsilon$ .

**Note.** Convergent sequence are [Cauchy](#).

**Definition 5.4.2 (Complete).** A [metric](#) space  $(Y, \rho)$  is called *complete* if every [Cauchy sequence](#) in  $Y$  converges.

**Example.** We first see some examples.

- (1) We see that  $\mathbb{Q}$  with  $\rho(x, y) = |x - y|$  is **not** [complete](#), but  $\mathbb{R}$  with the same [metric](#) is [complete](#).
- (2)  $C([0, 1])$  with  $\rho(f, g) = \|f - g\|_\infty$  is [complete](#), but with  $\rho(f, g) = \int |f - g|$  is not.

**Definition 5.4.3 (Banach space).** A *Banach space* is a [complete normed](#) vector space.

**Remark.** Namely, a vector space equipped with a [norm](#) whose [metric induced by the norm](#) is [complete](#).

**Theorem 5.4.1.** Let  $(V, \|\cdot\|)$  be a [normed](#) space. Then,

$V$  is [complete](#)  $\Leftrightarrow$  every absolutely convergent series is convergent.

i.e., if  $\sum_{i=1}^{\infty} \|v_i\| < \infty$ , then  $\left\{ \sum_{i=1}^N v_i \right\}_{N \in \mathbb{N}}$  converges to some  $s \in V$ .

Before we prove [Theorem 5.4.1](#), we first see one of the result based on this theorem.<sup>1</sup>

**Theorem 5.4.2 (Riesz-Fischer theorem).** For every  $1 \leq p \leq \infty$ , we have  $L^p(X, \mathcal{A}, \mu)$  is [complete](#), hence a [Banach space](#).

**Proof.** We prove this in two cases.

- We first prove this for  $1 \leq p < \infty$ . Suppose  $f_n \in L^p$  and  $\sum_{n=1}^{\infty} \|f_n\|_p < \infty$ .

We need to show that there is an  $F \in L^p$  such that  $\left\| \sum_{n=1}^N f_n - F \right\|_p \rightarrow 0$  as  $N \rightarrow \infty$ . i.e., we need to show the following.

1.  $\sum_{n=1}^{\infty} f_n(x)$  is convergent [a.e.](#) In fact, we can show this by showing the following.

<sup>1</sup>The proof can be found in [here](#).

**Claim.** We have

$$\int \sum_{n=1}^{\infty} |f_n(x)| < \infty.$$

**Proof.** Let  $G(x) = \sum_{n=1}^{\infty} |f_n(x)| = \sup_N \sum_{n=1}^N |f_n(x)|$ ,  $G: X \rightarrow [0, \infty]$ . Also, let  $G_N(x) = \sum_{n=1}^N |f_n(x)|$ . Then, we have

$$0 \leq G_1 \leq G_2 \leq \dots \leq G,$$

and  $G_N \rightarrow G$ . Furthermore,

$$0 \leq G_1^p \leq G_2^p \leq \dots \leq G^p,$$

and  $G_N^p \rightarrow G^p$ . From [monotone convergence theorem](#),

$$\int G^p = \lim_{N \rightarrow \infty} \int G_N^p.$$

From [Minkowski inequality](#), we further have

$$\|G_N\|_p \leq \sum_{n=1}^N \|f_n\|_p \leq \sum_{n=1}^{\infty} \|f_n\|_p := B < \infty.$$

Thus,

$$\int G(x)^p = \lim_{N \rightarrow \infty} \int G_N^p = \lim_{N \rightarrow \infty} \|G_N\|_p^p \leq B^p < \infty.$$

We see that  $G$  is finite [a.e.](#) as desired. This implies that  $\sum_{n=1}^{\infty} |f_n(x)| < \infty$  [a.e.](#), so  $\sum_{n=1}^{\infty} f_n(x)$  converges [a.e.](#) Now, we simply let

$$F(x) = \begin{cases} \sum_{n=1}^{\infty} f_n(x), & \text{if it converges;} \\ 0, & \text{otherwise.} \end{cases}$$

⊗

2.  $F \in L^p$ , where  $F(x) := \sum_{n=1}^{\infty} f_n(x)$  [a.e.](#) and say is zero elsewhere.

**Claim.**  $F$  defined in this way is indeed in  $L^p$ .

**Proof.** This is clear since

$$|F(x)| \leq G(x) \Rightarrow \int |F|^p \leq \int G^p < \infty,$$

hence  $F \in L^p$ .

⊗

3. We then show the last condition we need to check.

**Claim.**  $\left\| \sum_{n=1}^N f_n - F \right\|_p \rightarrow 0$  as  $N \rightarrow \infty$ .

**Proof.** We now see that

$$\left| \sum_{n=1}^N f_n(x) - F(x) \right|^p \leq \left( \sum_{n=1}^{\infty} |f_n(x)| + |F(x)| \right)^p \leq (2G(x))^p.$$

Since  $2G \in L^p$ , so  $2G^p \in L^1$ . Thus, by [dominated convergence theorem](#), we have

$$\lim_{N \rightarrow \infty} \int \left| \sum_{n=1}^N f_n(x) - F(x) \right|^p dx = 0.$$

This implies

$$\left\| \sum_{n=1}^N f_n - F \right\|_p \rightarrow 0$$

as  $N \rightarrow \infty$ . ⊗

- Now assume  $1 \leq p \leq \infty$ . DIY

■

## Lecture 26: Bounded Linear Transformations

We now prove [Theorem 5.4.1](#), completing the proof of [Theorem 5.4.2](#) since the latter relies on this result. 16 Mar. 11:00

**Proof of Theorem 5.4.1.** We prove it by proving two directions.

( $\Rightarrow$ ) Suppose  $V$  is [complete](#), and fix an absolutely convergent series  $\sum_n v_n$ . Define  $s_N = \sum_{n=1}^N v_n$ . It suffices to show the partial sums are a [Cauchy Sequence](#).

Fix  $\epsilon > 0$ , then because  $\sum_{n=1}^{\infty} \|v_n\| < \infty$ , there is a  $K \in \mathbb{N}$  so that

$$\sum_{n=K}^{\infty} \|v_n\| < \epsilon.$$

Now let  $M > N > K$ , we see that

$$\|s_M - s_N\| = \left\| \sum_{n=N+1}^M v_n \right\| \leq \sum_{n=N+1}^M \|v_n\| \leq \sum_{n=N}^{\infty} \|v_n\| < \epsilon,$$

so this is [Cauchy](#).

( $\Leftarrow$ ) Now suppose  $v_n, n \in \mathbb{N}$  is a [Cauchy sequence](#). For all  $j \in \mathbb{N}$ , there exists an  $N_j \in \mathbb{N}$  such that

$$\|v_n - v_m\| < \frac{1}{2^j}$$

for all  $n, m \geq N_j$ . Without loss of generality, we may assume  $N_1 < N_2 < \dots$

Let  $w_1 = v_{N_1}$ ,  $w_j = v_{N_j} - v_{N_{j-1}}$  for  $j \geq 2$ . Therefore,

$$\sum_{j=1}^{\infty} \|w_j\| \leq \|v_{N_1}\| + \sum_{j=2}^{\infty} \frac{1}{2^{j-1}} < \infty.$$

Thus,  $\sum_{j=1}^k w_j \rightarrow s \in V$  as  $k \rightarrow \infty$ . But by telescoping, we have

$$v_{N_k} = \sum_{j=1}^k w_j \rightarrow s.$$

Now we claim that since  $v_n$  is **Cauchy**, so  $v_n \rightarrow s$ .

Explicitly, take  $\epsilon > 0$ , and let  $k$  be large enough so that  $\|v_{N_k} - s\| < \epsilon$  and  $1/2^k < \epsilon$ . Then if  $n > N_k$  then

$$\|v_n - s\| \leq \|v_n - v_{N_k}\| + \|v_{N_k} - s\| < \epsilon + \epsilon = 2\epsilon.$$

Thus,  $v_n \rightarrow s$ . ■

## 5.5 Bounded Linear Transformations

**Definition 5.5.1** (Bounded linear transformation). Given two **normed** vector spaces  $(V, \|\cdot\|)$ ,  $(W, \|\cdot\|')$ , a linear map  $T: V \rightarrow W$  is called a *bounded map* if there exists  $c \geq 0$  such that

$$\|Tv\|' \leq c\|v\|$$

for all  $v \in V$ .

**Proposition 5.5.1.** Suppose  $T: (V, \|\cdot\|) \rightarrow (W, \|\cdot\|')$  is a linear map. Then the following are equivalent.

- (1)  $T$  is continuous.
- (2)  $T$  is continuous at 0.
- (3)  $T$  is a **bounded map**.

**Proof.** (1)  $\Rightarrow$  (2) is clear.

**Claim.** (2)  $\Rightarrow$  (3).

**Proof.** Take  $\epsilon = 1$ , then there exists a  $\delta > 0$  such that  $\|Tu\|' < 1$  if  $\|u\| < \delta$ .

Now take an arbitrary  $\|v\| \in V$ ,  $v \neq 0$ . Let  $u = \frac{\delta}{2\|v\|}v$ . Then  $\|u\| < \delta$ . Therefore,

$$\|Tu\|' < 1 \Rightarrow \frac{\delta}{2\|v\|} \|Tv\|' < 1 \Rightarrow \|Tv\|' < \frac{2}{\delta} \|v\|.$$

Then  $2/\delta$  is our constant. ⊗

**Claim.** (3)  $\Rightarrow$  (1).

**Proof.** Fix  $v_0 \in V$ . Then for some constant  $c$

$$\|Tv - Tv_0\|' = \|T(v - v_0)\|' \leq c\|v - v_0\|.$$

Thus,  $T$  is continuous, as when  $v \rightarrow v_0$  the right-hand side goes to zero, and so  $Tv \rightarrow Tv_0$ . ⊗

**Example.** Let's see some examples.

- (1) We can look at

$$\begin{aligned} T: \ell^1 &\rightarrow \ell^1 \\ (a_1, a_2, \dots) &\mapsto (a_2, a_3, \dots). \end{aligned}$$

Then clearly  $\|Ta\|_1 \leq \|a\|_1$ , so  $T$  is a **bounded linear transformation**.

- (2) We can also look at  $S: (C([-1, 1]), \|\cdot\|_1) \rightarrow \mathbb{C}$ , where  $Sf = f(0)$ .  $S$  is not a **bounded linear**

transformation, because we can make

$$\begin{cases} \|Sf\| &= |f(0)| = n \\ \|f\|_1 &= 1 \end{cases}$$

for every  $n \in \mathbb{N}$  (take  $f$ 's graph to be a skinny triangle shooting up to  $n$  at 0).

(3) But  $U: (C([-1, 1]), \|\cdot\|_\infty) \rightarrow \mathbb{C}$  defined by  $Uf = f(0)$  is a **bounded linear transformation**, because  $|f(0)| \leq \|f\|_\infty$ .

(4) Let  $A$  be an  $n \times m$  matrix. Then  $T: \mathbb{R}^m \rightarrow \mathbb{R}^n$  defined by  $v \mapsto Av$  is a **bounded linear transformation**.

Explicitly this is

$$(Tv)_i = (Av)_i = \sum_{j=1}^m A_{ij}v_j.$$

(5) Let  $K(x, y)$  be a continuous function on  $[0, 1] \times [0, 1]$ . We'll define

$$T: (C[0, 1], \|\cdot\|_\infty) \rightarrow (C[0, 1], \|\cdot\|_\infty)$$

by

$$(Tf)(x) = \int_0^1 K(x, y)f(y) dy.$$

This is an analogue of matrix multiplication ( $K$  is like a continuous matrix). This is a **bounded linear transformation**.

(6) Let us look at  $T: L^1(\mathbb{R}) \rightarrow (C(\mathbb{R}), \|\cdot\|_\infty)$  defined by

$$(Tf)(t) = \int_{-\infty}^{\infty} e^{-itx} f(x) dx$$

that is the Fourier transform of  $f$ .

(7)  $T: (C^\infty[0, 1], \|\cdot\|_\infty) \rightarrow (C^\infty[0, 1], \|\cdot\|_\infty)$ . Define

$$(Tf)(x) = f'(x).$$

This is not a **bounded linear transformation**. In contrast,  $S$ , defined on the same spaces

$$(Sf)(x) = \int_0^x f(t) dt$$

is bounded.

**Definition 5.5.2 (Operator norm).** Let  $L(V, W)$  be defined as a vector space such that

$$L(V, W) := \{T: V \rightarrow W \mid T \text{ is a bounded linear transformation}\}.$$

Then for  $T \in L(V, W)$ , the *operator norm* of  $T$  is

$$\begin{aligned} \|T\| &:= \inf\{c \geq 0 \mid \|Tv\|'' \leq c\|v\|' \text{ for all } v \in V\} \\ &= \sup\left\{\frac{\|Tv\|''}{\|v\|'} \mid v \neq 0, v \in V\right\} \\ &= \sup\{\|Tv\|'' \mid \|v\|' = 1, v \in V\}. \end{aligned}$$

**Lemma 5.5.1.** We have that



- (1) The [three definitions](#) of  $\|T\|$  above are all equal.
- (2)  $(L(V, W), \|\cdot\|)$  is indeed a [normed](#) space.

**Proof.** ■

DIY

## Lecture 27: Dual Space

As previously seen. From [Definition 5.5.2](#), we have that

18 Mar. 11:00

$$\|Tv\|'' \leq \|T\| \|v\|'.$$

**Remark.** Notice that this [Definition 5.5.2](#) is only for [bounded linear transformation](#).

**Theorem 5.5.1.** If  $W$  is [complete](#), then  $L(V, W)$  is [complete](#).

**Proof.** Suppose  $T_n$  is a [Cauchy sequence](#) in  $L(V, W)$ . Fix  $v \in V$  and let  $w_n := T_n v \in W$ , we then have

$$\|w_n - w_m\| = \|T_n v - T_m v\| = \|(T_n - T_m)v\| \leq \underbrace{\|T_n - T_m\|}_{\rightarrow 0} \underbrace{\|v\|}_{\text{fixed value}}.$$

Thus,  $w_n$  is [Cauchy](#), so it converges since  $W$  is [complete](#). We call its unique limit  $Tv$ . This makes  $T: V \rightarrow W$  into a function. We must show it is a [bounded linear transformation](#) and that  $\|T_n - T\| \rightarrow 0$ . ■

DIY

## 5.6 Dual of $L^p$ Spaces

**Example.** Let  $w \in \mathbb{R}^d$ , and denote the inner product between  $w$  and  $v \in \mathbb{R}^d$  by

$$v \cdot w := \langle v, w \rangle.$$

Then we can consider

$$\max\{v \cdot w \mid \|v\|_2 = 1\} = \|w\|_2.$$

If  $w \in \mathbb{C}^d$ , this is similar since

$$\max\{|v \cdot w| \mid \|v\|_2 = 1\} = \|w\|_2.$$

These maximums are achieved by  $v = \frac{\bar{w}}{\|w\|}$  if  $w \neq 0$ .

**Proposition 5.6.1.** Let  $1/p + 1/q = 1$  with  $1 \leq q < \infty$ . For every  $g \in L^q$ ,

$$\|g\|_q = \sup \left\{ \left| \int fg \right| \mid \|f\|_p = 1 \right\}.$$

Suppose  $\mu$  is [σ-finite](#), then the result also holds for  $q = \infty$ ,  $p = 1$ .

**Proof.** By [Hölder's inequality](#), we know that

$$\left| \int fg \right| \leq \int |fg| = \|fg\|_1 \leq \|f\|_p \|g\|_q = \|g\|_q.$$

Thus, the supremum is less or equal to  $\|g\|_q$ .

(1) Let

$$f(x) = \frac{|g(x)|^{q-1} \cdot \overline{\operatorname{sgn}(g(x))}}{\|g\|_q^{q-1}}$$

Then  $\int |f|^p = 1$ , and  $\int fg = \|g\|_q$ .

Check

**Note.** For  $\alpha \in \mathbb{C}$ ,  $\operatorname{sgn}(\alpha) := e^{i\theta}$  where  $\alpha = |\alpha| e^{i\theta}$ .

(2) The case that  $\mu$  is  $\sigma$ -finite and  $q = \infty, p = 1$  can be shown.

DIY

■

**Remark.** One could use the above to prove [Minkowski's inequality](#) (as it only uses [Hölder's inequality](#)).

**Definition 5.6.1** (Dual space). For a [normed](#) space  $(V, \|\cdot\|)$ , its *dual space* is  $V^* = L(V, \mathbb{R})$  or  $V^* = L(V, \mathbb{C})$ .

**Remark.** Namely, the [dual space](#) of  $V$  contains [bounded linear transformations](#) with codomain being the scalar field.

**Definition 5.6.2** (Linear functional). Given a [normed](#) space  $(V, \|\cdot\|)$ ,  $\ell \in V^*$  is called a *linear functional* on  $V$ . i.e.,

- $\ell: V \rightarrow \mathbb{R}$  (or  $\mathbb{C}$ ).
- $\ell$  is linear.
- There exists a  $c \geq 0$  such that  $|\ell(v)| = c\|v\|$ .

**Note.**  $V^*$  is always a [Banach space](#) (even if  $V$  is not [complete](#)).

**Corollary 5.6.1.** We have the following.

(1) Let  $1/p + 1/q = 1, 1 \leq q < \infty$ . For  $g \in L^q$  define  $\ell_g \in L^p \rightarrow \mathbb{C}$  by

$$\ell_g(f) = \int fg.$$

Then  $\ell_g \in (L^p)^*$ . Furthermore,  $\|\ell_g\| = \|g\|_q$ .

(2) If  $\mu$  is  $\sigma$ -finite, then this also holds for  $q = \infty, p = 1$ .

**Proof.**  $\ell_g$  is clearly linear in  $f$  because the integral is linear. Then [Proposition 5.6.1](#) gives in both (1) and (2) that

$$\|g\|_q = \sup\{|\ell_g(f)| \mid \|g\|_p = 1\} = \|\ell_g\|$$

and so  $\ell_g$  is a [bounded linear transformation](#) with the desired properties. ■

**Theorem 5.6.1.** We have the following.

(1) Let  $1/p + 1/q = 1, 1 \leq q < \infty$ . The map  $T: L^q \rightarrow (L^p)^*$  given by  $Tg = \ell_g$  is an isometric<sup>a</sup> linear isomorphism. This means that

- $T$  is a [bounded linear transformation](#).

- $T$  is bijective.
- $T$  is **norm**-preserving.

(2) If  $\mu$  is  **$\sigma$ -finite** then this also holds for  $q = \infty, p = 1$ .

<sup>a</sup>A map  $T$  is called isometric if for a given  $g$ ,  $\|Tg\| = \|g\|$ .

**Proof.** We have already proved this is isometric in [Corollary 5.6.1](#), it is clearly linear, and isometry implies injectivity.

We will prove that it is surjective later. ■

Fix!!!

**Note.** Even if  $\mu$  is  **$\sigma$ -finite** we might not have  $L^1 \cong (L^\infty)^*$ .

Also note that  $L^2 \cong (L^2)^*$ , and for all  $1 < p < \infty$  we have  $(L^p)^{**} \cong L^p$ .

# Chapter 6

## Signed and Complex Measures

### Lecture 28: Signed Measure

21 Mar. 11:00

**As previously seen.** Suppose  $f: X \rightarrow [0, \infty]$  is a measurable function on  $(X, \mathcal{A}, \mu)$ .

We can define  $\nu(E) = \int_E f d\mu$  for  $E \in \mathcal{A}$ , and  $\nu$  is a measure on  $(X, \mathcal{A})$ . This gives a map from the set of non-negative measurable functions on  $X$  to measures on  $X$ . This is injective if we identify functions which are equal almost everywhere. But it is not necessarily surjective. We can then think of measures as a generalization of functions.

For an example, think of a Dirac-Delta measure on  $\mathbb{R}$ . This is not the Lebesgue integral of any non-negative measurable function.

What if instead we took  $f: X \rightarrow \mathbb{R}, \overline{\mathbb{R}}$  or  $\mathbb{C}$ ? We could take the same construction to get  $\nu(E) = \int_E f d\mu$ , but this is no longer a measure as it can take  $\mathbb{R}, \overline{\mathbb{R}}$  or  $\mathbb{C}$  values.

### 6.1 Signed Measures

**Definition 6.1.1.** Let  $(X, \mathcal{A})$  be a measurable space. A signed measure is a function

$$\nu: \mathcal{A} \rightarrow [-\infty, \infty) \text{ or } \nu: \mathcal{A} \rightarrow (-\infty, \infty]$$

such that

- $\nu(\emptyset) = 0$ .
- If  $A_1, A_2, \dots \in \mathcal{A}$  are disjoint then

$$\nu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \nu(A_i)$$

where the series on the right-hand side converges absolutely if  $\nu(\bigcup_{i=1}^{\infty} A_i) \in (-\infty, \infty)$ .

**Remark.** This means the series does not depend on rearrangement if our function  $\nu$  takes finite value on the set  $\bigcup_i A_i$ .

**Example.** Consider

- (1)  $\nu$  is a positive measure (i.e., measure), then  $\nu$  is a signed measure.
- (2) If we have positive measures  $\mu_1, \mu_2$  such that either  $\mu_1(X) < \infty$  or  $\mu_2(X) < \infty$ , then  $\nu = \mu_1 - \mu_2$  is a signed measure.
- (3) If  $f: X \rightarrow \overline{\mathbb{R}}$  on a measure space  $(X, \mathcal{A}, \mu)$  such that  $\int_X f^+ d\mu < \infty$  or  $\int_X f^- d\mu < \infty$ , we can

define

$$\nu(E) = \int_E f \, d\mu$$

and this will be a [signed measure](#).

**Note.** The following weird things happen with [signed measures](#).

- (1)  $A \subseteq B$  does not imply  $\nu(A) \leq \nu(B)$ , as  $\nu(B) = \nu(A) + \nu(B \setminus A)$ , and  $\nu(B \setminus A)$  may be negative.
- (2) If  $A \subseteq B$  and  $\nu(A) = \infty$ , then  $\nu(B) = \infty$ , because  $\nu(B \setminus A) \in (-\infty, \infty]$ .
- (3) Similarly, if  $A \subseteq B$  and  $\nu(A) = -\infty$  then  $\nu(B) = -\infty$ .

**Lemma 6.1.1.** If  $\nu$  is a [signed measure](#) on  $(X, \mathcal{A})$ , then we have the following.

- (1) Continuity from below. If  $E_n \in \mathcal{A}$  and  $E_1 \subseteq E_2 \subseteq \dots$  then

$$\nu\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{N \rightarrow \infty} \nu(E_N).$$

- (2) Continuity from above. If  $E_n \in \mathcal{A}$ ,  $E_1 \supseteq E_2 \supseteq \dots$ , and  $-\infty < \nu(E_1) < \infty$  then

$$\nu\left(\bigcap_{n=1}^{\infty} E_n\right) = \lim_{N \rightarrow \infty} \nu(E_N).$$

**Proof.** Read [FF99]. ■

**Definition.** Let  $\nu$  be a [signed measure](#) on  $(X, \mathcal{A})$ . Let  $E \in \mathcal{A}$ , then we say that

**Definition 6.1.2** (Positive set for a signed measure).  $E$  is *positive* for  $\nu$  if for all  $F \subseteq E$ ,  $\nu(F) \geq 0$ .

**Definition 6.1.3** (Negative set for a signed measure).  $E$  is *negative* for  $\nu$  if for all  $F \subseteq E$ ,  $\nu(F) \leq 0$ .

**Definition 6.1.4** (Null set for a signed measure).  $E$  is *null* for  $\nu$  if for all  $F \subseteq E$ ,  $\nu(F) = 0$ .

**Note.** We see that

- (1) If  $E$  is a [positive set](#),  $F \subseteq E$ , then  $\nu(F) \leq \nu(E)$ .
- (2) If  $E$  is a [negative set](#),  $F \subseteq E$ , then  $\nu(F) \geq \nu(E)$ .

**Lemma 6.1.2.** Let  $\nu$  be a [signed measure](#) on  $(X, \mathcal{A})$ , then

- (1) If  $E$  is [positive](#),  $G \subseteq E$  is [measurable](#), then  $G$  is [positive](#).
- (2) If  $E$  is [negative](#),  $G \subseteq E$  is [measurable](#), then  $G$  is [negative](#).
- (3) If  $E$  is [null](#),  $G \subseteq E$  is [measurable](#), then  $G$  is [null](#).
- (4)  $E_1, E_2, \dots$  are [positive](#) sets, then  $\bigcup_{i=1}^{\infty} E_i$  is [positive](#).

**Proof.** The first three are trivial from their definition. For 4., if  $E_1, \dots$  are [positive sets](#), let

$F_n := E_n \setminus \bigcup_{j=1}^{n-1} E_j$ . Then  $F_n \subset E_n$ , so  $F_n$  is **positive sets** from 1., hence if  $E \subset \bigcup_{j=1}^{\infty} E_j$ , then

$$\nu(E) = \sum_{j=1}^{\infty} \nu(E \cap F_j) \geq 0$$

as desired. ■

**Lemma 6.1.3.** Suppose that  $\nu$  is a **signed measure** with  $\nu: \mathcal{A} \rightarrow [-\infty, \infty)$ . Suppose  $E \in \mathcal{A}$  and  $0 < \nu(E) < \infty$ , then there exists a **measurable**  $A \subseteq E$  such  $A$  is a **positive set** and  $\nu(A) > 0$ .

**Proof.** If  $E$  is **positive**, we're done. Otherwise, there exist **measurable** subsets with **negative** measure. Let  $n_1 \in \mathbb{N}$  be the least such  $n_1$  such that there exists  $E_1 \subseteq E$  with  $\nu(E_1) < -1/n_1$ .

If  $E \setminus E_1$  is **positive**, we're done. Else we can inductively define  $n_2, n_3, \dots$  as well as  $E_2, E_3, \dots$

Explicitly, if  $E \setminus \bigcup_{i=1}^{k-1} E_i$  is not **positive**, let  $n_k$  be the least such that there exists  $E_k \subseteq E \setminus \bigcup_{i=1}^{k-1} E_i$  with  $\nu(E_k) < -1/n_k$ .

Note then that if  $n_k \geq 2$ , for all  $B \subseteq E \setminus \bigcup_{i=1}^{k-1} E_i$  we have that  $\nu(B) \geq -\frac{1}{n_k-1}$ .

Now let  $A = E \setminus \bigcup_{i=1}^{\infty} E_i$ . Since  $E = A \cup (\bigcup_i E_i)$  we have by **countable additivity** that

$$0 < \nu(E) = \nu(A) + \sum_{k=1}^{\infty} \nu(E_k) < \nu(A).$$

Furthermore,  $\nu(E), \nu(A)$  are both in  $(0, \infty)$ , and we see that

$$0 < \nu(E) \leq \nu(A) - \sum_{k=1}^{\infty} \frac{1}{n_k}.$$

Therefore, the sum on the right-hand side must converge, meaning that  $1/n_k \rightarrow 0$  as  $k \rightarrow \infty$ . That is  $\lim_{k \rightarrow \infty} n_k = \infty$ .

Now if  $B \subseteq A$ , then  $B \subseteq E \setminus \bigcup_{i=1}^{\infty} E_i$ . Therefore,  $B \subseteq E \setminus \bigcup_{i=1}^{k-1} E_i$ . By the note above, for large enough  $k$  such that  $n_k \geq 2$  we have

$$\nu(B) \geq \frac{-1}{n_k - 1},$$

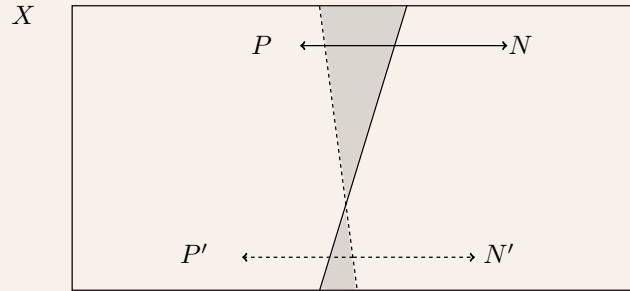
then taking  $k \rightarrow \infty$  we have  $\nu(B) \geq 0$ , and so  $A$  is a **positive set** as desired. ■

**Theorem 6.1.1 (Hahn decomposition theorem).** If  $\nu$  is a **signed measure** on  $(X, \mathcal{A})$ , then there exist  $P, N \in \mathcal{A}$  such that

$$P \cap N = \emptyset, \quad P \cup N = X,$$

where  $P$  is **positive for  $\nu$** ,  $N$  is **negative for  $\nu$** .

Furthermore, if  $P', N'$  are another such pair, then  $P \triangle P' (= N \triangle N')$  is **null** for  $\nu$ .



## Lecture 29: Hahn and Jordan Decomposition Theorem

We now prove [Theorem 6.1.1](#).

**Proof of Theorem 6.1.1.** We first show the uniqueness. We see that  $P \setminus P' \subseteq P, P \setminus P' \subseteq N'$ . Thus,  $P \setminus P' \subseteq P \cap N'$  is both [positive](#) and [negative](#), hence  $P \setminus P'$  is [null](#).

Similarly, for  $P' \setminus P$ , and then their union  $P \Delta P'$  is [null](#) as well.

To show the existence, without loss of generality suppose  $\nu: \mathcal{A} \rightarrow [-\infty, \infty)$ . If not, consider  $-\nu$ . Let

$$s := \sup\{\nu(E) \mid E \in \mathcal{A} \text{ is a } \text{positive set}\},$$

which is a nonempty supremum because  $\emptyset$  is [positive](#). Then there exist  $P_1, P_2, \dots$  [positive sets](#) such that  $\lim_{n \rightarrow \infty} \nu(P_n) = s$ .

Then we have that  $P = \bigcup_n P_n$  is [positive](#) by [Lemma 6.1.2](#). We then have  $\nu(P) \leq s$  and  $\nu(P) = \nu(P_n) + \nu(P \setminus P_n) \geq \nu(P_n)$ . Thus,

$$\nu(P) \geq \lim_{n \rightarrow \infty} \nu(P_n) = s.$$

Hence,  $\nu(P) = s$  and the supremum is in fact a max. We then know that  $s = \nu(P) < \infty$  because  $\nu$  does not attain the value infinity.

Now let  $N = X \setminus P$ . We claim that  $N$  is [negative](#). If not then there exists a [measurable](#)  $E \subseteq N$  with  $\nu(E) > 0$ . By assumption,  $\nu(E) < \infty$ . Then  $0 < \nu(E) < \infty$ , so by [Lemma 6.1.3](#) there exists a [measurable](#)  $A \subseteq E$  such that  $A$  is [positive](#) and  $\nu(A) > 0$ .

But we then know that

$$\nu(P \cup A) = \nu(P) + \nu(A) > \nu(P)$$

which is a contradiction since  $P \cup A$  is a [positive set](#), and  $\nu(P)$  is maximal. Therefore,  $N$  is [negative](#), and the theorem holds.

Finally, if  $P', N'$  is another pair of sets as in the statement of the theorem, we have

$$P \setminus P' \subseteq P, \quad P \setminus P' \subseteq N',$$

so that  $P \setminus P'$  is both positive and negative, hence null; likewise for  $P' \setminus P$ . ■

**Definition 6.1.5 (Singular).** If  $\mu, \nu$  are [signed measures](#) on  $(X, \mathcal{A})$ , then we say  $\mu$  and  $\nu$  are *singular to each other*, denoted as  $\mu \perp \nu$ , if there exists  $E, F \in \mathcal{A}$  such that  $E \cap F = \emptyset, E \cup F = X$ ,  $F$  is [null](#) for  $\mu$ ,  $E$  is [null](#) for  $\nu$ .

**Example.** Consider  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  with

- The [Lebesgue measure](#)  $m$ .
- The [Cantor measure](#)  $\mu_C$  induced by the [Cantor function](#).
- The [discrete measure](#)  $\mu_D = \delta_1 + 2\delta_{-1}$ .

We then see that (1)  $m \perp \mu_D$ . (2)  $m \perp \mu_C$ . (3)  $\mu_C \perp \mu_D$ .

**Proof.** We see them as follows.

- (1) Take  $E = \mathbb{R} \setminus \{-1, 1\}, F = \{-1, 1\}$  to see that  $m \perp \mu_D$ .
- (2) Take  $E = \mathbb{R} \setminus K$  and  $F = K$  where  $K$  is the [Cantor set](#) to see that  $m \perp \mu_C$ .
- (3) We can also see that  $\mu_C \perp \mu_D$ .

⊛

**Theorem 6.1.2 (Jordan decomposition theorem).** Let  $\nu$  be a [signed measure](#) on  $(X, \mathcal{A})$ . Then there

exists unique **positive measures**  $\nu^+, \nu^-$  on  $(X, \mathcal{A})$  such that for all  $E \in \mathcal{A}$  we have

$$\nu(E) = \nu^+(E) - \nu^-(E)$$

and  $\nu^+ \perp \nu^-$ .

**Proof.** For existence, we take  $\nu^+(E) := \nu(E \cap P), \nu^-(E) := -\nu(E \cap N)$  where  $P, N$  is the **Hahn decomposition** of  $X$ .

If there exists  $\mu^+, \mu^-$  such that  $\nu = \mu^+ + \mu^-$  and  $\mu^+ \perp \mu^-$ , let  $E, F \in \mathcal{A}$  be such that  $E \cap F = \emptyset$ ,  $E \cup F = X$ , and  $\mu^+(F) = \mu^-(E) = 0$ . Then we have that  $X = E \cup F$  is another **Hahn decomposition** for  $\nu$ , so  $P \triangle E$  is  $\nu$ -null. Therefore, for any  $A \in \mathcal{A}$ ,  $\mu^+(A) = \mu^+(A \cap E) = \nu(A \cap E) = \nu(A \cap P) = \nu^+(A)$ , hence  $\mu^+ = \nu^+$ . Likewise, we have  $\nu^- = \mu^-$ . ■

## Lecture 30: Absolutely Continuous Measures

**Example.** For an example of **Theorem 6.1.2**, let  $(X, \mathcal{A}, \mu)$  be a **measure space**,  $f: X \rightarrow \overline{\mathbb{R}}$ , and  $\nu(E) = \int_E f \, d\mu$ . Then

$$\nu^+(E) = \int_E f^+ \, d\mu, \quad \nu^-(E) = \int_E f^- \, d\mu.$$

**Definition.** Given a **signed measure**  $\nu$  on  $(X, \mathcal{A})$  and its **Jordan decomposition**  $\nu = \nu^+ - \nu^-$ .

**Definition 6.1.6** (Positive variation). We call  $\nu^+$  the *positive variation* of  $\nu$ .

**Definition 6.1.7** (Negative variation). We call  $\nu^-$  the *negative variation* of  $\nu$ .

**Definition 6.1.8** (Total variation). The *total variation measure* of  $\nu$ , denoted as  $|\nu|$ , is defined as  $|\nu| := \nu^+ + \nu^-$ .

**Remark.** There is always a **positive measure** on  $X$ .

**Proof.** Consider the **total variation**  $|\nu|$  for an arbitrary **signed measure**  $\nu$ . ⊗

**Example.** In the above **example**,  $|\nu|(E) = \int_E |f| \, d\mu$ .

**Lemma 6.1.4.** We have the following

- (1)  $|\nu(E)| \leq |\nu|(E)$ .
- (2)  $E$  is  $\nu$ -null if and only if  $E$  is  $|\nu|$ -null.
- (3) If  $\kappa$  is another **signed measure**, then  $\kappa \perp \nu$  if and only if  $\kappa \perp |\nu|$  if and only if  $\kappa \perp \nu^+$  and  $\kappa \perp \nu^-$ .

**Proof.** ■

DIY

**Definition 6.1.9** (Finite signed measure). A **signed measure**  $\nu$  is *finite* if  $|\nu|$  is a **finite measure**, and similarly for  $\sigma$ -finite.

**Remark.** This holds if and only if  $\nu^+, \nu^-$  are both **finite** (resp.  $\sigma$ -finite) **measures**.



## 6.2 Absolutely Continuous Measures

**Definition 6.2.1** (Absolutely continuous measure). Let  $\mu$  be a **positive measure**,  $\nu$  be a **signed measure**, both on  $(X, \mathcal{A})$ . We say that  $\nu$  is *absolutely continuous with respect to*  $\mu$ , denoted as  $\nu \ll \mu$ , provided that for all  $E \in \mathcal{A}$ ,  $\mu(E) = 0$  implies  $\nu(E) = 0$ .

**Remark.** This is equivalent to every  $\mu$ -null set being  $\nu$ -null.

**Example.** If  $(X, \mathcal{A}, \mu)$ ,  $f: X \rightarrow \overline{\mathbb{R}}$ ,  $\nu(E) = \int_E f d\mu$ , then  $\nu \ll \mu$ .

**Notation.**  $d\nu = f d\mu$  means  $\nu$  is a **signed measure** defined by

$$\nu(E) = \int_E f d\mu.$$

**Lemma 6.2.1.** If  $\mu$  is a **positive measure**,  $\nu$  is a **signed measure** on  $(X, \mathcal{A})$ , then

- (1)  $\nu \ll \mu$  if and only if  $|\nu| \ll \mu$  if and only if  $\nu^+ \ll \mu$  and  $\nu^- \ll \mu$ .
- (2)  $\nu \ll \mu$  and  $\nu \perp \mu$  implies  $\nu = 0$ .

**Proof.**

For (2), write  $X = A \cup B$ ,  $A \cap B = \emptyset$ ,  $A$   $\mu$ -null,  $B$   $\nu$ -null. Then

$$\nu(E) = \nu(E \cap A) + \nu(E \cap B) = \nu(E \cap A).$$

Then  $E \cap A \subseteq A$ , so  $\nu(E \cap A) = 0$ . By **absolute continuity**,  $\nu(E \cap A) = 0$ , thus  $\nu(E) = 0$ . ■

DIY (1)

**Theorem 6.2.1** (Radon-Nikodym theorem). Suppose  $\mu$  is a  **$\sigma$ -finite positive measure**,  $\nu$  is a  **$\sigma$ -finite signed measure**, and suppose  $\nu \ll \mu$ . Then there exists  $f: X \rightarrow \overline{\mathbb{R}}$  such that  $d\nu = f d\mu$ , in other words  $\nu(E) = \int_E f d\mu$ .

If  $g$  is another such function with  $d\nu = g d\mu$  then  $f = g$   $\mu$ -a.e..

**Proof.** We'll prove a more general form called **Lebesgue Radon Nikodym theorem**, which is a more general theorem compare to this theorem. ■

**Definition 6.2.2** (Radon-Nikodym derivative). Suppose  $\nu \ll \mu$ . The *Radon-Nikodym derivative of  $\nu$  with respect to  $\mu$*  is a function

$$\frac{d\nu}{d\mu}: X \rightarrow \overline{\mathbb{R}}$$

such that

$$\nu(E) = \int_E \frac{d\nu}{d\mu} d\mu$$

for all  $E \in \mathcal{A}$ .

**Remark.** i.e. we have  $d\nu = \frac{d\nu}{d\mu} d\mu$  in our notation.

**Note.** By **Theorem 6.2.1**, such a function exists and is unique up to equivalence  $\mu$ -a.e. in the  **$\sigma$ -finite** case.

**Example.** Let  $F(x) = e^{2x}: \mathbb{R} \rightarrow \mathbb{R}$ , then

$$\frac{dF}{dm} = 2e^{2x}.$$

**Proof.** Since  $F$  is continuous and strictly increasing, so we may define a [Lebesgue-Stieltjes measure](#)  $\mu_F$  on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

This is defined to be the unique [locally finite measure](#) satisfying

$$\mu_F([a, b]) = F(b) - F(a) = e^{2b} - e^{2a}.$$

Then one can check that

$$\mu_F(E) = \int_E 2e^{2x} dx.$$

By uniqueness and the classical [fundamental theorem of calculus](#), since the right-hand side is a [locally finite Borel measure](#), and  $\kappa([a, b]) = e^{2b} - e^{2a}$ , thus  $\mu_F = \kappa$ . Therefore,  $\mu_F \ll m$  and we have

$$\frac{d\mu_F}{dm} = 2e^{2x} = \frac{dF}{dx}.$$

⊛

**Example.** Let  $C(X): \mathbb{R} \rightarrow \mathbb{R}$  be the [Cantor function](#). Then  $C'(x) = 0$  outside the [Cantor set](#), but we don't always have

$$\mu_C(E) \neq \int_E 0 dx.$$

So the candidate derivative is 0, but this fails.

**Proof.** In particular,

$$C(b) - C(a) \neq \int_a^b C'(x) dx.$$

In fact,  $\mu_C \not\ll m$  because  $\mu_C \perp m$  and  $\mu_C \neq 0$ .

Thus, the existence of a derivative [almost everywhere](#) and continuity is not enough to guarantee a version of the [fundamental theorem of calculus](#) to hold. ⊛

## Lecture 31: Lebesgue-Radon-Nikodym Theorem

**Lemma 6.2.2.** Let  $\mu, \nu$  be [finite positive measures](#) on  $(X, \mathcal{A})$ . Then either

- $\nu \perp \mu$ .
- There exists an  $\epsilon > 0$ , an  $F \in \mathcal{A}$  such that  $\mu(F) > 0$  and  $F$  is a [positive set for the measure](#)  $\nu - \epsilon\mu$ , i.e., for all  $G \subseteq F$ ,  $\nu(G) \geq \epsilon\mu(G)$ .

**Proof.** Let  $\kappa_n = \nu - (1/n)\mu$ . By [Theorem 6.1.1](#) we have  $X = P_n \cup N_n$  for  $P_n$  [positive](#),  $N_n$  [negative](#) for  $\kappa_n$ . Also, we let  $P = \bigcup_n P_n$ ,  $N = \bigcap_n N_n = X \setminus P$ , then  $X = P \cup N$ .

We see that for any  $n$  we have  $\kappa_n(N) \leq 0$  because  $N \subseteq N_n$ . Thus,

$$0 \leq \nu(N) \leq \frac{1}{n}\mu(N),$$

which implies  $\nu(N) = 0$ . Because  $\nu$  is [positive](#) for any  $N' \subseteq N$  we have  $0 \leq \nu(N') \leq \nu(N)$ , and thus  $\nu(N') = 0$ . This shows  $N$  is [null](#) for  $\nu$ . Now, we see that

- If  $\mu(P) = 0$ , then  $\nu \perp \mu$ .
- If  $\mu(P) \neq 0$ , then we have  $\mu(P) > 0$  hence  $\mu(P_n) > 0$  for some  $n$ . With  $F = P_n$  and  $\epsilon = 1/n$ , then  $F$  is a [positive set](#) for  $\kappa_n = \nu - (1/n)\mu$  as desired.

28 Mar. 11:00



**Theorem 6.2.2** (Lebesgue-Radon-Nikodym theorem). Let  $\mu$  be a  $\sigma$ -finite positive measure,  $\nu$  a  $\sigma$ -finite signed measure on  $(X, \mathcal{A})$ . Then there are unique  $\sigma$ -finite signed measures  $\lambda, \rho$  on  $(X, \mathcal{A})$  such that

$$\lambda \perp \mu, \quad \rho \ll \mu, \quad \nu = \lambda + \rho.$$

Furthermore, there exists a measurable function  $f: X \rightarrow \overline{\mathbb{R}}$  such that  $d\rho = f d\mu$ .<sup>a</sup> And if there is another  $g$  such that  $d\rho = g d\mu$ , then  $f = g$   $\mu$ -a.e.

<sup>a</sup>That is for all  $E \in \mathcal{A}$ ,  $\rho(E) = \int_E f d\mu$ .

**Proof.** We prove it step by step.

- (1) Assume  $\mu, \nu$  are finite positive measures. We first prove the existence of  $\lambda, f$ , and  $d\rho = f d\mu$ . Let

$$\begin{aligned} \mathcal{F} &= \left\{ g: X \rightarrow [0, \infty] \mid \int_E g d\mu \leq \nu(E), \forall E \in \mathcal{A} \right\} \\ &= \{g: X \rightarrow [0, \infty] \mid d\nu - g d\mu \text{ is a positive measure}\}. \end{aligned}$$

This set is nonempty since  $g = 0 \in \mathcal{F}$ . Let  $s = \sup\{\int_X g d\mu \mid g \in \mathcal{F}\}$ .

**Claim.** There is an  $f \in \mathcal{F}$  such that  $s = \int_X f d\mu$ .

**Proof.** If  $g, h \in \mathcal{F}$ , we can define  $u(x) = \max\{g(x), h(x)\}$ , then  $u \in \mathcal{F}$ . This can be seen by letting  $A = \{x \mid g(x) \geq h(x)\}$ , then

$$\int_E u d\mu = \int_{E \cap A} g d\mu + \int_{E \cap A^c} h d\mu \leq \nu(E \cap A) + \nu(E \cap A^c) = \nu(E).$$

There exist measurable functions  $g_1, g_2, \dots \in \mathcal{F}$  such that

$$\lim_{n \rightarrow \infty} \int_X g_n d\mu = s.$$

We can replace  $g_2$  by  $\max(g_1, g_2)$ ,  $g_3$  by  $\max(g_1, g_2, g_3)$ . Generally,

$$g_n \leftarrow \max(g_1, g_2, \dots, g_n),$$

so that we may assume  $0 \leq g_1 \leq g_2 \leq \dots$

Then we still know that  $\lim_{n \rightarrow \infty} \int_X g_n d\mu = s$ , as all the relevant integrals are bounded above by  $s$ . Now let  $f(x) = \sup_n g_n(x) = \lim_{n \rightarrow \infty} g_n(x)$ , by monotone convergence theorem,

$$\int_E f d\mu = \lim_{n \rightarrow \infty} \int_E g_n d\mu \leq \nu(E).$$

Thus,  $f \in \mathcal{F}$ , and when  $E = X$  we get  $\int_X f d\mu = s$  as desired. ⊛

Let  $\rho(E) := \int_E f d\mu$ , then we of course have  $\rho \ll \mu$ , and also, we know

$$0 \leq \rho(X) = \int_X f d\mu \leq \nu(X) < \infty.$$

Thus,  $\rho$  is a finite positive measure, so we can define  $\lambda(E) := \nu(E) - \rho(E)$ , then

$$\lambda(E) = \nu(E) - \int_E f d\mu \geq 0$$

because  $f \in \mathcal{F}$ . Thus,  $\lambda$  is also a **positive measure**, and  $\lambda(X) \leq \nu(X) < \infty$ . It remains to show the following.

**Claim.**  $\lambda \perp \mu$ .

**Proof.** Suppose not, by [Lemma 6.2.2](#), there exists  $\epsilon > 0$ ,  $F \in \mathcal{A}$  such that  $\mu(F) > 0$  and  $F$  is a **positive set** for  $\lambda - \epsilon\mu$ .

Then this says that  $d\lambda - \epsilon\mathbb{1}_F d\mu$  is a **positive measure**, that is,

$$d\nu - f d\mu - \epsilon\mathbb{1}_F d\mu$$

is a **positive measure**. But, this will break maximality of  $f$ , specifically, let  $g(x) = f(x) + \epsilon\mathbb{1}_F(x)$ . Then for all  $E \in \mathcal{A}$  we have

$$\begin{aligned} \int_E g d\mu &= \int_E f d\mu + \epsilon\mu(E \cap F) \\ &= \nu(E) - \lambda(E) + \epsilon\mu(E \cap F) \\ &\leq \nu(E) - \lambda(E \cap F) + \epsilon\mu(E \cap F) \leq \nu(E) \end{aligned}$$

since  $\lambda(E \cap F) - \epsilon\mu(E \cap F) \geq 0$ . Thus,  $g \in \mathcal{F}$ . We then see that

$$s \geq \int_X g d\mu = \int_X f d\mu + \int_X \epsilon\mathbb{1}_F d\mu = s + \epsilon\mu(F) > s,$$

which is a contradiction. ⊗

We see that the existence of  $\lambda, f$ , and  $d\rho = f d\mu$  is proved. As for uniqueness, if there are  $\lambda'$  and  $f'$  such that  $d\nu = d\lambda' + f' d\mu$ , we then have

$$d\lambda - d\lambda' = (f' - f) d\mu.$$

But we see that  $\lambda - \lambda' \perp \mu$  while  $(f' - f) d\mu \ll \mu$ , hence

$$d\lambda - d\lambda' = (f' - f) d\mu = 0,$$

so  $\lambda = \lambda'$  and  $f = f'$   $\mu$ -a.e. by [Proposition 2.3.1](#).

- (2) Suppose that  $\mu$  and  $\nu$  are  **$\sigma$ -finite measures**. Then  $X$  is a countable disjoint union of  **$\mu$ -finite** sets and a countable disjoint union of  $\nu$ -finite sets. By taking intersections of these we obtain a disjoint sequence  $\{A_j\} \subset \mathcal{A}$  such that  $\mu(A_j)$  and  $\nu(A_j)$  are finite for all  $j$  and  $X = \bigcup_j A_j$ . Define  $\mu_j(E) = \mu(E \cap A_j)$  and  $\nu_j(E) = \nu(E \cap A_j)$ , then by the reasoning above, for each  $j$  we have

$$d\nu_j = d\lambda_j + f_j d\mu_j$$

where  $\lambda_j \perp \mu_j$ . Since  $\mu_j(A_j^c) = \nu_j(A_j^c) = 0$ , we have

$$\lambda_j(A_j^c) = \nu_j(A_j^c) - \int_{A_j^c} f_j d\mu_j = 0,$$

and we may assume that  $f_j = 0$  on  $A_j^c$ . Let  $\lambda = \sum_j \lambda_j$  and  $f = \sum_j f_j$ , we then have

$$d\nu = d\lambda + f d\mu, \quad \lambda \perp \mu,$$

and  $d\lambda$  and  $f d\mu$  are  **$\sigma$ -finite**, as desired. As for uniqueness, it's the same as for the first case.

- (3) We now consider the general case. If  $\nu$  is a **signed measure**, we apply the preceding argument to  $\nu^+$  and  $\nu^-$  and subtract the results.

■

Check!

**Remark.** Notationally, we may write  $d\nu = d\lambda + f d\mu$ , where  $d\lambda$  and  $d\mu$  are [singular](#) to each other.

## Lecture 32: Lebesgue Differentiation Theorem for Regular Borel Measures

We now do an example to illustrate [Theorem 6.2.2](#).

30 Mar. 11:00

**Example.** Let  $\mu = m$ ,  $\nu = \mu_F$  ([Lebesgue-Stieltjes measure](#) for  $F$ ). We'll define  $F(x)$  by

$$F(x) = \begin{cases} e^{3x}, & \text{if } x \leq 0; \\ 1, & \text{if } 0 < x < 1; \\ 5, & \text{if } x \geq 1. \end{cases}$$

Then we have

$$\mu_F(E) = \int_{E \cap \mathbb{R}_{<0}} 3e^{3x} dx + 4\delta_1(E).$$

**Proof.** It is enough to check on  $(-\infty, x]$  because these are [locally finite Borel measures](#) on  $\mathbb{R}$ . Then we have

$$\mu_F = d\rho + d\lambda = f dm + d\lambda$$

where  $f = \mathbb{1}_{\mathbb{R}_{<0}} 3e^{3x}$  and  $\lambda = 4\delta_1$ ,  $\lambda \perp m$ . \*

Specifically, we call such a decomposition *Lebesgue decomposition* of  $\nu$  with respect to  $\mu$ . Now, with the condition  $\nu \ll \mu$ , [Theorem 6.2.2](#) implies that  $d\nu = f d\mu$  for some  $f$ , which is exactly the statement of [Theorem 6.2.1](#). And, it should be clear now that the definition of [Radon Nikodym derivative](#) of  $\nu$  with respect to  $\mu$ , denoted as  $d\nu/d\mu$ , is just  $f$  in this case.

**As previously seen.** If  $\nu = \nu^+ - \nu^-$ , we defined the [total variation](#)  $|\nu| = \nu^+ + \nu^-$ . Then we have  $|\nu(E)| \leq |\nu|(E)$ .

## 6.3 Lebesgue Differentiation Theorem for Regular Borel Measures

**Definition 6.3.1 (Regular).** A Borel [signed measure](#)  $\nu$  on  $\mathbb{R}^d$  is called *regular* if

- (compact finite)  $|\nu|(K) < \infty$  for all compact  $K$ .
- (outer regularity) We have [outer regularity](#)

$$|\nu|(E) = \inf\{|\nu|(U) \mid \text{open } U \supseteq E\}$$

for every [Borel set](#)  $E$ .

**Example.** We see that

- (1) Any [Lebesgue-Stieltjes measure](#) on  $\mathbb{R}$  has this property from [Theorem 1.7.1](#), so is the difference between two of them (at least if one of them is [finite](#)).
- (2) The [Lebesgue measure](#) on  $\mathbb{R}^d$  is [regular](#).

**Note.** From [compact finiteness](#), if  $\nu$  is [regular](#) then it is  $\sigma$ -finite.

**Lemma 6.3.1.**  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  if and only if  $d\nu = f dm$  is [regular](#).

**Proof.** We prove this in two directions.

( $\Leftarrow$ ) Suppose  $d\nu = f dm$  is **regular**. Then

$$|\nu|(K) = \int_K |f| dm < \infty$$

for all compact  $K$ , thus  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ .

( $\Rightarrow$ ) Suppose  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ . This condition is clearly equivalent to **compact finiteness**. If this holds, then the **outer regularity** may be verified directly as follows. Suppose that  $E$  is a bounded **Borel set**. Given  $\delta > 0$ , by **Theorem 3.5.1**, there is a bounded open  $U \supset E$  such that  $m(U) < m(E) + \delta$  and hence  $m(U \setminus E) < \delta$ . But then, given  $\epsilon > 0$ , there is<sup>a</sup> an open  $U \supset E$  such that

$$\int_{U \setminus E} f dm < \epsilon$$

and hence

$$\int_U f dm < \int_E f dm + \epsilon.$$

The case of unbounded  $E$  follows easily by writing  $E = \bigcup_{j=1}^{\infty} E_j$  where  $E_j$  is bounded and finding an open  $U_j \supset E_j$  such that

$$\int_{U_j \setminus E_j} f dm < \epsilon 2^{-j}.$$

■

<sup>a</sup>This follows from [FF99] Corollary 3.6.

**As previously seen.** Recall the **Lebesgue differentiation theorem**, here we had that if  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$  implies that for Lebesgue **almost every**  $x$ ,

$$\lim_{r \rightarrow 0} \frac{1}{m(E_r)} \int_{E_r} f(y) dy = f(x)$$

for any  $\{E_r\}$  **shrinks nicely** to  $x$ .

**Corollary 6.3.1.** Let  $\rho$  be a **regular signed Borel measure** on  $\mathbb{R}^d$ . Suppose  $\rho \ll m$ . Then  $d\rho = f dm$  for some  $f \in L^1_{\text{loc}}(\mathbb{R}^d)$ . So then for Lebesgue **almost every**  $x$  we have

$$\lim_{r \rightarrow 0} \frac{1}{m(E_r)} \int_{E_r} f(y) dy = f(x).$$

Writing this nicely, using established notation, this is

$$\lim_{r \rightarrow 0} \frac{\rho(E_r)}{m(E_r)} = \frac{d\rho}{dm}(x)$$

for every  $\{E_r\}$  **shrinks nicely** to  $x$ .

**Proposition 6.3.1.** Let  $\lambda$  be a **regular positive Borel measure** on  $\mathbb{R}^d$ . Suppose  $\lambda \perp m$ . Then for **Lebesgue almost every**  $x$ , we have

$$\lim_{r \rightarrow 0} \frac{\lambda(E_r)}{m(E_r)} = 0$$

for every  $\{E_r\}$  **shrinking to  $x$  nicely** (equivalently, **shrinking to 0 nicely**).

**Proof.** It is enough to consider  $E_r = B(x, r)$ . We wish to prove that

$$G := \left\{ x \mid \limsup_{r \rightarrow 0} \frac{\lambda(E_r)}{m(E_r)} \neq 0 \right\} = \bigcup_{n=1}^{\infty} G_n$$

where

$$G_n := \left\{ x \mid \limsup_{r \rightarrow 0} \frac{\lambda(E_r)}{m(E_r)} > \frac{1}{n} \right\}$$

such that  $m(G) = 0$ . We see that it suffices to show  $m(G_n) = 0$  for all  $n$ . Since  $\lambda \perp m$ , so we know there exists  $A, B$  such that  $\mathbb{R}^d = A \cup B$  disjoint with  $\lambda(A) = 0$ ,  $m(B) = 0$ . Thus, it suffices to show  $m(G_n \cap A) = 0$ .

**Note.** Alternatively, we can simply define  $G_n$  over  $A$  instead of  $\mathbb{R}^d$ , as in Folland[FF99].

**Claim.** Given a  $A$  and  $B$  defined above induced from Theorem 6.2.2,  $m(G_n \cap A) = 0$  for all  $n$

**Proof.** Fix  $\epsilon > 0$ , since  $\lambda$  is regular, there exists an open set  $U \supseteq A$  such that  $\lambda(U) \leq \lambda(A) + \epsilon = \epsilon$ . We see that for every  $x \in G_n \cap A$ , there is an  $r_x > 0$  such that  $\lambda(B(x, r_x))/m(B(x, r_x)) > 1/n$  where  $B(x, r_x) \subseteq U$ .

Let  $K \subseteq G_n \cap A$ , compact. Then  $K \subseteq \bigcup_{x \in K} B(x, r_x)$ . By compactness, we can take a finite sub-cover, and then use Lemma 4.1.1 to find disjoint  $B_1, B_2, \dots, B_N$  such that each of  $B_i$  is in the form of  $B(x_i, r_{x_i})$  and  $K \subseteq \bigcup_i 3B_i$ . Therefore,

$$m(K) \leq 3^d \sum_{i=1}^N m(B_i) \leq 3^d n \sum_{i=1}^N \lambda(B_i) = 3^d n \lambda \left( \bigcup_{i=1}^N B_i \right) \leq 3^d n \lambda(U) = 3^d n \epsilon.$$

By inner regularity,  $m(G_n \cap A) \leq 3^d n \epsilon$  for any  $\epsilon > 0$ . Taking  $\epsilon \rightarrow 0$  yields  $m(G_n \cap A) = 0$ , so then  $m(G_n) = 0$  as desired.  $\otimes$

## Lecture 33: Monotone Differentiation Theorem

As previously seen. We have that if  $\rho \ll m$  is regular then

$$\lim_{r \rightarrow 0} \frac{\rho(E_r)}{m(E_r)} = \frac{d\rho}{dm}(x)$$

for Lebesgue almost every  $x$ , where  $\{E_r\}$  shrinks nicely to  $x$ . Likewise, if  $\lambda \perp m$  regular (positive measure) then

$$\lim_{r \rightarrow 0} \frac{\lambda(E_r)}{m(E_r)} = 0$$

for Lebesgue almost every  $x$ , where  $\{E_r\}$  shrinks nicely to  $x$ .

From this, we can easily deduce the following important result.

**Theorem 6.3.1** (Lebesgue differentiation theorem for regular measures). Let  $\nu$  be a regular Borel signed measure on  $\mathbb{R}^d$ . Then  $d\nu = d\lambda + f dm$ ,  $\lambda \perp m$  by Theorem 6.2.2. Then for Lebesgue almost every  $x$ ,

$$\lim_{r \rightarrow 0} \frac{\nu(E_r)}{m(E_r)} = f(x)$$

for every  $\{E_r\}$  shrinks nicely to  $x$ .

1 Apr. 11:00

**Proof.** It must be checked that  $\nu$  **regular** implies  $\lambda, f \, dm$  are **regular**. In particular, since  $f \in L^1_{\text{loc}}$ , so from [Theorem 4.2.1](#) and its corollary ([Corollary 4.2.1](#), [Corollary 4.2.2](#)), we see that it suffices to show that if  $\lambda$  is **regular** and  $\lambda \perp m$ , then for Lebesgue a.e.  $x$ ,

Check!

$$\lim_{r \rightarrow 0} \frac{\lambda(E_r)}{m(E_r)} \rightarrow 0$$

when  $\{E_r\}$  **shrinks nicely to  $x$** . It also suffices to take  $E_r = B(r, x)$  and to assume that  $\lambda$  is **positive**, since for some  $\alpha > 0$ , we have

$$\left| \frac{\lambda(E_r)}{m(E_r)} \right| \leq \frac{|\lambda|(E_r)}{m(E_r)} \leq \frac{|\lambda|(B(r, x))}{m(E_r)} \leq \frac{|\lambda|(B(r, x))}{\alpha m(B(r, x))}.$$

Therefore, if  $|\lambda|(E_r)/m(E_r) \rightarrow 0$ , so does  $|\lambda(E_r)/m(E_r)|$ , hence  $\lambda(E_r)/m(E_r)$ . We see that the result then follows directly from [Proposition 6.3.1](#). ■

## 6.4 Monotone Differentiation Theorem

We first formalize one ambiguous notation we used long time ago with discussing **distribution function**. Namely,  $F(x^+), F(x^-)$ .

**Definition.** For an  $F: \mathbb{R} \rightarrow \mathbb{R}$  that is monotonically increasing, we have the following.

**Definition 6.4.1** ( $F(x^+)$ ). We define  $F(x^+) = \lim_{y \rightarrow x^+} F(y)$ .

**Definition 6.4.2** ( $F(x^-)$ ). We define  $F(x^-) = \lim_{y \rightarrow x^-} F(y)$ .

**Remark.** If  $F$  is monotonically increasing, then  $F(x^+), F(x^-)$  exist and are

$$\inf_{y > x} F(y), \quad \sup_{y < x} F(y)$$

respectively since it's bounded below/above respectively by  $F(x)$ .

**Lemma 6.4.1.** If  $F: \mathbb{R} \rightarrow \mathbb{R}$  is monotonically increasing, then

$$D = \{x \in \mathbb{R} \mid F \text{ is discontinuous at } x\}$$

is a countable set.

**Proof.** Since  $F$  is monotonically increasing,  $x \in D$  if and only if  $F(x^+) > F(x^-)$ . For each  $x \in D$ , let  $I_x = (F(x^-), F(x^+))$ , not empty. Also, if  $x, y \in D$ ,  $x \neq y$ , then  $I_x, I_y$  are disjoint. Now, for  $|x| < N$ ,  $I_x$  lies in the interval  $(F(-N), F(N))$ . Hence,

$$\sum_{|x| < N} [F(x^+) - F(x^-)] \leq F(N) - F(-N) < \infty,$$

which implies that

$$D \cap (-N, N) = \{x \in (-N, N) \mid F(x^+) \neq F(x^-)\}$$

is countable. Since this is true for all  $N$ , the result follows. ■

**Theorem 6.4.1** (Monotone differentiation theorem). Let  $F$  be an increasing function, then



- $F$  is differentiable [Lebesgue almost everywhere](#).
- $G(x) := F(x^+)^a$  is differentiable [almost everywhere](#).
- $G' = F'$  [almost everywhere](#)

<sup>a</sup>Observe that  $G$  is increasing and right-continuous.

**Proof.** Start with  $G(x) := F(x^+)$ , which is increasing and right-continuous on  $\mathbb{R}$ . There is then a [Lebesgue-Stieltjes measure](#)  $\mu_G$  on  $\mathbb{R}$ , thus it is [regular](#) on  $\mathbb{R}$ . We see

$$\frac{G(x+h) - G(x)}{h} = \begin{cases} \frac{\mu_G((x, x+h])}{m((x, x+h])}, & \text{if } h > 0; \\ \frac{\mu_G((x+h, x])}{m((x+h, x])}, & \text{if } h < 0. \end{cases}$$

Note that both  $\{(x, x+h]\}$  and  $\{(x+h, x]\}$  [shrink nicely](#) to  $x$  as  $|h| \rightarrow 0$ . By [Theorem 6.3.1](#) (since these [shrink nicely](#)), we then know that these both converge for [Lebesgue almost every](#)  $x$  to some common limit  $f(x)$ . Hence,  $G'$  exists [Lebesgue almost everywhere](#). We now show that by defining  $H := G - F$ ,  $H'$  exists and equals zero [a.e.](#)

Observe that  $H(x) = G(x) - F(x) \geq 0$ , and we see that

$$\{x \mid H(x) > 0\} \subseteq \{x \mid F \text{ is discontinuous at } x\}.$$

The latter set is then countable by [Lemma 6.4.1](#), hence we can write  $\{x \mid H(x) > 0\} = \{x_n\}$ . Then let

$$\mu := \sum_n H(x_n) \delta_{x_n}.$$

This is a [Borel measure](#), so we check if it is [locally finite](#). Indeed, since

$$\mu((-N, N)) = \sum_{-N < x_n < N} H(x_n) \leq G(N) - F(-N) < \infty,$$

where checking the inequality just consists of seeing that the intervals  $(F(x_n), G(x_n))$  are disjoint and is a subset of  $(F(-N), G(N))$ , so

$$\sum_{-N < x_n < N} H(x_n) = \mu \left( \bigcup_n (F(x_n), G(x_n)) \right) \leq \mu((F(-N), G(N))).$$

Thus,  $\mu$  is a [Lebesgue-Stieltjes measure](#) on  $\mathbb{R}$ , so it is [regular](#).

**Remark.** Special to  $\mathbb{R}$ , we have that for a [measure](#),

$$\text{locally finite Borel} \Rightarrow \text{Lebesgue-Stieltjes} \Rightarrow \text{regular} \Rightarrow \text{outer regularity}.$$

Also, we have  $\mu \perp m$  since  $m(E) = \mu(E^c) = 0$  where  $E = \{x_n\}$ . Then we have that

$$\left| \frac{H(x+h) - H(x)}{h} \right| \leq \frac{H(x+h) + H(x)}{|h|} \leq \frac{\mu((x-2h, x+2h))}{|h|},$$

which goes to 0 for [Lebesgue almost every](#)  $x$  by [Theorem 6.3.1](#) and that  $\mu \perp m$ .

Thus,  $H$  is differentiable [almost everywhere](#) and  $H' = 0$  [almost everywhere](#), which implies  $F$  is differentiable [almost everywhere](#) and  $F' = G'$  [almost everywhere](#). ■

**Proposition 6.4.1.** Suppose  $F$  is an increasing function, then  $F'$  exists [almost everywhere](#) and is [measurable](#), then

$$\int_a^b F'(x) dx \leq F(b) - F(a).$$

**Example.** The inequality can't be made into equality in [Proposition 6.4.1](#) by the given condition.

**Proof.** Take  $F(x)$  as

$$F(x) := \begin{cases} 0, & \text{if } x \leq 0; \\ 1, & \text{if } x > 0. \end{cases}$$

Then  $F'(x) = 0$  [almost everywhere](#). So

$$\int_{-1}^1 F'(x) dx = 0 < 1 = F(1) - F(-1).$$

⊛

**Example.** The inequality can't be made into equality in [Proposition 6.4.1](#) even if  $F$  is continuous in addition.

**Proof.** Take  $F(x)$  to be the [Cantor function](#). Then  $F'(x) = 0$  [almost everywhere](#), but

$$\int_0^1 F'(x) dx = 0 < 1 = F(1) - F(0).$$

⊛

## Lecture 34: Functions of Bounded Variation

4 Apr. 11:00

**Proof of Proposition 6.4.1.** Let

$$G(x) := \begin{cases} F(a), & \text{if } x < a; \\ F(x), & \text{if } a \leq x \leq b; \\ F(b), & \text{if } x > b. \end{cases}$$

Then  $G$  is increasing. We define

$$g_n(x) = \frac{G(x + 1/n) - G(x)}{1/n} \rightarrow F'(x)$$

for almost every  $x \in [a, b]$ . We note that  $g_n(x) \geq 0$ . [Theorem 2.2.2](#) tells us that

$$\int_a^b F'(x) dx = \int_a^b \liminf_{n \rightarrow \infty} g_n(x) dx \leq \liminf_{n \rightarrow \infty} \int_a^b g_n(x) dx.$$

We then evaluate

$$\begin{aligned} \int_a^b g_n(x) dx &= n \left( \int_{a+1/n}^{b+1/n} G(x) dx - \int_a^b G(x) dx \right) \\ &= n \left( \int_b^{b+1/n} G(x) dx - \int_a^{a+1/n} G(x) dx \right) \\ &\leq n \left( G\left(b + \frac{1}{n}\right) \cdot \frac{1}{n} - G(a) \cdot \frac{1}{n} \right) = F(b) - F(a). \end{aligned}$$

Therefore,

$$\int_a^b F'(x) dx \leq F(b) - F(a).$$

■

## 6.5 Functions of Bounded Variation

**Definition 6.5.1** (Total variation function). For  $F: \mathbb{R} \rightarrow \mathbb{R}$ , the *total variation function* of  $F$  is  $T_F: \mathbb{R} \rightarrow [0, \infty]$  defined by

$$T_F(x) = \sup \left\{ \sum_{i=1}^n |F(x_i) - F(x_{i-1})| \mid n \in \mathbb{N}, -\infty < x_0 < x_1 < \dots < x_n = x \right\}.$$

**Lemma 6.5.1.**  $T_F(b)$  is equal to

$$T_F(a) + \sup \left\{ \sum_{i=1}^n |F(x_i) - F(x_{i-1})| \mid n \in \mathbb{N}, a = x_0 < x_1 < \dots < x_n = b \right\}$$

if  $a < b$ .

**Proof.** The idea is that the sums in the Definition 6.1.8 of  $T_F$  are made bigger if the additional subdivision points  $x_j$  are added. Hence, if  $a < b$ ,  $T_F(b)$  is unaffected if we assume that  $a$  is always one of the subdivision points. ■

**Remark.**  $T_F$  is increasing.

**Definition 6.5.2** (Bounded variation). We say that  $F$  is of *bounded variation*, denoted as  $F \in BV$ , provided that

$$T_F(\infty) = \lim_{x \rightarrow \infty} T_F(x) < \infty.$$

Similarly,  $F \in BV([a, b])$  means that

$$\sup \left\{ \sum_{i=1}^n |F(x_i) - F(x_{i-1})| \mid n \in \mathbb{N}, a = x_0 < x_1 < \dots < x_n = b \right\} < \infty.$$

**Remark.** We see the following.

- (1) If  $F$  is of *bounded variation*, then  $F$  is bounded.
- (2)  $F(x) = \sin x$  is not of *bounded variation*, but it is of *bounded variation* over any  $[a, b]$ .
- (3) For  $F(x)$  defined as

$$F(x) = \begin{cases} \frac{\sin x}{x}, & \text{if } x \neq 0; \\ 1, & \text{if } x = 0 \end{cases}$$

is not of *bounded variation* of  $[a, b]$  if  $a < 0 < b$  because the harmonic series does not converge.

Before we see more properties of *bounded variation* function, we introduce a useful characterization of a function.

**Definition 6.5.3** (Lipschitz). A function  $F: [a, b] \rightarrow \mathbb{C}$  is called *Lipschitz* if there exists an  $M \geq 0$  such that

$$|F(x) - F(y)| \leq M |x - y|.$$

**Remark.** We have the following.

- (1) If  $F, G$  are of *bounded variation*,  $\alpha F + \beta G$  are of *bounded variation* for  $\alpha, \beta \in \mathbb{C}$ .
- (2) If  $F$  is increasing and bounded, then  $F$  is a function of *bounded variation*.

- (3) If  $F$  is **Lipschitz** on  $[a, b]$ , then  $F \in BV([a, b])$ .
- (4) If  $F$  is differentiable, and  $F'$  is bounded on  $[a, b]$ , then  $F$  is **Lipschitz** from mean value theorem, so it is in  $BV([a, b])$ .

In particular, we have the following.

**Remark.** If  $F(x) = \int_{-\infty}^x f(t) dt$  for  $f \in L^1(\mathbb{R})$ , then  $F \in BV$ .

**Proof.** We see this by

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| = \sum_{i=1}^n \left| \int_{x_{i-1}}^{x_i} f(t) dt \right| \leq \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |f(t)| dt = \int_{x_0}^{x_n} |f(t)| dt \leq \int_{-\infty}^{\infty} |f(t)| dt,$$

which is finite since  $f \in L^1(\mathbb{R})$ . \*

**Lemma 6.5.2.** If  $F \in BV$ , then  $T_F$  is bounded, increasing,  $T_F(-\infty) = 0$ .

**Proof.** Fix  $\epsilon > 0$  and for a  $x \in \mathbb{R}$ , there exists  $x_0 < \dots < x_n = x$  such that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \geq T_F(x) - \epsilon.$$

From **Lemma 6.5.1**, we see that  $T_F(x) - T_F(x_0) \geq T_F(x) - \epsilon$ , and hence  $T_F(y) \leq \epsilon$  for  $y \leq x_0$ , so  $T_F(-\infty) = 0$ . ■

**Lemma 6.5.3.**  $F \in BV$ , then  $T_F \pm F$  are increasing and bounded and.

**Proof.** Let  $x < y$  and fix  $\epsilon > 0$ , then there are points  $x_0 < x_1 < \dots < x_n = x$  such that

$$T_F(x) \leq \sum_{i=1}^n |F(x_i) - F(x_{i-1})| + \epsilon.$$

Furthermore,

$$T_F(y) \geq \sum_{i=1}^n |F(x_i) - F(x_{i-1})| + |F(y) - F(x)|.$$

Then, since  $\pm(F(y) - F(x)) \leq |F(y) - F(x)|$ , we have<sup>a</sup>

$$T_F(y) \pm (F(y) - F(x)) \geq \sum_{i=1}^n |F(x_i) - F(x_{i-1})| \geq T_F(x) - \epsilon,$$

hence

$$T_F(y) \pm F(y) \geq T_F(x) \pm F(x) - \epsilon.$$

Taking  $\epsilon \rightarrow 0$  yields the result. ■

<sup>a</sup>To be precise, we note that  $\pm(F(y) - F(x)) \geq -|F(y) - F(x)|$  and proceed.

With this, we have the following useful decomposition for any  $F \in BV$ .

**Definition 6.5.4 (Jordan decomposition).** For any  $F \in BV$ , the *Jordan decomposition* of  $F$  is

$$F = \frac{T_F + F}{2} - \frac{T_F - F}{2},$$

which is a difference of increasing and bounded functions.

Specifically, we give special names for the terms in the **Jordan decomposition** of  $F$ .

**Definition 6.5.5** (Positive variation). The *positive variation* of  $F$  is  $\frac{1}{2}(T_F + F)$ .

**Definition 6.5.6** (Negative variation). The *negative variation* of  $F$  is  $\frac{1}{2}(T_F - F)$ .

**Proof.** It's clear that the above decomposition holds, we just need to check that both terms are increasing and bounded functions. They are increasing from [Lemma 6.5.3](#). To show they are bounded, for  $y > x$ ,

$$T_F(y) \pm F(y) \geq T_F(x) \pm F(x)$$

implies that

$$|F(y) - F(x)| \leq T_F(y) - T_F(x) \leq T_F(\infty) - T_F(-\infty) < \infty,$$

so that  $F$ , and hence  $T_F \pm F$  is all bounded.  $\circledast$

**Note.** Do not be confused between [Definition 6.5.4](#) and [Theorem 6.1.2](#) for *signed measure*! Also, the *positive variation of a function* and *positive variation of a measure* is confusing, same for the negative case.

**Theorem 6.5.1.**  $F$  is of *bounded variation* if and only if  $F = F_1 - F_2$  for  $F_1, F_2$  increasing and bounded.

**Proof.** The forward implication is given by [Lemma 6.5.3](#) and [the Jordan decomposition](#), and the other direction is easy to check.  $\blacksquare$

Check!

**Corollary 6.5.1.**  $F \in BV$  implies that  $F$  is differentiable *almost everywhere*. Furthermore,

- (1)  $F(x^+), F(x^-)$  exist for all  $x$  as do  $F(-\infty), F(\infty)$ .
- (2) The set of discontinuities of  $F$  is countable.
- (3)  $G(x) = F(x^+)$  is differentiable and  $G' = F'$  *almost everywhere*.
- (4)  $F' \in L^1(\mathbb{R}, m)$  (i.e.  $F \in L^1_{\text{loc}}(\mathbb{R})$ ) for every  $a < b$ .

**Proof.**  $\blacksquare$

DIY

## Lecture 35: Continue on Functions of Bounded Variation

**Definition 6.5.7** (Normalized bounded variation). A function  $G \in BV$  is said to have *normalized bounded variation*, denoted as  $G \in NBV$  provided that  $G$  is right continuous and  $G(-\infty) = 0$ .

6 Apr. 11:00

**Example.** If  $F$  is increasing and bounded,  $F$  right continuous,  $F(-\infty) = 0$ .

$F(x) = \int_{-\infty}^x f(t) dt, f \in L^1(\mathbb{R})$ . Midterm gave  $F$  is uniformly continuous.

**Lemma 6.5.4.** If  $F \in BV$  is right continuous, then  $T_F \in NBV$ .

**Proof.**  $T_F$  is bounded, increasing, and satisfies  $T_F(-\infty) = 0$  by [Lemma 6.5.2](#). Thus,  $T_F \in BV$ . Hence, we just need to check that  $T_F$  is right continuous. Suppose not, then there is a point  $a \in \mathbb{R}$  such that  $c := T_F(a^+) - T_F(a) > 0$  since  $T_F$  is increasing.

Fix  $\epsilon > 0$ , since  $F(x)$  and  $g(x) := T_F(x^+)$  are right continuous, there exists a  $\delta > 0$  such that for  $y \in (a, a + \delta]$  we have  $|F(y) - F(a)| < \epsilon$  and  $|g(y) - g(a)| < \epsilon$ . We then have that

$$T_F(y) - T_F(a^+) \leq \underbrace{T_F(y^+) - T_F(a^+)}_{|g(y) - g(a)|} < \epsilon.$$

There exist  $a = x_0 < x_1 < \cdots < x_n = a + \delta$  such that

$$\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \geq T_F(a + \delta) - T_F(a) - \frac{c}{4} \geq T_F(a^+) - T_F(a) - \frac{c}{4} = \frac{3c}{4}.$$

From  $|F(x_1) - F(a)| < \epsilon$ , we have

$$\sum_{i=2}^n |F(x_i) - F(x_{i-1})| \geq \frac{3}{4}c - \epsilon.$$

From the same reason, there exist  $a = t_0 < \cdots < t_k = x_1$  such that

$$\sum_{i=1}^k |F(t_i) - F(t_{i-1})| \geq T_F(x_1) - T_F(a) - \frac{c}{4} \geq \frac{3}{4}c$$

since  $T_F$  is increasing and  $x_1 > a$  strictly. Then as  $[a, a + \delta] = [a, x_1] \cup [x_1, a + \delta]$  we see that

$$T_F(a + \delta) - T_F(a) \geq \sum_{j=1}^k |F(t_j) - F(t_{j-1})| + \sum_{i=2}^n |F(x_i) - F(x_{i-1})| \geq \frac{3}{4}c + \left(\frac{3}{4}c - \epsilon\right) = \frac{3}{2}c - \epsilon.$$

Combine these all together, we get

$$\epsilon + c \geq \underbrace{T_F(a + \delta) - T_F(a^+)}_{\leq \epsilon} + \underbrace{T_F(a^+) - T_F(a)}_{\leq c} = \underbrace{T_F(a + \delta) - T_F(a)}_{\text{above}} \geq \frac{3}{2}c - \epsilon \Rightarrow c \leq 4\epsilon.$$

By taking  $\epsilon \rightarrow 0$  yields  $c = 0$ , which is a contradiction. ■

**Corollary 6.5.2.**  $F \in NBV$  if and only if  $F = F_1 - F_2$ ,  $F_1, F_2 \in NBV$  and increasing.

**Proof.** Consider the [Jordan decomposition](#)

$$F = \frac{T_F + F}{2} - \frac{T_F - F}{2}$$

and proceed similar as in [Theorem 6.5.1](#). ■

**Theorem 6.5.2.** We have the following.

- (1) Suppose that  $\mu$  is a [finite signed Borel measure](#) on  $\mathbb{R}$ , then  $F(x) = \mu((-\infty, x]) \in NBV$ .
- (2)  $F \in NBV$  implies that there exists a unique [finite signed Borel measure](#) on  $\mathbb{R}$  satisfying  $\mu_F((-\infty, x]) = F(x)$ .

**Proof.** We prove this one by one.

- (1) Let  $\mu = \mu^+ - \mu^-$ , then  $F = F^+ - F^-$ , where  $F^\pm(x) = \mu^\pm((-\infty, x])$ , which are bounded, right continuous and  $F^\pm(-\infty) = 0$ , so  $F^\pm \in NBV$ .
- (2) Let  $F \in NBV$ , then  $F = F_1 - F_2$ ,  $F_1, F_2 \in NBV$  and increasing from [Corollary 6.5.2](#). Then define  $\mu_{F_1}, \mu_{F_2}$  as the [Lebesgue-Stieltjes measure](#) correspond to  $F_1$  and  $F_2$ , and set  $\mu_F := \mu_{F_1} - \mu_{F_2}$ .

Uniqueness  
is shown  
in [Appendix A.6.1](#).

**Remark (Jordan decomposition).** We see that why both  $\mu^\pm$  and  $F^\pm$  all has the name *Jordan decomposition* from the above proof of (1).

The next natural arising question is the following.

**Problem.** Which functions in  $NBV$  correspond to measures  $\mu$  such that  $\mu \perp m$  or  $\mu \ll m$ ?

One answer is given in the following proposition.

**Proposition 6.5.1.** We have the following.

- (1) If  $F \in NBV$ , then  $F$  is differentiable **almost everywhere** and  $F' \in L^1(\mathbb{R}, m)$ .
- (2)  $\mu_F = \lambda + F' m$  for some measure  $\lambda$  satisfying  $\lambda \perp m$ .
- (3)  $\mu_F \perp m$  if and only if  $F' = 0$  Lebesgue **almost everywhere**.
- (4)  $\mu_F \ll m$  if and only if  $\int_{-\infty}^x F'(t) dt = F(x) - F(-\infty) = F(x)$ .

**Proof.**

For (4), we have

$$\begin{aligned}
 \mu_F \ll m &\Leftrightarrow \lambda = 0 \\
 &\Leftrightarrow \mu_F = F' m \\
 &\Leftrightarrow \mu_F(E) = \int_E F' m \quad \forall \text{ Borel } E \\
 &\Leftrightarrow F(x) = \mu_F((-\infty, x]) = \int_{-\infty}^x F'(t) dt, \quad \forall x \in \mathbb{R}.
 \end{aligned}$$

The last converse comes from the uniqueness of [Theorem 6.5.2](#) above. ■

Check (1),  
(2), (3)

## Lecture 36: Absolutely Continuous Functions

### 6.5.1 Absolutely Continuous Functions

8 Apr. 11:00

We start with a definition.

**Definition 6.5.8** (Absolutely continuous function). We say that  $F: \mathbb{R} \rightarrow \mathbb{R}$  is *absolutely continuous*, denoted as  $F \in AC$ , if for all  $\epsilon > 0$ , there exists a  $\delta > 0$  such that if  $(a_1, b_1), \dots, (a_N, b_N)$  are finitely many disjoint open intervals satisfying  $\sum_{n=1}^N (b_n - a_n) < \delta$ , then

$$\sum_{n=1}^N |F(b_n) - F(a_n)| < \epsilon.$$

**Note.** Do not be confused between [Definition 6.5.8](#) and [Definition 6.2.1](#).

**Lemma 6.5.5.** We have the following.

- (1) If  $F$  is **absolutely continuous**, then it is uniformly continuous.
- (2) If  $F$  is **Lipschitz**, then  $F$  is **absolutely continuous**.
- (3) Given  $f \in L^1$ ,  $F(x) = \int_{-\infty}^x f(t) dt$  is **absolutely continuous**.

**Proof.** We prove this one by one.

- (1) We simply take  $N = 1$  in [Definition 6.5.8](#).
- (2) This is trivial.

(3) We write this out as

$$\sum_{n=1}^N |F(b_n) - F(a_n)| = \sum_{n=1}^N \left| \int_{a_n}^{b_n} f(t) dt \right| \leq \sum_{n=1}^N \int_{a_n}^{b_n} |f(t)| dt = \int_E |f(t)| dt$$

where  $E = \bigcup_{n=1}^N (a_n, b_n)$ , so  $m(E) = \sum_{n=1}^N (b_n - a_n)$ . From Midterm Q1, if  $f \in L^1(X, \mu)$ , for all  $\epsilon > 0$ , there is a  $\delta > 0$  such that  $\mu(E) < \delta$  implies  $\int_E |f| < \epsilon$ . This directly implies that this function is **absolutely continuous**. ■

**Example.** The **Cantor function**  $F$  is uniformly continuous. However, we will see that it is not **absolutely continuous**.

**Proposition 6.5.2.** Suppose  $F \in NBV$ , then  $F$  is **absolutely continuous** if and only if  $\mu_F \ll m$ .

**Proof.** We prove two directions.

( $\Leftarrow$ ) Suppose  $\mu_F \ll m$ . Then  $F(x) = \int_{-\infty}^x F'(t) dt$ , and  $F' \in L^1(\mathbb{R}, m)$ , by **Proposition 6.5.1**. Therefore,  $F \in AC$ .

( $\Rightarrow$ ) Now suppose  $F \in AC$ . Note that since  $F$  is continuous,

$$\mu_F((a, b)) = \lim_{n \rightarrow \infty} \mu_F((a, b - 1/n]) = \lim_{n \rightarrow \infty} F(b - 1/n) - F(a) = F(b) - F(a).$$

We let  $E$  be a **Borel set** with  $m(E) = 0$ . Fix  $\epsilon > 0$ , we will show  $|\mu_F(E)| \leq \epsilon$ . Let  $\delta > 0$  be the constant from  $F \in AC$ .

We know that there exists open  $U_1 \supseteq U_2 \supseteq \dots \supseteq E$  such that  $\lim_{n \rightarrow \infty} m(U_n) = m(E) = 0$ , and open  $V_1 \supseteq V_2 \supseteq \dots \supseteq E$  such that  $\lim_{n \rightarrow \infty} \mu_F(V_n) = \mu_F(E)$  by regularity.

Let  $O_n = U_n \cap V_n$ , then  $O_1 \supseteq O_2 \supseteq \dots \supseteq E$ , and by monotonicity (for  $\mu_F$  decomposing into pos/neg first)

$$\lim_{n \rightarrow \infty} m(O_n) = m(E) = 0, \quad \lim_{n \rightarrow \infty} \mu_F(O_n) = \mu_F(E).$$

Thus without loss of generality, we may assume  $m(O_1) < \delta$ . Each  $O_n$  is a countable union of disjoint intervals

$$O_n = \bigcup_{k=1}^{\infty} (a_k^n, b_k^n).$$

For any  $N$  we also have

$$\sum_{k=1}^N (b_k^n - a_k^n) \leq m(O_n) \leq m(O_1) < \delta.$$

Therefore,

$$\left| \mu_F \left( \bigcup_{k=1}^N (a_k^n, b_k^n) \right) \right| = \left| \sum_{k=1}^N \mu_F((a_k^n, b_k^n)) \right| \leq \sum_{k=1}^N |\mu_F((a_k^n, b_k^n))| \leq \sum_{k=1}^N |F(b_k^n) - F(a_k^n)| < \epsilon,$$

which implies

$$|\mu_F(O_n)| = \lim_{N \rightarrow \infty} \left| \mu_F \left( \bigcup_{k=1}^N (a_k^n, b_k^n) \right) \right| \leq \epsilon$$



and

$$|\mu_F(E)| = \lim_{n \rightarrow \infty} |\mu_F(O_n)| \leq \epsilon.$$

By taking  $\epsilon \rightarrow 0$  we have  $\mu_F(E) = 0$ , hence  $\mu_F \ll m$ . ■

**Corollary 6.5.3.**  $F \in NBV \cap AC$  if and only if  $F(x) = \int_{-\infty}^x f(t) dt$  for some  $f \in L^1(\mathbb{R}, m)$ . If this holds, we have  $f = F'$  **Lebesgue almost everywhere**.

**Lemma 6.5.6.** If  $F \in AC([a, b])$ , then  $F \in NBV([a, b])$ .

**Proof.** ■

DIY

**Theorem 6.5.3** (Fundamental Theorem of Calculus). For  $F \in [a, b] \rightarrow \mathbb{R}$ , the following are equivalent.

- (1)  $F \in AC([a, b])$ .
- (2)  $F(x) - F(a) = \int_a^x f(t) dt$  for some  $f \in L^1([a, b], m)$ .
- (3)  $F$  is differentiable **almost everywhere** on  $[a, b]$  and  $F(x) - F(a) = \int_a^x F'(t) dt$ .

**Proof.** This follows directly from ??.

Figure out the corresponding references.

**Definition.** Let  $\mu$  be a **finite signed Borel measure** on  $\mathbb{R}$ .

**Definition 6.5.9** (Discrete measure).  $\mu$  is called a *discrete measure* if there is a countable set  $\{x_n\}$  and  $c_n \neq 0$  such that  $\sum_{n=1}^{\infty} |c_n| < \infty$  and  $\mu = \sum_n c_n \delta_{x_n}$ .

**Note.**  $\delta_{x_n}$  is the **Dirac delta measure** at  $x_n$ .

**Definition 6.5.10** (Continuous measure).  $\mu$  is called a *continuous measure* if  $\mu(\{a\}) = 0$  for all  $a \in \mathbb{R}$ .

**Lemma 6.5.7.** Given a **finite signed Borel measure**  $\mu$ ,

- (1) Any  $\mu = \mu_d + \mu_c$ , where  $\mu_d$  is **discrete**,  $\mu_c$  is **continuous** are uniquely determined.
- (2)  $\mu$  **discrete** implies  $\mu \perp m$ .
- (3)  $\mu \ll m$  implies  $\mu$  is **continuous**.

**Proof.** We see that (2) and (3) are obvious, so we prove (1) only. Let  $E = \{x \in \mathbb{R} \mid \mu(\{x\}) \neq 0\}$ , then for any countable subset  $F$  of  $E$ , the series  $\sum_{x \in F} \mu(\{x\})$  converges absolutely to  $\mu(F)$ , so

$$\{x \in E \mid |\mu(\{x\})| > k^{-1}\}$$

is finite for all  $k$ , and it follows that  $E$  itself is countable. Hence,  $\mu_d(A) = \mu(A \cap E)$  is **discrete** and  $\mu_c(A) = \mu(A \setminus E)$  is **continuous**. ■

**Corollary 6.5.4.** For  $\mu$  a **finite signed Borel measure** on  $\mathbb{R}$ , we have that

$$\mu = \mu_d + \mu_{ac} + \mu_{sc}$$

where  $\mu_d$  is **discrete**,  $\mu_{ac}$  is **absolutely continuous** respect to  $m$ , and  $\mu_{sc}$  is **continuous** and **singularly** to  $m$ .

# Chapter 7

## Hilbert Spaces

### Lecture 37: Hilbert Spaces

#### 7.1 Inner Product Spaces

11 Apr. 11:00

**Definition 7.1.1 (Inner product).** Let  $V$  be a (complex) vector space. An *inner product* is a function

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$$

such that

- $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$  for all  $x, y, z \in V$ , and  $\alpha, \beta \in \mathbb{C}$ .
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$  for every  $x, y \in V$ .
- $\langle x, x \rangle \in [0, \infty)$ , and  $\langle x, x \rangle = 0$  if and only if  $x = 0$ .

**Note.** Note that we have conjugate linearity in the second argument, i.e.,

$$\langle x, \alpha y + \beta z \rangle = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle$$

for any  $x, y, z \in V$  and  $\alpha, \beta \in \mathbb{C}$ .

**Proof.** This follows from

$$\langle x, \alpha y + \beta z \rangle = \overline{\langle \alpha y + \beta z, x \rangle} = \overline{\alpha \langle y, x \rangle + \beta \langle z, x \rangle} = \overline{\alpha} \overline{\langle y, x \rangle} + \overline{\beta} \overline{\langle z, x \rangle} = \overline{\alpha} \langle x, y \rangle + \overline{\beta} \langle x, z \rangle.$$

⊛

**Example.** We have the following examples.

- $\mathbb{R}^d$  with  $\langle x, y \rangle = x \cdot y = \sum_{i=1}^d x_i y_i$ .
- $\mathbb{C}^d$  with  $\langle x, y \rangle = \sum_{i=1}^d x_i \overline{y_i}$ .
- $L^2(X, \mu)$  with  $\langle f, g \rangle = \int_X f \overline{g} d\mu$ . Note by [Theorem 5.2.1](#),

$$\left| \int_X f \overline{g} \right| \leq \|f \overline{g}\|_1 \leq \|f\|_2 \|g\|_2 < \infty$$

because  $1/2 + 1/2 = 1$ .

- A special case is  $\ell^2$ , where we have

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i}.$$

**Note.** Note that

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle = \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2$$

**Theorem 7.1.1 (Cauchy-Schwarz Inequality).** Given an **inner product** space,  $|\langle x, y \rangle| \leq \|x\| \|y\|$ .

**Proof.** This is clear if  $\langle x, y \rangle = 0$ . Assume  $\langle x, y \rangle \neq 0$ , then for every  $\alpha \in \mathbb{C}$ , we know that

$$0 \leq \|\alpha x - y\|^2 = |\alpha|^2 \|x\|^2 - 2 \operatorname{Re} \alpha \langle x, y \rangle + \|y\|^2.$$

Write  $\langle x, y \rangle = |\langle x, y \rangle| e^{i\theta}$ , and take  $\alpha = e^{-i\theta} t$  for arbitrary  $t \in \mathbb{R}$ . Then, the right-hand side gives

$$0 \leq \|x\|^2 t^2 - 2 |\langle x, y \rangle| t + \|y\|^2.$$

Note this is a real quadratic function of  $t$ , with at most one real root. Thus, the discriminant  $\Delta \leq 0$ . Specifically, we have

$$\Delta = 4 |\langle x, y \rangle|^2 - 4 \|x\|^2 \|y\|^2 \leq 0 \Leftrightarrow |\langle x, y \rangle|^2 \leq \|x\|^2 \|y\|^2 \Leftrightarrow |\langle x, y \rangle| \leq \|x\| \|y\|.$$

■

**Definition 7.1.2 (Induced norm from inner product).** Given an **inner product** space  $V$ , let

$$\|x\| := \sqrt{\langle x, x \rangle},$$

which is so-call the *norm induced from the inner product*.

**Proof.** We need to check that this actually defines a **norm**. We check the following.

**Claim.**  $\|x\| = 0 \Leftrightarrow x = 0$  for all  $x \in V$ .

**Proof.** This follows from the **definition of an inner product**. ⊗

**Claim.**  $\|\alpha x\| = |\alpha| \|x\|$  for all  $x \in V$ .

**Proof.** This follows from

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{\alpha \overline{\alpha} \langle x, x \rangle} = |\alpha| \|x\|.$$

⊗

**Claim (Triangle inequality).**  $\|x + y\| \leq \|x\| + \|y\|$  for all  $x, y \in V$ .

**Proof.** The triangle inequality is less obvious, and comes from [Theorem 7.1.1](#). Namely,

$$\begin{aligned}\|x + y\|^2 &= \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2 |\langle x, y \rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2\end{aligned}$$

Taking square root on both sides, we have

$$\|x + y\| \leq \|x\| + \|y\|.$$

⊗

■

**Theorem 7.1.2** (Parallelogram law). Let  $V$  be a [normed vector space](#). Then,  $\|\cdot\|$  is [induced by an inner product](#) if and only if

$$\|x + y\|^2 + \|x - y\|^2 = 2 \|x\|^2 + 2 \|y\|^2$$

for all  $x, y \in V$ .

**Proof.** We show two directions.

( $\Rightarrow$ ) This follows from

$$\|x \pm y\|^2 = \|x\|^2 \pm 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2$$

and

$$\|x \pm iy\|^2 = \|x\|^2 \pm 2 \operatorname{Im} \langle x, y \rangle + \|y\|^2.$$

( $\Leftarrow$ ) Firstly, we define

$$\langle x, y \rangle = \frac{1}{4} \left( \|x + y\|^2 - \|x - y\|^2 + \|x + iy\|^2 - i \|x - iy\|^2 \right)$$

as motivated by the above relationship.

**Exercise.** Check this [inner product](#) is indeed inducing the desired norm.

■

**Example.** Consider  $L^p(\mathbb{R}, m)$ ,  $f = \mathbb{1}_{(0,1)}$ ,  $g = \mathbb{1}_{(1,2)}$ . We see the [parallelogram law](#) is satisfied only when  $p = 2$ .

**Remark.** Hence,  $L^p(\mathbb{R}, m)$  is only an [inner product](#) space when  $p = 2$ .

Since we're doing real analysis, we want to deal with limits. It turns out that with an [inner product](#) space, we can say something more compare to the case of a [normed](#) vector space. We now illustrate this.

**Definition.** Given a vector space  $V$  with either a [norm](#) or an [inner product](#), we have the followings.

**Definition 7.1.3 (Strong convergence).** We say that  $x_n \in V$  converges to  $x \in V$  strongly if

$$\lim_{n \rightarrow \infty} \|x_n - x\| = 0.$$

**Definition 7.1.4 (Weak convergence).** We say that  $x_n \in V$  converges to  $x \in V$  weakly if for any fixed  $y \in V$ ,

$$\lim_{n \rightarrow \infty} \langle x_n - x, y \rangle = 0.$$

**Lemma 7.1.1 (Strong convergence implies weak convergence).** Suppose  $V$  is an inner product space. If  $x_n \rightarrow x$  strongly, then  $x_n \rightarrow x$  weakly.

**Proof.** By Cauchy-Schwarz inequality,

$$0 \leq |\langle x_n - x, y \rangle| \leq \|x_n - x\| \cdot \|y\|.$$

Since  $\|x_n - x\| \rightarrow 0$  and  $\|y\|$  is constant in  $n$ , from the Squeeze theorem, we have

$$\langle x_n - x, y \rangle \rightarrow 0$$

as  $n \rightarrow \infty$ . ■

**Example.** Consider  $\ell^2$ ,  $x_n = (0, \dots, 0, 1, 0, \dots)$  and  $x = 0$ . Then  $x_n$  does not converge strongly to any vector, but it does converge to 0 weakly.

**Proof.** If we fix  $y \in \ell^2$ , then

$$\langle x_n - x, y \rangle = \bar{y}_n$$

which goes to 0 as  $n \rightarrow \infty$  because  $\sum_n |y_n|^2 < \infty$ . Therefore,  $x_n \rightarrow 0$  weakly, but we see that

$$\|x_n - 0\| = \|x_n\| = 1.$$

Thus,  $x_n \not\rightarrow 0$  strongly. ⊛

## 7.2 Orthonormal Bases

**Definition 7.2.1 (Orthogonal).** Two vectors  $x, y$  are orthogonal if  $\langle x, y \rangle = 0$ , denoted as  $x \perp y$ .

**Remark.** Do not confuse between this notation and Definition 6.1.5.

**Lemma 7.2.1 (Pythagorean Theorem).** If  $x_1, \dots, x_n \in V$ ,  $\langle x_i, x_j \rangle = 0$  for all  $i \neq j$ , then

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2.$$

**Proof.** Use that  $\|x + y\|^2 = \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2$  and induct. ■

**Definition 7.2.2 (Orthonormal set).** We call  $\{e_i\}_{i \in I}$  an orthonormal set if

$$\langle e_i, e_j \rangle = \begin{cases} 0, & \text{if } i \neq j; \\ 1, & \text{if } i = j. \end{cases}$$

## Lecture 38: Orthonormal Basis

Let's start with a lemma.

13 Apr. 11:00

**Lemma 7.2.2** (Best approximation). Let  $e_1, \dots, e_N$  be [orthonormal](#), then for  $x \in V$ , let  $\alpha_i = \langle x, e_i \rangle$ , we have

$$\left\| x - \sum_{i=1}^N \alpha_i e_i \right\| \leq \left\| x - \sum_{i=1}^N \beta_i e_i \right\|$$

for all  $\beta_1, \dots, \beta_N \in \mathbb{C}$ . Namely,  $\sum_{i=1}^N \alpha_i e_i$  is the *best approximation* to  $x$  within the span of  $e_1, \dots, e_N$ .

**Note.** We can also think of it as an [orthogonal](#) projection.

**Proof.** Let  $z = x - \sum_{i=1}^N \alpha_i e_i$ ,  $w = \sum_{i=1}^N (\alpha_i - \beta_i) e_i$  for any  $\{\beta_i\}$ . Note that for all  $n = 1, \dots, N$ , we have

$$\langle z, e_n \rangle = \langle x, e_n \rangle - \alpha_n = 0.$$

Thus,  $\langle z, w \rangle = 0$ . So by the [Pythagorean theorem](#),

$$\|z + w\|^2 = \|z\|^2 + \|w\|^2 \geq \|z\|^2,$$

which proves the result. ■

**Lemma 7.2.3** (Bessel's inequality). Let  $\{e_i\}_1^\infty$  be an [orthonormal set](#). For  $x \in V$ , let  $\alpha_i = \langle x, e_i \rangle$ . Then we have

$$\sum_{i=1}^{\infty} |\alpha_i|^2 \leq \|x\|^2,$$

which is so-called the *Bessel's inequality*. More generally, we have

$$\|x\|^2 = \left\| x - \sum_{i=1}^N \alpha_i e_i \right\|^2 + \sum_{i=1}^N |\alpha_i|^2$$

for all  $N \in \mathbb{N}$ .

**Proof.** We see that

$$\begin{aligned} \left\| x - \sum_{i=1}^N \alpha_i e_i \right\|^2 &= \|x\|^2 - 2 \operatorname{Re} \left\langle x, \sum_{i=1}^N \alpha_i e_i \right\rangle + \left\| \sum_{i=1}^N \alpha_i e_i \right\|^2 \\ &= \|x\|^2 - 2 \sum_{i=1}^N \operatorname{Re} \overline{\alpha_i} \langle x, e_i \rangle + \sum_{i=1}^N |\alpha_i|^2 \\ &= \|x\|^2 - 2 \sum_{i=1}^N |\alpha_i|^2 + \sum_{i=1}^N |\alpha_i|^2 \\ &= \|x\|^2 - \sum_{i=1}^N |\alpha_i|^2. \end{aligned}$$

From this, we see that the inequality holds trivially. ■

**Remark.** [Lemma 7.2.3](#) actually holds even for an uncountable collection.

**Definition 7.2.3** (Orthonormal basis). An [orthonormal set](#)  $\{e_i\}$  is said to be an *orthonormal basis* of

$V$  if  $\overline{W} = V$ , where

$$W = \left\{ \sum_{i=1}^N \beta_i e_i \mid N \in \mathbb{N}, \beta_1, \dots, \beta_N \in \mathbb{C} \right\}$$

is the subspace of finite linear combinations. In other words, for all  $x \in V$  and for every  $\epsilon > 0$ , there exists  $w \in W$  such that  $\|x - w\| < \epsilon$ .

**Example.** We have the following.

- (1) For  $\mathbb{C}^d$ , the orthonormal basis is  $e_i = (0, \dots, 0, 1, 0, \dots, 0)$  for  $i = 1, \dots, d$ .
- (2) For  $\ell^2$ , the orthonormal basis is the countably many  $e_i = (0, \dots, 0, 1, 0, \dots)$  for  $i \in \mathbb{N}$ .

**Definition 7.2.4 (Hilbert Space).** A Hilbert space is a complete inner product space.

**Note.** Namely, a Hilbert space is a Banach space with an inner product.

**Example.**  $\mathbb{R}^d$ ,  $\mathbb{C}^d$ ,  $L^2(X, \mathcal{A}, \mu)$ ,  $\ell^2$  are all Hilbert spaces.

**Example.**  $C([0, 1]) \subseteq L^2(X, \mathcal{A}, \mu)$  is not a Hilbert space.

**Proof.** We see this by the fact that it is not complete. Take a function  $f_n$  so that  $f_n$  is zero from 0 to  $1/2$  and 1 from  $1/2 + 1/n$  to 1, connected continuously by a line. We then see that  $f_n$  is Cauchy, but its natural limit is discontinuous.  $\otimes$

**Theorem 7.2.1.** Let  $\mathcal{H}$  be a Hilbert space. Let  $\{e_i\}_{i=1}^\infty$  be an orthonormal set. The following are equivalent.

- (1)  $\{e_i\}_{i=1}^\infty$  is an orthonormal basis.
- (2) If  $x \in \mathcal{H}$  and  $\langle x, e_i \rangle = 0$  for all  $i$ , then  $x = 0$ .
- (3) If  $x \in \mathcal{H}$ , then  $s_N := \sum_{i=1}^N \alpha_i e_i \rightarrow x$  strongly where  $\alpha_i = \langle x, e_i \rangle$ .
- (4) If  $x \in \mathcal{H}$ , then  $\|x\|^2 = \sum_{i=1}^\infty |\alpha_i|^2$ .

**Proof.** We prove this as follows.

**Claim.** (3)  $\Rightarrow$  (4).

**Proof.** From Lemma 7.2.3, we have

$$\|x\|^2 = \|x - s_N\|^2 + \sum_{i=1}^N |\alpha_i|^2.$$

Taking  $N \rightarrow \infty$  and noting  $s_N \rightarrow x$  strongly gives

$$\|x\|^2 = \lim_{N \rightarrow \infty} \sum_{i=1}^N |\alpha_i|^2 = \sum_{i=1}^\infty |\alpha_i|^2.$$

$\otimes$

**Claim.** (4)  $\Rightarrow$  (1).

**Proof.** Using the same equality from [Lemma 7.2.3](#), we have  $\|x\|^2 = \|x - s_N\|^2 + \sum_{i=1}^N |\alpha_i|^2$ . Now, take  $N \rightarrow \infty$  yields  $\|x - s_N\|^2 \rightarrow 0$  so  $\|x - s_N\| \rightarrow 0$ . Therefore,  $s_N \rightarrow x$  strongly, yielding that  $x$  can be approximated by finite linear combinations as desired.  $\otimes$

**Claim.** (1)  $\Rightarrow$  (2).

**Proof.** Fix  $x \in \mathcal{H}$ , and fix  $\epsilon > 0$ . Then by (1), there exists a  $y = \sum_{i=1}^k \beta_i e_i$  such that  $\|x - y\| < \epsilon$ .

By the [best approximation lemma](#),  $\|x - s_k\| \leq \|x - y\| < \epsilon$ . If  $\langle x, e_i \rangle = 0$  for all  $i$ , then  $s_k = 0$ , so  $\|x\| < \epsilon$ . Taking  $\epsilon \rightarrow 0$  would yield  $\|x\| = 0$ , implying  $x = 0$ .  $\otimes$

**Claim.** (2)  $\Rightarrow$  (3).

**Proof.** [Bessel's inequality](#) gives

$$\sum_{i=1}^{\infty} |\alpha_i|^2 \leq \|x\|^2 < \infty.$$

We now see that for  $N > M$ ,

$$\|s_N - s_M\|^2 = \left\| x - \sum_{i=M+1}^N \alpha_i e_i \right\|^2 = \sum_{i=M+1}^N |\alpha_i|^2 \rightarrow 0$$

as  $N > M \rightarrow \infty$ , by convergence of the series. This implies that  $\{s_N\}_{N=1}^{\infty}$  is a [Cauchy sequence](#) in  $\mathcal{H}$ . Since  $\mathcal{H}$  is [complete](#), there is a vector  $y$  such that  $s_N \rightarrow y$  strongly.

**Problem.** Is  $y = x$ ?

**Answer.** Fix  $i \in \mathbb{N}$ , consider  $\langle y - x, e_i \rangle$ . We see that

$$\langle y - x, e_i \rangle = \langle y - s_N, e_i \rangle + \langle s_N - x, e_i \rangle.$$

We can compute that for  $N > i$ , that

$$\langle s_N - x, e_i \rangle = \alpha_i - \langle x, e_i \rangle = 0.$$

Hence,  $\langle y - x, e_i \rangle = \langle y - s_N, e_i \rangle$ . Because [strong convergence implies weak convergence](#), taking  $N \rightarrow \infty$  yields that  $\langle y - x, e_i \rangle = 0$  for all  $i \in \mathbb{N}$ . Therefore, by the assumption of (2),  $y - x = 0$ , so  $x = y$  and we're done.  $\otimes$

$\otimes$

**Remark.** Note that we did not use the [Hilbert space](#) properties for everything except (2)  $\Rightarrow$  (3). When  $\mathcal{H}$  is replaced by any [inner product](#) space  $V$  we only have

$$(3) \Rightarrow (4) \Rightarrow (1) \Rightarrow (2).$$

**Note** (Planchenel identity). The (4) in [Theorem 7.2.1](#) is called the *Planchenel identity*.

**Definition 7.2.5** (Separable). A [metric](#) space is called *separable* if there exists a countable dense subset.



**Example.**  $\mathbb{R}^d \supseteq \mathbb{Q}^d$  and  $\ell^p$  with  $1 \leq p < \infty$  are all **separable**, but not for  $p = \infty$ .

**Proof.** To do this, consider sequences of rational numbers. ⊛

**Example.**  $L^p(\mathbb{R}, m)$  is **separable** for  $1 \leq p < \infty$ .

**Proof.** Take **step functions** with rational heights and rational endpoints to intervals. ⊛

**Theorem 7.2.2.** Every **separable Hilbert space** has a countable **orthonormal basis**.

**Proof.** Gram-Schmidt algorithm<sup>a</sup> will construct such an **orthonormal basis** explicitly. ■

<sup>a</sup>[https://en.wikipedia.org/wiki/Gram-Schmidt\\_process](https://en.wikipedia.org/wiki/Gram-Schmidt_process)

**Note** (Dimension of a Hilbert space). The cardinality of an **orthonormal basis** is determined by the space, and we can call this the *dimension* of a **Hilbert space**.

## Lecture 39: Introduction to Fourier Analysis

**As previously seen.** **Planchenel identity** says that if  $e_i$  is an **orthonormal basis**, then

15 Apr. 11:00

$$\langle x, y \rangle = \sum_{i=1}^{\infty} \alpha_i \overline{\beta_i}$$

where  $\alpha_n = \langle x, e_n \rangle$ ,  $\beta_n = \langle y, e_n \rangle$ .

# Chapter 8

## Introduction to Fourier Analysis

### 8.1 Fourier Series

We'll be considering the [Hilbert space](#)  $L^2([-\pi, \pi])$ , which by scaling is equivalent to any finite interval. Let's start with a lemma.

**Lemma 8.1.1.** The set

$$\left\{ e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx} = \frac{1}{\sqrt{2\pi}} (\cos(nx) + i \sin(nx)) \right\}$$

is an [orthonormal set](#) in  $L^2([-\pi, \pi])$ .

**Proof.** We see that

$$\langle e_n, e_m \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(m-n)x} dx = \begin{cases} 1, & \text{if } m = n; \\ 0, & \text{if } m \neq n, \end{cases}$$

hence, the result follows directly. ■

Now, a natural problem arises:

**Problem.** Is  $\{e_n\}$  defined in [Lemma 8.1.1](#) an [orthonormal basis](#)?

**Answer.** By [Hölder's inequality](#) on  $L^2([-\pi, \pi])$ , we have

$$\|f\|_1 \leq \|1\|_2 \|f\|_2 = \sqrt{2\pi} \cdot \|f\|_2 < \infty.$$

Similarly, we have

$$\|f\|_2 = \sqrt{\int_{-\pi}^{\pi} |f(t)|^2 dt} \leq \sqrt{\|f\|_{\infty}^2 2\pi},$$

which implies

$$\|f\|_1 \leq \sqrt{2\pi} \|f\|_2 \leq 2\pi \|f\|_{\infty}.$$

⊛

**Definition 8.1.1** (Fourier coefficient). For  $f \in L^1([-\pi, \pi])$ , its  $n^{\text{th}}$  Fourier coefficient is

$$\hat{f}_n := \langle f, e_n \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(y) e^{-iny} dy.$$

We can show that [Definition 8.1.1](#) form an [orthonormal basis](#), we show that

$$\sum_{n=-M}^N \hat{f}_n e_n(x)$$

converge strongly to  $f(x)$  as  $M, N \rightarrow \infty$  Explicitly, this is true since

$$\sum_{n=-M}^N \hat{f}_n e_n(x) = \frac{1}{\sqrt{2\pi}} \sum_{n=-M}^N \left[ \int_{-\pi}^{\pi} f(y) e^{-iny} dy \right] e^{inx} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(y) \left( \sum_{n=-M}^N e^{in(x-y)} \right) dy.$$

**Definition 8.1.2** (Poisson kernel). For  $0 \leq r < 1$ , the *Poisson kernel* is defined as

$$P_r(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{int} r^{|n|}.$$

Explicitly, summing as a geometric series gives us

$$P_r(t) = \frac{1 - r^2}{2\pi(1 - 2r \cos t + r^2)}.$$

**Lemma 8.1.2.** For  $f \in L^1([-\pi, \pi])$  and  $0 \leq r < 1$ , for  $x \in [-\pi, \pi]$ ,

$$\sum_{n=-\infty}^{\infty} \hat{f}_n e_n(x) r^{|n|}$$

converge absolutely and uniformly to

$$\int_{-\pi}^{\pi} P_r(x - y) f(y) dy.$$

**Proof.** We have

$$\sum_{n=-\infty}^{\infty} \left| \hat{f}_n e_n(x) r^{|n|} \right| \leq \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left( \int_{-\pi}^{\pi} |f(y) e^{-iny}| dy \right) |e_n(x)| r^{|n|} = \frac{\|f\|_1}{2\pi} \sum_{n=-\infty}^{\infty} r^{|n|} < \infty.$$

Therefore, [Corollary 2.3.1](#) applies, and we further have

$$\begin{aligned} \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \left( \int_{-\pi}^{\pi} f(y) e^{-iny} dy \right) e^{inx} r^{|n|} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left( f(y) \sum_{n=-\infty}^{\infty} e^{in(x-y)} r^{|n|} \right) dy \\ &= \int_{-\pi}^{\pi} P_r(x - y) f(y) dy. \end{aligned}$$

Uniform convergence takes a small bit more work. ■

DIY

**Note.** Note that

$$P_r(0) = \frac{1 - r^2}{2\pi(1 - r)^2} = \frac{1 + r}{2\pi(1 - r)} \rightarrow 0 \text{ as } r \rightarrow 1.$$

For any  $t \neq 0$  we have

$$\frac{1 - r^2}{1\pi(1 - 2r \cos t + r^2)} \rightarrow 0$$

as the bottom is always nonzero and finite.

**Definition 8.1.3** (Good kernel). A kernel  $K_r$  is said to be in a *family<sup>a</sup> of good kernels* if

- (1)  $K_r(t) \geq 0$ .
- (2)  $\int_{-\pi}^{\pi} K_r(t) dt = 1$ .

(3) For every  $\delta > 0$ ,

$$\int_{[-\pi, \pi] \setminus [-\delta, \delta]} K_r(t) dt \rightarrow 0$$

as  $r \rightarrow 1$ .<sup>b</sup>

<sup>a</sup>Respect to  $r$ .

<sup>b</sup>This definition is not universal. Some text uses other convention.

**Lemma 8.1.3.**  $P_r(t)$  form a family of good kernels.

**Proof.** For (2), we use the first formula with Fubini's theorem. For (1) and (3), we use the second formula, namely

$$\int_{[-\pi, \pi] \setminus [-\delta, \delta]} P_r(t) dt \leq \frac{1 - r^2}{2\pi(1 - 2r \cos \delta + r^2)} 2\pi \rightarrow 0$$

as  $r \rightarrow 1$ . ■

**Lemma 8.1.4.** For  $f \in C([-\pi, \pi])$  satisfying  $f(-\pi) = f(\pi)$ , we have

$$\lim_{r \nearrow 1} \int_{-\pi}^{\pi} P_r(x - y) f(y) dy = f(x)$$

uniformly for  $x \in [-\pi, \pi]$ .

**Proof.** We extend  $f$  to  $f: \mathbb{R} \rightarrow \mathbb{C}$  by setting  $f(x + 2\pi) = f(x)$ , then  $f$  is uniformly continuous and bounded. Also, we have

$$\begin{aligned} & \int_{-\pi}^{\pi} P_r(x - y) f(y) dy - f(x) \\ &= \int_{-\pi}^{\pi} P_r f(x - y) dy - f(x) \\ &= \int_{-\pi}^{\pi} P_r(y) f(x - y) dy - f(x) \int_{-\pi}^{\pi} P_r(y) dy \\ &= \int_{-\delta}^{\delta} P_r(y) (f(x - y) - f(x)) dy + \int_{[-\pi, \pi] \setminus [-\delta, \delta]} P_r(y) (f(x - y) - f(x)) dy. \end{aligned}$$

Now, fix  $\epsilon > 0$ , then  $f$  is uniformly continuous, so we can choose a  $\delta > 0$  so that  $|f(x - y) - f(x)| < \epsilon$  for any choice of  $x$ . Then, the left-hand side is bounded by  $\epsilon$ . As for the right-hand side, we note that  $f$  is bounded, so for some  $M$ , we have

$$\left| \int_{[-\pi, \pi] \setminus [-\delta, \delta]} P_r(y) (f(x - y) - f(x)) dy \right| \leq M \int_{[-\pi, \pi] \setminus [-\delta, \delta]} P_r(y) dy.$$

Therefore, sending  $r \nearrow 1$  will give

$$\int_{-\pi}^{\pi} P_r(x - y) f(y) dy - f(x) \rightarrow 0,$$

as we desired. ■

## Lecture 40: Fourier Analysis

18 Apr. 11:00

**As previously seen.** We have from last time that for  $0 \leq r < 1$ ,

$$\begin{aligned} e_n(x) &:= \frac{1}{\sqrt{2\pi}} e^{inx} \\ P_r(t) &:= \sum_{n=-\infty}^{\infty} e_n(t) r^{|n|} \\ f \in L^1([-\pi, \pi]) &\Rightarrow \sum_{n=-\infty}^{\infty} \hat{f}_n e_n(x) r^{|n|} = \int_{-\pi}^{\pi} P_r(x-y) f(y) dy \end{aligned}$$

uniformly in  $x \in [-\pi, \pi]$ . Furthermore, if  $f \in C([-\pi, \pi])$  and  $f(\pi) = f(-\pi)$  then

$$\lim_{r \rightarrow 1} \int_{-\pi}^{\pi} P_r(x-y) f(y) dy = f(x)$$

uniformly in  $x \in [-\pi, \pi]$ .

Finally, we also have

$$\|f\|_1 \leq \sqrt{2\pi} \|f\|_2 \leq 2\pi \|f\|_{\infty}.$$

**Theorem 8.1.1 (Fourier Series).** The set of  $e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}$  for  $n \in \mathbb{Z}$  is an **orthonormal basis** of  $L^2([-\pi, \pi])$ .

**Proof.** From **Lemma 8.1.1**, this is an **orthonormal set**. Thus, we just need to show that for all  $f \in L^2([-\pi, \pi])$ , that there exists  $h = \sum_{n=-M}^N \beta_n e_n(x)$  such that  $\|f - h\|_2 < \epsilon$ . That is we need to show that the span of the  $e_n$  is dense.

Let  $f \in L^2([-\pi, \pi])$  and fix an  $\epsilon > 0$ . Then there is a function  $g \in C([-\pi, \pi])$  with  $g(\pi) = g(-\pi)$  such that  $\|f - g\|_2 < \frac{\epsilon}{3}$  which follows from a simple density argument.

Let  $g_r(x) = \int_{-\pi}^{\pi} P_r(x-y) g(y) dy$ . By the above, there exists an  $r \in [0, 1)$  such that  $\|g_r - g\|_{\infty} < \frac{\epsilon}{3\sqrt{2\pi}}$ .

Therefore,  $\|g_r - g\|_2 < \frac{\epsilon}{3}$ . Consider  $g_{r,N}(x) = \sum_{n=-N}^N \hat{g}_n e_n(x) r^{|n|}$ . By the above there exists an  $N \in \mathbb{N}$  such that  $\|g_{r,N} - g_r\|_{\infty} < \frac{\epsilon}{3\sqrt{2\pi}}$ , thus  $\|g_{r,N} - g_r\|_2 < \frac{\epsilon}{3}$ . Therefore,

$$\|f - g_{r,N}\| < \epsilon$$

where  $g_{r,N}$  is a finite linear combination of  $e_n$ 's, so these form an **orthonormal basis** as desired. ■

**Example (Plancherel's identity).** We have  $\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\hat{f}_n|^2$ .

**Proof.** For  $f(x) = x$ , we have

$$\hat{f}_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} x e^{-inx} dx = \begin{cases} 0, & \text{if } n = 0; \\ \frac{(-1)^n i \sqrt{2\pi}}{n}, & \text{if } n \geq 1. \end{cases}$$

Together these imply that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

⊛

**Example (Isoperimetric inequality).** Consider a parametric curve  $(x(t), y(t)) \in \mathbb{R}^2$  with  $t \in [-\pi, \pi]$ . Assume that

- (1) This is a closed curve, so  $(x(-\pi), y(-\pi)) = (x(\pi), y(\pi))$ .

(2) We assume these are smooth, but in fact we just need  $x, y$  are  $C^1$  functions.

(3) The curve is simple.

Suppose that

$$L = \int_{-\pi}^{\pi} \sqrt{x'(t)^2 + y'(t)^2} dt = 2\pi,$$

then the largest area  $A$  this curve can enclose is  $\pi$ .

**Proof.** By Green's Theorem<sup>a</sup>

$$A = \frac{1}{2} \oint_C (x dy - y dx) = \frac{1}{2} \int_{-\pi}^{\pi} (x(t)y'(t) - x'(t)y(t)) dt.$$

Arc length parametrization, so that  $x'(t)^2 + y'(t)^2 = 1$  for all  $t$ . Then the condition  $L = 2\pi$  is automatically satisfied and can be written as

$$\int_{-\pi}^{\pi} (x'(t)^2 + y'(t)^2) dt = 2\pi.$$

For ease of computation, rewrite using complex numbers, i.e.,  $z(t) = x(t) + iy(t)$ . This is then subject to

$$\int_{-\pi}^{\pi} |z'(t)|^2 dt = 2\pi.$$

Rewriting our above formula for area, we need to find the maximum of

$$A = \frac{1}{4i} \int_{-\pi}^{\pi} (\overline{z(t)}z'(t) - z(t)\overline{z'(t)}) dt.$$

Note  $z$  is  $C^1$  and  $z(\pi) = z(-\pi)$ . Denote  $\hat{z}_n = \alpha_n$ , then we now have

$$\hat{z}'_n = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} z'(t)e^{-int} dt = (z(t)e^{-int})_{\pi}^{\pi} + in \int_{-\pi}^{\pi} z(t)e^{-int} dt = in\alpha_n.$$

By [Plancherel's identity](#), we have that

$$\int_{-\pi}^{\pi} |z'(t)|^2 dt = \sum_{n=-\infty}^{\infty} |in\alpha_n|^2 = \sum_{n=-\infty}^{\infty} n^2 |\alpha_n|^2 = 2\pi.$$

We know [Plancherel's identity](#) says  $\langle f, g \rangle = \sum_{n=-\infty}^{\infty} \hat{f}_n \overline{\hat{g}_n}$ . Therefore,

$$A = \frac{1}{4i} \sum_{n=-\infty}^{\infty} \overline{\alpha_n}(in\alpha_n) - \alpha_n \overline{(in\alpha_n)} = \frac{1}{2} \sum_{n=-\infty}^{\infty} n |\alpha_n|^2.$$

Then the question becomes what is the maximum of  $\frac{1}{2} \sum_{n=-\infty}^{\infty} n |\alpha_n|^2$  subject to  $\sum_{n=-\infty}^{\infty} n^2 |\alpha_n|^2 = 2\pi$ .

We will show that

$$2\pi - \sum_{n=-\infty}^{\infty} n |\alpha_n|^2 \geq 0$$

as we guess the maximum should be  $\pi$ , by virtue of the circle being theoretically optimal. Then we have

$$2\pi - \sum_{n=-\infty}^{\infty} n |\alpha_n|^2 = \sum_{n=-\infty}^{\infty} (n^2 - n) |\alpha_n|^2 \geq 0$$

because term by term every term is non-negative.

---

Thus, the area cannot be more than  $\pi$ . We have  $A = \pi$  if and only if equality holds above, that is  $\alpha_n = 0$  for  $n \neq 0, 1$ . This means  $z(t) = \alpha_0 + \alpha_1 e_1(x)$ , that is

$$z(t) = \alpha_0 + C e^{it}.$$

This means that  $z(t)$  is a circle about  $\alpha_0$ . ⊛

---

<sup>a</sup>[https://en.wikipedia.org/wiki/Green's\\_theorem](https://en.wikipedia.org/wiki/Green's_theorem)

**Note.** Professor Baik likes Dym and McKean's Fourier Series & Integrals[DM72]!

# Appendix



# Appendix A

## Additional Proofs

### A.1 Measure

This section gives all additional proofs in [chapter 1](#).

**Theorem A.1.1** ([Theorem 1.3.2 3.](#)). Under the setup of [Theorem 1.3.2](#),  $(X, \mathcal{A}, \mu)$  is a [complete measure space](#).

**Proof.** We see this in two parts.

**Claim.** If  $A \subset X$  satisfies  $\mu^*(A) = 0$ , then  $A$  is [Carathéodory measurable](#) with respect to  $\mu^*$ .

**Proof.** If  $A \subset X$  and  $\mu^*(A) = 0$ , where  $\mu^*$  is an outer measure on  $X$ , we'll show that  $A$  is [Carathéodory measurable](#) with respect to  $\mu^*$ .

Equivalently, we want to show that for any  $E \subset X$ ,

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A).$$

Firstly, noting that  $(E \cap A) \subset A$ , and by [monotonicity](#) of  $\mu^*$ , we see that

$$\mu^*(E \cap A) \leq \mu^*(A) = 0,$$

and since  $\mu^* \geq 0$ , hence  $\mu^*(E \cap A) = 0$ . Now, we only need to show that

$$\mu^*(E) = \mu^*(E \setminus A).$$

Since  $E \setminus A = E \cap A^c$ , and hence we have  $E \cap A^c \subset E$ , so

$$\mu^*(E) \geq \mu^*(E \setminus A).$$

To show another direction, we note that

$$\mu^*(E) \leq \mu^*(E \cup A) = \mu^*((E \setminus A) \cup A) \leq \mu^*(E \setminus A),$$

hence we conclude that  $A$  is [Carathéodory measurable](#) with respect to  $\mu^*$  if  $\mu^*(A) = 0$ .  $\otimes$

**Claim.** If  $A$  is [μ-subnull](#), then  $A \in \mathcal{A}$ .

**Proof.** Let  $\mathcal{A}$  denotes the [Carathéodory  \$\sigma\$ -algebra](#), and  $\mu := \mu^*|_{\mathcal{A}}$ . We want to show if  $A$  is  [\$\mu\$ -subnull](#), then  $A \in \mathcal{A}$ .

Firstly, if  $A$  is  [\$\mu\$ -subnull](#), then there exists a  $B \in \mathcal{A}$  such that  $A \subset B$  and  $\mu(B) = 0$ . But since from the [monotonicity](#) of  $\mu^*$ , we further have

$$0 = \mu(B) = \mu^*(B) \geq \mu^*(A),$$

hence  $\mu^*(A) = 0$ .

From the first claim, we immediately see that  $A$  is [Carathéodory measurable](#) with respect to  $\mu^*$ , which implies  $A$  is in [Carathéodory  \$\sigma\$ -algebra](#), hence  $A \in \mathcal{A}$ . ⊗

We see that the second claim directly proves that  $(X, \mathcal{A}, \mu)$  is a [complete measure space](#). ■

**Lemma A.1.1.** The function  $F$  defined in [this example](#) is a [distribution function](#)

**Proof.** We define

$$F_n(x) = \begin{cases} 1, & \text{if } x \geq r_n; \\ 0, & \text{if } x < r_n \end{cases}$$

where  $\{r_1, r_2, \dots\} = \mathbb{Q}$ , and

$$F(x) = \sum_{n=1}^{\infty} \frac{F_n(x)}{2^n} = \sum_{n; r_n \leq x} \frac{1}{2^n}$$

is both increasing and right-continuous.

- Increasing. Consider  $x < y$ . We see that

$$F(x) = \sum_{n; r_n \leq x} \frac{1}{2^n} \leq \sum_{n; r_n \leq y} \frac{1}{2^n} = F(y)$$

clearly.<sup>a</sup>

- Right-continuous. We want to show  $F(x^+) = F(x)$ . Let  $x^+(\epsilon) := x + \epsilon$  with  $\epsilon > 0$ , we'll show that

$$\lim_{\epsilon \rightarrow 0} F(x^+(\epsilon)) = \lim_{\epsilon \rightarrow 0} F(x + \epsilon) = F(x).$$

Firstly, we have

$$F(x^+(\epsilon)) - F(x) = \sum_{n; r_n \leq x+\epsilon} \frac{1}{2^n} - \sum_{n; r_n \leq x} \frac{1}{2^n} = \sum_{n; x < r_n \leq x+\epsilon} \frac{1}{2^n},$$

and we want to show

$$\lim_{\epsilon \rightarrow 0} F(x^+(\epsilon)) - F(x) = \lim_{\epsilon \rightarrow 0} \sum_{n; x < r_n \leq x+\epsilon} \frac{1}{2^n} = 0.$$

**Remark.** The strict is crucial to show the result, since if  $x = r_k$  for some fixed  $k$ , then we can't make the summation arbitrarily small.

<sup>a</sup>This is trivial since we're always going to sum more strictly positive terms in  $F(y)$  than in  $F(x)$ .

Before we show how we choose  $\epsilon$ ,<sup>b</sup> we see that

$$\sum_{n=k}^{\infty} \frac{1}{2^n} = 2^{1-k}.$$

Now, we observe that

$$\sum_{n; x < r_n \leq x + \epsilon} \frac{1}{2^n} \leq \sum_{n = \arg \min_k x < r_k \leq x + \epsilon}^{\infty} \frac{1}{2^n} = 2^{1-k}.$$

With this observation, it should be fairly easy to see that we can choose  $\epsilon$  based on how small we want to make  $2^{1-k}$  be,<sup>c</sup> and we indeed see that

$$\lim_{k \rightarrow \infty} 2^{1-k} = 0,$$

which implies that  $F$  is right-continuous by squeeze theorem. ■

<sup>b</sup>To be precise, how  $\epsilon$  depends on  $r_n$ .

<sup>c</sup>We're referring to  $\epsilon - \delta$  proof approach.

**Lemma A.1.2.** The function  $F$  defined in [this example](#) satisfies

- $\mu_F(\{r_i\}) > 0$  for all  $r_i \in \mathbb{Q}$ .
- $\mu_F(\mathbb{R} \setminus \mathbb{Q}) = 0$

given in [this example](#).

**Proof.** We prove them one by one. And notice that  $F$  is indeed a distribution function as we proved in [Lemma A.1.1](#).

(1) To show  $\mu_F(\{r\}) > 0$  for every  $r \in \mathbb{Q}$ , we first note that

$$\{r\} = \bigcap_{a-1 \leq x < r} (x, r].$$

Then, we see that

$$\mu_F(\{r\}) = \mu_F\left(\bigcap_{a-1 \leq x < a} (x, r]\right),$$

where each  $(x, r] \in \mathcal{A}$  and  $(x, r] \supset (y, r]$  whenever  $r-1 \leq x \leq y < r$ . Notice that we implicitly assign the order of the index by the order of  $x$ . Then, we see that  $\mu_F(r-1, r] < \infty$ .<sup>a</sup> Then, from [continuity from above](#), we see that

$$\mu_F(\{r\}) = \lim_{i \rightarrow \infty} \mu_F((x_i, r]),$$

where we again implicitly assign an order to  $x$  as the usual order on  $\mathbb{R}$  by given index  $i$ . It's then clear that as  $i \rightarrow \infty$ ,  $x_i \rightarrow r$ . From the definition of  $F$ , we see that

$$F((x_i, r]) = F(r) - F(x_i) = \sum_{n; r_n \leq r} \frac{1}{2^n} - \sum_{n; r_n \leq x_i} \frac{1}{2^n} = \sum_{n; x_i < r_n \leq r} \frac{1}{2^n}.$$

It's then clear that since  $r \in \mathbb{Q}$ , there exists an  $i'$  such that  $r_{i'} = r$ . Then, we immediately see that no matter how close  $x_i \rightarrow r$ , this sum is at least

$$\frac{1}{2^{i'}}$$

for a fixed  $i'$ . Hence, we conclude that  $\mu_F(\{r\}) > 0$  for every  $r \in \mathbb{Q}$ .

(2) Now, we show  $\mu_F(\mathbb{R} \setminus \mathbb{Q}) = 0$ . Firstly, we claim that

$$\mu_F(\mathbb{Q}) = 1$$

and

$$\mu_F(\mathbb{R}) = 1$$

as well. Since  $\mu_F(\mathbb{Q}) = 1$  is clear, we note that the second one essentially follows from the fact that we can write

$$\mathbb{R} = \lim_{N \rightarrow \infty} \bigcup_{i=1}^N (a - i, a + i]$$

for any  $a \in \mathbb{R}$ , say 0. From [continuity from below](#), we have

$$\mu_F \left( \bigcup_{i=1}^{\infty} (-i, +i] \right) = \lim_{n \rightarrow \infty} \mu_F((-n, n]) = \sum_{n: r_n \in \mathbb{Q}} \frac{1}{2^n} = 1.$$

Given the above, from countable additivity of  $\mu_F$ , we have

$$\mu_F(\mathbb{R} \setminus \mathbb{Q}) + \underbrace{\mu_F(\mathbb{Q})}_1 = \underbrace{\mu_F(\mathbb{R})}_1 \Rightarrow \mu_F(\mathbb{R} \setminus \mathbb{Q}) = 0$$

as we desired. ■

---

<sup>a</sup>This will be  $\mu(A_1)$  in the condition of [continuity from above](#). Furthermore, since  $\mathbb{Q}$  is countable, hence  $F(x) < \infty$  is promised.

**Lemma A.1.3** (Cantor set has measure 0). Let  $C$  denotes the [middle thirds Cantor set](#), then the [Lebesgue measure](#) of  $C$  is 0. i.e.,

$$m(C) = 0.$$

**Proof.** Since we're removing  $\frac{1}{3}$  of the whole interval at each  $n$ , we see that the measure of those removing parts, denoted by  $r$ , is

$$m(r) = \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = \frac{1}{2} \sum_{n=1}^{\infty} \left( \frac{2}{3} \right)^n = 1.$$

Then, by [countable additivity](#) of  $m$ , we see that

$$m(C) = m([0, 1]) - m(r) = 1 - 1 = 0. \quad \blacksquare$$

## A.2 Integration

## A.3 Produce Measure

## A.4 Differentiation on Euclidean Space

## A.5 Normed-Vector-Space

## A.6 Signed and Complex Measures

### A.6.1 Uniqueness of [Theorem 6.5.2](#)

**Claim.** We show that if  $F \in NBV$ , then the unique

**Proof of Uniqueness in Theorem 6.5.2.** We observe that

$$\begin{aligned}\mu((-\infty, x]) &= \mu^+((-\infty, x]) - \mu^-((-\infty, x]) \\ &= \nu^+((-\infty, x]) - \nu^-((-\infty, x]) = \nu((-\infty, x]).\end{aligned}$$

Then, since we know that  $\mu^\pm$  is uniquely defined, hence we define

$$F^\pm(x) := \mu^\pm((-\infty, x]), \quad G^\pm(x) := \nu^\pm((-\infty, x]).$$

Then we see that

$$F^+ + G^- = G^+ + F^-,$$

while both left-hand side and right-hand sides are right-continuous and increasing. We hence associate it with the Lebesgue-Stieljes measure, which is uniquely defined, denoted it as  $\lambda$ . Hence, we see that

$$\lambda((a, b]) = (F^+ + G^-)(b) - (F^+ + G^-)(a) = (G^+ + F^-)(b) - (G^+ + F^-)(a)$$

which implies

$$\begin{aligned}(\mu^+((-\infty, b]) + \nu^-((-\infty, b])) - (\mu^+((-\infty, a]) + \nu^-((-\infty, a])) \\ = (\nu^+((-\infty, b]) + \mu^-((-\infty, b])) - (\nu^+((-\infty, a]) + \mu^-((-\infty, a])),\end{aligned}$$

i.e.,

$$\mu^+((a, b]) + \nu^-((a, b]) = \nu^+((a, b]) + \mu^-((a, b])$$

which further implies

$$\underbrace{\mu^+((a, b]) - \mu^-((a, b])}_{\mu((a, b])} = \underbrace{\nu^+((a, b]) - \nu^-((a, b])}_{\nu((a, b])}.$$

Since this holds for all  $(a, b]$ , while  $(a, b]$  generates  $\mathcal{B}(\mathbb{R})$ , hence we have that  $\mu = \nu$ . ■

## A.7 Hilbert Spaces

## A.8 Introduction to Fourier Analysis

# Bibliography

- [Axl19] S. Axler. *Measure, Integration & Real Analysis*. Graduate Texts in Mathematics. Springer International Publishing, 2019. ISBN: 9783030331429. URL: <https://books.google.com/books?id=8hCDyQEACAAJ>.
- [DM72] H. Dym and H.P. McKean. *Fourier Series and Integrals*. Probability and Mathematical Statistics. Elsevier Science, 1972. ISBN: 9780122264511. URL: <https://books.google.com/books?id=atkQAQAAIAAJ>.
- [FF99] G.B. Folland and G.B.A. FOLLAND. *Real Analysis: Modern Techniques and Their Applications*. A Wiley-Interscience publication. Wiley, 1999. ISBN: 9780471317166. URL: <https://books.google.com/books?id=uPkYAQAAIAAJ>.
- [Tao13] T. Tao. *An Introduction to Measure Theory*. Graduate studies in mathematics. American Mathematical Society, 2013. ISBN: 9781470409227. URL: <https://books.google.com/books?id=SPGJjwEACAAJ>.