

# MAT4400: Detailed Notes on Linear analysis

April 2, 2024

## 1 $\sigma$ -Algebras (3, [Schilling(2017)])

**Definition 1.1** ( $\sigma$ -Algebra). A family  $\mathcal{A}$  of subsets of  $X$  with:

- (i)  $X \in \mathcal{A}$ ,
- (ii)  $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ ,
- (iii)  $(A_n)_{n \in \mathbb{N}} \in \mathcal{A} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$

**Theorem 1.2** (and Definition).

- (i) The intersection of arbitrarily many  $\sigma$ -algebras in  $X$  is again a  $\sigma$ -algebra in  $X$ .
- (ii) For every system of sets  $p \subset \mathcal{P}(X)$  there exists a smallest  $\sigma$ -algebra containing  $p$ . This is the  $\sigma$ -algebra generated by  $p$ , denoted  $\sigma(p)$ , and  $\sigma(p)$  is called its generator.

**Definition 1.3** (Borel). The  $\sigma$ -algebra  $\sigma(\mathcal{O})$  generated by the open sets  $\mathcal{O} = \mathcal{O}_{\mathbb{R}^n}$  of  $\mathbb{R}^n$  is called **Borel  $\sigma$ -algebra**, and its members are called **Borel sets** or **Borel measurable sets**.

## 3 Uniqueness of Measures (5, [Schilling(2017)])

**Lemma 3.1.** A Dynkin system  $D$  is a  $\sigma$ -algebra iff it is stable under finite intersections, i.e.  $A, B \in D \Rightarrow A \cap B \in D$ .

**Theorem 3.2** (Dynkin). Assume  $X$  is a set,  $S$  is a collection of subsets of  $X$  closed under finite intersections, that is, if  $A, B \in S \Rightarrow A \cap B \in S$ . Then  $D(S) = \sigma(S)$ .

*Proof.* We clearly have that  $D(S) \subset \sigma(S)$ . If we can show that  $D(S)$  is a  $\sigma$ -algebra, that is, that a Dynkin system generated by a subset  $S \subset X$  (where  $S$  is  $\cap$ -stable) is a  $\sigma$ -algebra, then the inverse conclusion  $D(S) \supset \sigma(S)$  follows logically. This is the case because the  $\sigma$ -algebra  $\sigma(S)$  is the smallest  $\sigma$ -algebra containing  $S$ , and so if  $D(S)$  is a  $\sigma$ -algebra it must be a greater or equal (in some sense) than  $\sigma(S)$ .

Using Lemma 3.1 we only need to show that  $D(S)$  is stable under finite intersections, to prove that  $D(S)$  is a  $\sigma$ -algebra. Consider:

$$D_A := \{B \subset X : B \cap A \in D(S)\},$$

for some  $A \in D(S)$ . Notice that this set is  $\cap$ -stable, and so if we can show that  $D_A = D(S)$  we must have that (by Lemma 3.1)  $D(S)$  is a  $\sigma$ -algebra. Firstly, however, let us show that  $D_A$  is a Dynkin system.

1.  $\emptyset$  must be in  $D_A$ , since  $\emptyset \cap A = \emptyset \in D(S)$ .

2. Let  $B \in D_A$ . Then

$$A \cap B^c = A \setminus (A \cap B) = (A^c \cup (A \cap B))^c,$$

here  $A \cap B$  and  $A^c$  must be in  $D(S)$ . Furthermore, since disjoint unions of set from  $D(S)$  are still in  $D(S)$ , we must have  $A^c \in D_A$ .

3. Assume that  $(B_n)_{n \in \mathbb{N}} \subset D_A$  is a pairwise disjoint sequence. Then

$$\begin{aligned} (B_n \cap A)_{n \in \mathbb{N}} &\in D(S) \quad (\text{by def. of } D_A) \\ \Rightarrow \bigcup_{n \in \mathbb{N}} (B_n \cap A) &= \left( \bigcup_{n \in \mathbb{N}} B_n \right) \cap A \in D(S) \\ \Rightarrow \bigcup_{n \in \mathbb{N}} B_n &\in D_A. \end{aligned}$$

So  $D_A$  is indeed a Dynkin system.

We now want to show that  $D(S)$  is  $\cap$ -stable, we have:

$$\begin{aligned} S &\subset D_A \quad \forall A \in S \\ \Rightarrow D(S) &\subset D_A \quad \forall A \in S \\ \Rightarrow B \cap A &\in D(S) \quad \forall B \in S, \forall A \in D(S) \\ \Rightarrow B &\in D_A \quad \forall B \in S, \forall A \in D(S) \\ \Rightarrow S &\subset D_A \quad \forall A \in D(S) \\ \Rightarrow D(S) &\subset D_A \quad \forall A \in D(S) \\ \Rightarrow A \cap B &\in D(S) \quad \forall A, B \in D(S), \end{aligned}$$

and so  $D(S)$  is  $\cap$ -stable and then  $D(S) \supset \sigma(S) \Rightarrow D(S) = \sigma(S)$ .  $\square$

**Theorem 3.3** (uniqueness of measures). Let  $(X, \mathcal{B})$  be a measurable space, and  $S \subset \mathcal{P}(X)$  be the generator of  $\mathcal{B}$ , i.e.  $\mathcal{B} = \sigma(S)$ . If  $S$  satisfies the following conditions:

1.  $S$  is stable under finite intersections ( $\cap$ -stable), i.e.  $A, C \in S \Rightarrow A \cap C \in S$ .
2. There exists an exhausting sequence  $(G_n)_{n \in \mathbb{N}} \subset S$  with  $G_n \uparrow X$ . Assume also that there are two measures  $\mu, \nu$  satisfying:
3.  $\mu(A) = \nu(A), \forall A \in S$ .
4.  $\mu(G_n) = \nu(G_n) < \infty$ .

Then  $\mu = \nu$ .

*Proof (outline).* Define

$$D_n := \{A \in \mathcal{B} : \mu(G_n \cap A) = \nu(G_n \cap A) (< \infty)\},$$

and show that it is a Dynkin system. Then, use the fact that  $S$  is  $\cap$ -stable and Theorem 3.2 to argue that  $D(S) = \sigma(S) \dots \rightarrow \dots \mathcal{B} = D_n$ .  $\square$

## 4 Existence of Measures (6, [Schilling(2017)])

**Theorem 4.1** (Carathéodory). *Let  $S \subset P(X)$  be a semi-ring and  $\mu : S \rightarrow [0, \infty)$  a pre-measure. Then  $\mu$  has an extension to a measure  $\mu^*$  on  $\sigma(S)$ , i.e. that  $\mu(s) = \mu^*(s)$ ,  $\forall s \in \sigma(S)$ .*

*Also, if  $S$  contains an exhausting sequence,  $S_n \uparrow X$ , s.t.  $\mu(S_n) < \infty$ , then the extension is unique.*

*Outline of proof:* Firstly, let us define an outer measure.

**Definition 4.2** (Outer measure). An outer measure is a function  $\mu^* : P(X) \rightarrow [0, \infty)$  with the following properties:

1.  $\mu^*(\emptyset) = 0$ ,
2.  $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$ ,
3.  $\mu^*\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu^*(A_n)$ ,

and define for each  $A \subset X$  the family of countable  $S$ -coverings:

$$C(A) := \left\{ (S_n)_{n \in \mathbb{N}} \subset S : \bigcup_{n \in \mathbb{N}} S_n \supset A \right\},$$

and the set function

$$\mu^*(A) := \inf \left\{ \sum_{n \in \mathbb{N}} \mu(S_n) : (S_n)_{n \in \mathbb{N}} \in C(A) \right\}.$$

**Step 1: Claim:**  $\mu^*(A)$  is an outer measure.

*Proof.*

1.  $C(\emptyset) = \{\text{any sequence in } S \text{ containing } \emptyset\} \Rightarrow \mu^*(\emptyset) = 0$ .
2. Assume  $A \subset B$ . Then  $C(A) \subset C(B) \Rightarrow \mu^*(A) \leq \mu^*(B)$ .
3. If  $\mu^*(A_n) = \infty$  for some  $n$ , then there is nothing to prove. Thus, assume  $\mu^*(A_n) < \infty \forall n$ . Fix  $\epsilon > 0$ , and for every  $n$  choose  $A_{n_k} \in S$  s.t.

$$A_n \subset \bigcup_{k \in \mathbb{N}} A_{n_k},$$

$$\sum_{k \in \mathbb{N}} \mu^*(A_{n_k}) < \mu^*(A_n) + \frac{\epsilon}{2^n}.$$

Then

$$\bigcup_{n \in \mathbb{N}} A_n \subset \bigcup_{k \in \mathbb{N}} \bigcup_{n \in \mathbb{N}} A_{n_k},$$

so

$$\begin{aligned} \mu^*\left(\bigcup_{n \in \mathbb{N}} A_n\right) &\leq \sum_{n \in \mathbb{N}} \sum_{k \in \mathbb{N}} \mu(A_{n_k}) \\ &< \sum_{n \in \mathbb{N}} \left(\mu^*(A_n) + \frac{\epsilon}{2^n}\right) \\ &= \sum_{n \in \mathbb{N}} \mu^*(A_n) + \epsilon. \end{aligned}$$

As  $\epsilon$  was arbitrarily, we get that

$$\mu^*\left(\bigcup_{n \in \mathbb{N}} A_n\right) \leq \sum_{n \in \mathbb{N}} \mu^*(A_n),$$

so  $\mu^*$  fulfills all the conditions for being an outer measure.  $\square$

**Step 2:** Showing that  $\mu^*$  extends  $\mu$ , i.e.  $\mu^*(s) = \mu(s) \forall s \in S$ .

**Step 3:** Define  $\mu^*$ -measurable sets

$$\Sigma^* := \left\{ A \subset X : \mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \setminus A) \right. \\ \left. \forall Q \subset X \right\}$$

**Step 4:** Show that  $\mu|_{\Sigma^*}$  is a measure. In particular,  $\mu|_{\sigma(S)}$  is a measure which extends  $\mu$ .

## 5 Measurable Mappings (7, [Schilling(2017)])

We consider maps  $T : X \rightarrow X'$  between two measurable spaces  $(X, \mathcal{A})$  and  $(X', \mathcal{A}')$  which respects the measurable structures, the  $\sigma$ -algebras on  $X$  and  $X'$ . These maps are useful as we can transport a measure  $\mu$ , defined on  $(X, \mathcal{A})$ , to  $(X', \mathcal{A}')$ .

**Definition 5.1.** Let  $(X, \mathcal{A})$ ,  $(X', \mathcal{A}')$  be measurable spaces. A map  $T : X \rightarrow X'$  is called  $\mathcal{A}/\mathcal{A}'$ -measurable if the pre-image of every measurable set is a measurable set:

$$T^{-1}(A') \in \mathcal{A}, \quad \forall A' \in \mathcal{A}'.$$

- A  $\mathcal{B}(\mathbb{R}^n)/\mathcal{B}(\mathbb{R}^m)$  measurable map is often called a Borel map.
- The notation  $T : (X, \mathcal{A}) \rightarrow (X', \mathcal{A}')$  is often used to indicate measurability of the map  $T$ .

**Lemma 5.2.** Let  $(X, \mathcal{A})$ ,  $(X', \mathcal{A}')$  be measurable spaces and let  $\mathcal{A}' = \sigma(\mathcal{G}')$ . Then  $T : X \rightarrow X'$  is  $\mathcal{A}/\mathcal{A}'$ -measurable iff  $T^{-1}(\mathcal{G}') \subset \mathcal{A}$ , i.e. if

$$T^{-1}(G') \in \mathcal{A}, \quad \forall G' \in \mathcal{G}'.$$

**Theorem 5.3.** Let  $(X_i, \mathcal{A}_i)$ ,  $i = 1, 2, 3$ , be measurable spaces and  $T : X_1 \rightarrow X_2$ ,  $S : X_2 \rightarrow X_3$  be  $\mathcal{A}_1/\mathcal{A}_2$  and  $\mathcal{A}_2/\mathcal{A}_3$ -measurable maps respectively. Then  $S \circ T : X_1 \rightarrow X_3$  is  $\mathcal{A}_1/\mathcal{A}_3$ -measurable.

**Corollary 5.4.** Every continuous map between metric spaces is a Borel map.

**Definition 5.5. (and lemma)** Let  $(T_i)_{i \in I}$ ,  $T_i : X \rightarrow X_i$ , be arbitrarily many mappings from the same space  $X$  into measurable spaces  $(X_i, \mathcal{A}_i)$ . The smallest  $\sigma$ -algebra on  $X$  that makes all  $T_i$  simultaneously measurable is

$$\sigma(T_i : i \in I) := \sigma\left(\bigcup_{i \in I} T_i^{-1}(\mathcal{A}_i)\right)$$

**Corollary 5.6.** A function  $f : (X, \mathcal{B}) \rightarrow \mathbb{R}$  is measurable if  $f((a, +\infty)) \in \mathcal{B}$ ,  $\forall a \in \mathbb{R}$ .

**Corollary 5.7.** Assume  $(X, \mathcal{B})$  is a measurable space,  $(Y, d)$  is a metric space, and  $(f_n : (X, \mathcal{B}) \rightarrow Y)_{n=1}^{\infty}$  is a sequence of measurable maps. Assume this sequence of images  $(f_n(x))_{n=1}^{\infty}$  is convergent in  $Y \forall x \in X$ . Define

$$f : X \rightarrow Y, \quad \text{by } f(x) = \lim_{n \rightarrow \infty} f_n(x).$$

Then  $f$  is measurable.

**Theorem 5.8.** Let  $(X, \mathcal{A})$ ,  $(X', \mathcal{A}')$  be measurable spaces and  $T : X \rightarrow X'$  be an  $\mathcal{A}/\mathcal{A}'$ -measurable map. For every measurable  $\mu$  on  $(X, \mathcal{A})$ ,

$$\mu'(A') := \mu(T^{-1}(A')), \quad A' \in \mathcal{A}',$$

defines a measure on  $(X', \mathcal{A}')$ .

**Definition 5.9.** The measure  $\mu'(\cdot)$  in the above theorem is called the push forward or image measure of  $\mu$  under  $T$  and it is denoted as  $T(\mu)(\cdot)$ ,  $T_{*\mu}(\cdot)$  or  $\mu \circ T^{-1}(\cdot)$ .

**Theorem 5.10.** If  $T \in \mathbb{R}^{n \times n}$  is an orthogonal matrix, then  $\lambda^n = T(\lambda^n)$ .

**Theorem 5.11.** Let  $S \in \mathbb{R}^{n \times n}$  be an invertible matrix. Then

$$S(\lambda^n) = |\det S|^{-1} \lambda^n = |\det S|^{-1} \lambda^n.$$

**Corollary 5.12.** Lebesgue measure is invariant under motions:  $\lambda^n = M(\lambda^n)$  for all motions  $M$  in  $\mathbb{R}^n$ . In particular, congruent sets have the same measure. Two sets of points are called congruent if, and only if, one can be transformed into the other by an isometry

## Measurable Functions (8, [Schilling(2017)])

A measurable function is a measurable map  $u : X \rightarrow \mathbb{R}$  from some measurable space  $(X, \mathcal{A})$  to  $(\mathbb{R}, \mathcal{B}(\mathbb{R}^1))$ . They play central roles in the theory of integration.

We recall that  $u : X \rightarrow \mathbb{R}$  is  $\mathcal{A}/\mathcal{B}(\mathbb{R}^1)$ -measurable if

$$u^{-1}(B) \in \mathcal{A}, \quad \forall B \in \mathcal{B}(\mathbb{R}^1).$$

Moreover from a lemma from chapter 7, we actually only need to show that

$$u^{-1}(G) \in \mathcal{A}, \quad \forall G \in \mathcal{G} \text{ where } \mathcal{G} \text{ generates } \mathcal{B}(\mathbb{R}^1).$$

**Proposition 5.13.**

- 1 If  $f, g : (X, \mathcal{B}) \rightarrow \mathbb{C}$  are measurable, then the function  $f + g$ ,  $f \cdot g$ ,  $cf$ , ( $c \in \mathbb{C}$ ) are measurable.
- 2 If  $b : \mathbb{C} \rightarrow \mathbb{C}$  is Borel and  $b : (\mathbb{C}, \mathcal{B}) \rightarrow \mathbb{C}$  is measurable, then  $b \circ f$  is measurable.
- 3 If  $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ ,  $x \in X$  and  $f_n$  are measurable, then  $f$  is measurable.
- 4 If  $X = \bigcup_{n=1}^{\infty} A_n$ , ( $A_n \in \mathcal{B}$ ),  $f|_{A_n} : (A_n, \mathcal{B}_{A_n}) \rightarrow \mathbb{C}$  is measurable  $\forall n$ , then  $f$  is measurable.

**Definition 5.14.** Given a measurable space  $(X, \mathcal{B})$ , a measurable function  $f : (X, \mathcal{B}) \rightarrow \mathbb{C}$  is called simple if

$$f(x) = \sum_{k=1}^N c_k \mathbb{1}_{A_k}(x),$$

for some  $c_k \in \mathbb{C}$ ,  $A_k \in \mathcal{B}$ , where  $\mathbb{1}$  is the characteristic function,

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

The representation of simple function is **not** unique. We denote the standard representation of  $f$  by

$$f(x) = \sum_{n=0}^N z_n \mathbb{1}_{B_n}(x),$$

for  $N \in \mathbb{N}$ ,  $z_n \in \mathbb{R}$ ,  $B_n \in \mathcal{A}$ , and

$$X = \bigcup_{n=1}^N B_n,$$

for  $B_n \cap B_m = \emptyset$ ,  $n \neq m$ . The set of simple functions is denoted  $\mathcal{E}(\mathcal{A})$  of  $\mathcal{E}$ .

**Definition 5.15.** Assume  $\mu$  is a measure on  $(X, \mathcal{B})$ . Given a positive simple function

$$f = \sum_{k=1}^N c_k \mathbb{1}_{A_k}, \quad (c_k \geq 0).$$

We define

$$\int_X f d\mu = \sum_{k=1}^N c_k \mu(A_k) \in [0, +\infty].$$

We also denote this by  $I_\mu(f)$ .

**Lemma 5.16.** This is well defined, that is,  $\int_X f d\mu$  does not depend on the presentation of the simple function  $f$ .

**Properties 5.17.** For every positive simple function

- 1  $\int_X cf d\mu = c \int_X f d\mu$ , for only  $c \geq 0$
- 2  $\int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu$ .

**Corollary 5.18.** If  $f \geq g \geq 0$  are simple functions, then

$$\int_X f d\mu \geq \int_X g d\mu.$$

**Definition 5.19.** If  $f : X \rightarrow [0, +\infty)$  is measurable, then we define

$$\int_X f d\mu = \sup \left\{ \int_X g d\mu : f \geq g \geq 0, \quad g \text{ is simple} \right\}$$

**Remark.** This means that any measurable function can be approximated by simple functions.

**Properties 5.20.** Measurable functions like this have the following properties

- 1  $\int_X cf d\mu = c \int_X f d\mu$ ,  $\forall c \geq 0$ .
- 2 If  $f \geq g \geq 0$ , then  $\int_X f d\mu \geq \int_X g d\mu$  for any measurable  $g, f$ .
- 3 If  $f \geq 0$  is simple, then  $\int_X f d\mu$  is the same value as obtained before.

To advance in measure theory we consider measurable functions

$$f : X \rightarrow [0, +\infty].$$

Measurability is understood w.r.t the  $\sigma$ -algebra  $\mathcal{B}([0, +\infty])$  generated by  $\mathcal{B}([0, +\infty))$  and  $\{+\infty\}$ . In other words,  $A \subset [0, +\infty] \in \mathcal{B}([0, +\infty])$  iff  $A \cap [0, +\infty) \in \mathcal{B}([0, +\infty))$ .

**Remark.** Hence  $f : X \rightarrow [0, +\infty]$  is measurable iff  $f^{-1}(A)$  is measurable  $\forall A \in \mathcal{B}([0, +\infty])$ .

**Definition 5.21.** For measurable functions  $f_X \rightarrow [0, +\infty]$ , we define

$$\int_X f d\mu = \sup \left\{ \int_X g d\mu : f \geq g \geq 0 : g \text{ is simple} \right\} \in [0, +\infty].$$

**Theorem 5.22. Monotone convergence theorem** Assume  $(X, \mathcal{B}, \mu)$  is a measure space,  $(f)_{n=1}^{\infty}$  is an increasing sequence of measurable positive functions  $f_n : X \rightarrow [0, +\infty]$ . Define  $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ . Then  $f$  is measurable and

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

**Theorem 5.23.** Assume  $(X, \mathcal{B})$  is a measurable space and  $f : X \rightarrow [0, +\infty]$  is measurable. Then there are simple functions  $g_n$ , s.t.

$$0 \leq g_1 \leq g_2 \leq \dots, \quad g_n(x) \rightarrow f(x), \quad \forall x \in X.$$

Moreover, if  $f$  is bounded, we can choose  $g_n$  s.t. the convergence is uniform, that is,

$$\lim_{n \rightarrow \infty} \sup_{x \in X} |g_n(x) - f(x)| = 0.$$

## 6 Integration of Measurable Functions

(9, [Schilling(2017)])

Through this chapter  $(X, \mathcal{A}, \mu)$  will be some measure space. Recall that  $\mathcal{M}^+(\mathcal{A})$  [ $\mathcal{M}_{\mathbb{R}}^+(\mathcal{A})$ ] are the  $\mathcal{A}$ -measurable positive functions and  $\mathcal{E}(\mathcal{A})$  [ $\mathcal{E}_{\mathbb{R}}^+(\mathcal{A})$ ] are the positive and simple functions.

The fundamental idea of *Integration* is to measure the area between the graph of the function and the abscissa. For positive simple functions  $f \in \mathcal{E}^+(\mathcal{A})$  in standard representation, this is done easily

$$\text{if } f = \sum_{i=0}^M y_i \mathbb{1}_{A_i} \in \mathcal{E}^+(\mathcal{A}) \quad \text{then} \quad \sum_{i=0}^M y_i \mu(A_i) \quad (1)$$

would be the  $\mu$ -area enclosed by the graph and the abscissa. We note that the representation of  $f$  should not impact the integral of  $f$ .

**Lemma 6.1.** Let  $\sum_{i=0}^M y_i \mathbb{1}_{A_i} = \sum_{k=0}^N z_k \mathbb{1}_{B_k}$  be two standard representations of the same function  $f \in \mathcal{E}^+(\mathcal{A})$ . Then

$$\sum_{i=0}^M y_i \mu(A_i) = \sum_{k=0}^N z_k \mu(B_k). \quad (2)$$

**Definition 6.2.** Let  $f = \sum_{i=0}^M y_i \mathbb{1}_{A_i} \in \mathcal{E}^+(\mathcal{A})$  be a simple function in standard representation. Then the number

$$I_{\mu}(f) = \sum_{i=0}^M y_i \mu(A_i) \in [0, \infty] \quad (3)$$

(which is independent of the representation of  $f$ ) is called the  $\mu$ -integral of  $f$ .

**Proposition 6.3.** Let  $f, g \in \mathcal{E}^+(\mathcal{A})$ . Then

- (i)  $I_{\mu}(\mathbb{1}_A) = \mu(A) \quad \forall A \in \mathcal{A}$ .
- (ii)  $I_{\mu}(\lambda f) = \lambda I_{\mu}(f) \quad \forall \lambda \geq 0$ .
- (iii)  $I_{\mu}(f + g) = I_{\mu}(f) + I_{\mu}(g)$ .
- (iv)  $f \leq g \Rightarrow I_{\mu}(f) \leq I_{\mu}(g)$ .

In theorem 8.8 we saw that we could for every  $u \in \mathcal{M}^+(\mathcal{A})$  write it as an increasing limit of simple functions. By corollary

8.10, the suprema of simple functions are again measurable, so that

$$u \in \mathcal{M}^+(\mathcal{A}) \Leftrightarrow u = \sup_{n \in \mathbb{N}} f_n, f \in \mathcal{E}^+(\mathcal{A}), \\ f_n \leq f_{n+1} \leq \dots$$

We will use this to "inscribe" simple functions (which we know how to integrate) below the graph of a positive measurable function  $u$  and exhaust the  $\mu$ -area below  $u$ .

**Definition 6.4.** Let  $(X, \mathcal{A}, \mu)$  be a measure space. The  $(\mu)$ -integral of a positive function  $u \in \mathcal{M}_{\mathbb{R}}^+(\mathcal{A})$  is given by

$$\int u d\mu = \sup \{ I_{\mu}(g) : g \leq u, g \in \mathcal{E}^+(\mathcal{A}) \}, \quad (4)$$

with  $\int u d\mu \in [0, +\infty]$ . If we need to emphasize the *integration variable*, we write  $\int u(x) \mu(dx)$ . The key observation is that the integral  $\int \dots d\mu$  extends  $I_{\mu}$ .

**Lemma 6.5.** For all  $f \in \mathcal{E}^+(\mathcal{A})$  we have  $\int f d\mu = I_{\mu}(f)$ .

The next theorem is one of many convergence theorems. It shows that we could have defined 4 using any increasing sequence  $f_n \uparrow u$  of simple functions  $f_n \in \mathcal{E}^+(\mathcal{A})$ .

**Theorem 6.6. (Beppo Levi)** Let  $(X, \mathcal{A}, \mu)$  be a measure space. For an increasing sequence of functions  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{M}_{\mathbb{R}}^+(\mathcal{A})$ ,  $0 \leq u_n \leq u_{n+1} \leq \dots$ , we have for the supremum  $u = \sup_{n \in \mathbb{N}} u_n \in \mathcal{M}_{\mathbb{R}}^+(\mathcal{A})$  and

$$\int \sup_{n \in \mathbb{N}} u_n d\mu = \sup_{n \in \mathbb{N}} \int u_n d\mu. \quad (5)$$

Note we can write  $\lim_{n \rightarrow \infty}$  instead of  $\sup_{n \in \mathbb{N}}$  as the supremum of an increasing sequence is its limit. Moreover, this theorem holds in  $[0, +\infty]$ , so the case  $+\infty = +\infty$  is possible.

**Corollary 6.7.** Let  $u \in \mathcal{M}_{\mathbb{R}}^+(\mathcal{A})$ . Then

$$\int u d\mu = \lim_{n \rightarrow \infty} \int f_n d\mu$$

holds for every sequence  $(f_n)_{n \in \mathbb{N}} \subset \mathcal{E}^+(\mathcal{A})$  with  $\lim_{n \rightarrow \infty} f_n = u$ .

**Proposition 6.8. (of integral)** Let  $u, v \in \mathcal{M}_{\mathbb{R}}^+(\mathcal{A})$ . Then

- (i)  $\int \mathbb{1}_A d\mu = \mu(A) \quad \forall A \in \mathcal{A}$ .
- (ii)  $\int \alpha u d\mu = \alpha \int u d\mu \quad \forall \alpha \geq 0$ .
- (iii)  $\int u + v d\mu = \int u d\mu + \int v d\mu$ .
- (iv)  $u \leq v \Rightarrow \int u d\mu \leq \int v d\mu$ .

**Corollary 6.9.** Let  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{M}_{\mathbb{R}}^+(\mathcal{A})$ . Then  $\sum_{n=1}^{\infty} u_n$  is measurable and we have

$$\int \sum_{n=1}^{\infty} u_n d\mu = \sum_{n=1}^{\infty} \int u_n d\mu$$

(including the possibility  $+\infty = +\infty$ .)

**Theorem 6.10. (Fatou)** Let  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{M}_{\mathbb{R}}^+(\mathcal{A})$  be a sequence of positive measurable functions. Then  $u = \liminf_{n \rightarrow \infty} u_n$  is measurable and

$$\int \liminf_{n \rightarrow \infty} u_n d\mu \leq \liminf_{n \rightarrow \infty} \int u_n d\mu \quad (6)$$

## 7 Integrals of Measurable Functions

(10, [Schilling(2017)])

We have defined our integral for positive measurable functions, i.e. functions in  $\mathcal{M}^+(\mathcal{A})$ . To extend our integral to not only functions in  $\mathcal{M}^+(\mathcal{A})$  we first notice that

$$u \in \mathcal{M}_{\overline{\mathbb{R}}}(\mathcal{A}) \Leftrightarrow u = u^+ - u^-, \quad u^+, u^- \in \mathcal{M}_{\overline{\mathbb{R}}}^+,$$

i.e. that every measurable function can be written as a sum of **positive** measurable functions.

**Definition 7.1** ( $\mu$ -integrable). A function  $u : X \rightarrow \overline{\mathbb{R}}$  on  $(X, \mathcal{A}, \mu)$  is  $\mu$ -integrable, if it is  $\mathcal{A}/\mathcal{B}(\overline{\mathbb{R}})$ -measurable and if  $\int u^+ d\mu, \int u^- d\mu < \infty$  (recall the definition for the integral of positive measurable functions). Then

$$\int u d\mu := \int u^+ d\mu - \int u^- d\mu \in (-\infty, \infty)$$

is the ( $\mu$ -)integral of  $u$ . We write  $\mathcal{L}^1(\mu)$  for the set of all real-valued  $\mu$ -integrable functions <sup>1</sup>.

**Theorem 7.2.** Let  $u \in \mathcal{M}_{\overline{\mathbb{R}}}(\mathcal{A})$ , then the following conditions are equivalent:

- (i)  $u \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$ .
- (ii)  $u^+, u^- \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$ .
- (iii)  $|u| \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$ .
- (iv)  $\exists w \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$  with  $w \geq 0$  s.t.  $|u| \leq w$ .

**Theorem 7.3** (Properties of the  $\mu$ -integral). The  $\mu$ -integral is: **homogeneous, additive, and:**

- (i)  $\min\{u, v\}, \max\{u, v\} \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$  (lattice property)
- (ii)  $u \leq v \Rightarrow \int u d\mu \leq \int v d\mu$  (monotone)
- (iii)  $\left| \int u d\mu \right| \leq \int |u| d\mu$  (triangle inequality)

**Remark.** If  $u(x) \pm v(x)$  is defined in  $\overline{\mathbb{R}}$  for all  $x \in X$  then we can exclude  $\infty - \infty$  and the theorem above just says that the integral is linear:

$$\int (au + bv) d\mu = a \int u d\mu + b \int v d\mu.$$

This is always true for real-valued  $u, v \in \mathcal{L}^1(\mu) = \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$ , making  $\mathcal{L}^1(\mu)$  a vector space with addition and scalar multiplication defined by

$$(u + v)(x) := u(x) + v(x), \quad (a \cdot u)(x) := a \cdot u(x),$$

and

$$\int \dots d\mu : \mathcal{L}^1(\mu) \rightarrow \mathbb{R}, \quad u \mapsto \int u d\mu,$$

is a **positive linear functional**.

<sup>1</sup>In words, we extend our integral to ~~positive~~ measurable functions by noticing that we can write every measurable function as a sum of positive measurable functions, something that we do know how to integrate. We don't want to run into the problem of  $\infty - \infty$ , thus we require the integral of the positive and negative parts to both (separately) be less than infinity.

## 8 Null sets and the Almost Everywhere

(11, [Schilling(2017)])

**Definition 8.1.** A ( $\mu$ -)null set  $N \in \mathcal{N}_{\mu}$  is a measurable set  $N \in \mathcal{A}$  satisfying

$$N \in \mathcal{N}_{\mu} \Leftrightarrow N \in \mathcal{A} \text{ and } \mu(N) = 0.$$

This can be used generally about a 'statement' or 'property', but we will be interested in questions like 'when is  $u(x)$  equal to  $v(x)$ ', and we answer this by saying

$$u = v \text{ a.e.} \Leftrightarrow \{x : u(x) \neq v(x)\} \text{ is (contained in) a } \mu\text{-null set,}$$

i.e.

$$u = v \quad \mu\text{-a.e.} \Leftrightarrow \mu(\{x : u(x) \neq v(x)\}) = 0.$$

The last phrasing should of course include that the set  $\{x : u(x) \neq v(x)\}$  is in  $\mathcal{A}$ .

**Theorem 8.2.** Let  $u \in \mathcal{M}_{\overline{\mathbb{R}}}(\mathcal{A})$ , then:

- (i)  $\int |u| d\mu = 0 \Leftrightarrow |u| = 0 \text{ a.e.} \Leftrightarrow \mu\{u \neq 0\} = 0$ ,
- (ii)  $\mathbb{1}_N u \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu) \quad \forall N \in \mathcal{N}_{\mu}$ ,
- (iii)  $\int_N u d\mu = 0$ .

(i) is really useful, later we will define  $\mathcal{L}^p$  and the  $\|\cdot\|_p$ -(semi)norm. Then (i) means that if we have a sequence  $u_n$  converging to  $u$  in the  $\|\cdot\|_p$ -norm then  $u_n(x) = u(x)$  a.e.

**Corollary 8.3.** Let  $u = v$   $\mu$ -a.e. Then

- (i)  $u, v \geq 0 \Rightarrow \int u d\mu = \int v d\mu$ ,
- (ii)  $u \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu) \Rightarrow v \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$  and  $\int u d\mu = \int v d\mu$ .

**Corollary 8.4.** If  $u \in \mathcal{M}_{\overline{\mathbb{R}}}(\mathcal{A})$ ,  $v \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$  and  $v \geq 0$  then

$$|u| \leq v \text{ a.e.} \Rightarrow u \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu).$$

**Proposition 8.5** (Markow inequality). For all  $u \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$ ,  $A \in \mathcal{A}$  and  $c > 0$

$$\mu(\{|u| \geq c\} \cap A) \leq \frac{1}{c} \int_A |u| d\mu,$$

if  $A = X$ , then (obviously)

$$\mu\{|u| \geq c\} \leq \frac{1}{c} \int |u| d\mu.$$

**Corollary 8.6.** If  $u \in \mathcal{L}_{\overline{\mathbb{R}}}^1(\mu)$ , then  $\mu$  is a.e.  $\mathbb{R}$ -valued. In particular, we can find a version  $\tilde{u} \in \mathcal{L}^1(\mu)$  s.t.  $\tilde{u} = u$  a.e. and  $\int \tilde{u} d\mu = \int u d\mu$



## Completions of measure spaces

**Definition 8.7.** A measure space  $(X, \mathcal{B}, \mu)$  is called **complete** if whenever  $A \in \mathcal{B}$  and  $\mu(A) = 0$ , we have  $B \in \mathcal{B} \forall B \subset A$ .

**Remark.** Any measure space can be completed as follows: Let  $\bar{\mathcal{B}}$  be the  $\sigma$ -algebra generated by  $\mathcal{B}$  and all sets  $B \subset X$  s.t. there exists  $A \in \mathcal{B}$  with  $B \subset A$  and  $\mu(A) = 0$ .

**Proposition 8.8.** The  $\sigma$ -algebra  $\bar{\mathcal{B}}$  can also be described as follows:

$$\bar{\mathcal{B}} := \left\{ B \subset X : A_1 \subset B \subset A_2 \text{ for some } A_1, A_2 \in \mathcal{B} \text{ with } \mu(A_2 \setminus A_1) = 0 \right\},$$

with  $B, A_1, A_2$  as above, we define

$$\bar{\mu} := \mu(A_1) = \mu(A_2)$$

Then  $(X, \bar{\mathcal{B}}, \bar{\mu})$  is a complete measure space.

**Definition 8.9.** If  $\mu$  is a Borel measure on a **metric** space  $(X, d)$ , then the completion  $\bar{\mathcal{B}}(X)$  of the Borel  $\sigma$ -algebra with respect to  $\mu$  is called the  $\sigma$ -algebra of  $\mu$ -measurable sets.

**Remark.** For  $\mu = \lambda_n$  on  $\mathbb{R}^n$  we talk about the  $\sigma$ -algebra of **Lebesgue measurable sets**. Instead of  $\bar{\lambda}_n$  we still write  $\lambda_n$  and call it the **Lebesgue measure**. A function  $f : \mathbb{R}^n \rightarrow \mathbb{C}$ , measurable w.r.t. the  $\sigma$ -algebra of Lebesgue measurable sets is called the **Lebesgue measurable**.

The following result shows that any Lebesgue measurable function coincides with a Borel function a.e.

**Proposition 8.10.** Assume  $(X, \mathcal{B}, \mu)$  is a measure space and consider its completion  $(X, \bar{\mathcal{B}}, \bar{\mu})$ . Assume  $f : X \rightarrow \mathbb{C}$  is  $\bar{\mathcal{B}}$ -measurable. Then there is a  $\mathcal{B}$ -measurable function  $g : X \rightarrow \mathbb{C}$  s.t.  $f = g$   $\bar{\mu}$ -a.e.

## 9 Convergence Theorems and their Applications (12, [Schilling(2017)])

- To interchange limits and integrals in **Riemann integrals** one typically has to assume uniform convergence. - The set of Riemann integrable functions is somewhat limited, see theorem 9.7

**Theorem 9.1** (Generalization of Beppo Levi, monotone convergence).

(i) Let  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{L}^1(\mu)$  be s.t.  $u_1 \leq u_2 \leq \dots$  with limit  $u := \sup_{n \in \mathbb{N}} u_n = \lim_{n \rightarrow \infty} u_n$ . Then  $u \in \mathcal{L}^1(\mu)$  **iff**

$$\sup_{n \in \mathbb{N}} \int u_n d\mu < +\infty,$$

in which case

$$\sup_{n \in \mathbb{N}} \int u_n d\mu = \int \sup_{n \in \mathbb{N}} u_n d\mu.$$

(ii) Same thing only with a decreasing sequence  $\dots > -\infty$  in which case

$$\inf_{n \in \mathbb{N}} \int u_n d\mu = \int \inf_{n \in \mathbb{N}} u_n d\mu.$$

**Theorem 9.2** (Lebesgue; dominated convergence). Let  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{L}^1(\mu)$  s.t.

$$(a) |u_n|(x) \leq w(x), w \in \mathcal{L}^1(\mu),$$

$$(b) u(x) = \lim_{n \rightarrow \infty} u_n(x) \text{ exists in } \bar{\mathbb{R}},$$

then  $u \in \mathcal{L}^1(\mu)$  and we have

$$(i) \lim_{n \rightarrow \infty} \int |u_n - u| d\mu = 0;$$

$$(ii) \lim_{n \rightarrow \infty} \int u_n d\mu = \int \lim_{n \rightarrow \infty} u_n d\mu = \int u d\mu;$$

## Application 1: Parameter-Dependent Integrals

- We are interested in questions of the sort, when is

$$U(t) := \int u(t, x) \mu(dx), \quad t \in (a, b),$$

again a smooth function of  $t$ ? The answer involves interchange of limits and integration. Also, it turns out to better understand Riemann integrability, we need the Lebesgue integral.

**Theorem 9.3** (continuity lemma). Let  $\emptyset \neq (a, b) \subset \mathbb{R}$  be a non-degenerate open interval and  $u : (a, b) \times X \rightarrow \mathbb{R}$  satisfy

$$(a) x \mapsto u(t, x) \text{ is in } \mathcal{L}^1(\mu) \text{ for every fixed } t \in (a, b);$$

$$(b) t \mapsto u(t, x) \text{ is continuous for every fixed } x \in X;$$

$$(c) |u(t, x)| \leq w(x) \text{ for all } (t, x) \in (a, b) \times X \text{ and some } w \in \mathcal{L}^1(\mu).$$

Then the function  $U : (a, b) \rightarrow \mathbb{R}$  given by

$$t \mapsto U(t) := \int u(t, x) \mu(dx) \quad (7)$$

is continuous.

**Theorem 9.4** (differentiability lemma). Let  $\emptyset \leq (a, b) \subset \mathbb{R}$  be a non-degenerate open interval and  $u : (a, b) \times X \rightarrow \mathbb{R}$  satisfy

$$(a) \text{ Same}$$

$$(b) \text{ Same}$$

$$(c) |\partial_t u(t, x)| \leq w(x) \text{ for all } (t, x) \in (a, b) \times X \text{ and some } w \in \mathcal{L}^1(\mu).$$

Then the function in 7 is differentiable and its derivative is

$$\frac{d}{dt} U(t) = \frac{d}{dt} \int u(t, x) \mu(dx) = \int \frac{\partial}{\partial t} u(t, x) \mu(dx). \quad (8)$$

## Application 2: Riemann vs Lebesgue Integration

Consider only  $(X, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ .

**Definition 9.5** (The Riemann Integral). Consider on the finite interval  $[a, b] \subset \mathbb{R}$  the partition

$$\Pi := \{a = t_0 < t_1 < \dots < t_k < b\}, k = k(\Pi), \quad (9)$$

and introduce

$$S_\Pi[u] := \sum_{i=1}^{k(\Pi)} m_i(t_i - t_{i-1}), \quad m_i := \inf_{x \in [t_{i-1}, t_i]} u(x), \quad (10)$$

$$S^\Pi[u] := \sum_{i=1}^{k(\Pi)} M_i(t_i - t_{i-1}), \quad M_i := \sup_{x \in [t_{i-1}, t_i]} u(x). \quad (11)$$

$$(12)$$

A bounded function  $u : [a, b] \rightarrow \mathbb{R}$  is said to be **Riemann integrable** if the values

$$\int u := \sup_{\Pi} S_{\Pi}[u] = \inf_{\Pi} S^{\Pi}[u] =: \int u \quad (13)$$

coincide and are finite. Their common value is called the **Riemann integral** of  $u$  and denoted by  $(R) \int_a^b u(x)dx$  or  $\int_a^b u(x)dx$ .

**Theorem 9.6.** Let  $u : [a, b] \rightarrow \mathbb{R}$  be a **measurable** and **Riemann integrable** function. Then

$$u \in \mathcal{L}^1(\lambda) \text{ and } \int_{[a,b]} u d\lambda = \int_a^b u(x)dx. \quad (14)$$

**Theorem 9.7.** Let  $u : [a, b] \rightarrow \mathbb{R}$  be a bounded function, it is Riemann integrable **iff** the points in  $(a, b)$  where  $u$  is discontinuous are a (subset of) Borel measurable null set.

## Improper Riemann Integrals

- The Lebesgue integral extends the (*proper*) Riemann integral. However, there is a further extension of the Riemann integral which cannot be captured by Lebesgue's theory.  $u$  is Lebesgue integrable **iff**  $|u|$  has finite Lebesgue integral. - The Lebesgue integral does not respect sign-changes and cancellations. However, the following *improper Riemann integral* does:

$$(R) \int_0^{\infty} u(x)dx := \lim_{n \rightarrow \infty} (R) \int_0^n u(x)dx. \quad (15)$$

**Corollary 9.8.** Let  $u : [0, \infty) \rightarrow \mathbb{R}$  be a measurable, Riemann integrable function for every interval  $[0, N]$ ,  $N \in \mathbb{N}$ . Then  $u \in \mathcal{L}^1[0, \infty)$  **iff**

$$\lim_{N \rightarrow \infty} (R) \int_0^N |u(x)|dx < \infty. \quad (16)$$

In this case,  $(R) \int_0^{\infty} u(x)dx = \int_{[0, \infty)} u d\lambda$

**Example** of a function which is *improperly Riemann integrable* but **not** Lebesgue integrable:

$$f(x) = \frac{\sin(x)}{x}. \quad (17)$$

**Proposition 9.9** (appearing as example 12.13 in Schilling). Let  $f_{\alpha}(x) := x^{\alpha}$ ,  $x > 0$  and  $\alpha \in \mathbb{R}$ . Then

$$(i) f_{\alpha} \in \mathcal{L}^1(0, 1) \Leftrightarrow \alpha > -1.$$

$$(ii) f_{\alpha} \in \mathcal{L}^1[1, \infty) \Leftrightarrow \alpha < -1.$$

## 10 Regularity of Measures (App. H, [Schilling(2017)])

We let  $(X, d)$  be a metric space and denote by  $\mathcal{O}$  the open, by  $\mathcal{C}$  the closed and  $\mathcal{B}(X) = \sigma(\mathcal{O})$  the Borel set of  $X$ .

**Definition 10.1.** A measure  $\mu$  on  $(X, d, \mathcal{B}(X))$  is called outer regular, if

$$\mu(B) = \inf \{ \mu(U) \mid B \subset U, U \text{ open} \} \quad (18)$$

and inner regular, if  $\mu(K) < \infty$  for all compact sets  $K \subset X$  and

$$\mu(U) = \sup \{ \mu(K) \mid K \subset U, K \text{ compact} \}. \quad (19)$$

A measure which is both inner and outer regular is called **regular**. We write  $\mathfrak{m}_r^+(X)$  for the family of regular measures on  $(X, \mathcal{B}(X))$ .

**Remark.** The space  $X$  is called  $\sigma$ -compact if there is a sequence of compact sets  $K_n \uparrow X$ . A typical example of such a space is a locally compact, separable metric space.

**Theorem 10.2.** Let  $(X, d)$  be a metric space. Every finite measure  $\mu$  on  $(X, \mathcal{B}(X))$  is outer regular. If  $X$  is  $\sigma$ -compact, then  $\mu$  is also inner regular, hence regular.

**Theorem 10.3.** Let  $(X, d)$  be a metric space and  $\mu$  be a measure on  $(X, \mathcal{B}(X))$  such that  $\mu(K) < \infty$  for all compact sets  $K \subset X$ .

1 If  $X$  is  $\sigma$ -compact, then  $\mu$  is inner regular.

2 If there exists a sequence  $G_n \in \mathcal{O}$ ,  $G_n \uparrow X$  such that  $\mu(G_n) < \infty$ , then  $\mu$  is outer regular.

## 11 The Function Spaces $\mathcal{L}^p$ (13, [Schilling(2017)])

Assume  $V$  is a vector space over  $\mathbb{K} \in \{\mathbb{C}, \mathbb{R}\}$ .

**Definition 11.1.** A seminorm on  $V$  is a map  $p : V \rightarrow [0, +\infty)$  s.t.

$$(1) p(cx) = |c|p(x) \quad \forall x \in V, \forall c \in \mathbb{K}.$$

$$(2) p(x+y) \leq p(x) + p(y) \quad \forall x, y \in V. \quad \text{triangle inequality.}$$

A seminorm is called a norm if we also have

$$p(x) = 0 \iff x = 0.$$

A norm is commonly denoted  $\|x\|$ , and a vectorspace equipped with a norm is called a **normed space**.

**Definition 11.2.** Assume  $(X, d)$  is a measure space. Fix  $1 \leq p \leq \infty$ . For every measurable function  $f : X \rightarrow \mathbb{C}$  we define the following

$$\|f\|_p = \left( \int_X |f|^p d\mu \right)^{1/p} \in [0, +\infty]. \quad (20)$$

We can see that  $\|cf\|_p = |c|\|f\|_p \quad \forall c \in \mathbb{C}$ .

Notice that by Theorem 8.2(i) we have that  $\|f\|_p = 0 \Rightarrow f = 0$  a.e. Consider for example  $\lim_{n \rightarrow \infty} \|f_n - f\|_p = 0$ , then we can find a subsequence s.t.  $\lim_{k \rightarrow \infty} |f_{n(k)} - f| = 0$  a.e., i.e.  $\lim_{k \rightarrow \infty} f_{n(k)} = f$  a.e.

**Lemma 11.3.**

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p. \quad (21)$$

**Definition 11.4.** We define

$$\mathcal{L}^p(X, d\mu) = \{ f : X \rightarrow \mathbb{C} \mid f \text{ is measurable and } \|f\|_p < \infty \}.$$

This is a vectorspace with seminorm  $f \mapsto \|f\|_p$ . And in general this is not a normed space, since  $\|f\|_p = 0 \iff f = 0$  a.e.

Generally, if  $p$  is a seminorm on a vectorspace  $V$ , then

$$V_0 = \{ x \in V \mid p(x) = 0 \} \quad (22)$$

which is a subspace of  $V$ . Then we consider the quotient/factor space  $V/V_0$ .

**Definition 11.5.** For  $x, y \in V$ , define

$$x \sim y \iff x - y \in V_0. \quad (23)$$

This is an equivalence relation on  $V$ . The representation class of  $V$  is defined by  $[x]$  or  $x + V_0$ .

Then  $V/V_0$  equals the set of equivalence classes. We can show that it is a normed space.

$$[x] + [y] = [x + y], \quad c[x] = [cx], \quad \|[x]\| = p(x).$$

Applying this to  $\mathcal{L}^p(X, d\mu)$  we get the normed space

$$L^p(X, d\mu) := \mathcal{L}^p(X, d\mu)/\mathcal{N} = \mathcal{L}^p(X, d\mu)/\sim. \quad (24)$$

Where  $\mathcal{N}$  is the space of measurable functions  $f$  s.t.  $f = 0$  a.e. The equivalence relation  $\sim$  is defined by

$$u \sim v \iff \{u \neq v\} \in \mathcal{N}_\mu \iff \mu\{u \neq v\} = 0,$$

and so  $L^p(X, d\mu)$  consists of all equivalence classes  $[u]_p = \{v \in \mathcal{L}^p \mid u \sim v\}$ . So for every  $u \in L^p$  there is no  $v \in L^p$  such that  $\mu\{u \neq v\} \neq 0$ .

We will further continue to denote the norm by  $\|\cdot\|_p$ , and we will normally **not** distinguish between  $f \in \mathcal{L}^p(X, d\mu)$  and the vector in  $L^p(X, d\mu)$  that  $f$  defines.

**Definition 11.6.** A normed space  $(X, \|\cdot\|)$  is called a Banach space if  $V$  is complete w.r.t the metric  $d(x, y) = \|x - y\|$ .

**Theorem 11.7.** If  $(X, \mathcal{B}, \mu)$  is a measure space,  $1 \leq p \leq \infty$ , then  $L^p(X, d\mu)$  is a Banach space.

**Definition 11.8.** A measurable function  $f : X \rightarrow \mathbb{C}$  is called **essentially bounded** if there is  $c \geq 0$  s.t.

$$\mu(\{x : |f(x)| > c\}) = 0. \quad (25)$$

That is  $|f| \leq c$  a.e. The smallest such  $c$  is called the essential supremum of  $f$  and is denoted by  $\|f\|_\infty$ .

**Definition 11.9.**

$$\mathcal{L}^\infty(X, d\mu) = \{f : X \rightarrow \mathbb{C} \mid f \text{ is measurable and } \|f\|_\infty < \infty\}.$$

$$L^\infty(X, d\mu) = \mathcal{L}^\infty(X, d\mu)/\mathcal{N}.$$

Where by the previous definition these spaces become the spaces of all essentially bounded functions.

**Theorem 11.10.** If  $(X, \mathcal{B}, \mu)$  is a  $\sigma$ -finite measure space, then  $L^\infty(X, d\mu)$  is a Banach space.

## Convergence in $\mathcal{L}^p$ and completeness

**Lemma 11.11.** For any sequence  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{L}^p(\mu)$ ,  $p \in [1, \infty)$ , of positive functions  $u_n \geq 0$  we have

$$\left\| \sum_{n=1}^{\infty} u_n \right\|_p \leq \sum_{n=1}^{\infty} \|u_n\|_p.$$

**Theorem 11.12** (Riesz-Fischer). The spaces  $\mathcal{L}^p(\mu)$ ,  $p \in [1, \infty)$ , are complete, i.e. every Cauchy sequence  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{L}^p(\mu)$  converges to some limit  $u \in \mathcal{L}^p(\mu)$

**Corollary 11.13.** Let  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{L}^p(\mu)$ ,  $p \in [1, \infty)$  with  $\mathcal{L}^p - \lim_{n \rightarrow \infty} u_n = u$ . Then there exists a subsequence  $(u_{n_k})_{k \in \mathbb{N}}$  s.t.  $\lim_{k \rightarrow \infty} u_{n_k}(x) = u(x)$  holds for almost every  $x \in X$ .

**Theorem 11.14.** Let  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{L}^p(\mu)$ ,  $p \in [1, \infty)$ , be a sequence of functions s.t.  $|u_n| \leq w \forall n \in \mathbb{N}$  and some  $w \in \mathcal{L}^p(\mu)$ . If  $u(x) = \lim_{n \rightarrow \infty} u_n(x)$  exists for (almost) every  $x \in X$ , then

$$u \in \mathcal{L}^p \text{ and } \lim_{n \rightarrow \infty} \|u - u_n\|_p = 0.$$

**Theorem 11.15** (F. Riesz (convergence theorem)). Let  $(u_n)_{n \in \mathbb{N}} \subset \mathcal{L}^p(\mu)$ ,  $p \in [1, \infty)$ , be a sequence s.t.  $\lim_{n \rightarrow \infty} u_n(x) = u(x)$  for almost every  $x \in X$  and some  $u \in \mathcal{L}^p(\mu)$ . Then

$$\lim_{n \rightarrow \infty} \|u_n - u\|_p = 0 \iff \lim_{n \rightarrow \infty} \|u_n\|_p = \|u\|_p.$$

## 12 Dense and Determining Sets (17, [Schilling(2017)])

**Definition 12.1** (Dense Sets). A set  $\mathcal{D} \subset \mathcal{L}^p(\mu)$ ,  $p \in [0, \infty]$ , is called *dense* if for every  $u \in \mathcal{L}^p(\mu)$  there exist a sequence  $(f_n)_{n \in \mathbb{N}} \subset \mathcal{D}$  s.t.  $\lim_{n \rightarrow \infty} \|u - f_n\|_p = 0$ .

**Theorem 12.2.** Assume  $X, d$  is a metric space and  $\mu$  is a Borel measure that is finite on every ball  $1 \leq p < \infty$ . Then the space of bounded continuous functions with bounded support is dense in  $\mathcal{L}^p(X, d\mu)$ . Where bounded support means that  $f$  vanishes outside some ball.

*Proof.* We want to approximate  $f \in \mathcal{L}^p(X, d\mu)$  by bounded continuous functions with bounded support. By considering separately  $(\operatorname{Re}(f))_I$  and  $(\operatorname{Im}(f))_I$  we may assume that  $f \geq 0$ . Then we can find simple functions  $f_n$  s.t.  $0 \leq f_n \leq f$ ,  $f_n \rightarrow f$  pointwise. As  $|f - f_n|^p \leq |f|^p$ , by the dominated convergence theorem we have  $f_n \rightarrow f \in \mathcal{L}^p(X, d\mu)$ . Hence, it suffices to consider simple  $f$ , but then it suffices to approximate  $f = \pi_A$ . Note that  $\pi_A \in \mathcal{L}^p(X, d\mu)$  iff  $\mu(A) < \infty$ .

Fix  $x_0 \in X$ . Then  $\pi_{A \cap B_n(x_0)} \nearrow \pi_A$  pointwise, hence  $\pi_{A \cap B_n(x_0)} \rightarrow \pi_A \in \mathcal{L}^p(X, d\mu)$ , again by the dominated convergence theorem.

Therefor it suffices to consider  $A \subset B_n(x_0)$ . As  $\mu$  is outer regular, we have

$$\mu(A) = \inf_{\substack{A \subset U \subset B_n(x_0) \\ U \text{ is open}}} \mu(U).$$

Note that  $\|\pi_U - \pi_A\|_p = \mu(U \setminus A)^{1/p}$ . Hence, we can choose  $U_k \subset B_n(x_0)$  s.t.  $A \subset U_k$ ,  $U_k$  is open,  $\pi_{U_k} \rightarrow \pi_A \in \mathcal{L}^p(X, d\mu)$ .

Therefor it suffices to approximate  $\pi_U$  for open  $U \subset B_n(x_0)$ . Consider the functions

$$f_k(x) = \frac{kd(x, U^c)}{1 + kd(x, U^c)}.$$

Then  $0 \leq f_k \leq 1$ ,  $f_k$  is continuous, supported on  $\bar{U} \subset \bar{B}_n(x_0)$  and  $f_k \nearrow \pi_U$  pointwise, hence  $f_k \xrightarrow[k \rightarrow \infty]{} \pi_U \in \mathcal{L}^p(X, d\mu)$ .  $\square$

**Theorem 12.3.** Assume  $(X, d)$  is a separable locally compact metric space and  $\mu$  is a Borel Measure on  $X$  s.t.  $\mu(K) < \infty \forall$  compact  $K \subset X$ . Then the space  $C_c(X)$  of continuous compactly supported functions is dense in  $\mathcal{L}^p(X, d\mu)$ .

Recall that the support of a function  $f$  is  $\operatorname{supp}(f) = \{x \in X : f(x) \neq 0\}$ , *closed support* is the closure of  $\operatorname{supp}(f)$  (i.e. boundary points are included), often just written as  $\operatorname{supp}(f)$ , and a function is said to have *compact support* if  $\operatorname{supp}(f)$  is compact.

In particular, either theorem shows that if  $\mu$  is a Borel measure on  $\mathbb{R}^n$  s.t. the measure of every ball is finite, then  $C_c(\mathbb{R}^n)$  is dense in  $\mathcal{L}^p(\mathbb{R}^n, d\mu)$ ,  $1 \leq p < \infty$ . Later we will see that even  $C_c^\infty(\mathbb{R}^n)$  is dense in  $L^p(\mathbb{R}^n, d\mu)$ .



**Remark.** These results do not extend to  $p = \infty$  in general.

For  $\mu = \lambda_n$  we write simply  $\mathcal{L}^p(\mathbb{R}^n)$ .

**Remark.** Theorem 17.8 in the book is *WRONG*. For example,  $X = \mathbb{Q}$  with the usual metric is  $\sigma$ -compact, supports nonzero finite measure, but  $C_c(\mathbb{Q}) = 0$ .

## Modes of Convergence

(mixture of ex. 11.12 and ch. 22 p. 258-261. in [Schilling(2017)])

**Definition 12.4** (convergence in measure). A sequence of measurable functions  $u_n : X \rightarrow \bar{\mathbb{R}}$  converges in measure if

$$\forall \epsilon > 0 \forall A \in \mathcal{A}, \mu(A) < \infty : \lim_{n \rightarrow \infty} \mu(\{|u_n - u| > \epsilon\} \cap A) = 0$$

holds for some  $u \in \mathcal{M}(\mathcal{A})$ . We write  $\mu\text{-}\lim_{n \rightarrow \infty} u_n = u$  or  $u_n \xrightarrow{\mu} u$ .

Assume  $(X, \mathcal{B}, \mu)$  is a measure space. Given measurable functions  $f_n, f : X \rightarrow \mathbb{C}$ , recall that

$$f_n \rightarrow f \text{ a.e.}$$

means that  $f_n(x) \xrightarrow[n \rightarrow \infty]{} f(x)$  for all  $x$  outside a set of measure zero.

**Theorem 12.5** (Egorov). Assume  $\mu(X) < \infty$  and  $f_n \rightarrow f$  a.e. Then,  $\forall \epsilon > 0$ , there exists  $X_\epsilon \in \mathcal{B}$  s.t.  $\mu(X_\epsilon) < \epsilon$  and  $f_n \rightarrow f$  uniformly on  $X \setminus X_\epsilon$ .

In addition to pointwise and uniform convergence we also consider the following:

$f_n \rightarrow f$  in the  $p$ -th mean if  $\|f_n - f\|_p \xrightarrow[n \rightarrow \infty]{} 0$ . For  $p = 1$  we say in mean, for  $p = 2$  we say in quadratic mean.

$f_n \rightarrow f$  in measure if  $\forall \epsilon > 0$  we have

$$\mu(\{x : |f_n(x) - f(x)| \geq \epsilon\}) \xrightarrow[n \rightarrow \infty]{} 0.$$

**Theorem 12.6** (Lemma 22.4 in the book?). Assume  $(X, \mathcal{B}, d\mu)$  is a measure space,  $1 \leq p < \infty$ ,  $f_n, f : X \rightarrow \mathbb{C}$  are measurable functions. Then

- (i) If  $f_n \rightarrow f$  in the  $p$ -th mean, then  $f_n \rightarrow f$  in measure.
- (ii) If  $f_n \rightarrow f$  in measure, then there is a subsequence  $(f_{n_k})_{k=1}^\infty$  s.t.  $f_{n_k} \rightarrow f$  a.e.
- (iii) If  $f_n \rightarrow f$  a.e. and  $\mu(X) < \infty$ , then  $f_n \rightarrow f$  in measure.

In particular, if  $f_n \rightarrow f$  in the  $p$ -th mean, then  $f_{n_k} \rightarrow f$  a.e. for a subsequence  $(f_{n_k})_k$ .

## 13 Abstract Hilbert Spaces (26, [Schilling(2017)])

Assume  $\mathcal{H}$  is a vector space over  $\mathbb{C}$ .

**Definition 13.1.** A pre-inner product on  $\mathcal{H}$  is a map  $(\cdot, \cdot) : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{C}$  which is

- (i) Sesquilinear: linear in the first variable and antilinear in the second:

$$(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w),$$

$$(w, \alpha u + \beta v) = \bar{\alpha}(w, u) + \bar{\beta}(w, v), \quad u, v, w \in \mathcal{H} \text{ and } \alpha, \beta \in \mathbb{C}.$$

- (ii) Hermitian:  $(u, v) = \overline{(v, u)}$ .

- (iii) Positive semidefinite:  $(u, u) \geq 0$ .

It is called an **inner product**, or a scalar product, if instead of (iii) the map is positive definite;  $(u, u) > 0$ . This definition also works for  $\mathbb{R}$  instead of  $\mathbb{C}$ .

**Cauchy-Schwartz inequality** If  $(\cdot, \cdot)$  is a pre-inner product, then  $|(u, v)| \leq (u, u)^{1/2}(v, v)^{1/2}$ .

**Corollary 13.2.** Assume we have a seminorm  $\|u\| := (u, u)^{1/2}$ . It is a norm iff  $(\cdot, \cdot)$  is an inner product.

**Definition 13.3** (Hilbert space). A Hilbert space is a complex vector space  $\mathcal{H}$  with an inner product  $(\cdot, \cdot)$  s.t.  $\mathcal{H}$  is complete with respect to the norm  $\|u\| = (u, u)^{1/2}$ .

- The norm on a Hilbert space is determined by the inner product, but the inner product can also be recovered by the norm by the *polarization identity*:  $(u, v) = \frac{\|u+v\|^2 - \|u-v\|^2}{4} + i \frac{\|u+iv\|^2 - \|u-iv\|^2}{4}$ .
- Parallelogram law*:  $\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2$ .
- A norm on a vector space is given by an inner product iff it satisfies the parallelogram law, and then the scalar product is uniquely determined by the polarization identity.

**Example 13.4.** Assume  $(X, \mathcal{B}, \mu)$  is a measure space. Then  $\mathcal{L}^2(X, d\mu)$  is a Hilbert space with inner product

$$(f, g) = \int_X f \bar{g} d\mu.$$

This is well-defined, as  $|f \bar{g}| \leq \frac{1}{2}(|f|^2 + |g|^2)$ .

In particular, if  $\mathcal{B} = \mathcal{P}(X)$  and  $\mu$  is the counting measure, then  $L^2(X, d\mu)$  is denoted by  $l^2(X)$ ; for  $X = \mathbb{N}$  we write simply  $l^2$ . Note that in this case for  $f : X \rightarrow [0, +\infty]$  we have

$$\int_X f d\mu = \sum_{x \in X} f(x) := \sup_{\substack{F \subset X \\ F \text{ is finite}}} \sum_{x \in F} f(x),$$

and if  $\sum_{x \in X} f(x) < \infty$ , then  $\{x : f(x) > 0\}$  is at most countable, so  $\sum_{x \in X} f(x) = \sum_{x: f(x) > 0} f(x)$  is the usual sum of a series.

Recall that a subset  $\mathcal{C}$  of a vector space is called *convex* if

$$u, w \in \mathcal{C} \rightarrow tu + (1-t)w \in \mathcal{C} \quad \forall t \in (0, 1).$$

The following is one of the key properties of the Hilbert space

**Theorem 13.5** (projection theorem). Assume  $\mathcal{H}$  is a Hilbert space and  $\mathcal{C} \subset \mathcal{H}$  is a closed convex subset. Then for every  $u \in \mathcal{H}$  there is a unique  $u_0 \in \mathcal{C}$  (minimizer) s.t.

$$\|u - u_0\| = d(u, \mathcal{C}) (= \inf_{x \in \mathcal{C}} \|u - x\|).$$

*Proof.* Let  $d = d(u, \mathcal{C})$ . Choose  $u_n \in \mathcal{C}$  s.t.  $\|u - u_n\| \rightarrow d$ . We claim that  $(u_n)_{n=1}^\infty$  is a Cauchy sequence. As  $\frac{u_n + u_m}{2} \in \mathcal{C}$ , we have

$$\begin{aligned} d^2 &\leq \|u - \frac{u_n + u_m}{2}\|^2 = \frac{1}{4} \|(u - u_n) + (u - u_m)\|^2 \\ &= \frac{1}{4} (2\|u - u_n\|^2 + 2\|u - u_m\|^2) \end{aligned}$$

parallelogram law

so

$$\|u_n - u_m\|^2 \leq 2(\|u - u_n\|^2 - d^2) + 2(\|u - u_m\|^2 - d^2).$$

Thus,  $(u_n)_{n \in \mathbb{N}}$  is indeed Cauchy, hence  $u_n \rightarrow u_0 \in \mathcal{C}$  for some  $u_0$  and

$$\|u - u_0\| = \lim_{n \rightarrow \infty} \|u - u_n\| = d = d(u, \mathcal{C}).$$

If  $u'_0$  is another such point, we can take  $u_{2n} = u_0, u_{2n+1} = u'_0$  and conclude that  $u_0 = u'_0$ .  $\square$

For a Hilbert space  $\mathcal{H}$  and a subset  $A \subset H$ , let

$$A^\perp := \{x \in H : x \perp y \ \forall y \in A\},$$

where  $x \perp y$  means that  $(x, y) = 0$ .  $A^\perp$  is a closed subspace of  $\mathcal{H}$ .

**Proposition 14.1.** Assume  $\mathcal{H}_0$  is a closed subspace of a Hilbert space  $\mathcal{H}$ . Then every  $u \in H$  uniquely decomposes as

$$u = u_0 + u_1, \text{ with } u_0 \in \mathcal{H}_0 \text{ and } u_1 \in \mathcal{H}_0^\perp.$$

Moreover,  $\|u - u_0\| = d(u, \mathcal{H}_0)$  and  $\|u\|^2 = \|u_0\|^2 + \|u_1\|^2$ .

For a closed subspace  $\mathcal{H}_0 \subset \mathcal{H}$ , consider the map  $P : H \rightarrow \mathcal{H}_0$  s.t.  $Pu \in \mathcal{H}_0$  is the unique element satisfying  $u - Pu = H_0^\perp$ . The operator  $P$  is linear. It is also contractive, meaning that  $\|Pu\| \leq \|u\|$ , since  $\|u\|^2 = \|Pu\|^2 + \|u - Pu\|^2$ . It is called the orthogonal projection onto  $\mathcal{H}_0$ .

If  $\mathcal{H}_0$  is finite dimensional with an orthonormal basis  $u_1, \dots, u_n$  then

$$Pu = \sum_{k=1}^n (u, u_k) u_k.$$

Orthonormal bases can be defined for arbitrary Hilbert spaces.

**Definition 14.2** (orthonormal system). An orthonormal system in  $\mathcal{H}$  is a collection of vectors  $u_i \in H$  ( $i \in I$ ) s.t.

$$(u_i, u_j) = \delta_{ij} \ \forall i, j \in I.$$

It is called an *orthonormal basis* if  $\text{span}\{u_i\}_{i \in I}$  denotes the linear span of  $\{u_i\}_{i \in I}$ , the space of finite linear combinations of the vectors  $u_i$ .

**Definition 14.3.** A Hilbert space  $\mathcal{H}$  is said to be *separable* if  $\mathcal{H}$  contains a countable dense subset  $G \subset \mathcal{H}$ .

**Theorem 14.4.** Every Hilbert space  $\mathcal{H}$  has an orthonormal basis. If  $\mathcal{H}$  is separable, then there is a countable orthonormal basis.

**Proposition 14.5.** Assume  $\{u_i\}_{i \in I}$  is an orthonormal system in a Hilbert space  $H$ . Take  $u \in \mathcal{H}$ . Then

- (i) Bessel's inequality:  $\sum_{i \in I} |(u, u_i)|^2 \leq \|u\|^2$ , in particular,  $\{i : (u, u_i) \neq 0\}$  is countable.
- (ii) Parseval's identity: If  $\{u_i\}_{i \in I}$  is an orthonormal basis, then  $\sum_{i \in I} |(u, u_i)|^2 = \|u\|^2$ .

If  $(u_i)_{i \in I}$  is an orthonormal basis, then the numbers  $(u, u_i)$  are called the **Fourier coefficients** of  $u$  with respect to  $(u_i)_{i \in I}$ . The Parseval identity then suggests that  $u$  is determined by its Fourier coefficients. This is true, and even more, we have:

**Proposition 14.6.** Assume  $(u_i)_{i \in I}$  is an orthonormal basis in a Hilbert space  $\mathcal{H}$ . Then for every vector  $(c_i)_{i \in I} \in l^2(I)$  there is a unique vector  $u \in \mathcal{H}$  with Fourier coefficients  $c_i$ , and we write

$$u = \sum_{i \in I} c_i u_i.$$

**Remark.** Equivalently, the element  $u = \sum_{i \in I} c_i u_i$  can be described as the unique element in  $\mathcal{H}$  s.t.  $\forall \epsilon > 0$  there is a finite  $F_0 \subset I$  s.t.  $\|u - \sum_{i \in F} c_i u_i\| < \epsilon \ \forall \text{ finite } F \supset F_0$ .

**Corollary 14.7.** We have a linear isomorphism  $U : l^2(I) \xrightarrow{\sim} \mathcal{H}$ ,  $U((c_i)_{i \in I}) = \sum_{i \in I} c_i u_i$ . By Parseval's identity this isomorphism is isometric, that is,  $\|Ux\| = \|x\| \ \forall x \in l^2(I)$ . By the polarization identity this is equivalent to

$$(Ux, Uy) = (x, y) \ \forall x, y \in l^2(I).$$

Therefor  $U$  is unitary.

**Corollary 14.8.** Up to a unitary isomorphism, there is only one infinite dimensional separable Hilbert space, namely,  $l^2$ .

## 15 Dual spaces (26, [Schilling(2017)])

Given two orthonormal bases  $(u_i)_{i \in I}$  and  $(v_i)_{i \in I}$  in a Hilbert space  $\mathcal{H}$ , we can decompose

$$u_i = \sum_{j \in I} (u_i, v_j) v_j$$

and using that the sets  $\{j : (u_i, v_j) \neq 0\}$  are countable prove the following:

*Claim:* Any two orthonormal bases in a Hilbert space have the same cardinality.

**Example 15.1** (classical Fourier series). Consider  $\mathcal{H} = L^2(0, 2\bar{\mu}) = L^2((0, 2\bar{\mu}), d\lambda)$ . For  $n \in \mathbb{Z}$ , define  $e_n(t) = \frac{1}{\sqrt{2\bar{\mu}}} e^{int}$ . By a version of Weierstrass' theorem it is known that  $\text{span}\{e_n\}_{n \in \mathbb{Z}}$  is dense in the supremum-norm in

$$\{f \in C[0, 2\bar{\mu}] : f(0) = f(2\bar{\mu})\}.$$

As  $C[0, 2\bar{\mu}]$  is dense in  $L^2(0, 2\bar{\mu})$ , from this one can deduce that  $\text{span}\{e_n\}_{n \in \mathbb{Z}}$  is dense in  $L^2(0, 2\bar{\mu})$ . We then see that  $(e_n)_{n \in \mathbb{Z}}$  is an orthonormal basis in  $L^2(0, 2\bar{\mu})$ . We therefor have a unitary isomorphism

$$l^2(\mathbb{Z}) \xrightarrow{\sim} L^2(0, 2\bar{\mu}), (c_n)_{n \in \mathbb{Z}} \mapsto \sum_{n \in \mathbb{Z}} c_n e_n.$$

The Fourier coefficients at  $f \in L^2(0, 2\bar{\mu})$  are denoted by  $\hat{f}(n)$ , so

$$\hat{f}(n) = \frac{1}{\sqrt{2\bar{\mu}}} \int_0^{2\bar{\mu}} f(t) e^{-int} dt,$$

more practically

$$\hat{f}(n) = \frac{1}{\sqrt{2\bar{\mu}}} \int_{[0, 2\bar{\mu})} f(t) e^{-int} d\lambda(t).$$

Therefor we have  $f = \sum_{n \in \mathbb{Z}} \hat{f}(n) e_n$  in  $L^2(0, 2\bar{\mu})$ .

*Fact:* For every  $f \in L^2(0, 2\bar{\mu})$ , we have  $\frac{1}{\sqrt{2\bar{\mu}}} \sum_{n=-N}^N \hat{f}(n) e^{int} \xrightarrow[N \rightarrow \infty]{} f(t)$  for a.e.  $t$ .

**Lemma 15.2.** Assume  $V$  is a normed space over  $K = \mathbb{R}$  or  $K = \mathbb{C}$ . Consider a linear functional  $f : V \rightarrow K$ . The following are equivalent (TFAE):

- (i)  $f$  is continuous;
- (ii)  $f$  is continuous at 0;
- (iii) There is a  $c \geq 0$  s.t.  $|f(x)| \leq c\|x\| \ \forall x \in V$ .

If (i)-(iii) are satisfied, then  $f$  is called a *bounded linear functional*. The constant  $c$  in (iii) is denoted by  $\|f\|$ . We have  $\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)| = \sup_{\|x\| \leq 1} |f(x)|$ .

**Proposition 15.3.** For every normed vector space  $V$  over  $K = \mathbb{R}$  or  $K = \mathbb{C}$ , the bounded linear functionals on  $V$  form a Banach space  $V^*$ .

**Remark.** The sequence  $\{\|f_n - f_m\|\}_{m=1}^\infty$  actually converges, since

$$\|f_n - f_m\| \leq \|f_m - f_n\|.$$

When we study/use normed spaces, it is often important to understand the dual spaces. For Hilbert spaces this is particularly easy:

**Theorem 15.4** (Riesz). Assume  $\mathcal{H}$  is a Hilbert space. Then every  $f \in \mathcal{H}^*$  has the form

$$f(x) = (x, y),$$

for a uniquely defined  $y \in \mathcal{H}$ . Moreover, we have  $\|f\| = \|y\|$ .

For every Hilbert space  $\mathcal{H}$  we can define the *conjugate Hilbert space*  $\bar{\mathcal{H}}$ , which has its elements as the symbols  $\bar{x}$  for  $x \in \mathcal{H}$ , with the linear structure and inner product defined by

$$\bar{x} + \bar{y} = \overline{x + y}, c \cdot \bar{x} = \overline{c x}, (\bar{x}, \bar{y}) = \overline{(x, y)} = (y, x).$$

**Corollary 15.5.** For every Hilbert space  $\mathcal{H}$ , we have an isometric isomorphism  $\bar{\mathcal{H}} \xrightarrow{\sim} \mathcal{H}^*$ ,  $\bar{x} \mapsto (\cdot, x)$ .

## 16 Hahn-Banach Theorem (4.2, [Teschl(2010)])

**Theorem 16.1** (Hahn-Banach). Assume  $V$  is a real vector space,  $V_0 \subset V$  a subspace,  $e : V \rightarrow \mathbb{R}$  a convex function and  $f : V_0 \rightarrow \mathbb{R}$  a linear functional s.t.  $f \leq e$  on  $V_0$ . Then  $f$  can be extended to a linear functional  $F$  on  $V$  s.t.  $F \leq e$ .

Assume first that  $V = V_0 + \mathbb{R}x$  for some  $x \in V \setminus V_0$ . To define  $F$  we need to specify  $F(x)$ . The condition  $F \leq e$  means that we need

$$F(y \pm tx) \leq e(y \pm tx) \quad \forall y \in V_0, t \geq 0,$$

that is

$$\pm tF(x) + f(y) \leq e(y \pm tx).$$

Dividing by  $t$  this is equivalent to

$$\begin{cases} F(x) \leq \frac{e(y+tx) - f(y)}{t}, \\ F(x) \geq \frac{f(y) - e(y-tx)}{t}, \end{cases}$$

$\forall y \in V_0, t \geq 0$ . To show that there is a number satisfying these inequalities, we need to check that

$$\sup_{\substack{y \in V_0 \\ t > 0}} \frac{f(y) - e(y - tx)}{t} \leq \inf_{\substack{y \in V_0 \\ t > 0}} \frac{e(y + tx) - f(y)}{t}.$$

Then as  $F(x)$  we can take any number in the interval  $[\sup, \inf]$ . In other words, we need to check that

$$\frac{f(y) - e(y - tx)}{t} \leq \frac{e(z + sx) - f(z)}{s},$$

$\forall y, z \in V_0, t, s \geq 0$ . Equivalently,

$$sf(y) + tf(z) \leq se(y - tx) + te(z + sx).$$

We have

$$\begin{aligned} \frac{sf(y) + tf(z)}{s+t} &= f\left(\frac{s}{s+t}y + \frac{t}{s+t}z\right) \\ &\leq e\left(\frac{s}{s+t}y + \frac{t}{s+t}z\right) \\ &\leq e\left(\frac{s}{s+t}(y - tx) + \frac{t}{s+t}(z + sx)\right) \\ &\leq \frac{s}{s+t}e(y - tx) + \frac{t}{s+t}e(z + sx), \end{aligned}$$

which is what we need.

The general case is deduced from this using transitive induction (Zorn's lemma), as follows. Consider the set  $X$  of pairs  $(W, F)$ , where  $W \subset V$  is a subspace,  $F : W \rightarrow \mathbb{R}$  a linear functional,  $V_0 \subset W$  and  $F|_{V_0} = f$ . Define a partial order on  $X$  by  $F \leq e$  on  $W$ .

$$(W_1, F_1) \leq (W_2, F_2) \text{ iff } W_1 \subset W_2 \text{ and } F_1 = F_2|_{W_1}.$$

If  $\mathcal{C}$  is a chain in  $X$ , then it has an upper bound  $(W, F)$  defined by

$$W = \bigcup_{(Y, \rho) \in \mathcal{C}} Y, F(y) = \rho(y) \text{ if } y \in Y, (Y, \rho) \in \mathcal{C}.$$

Hence,  $X$  has a maximal element  $(W, F)$  by Zorn's lemma. If  $W \neq V$ , then we can take  $x \in V \setminus W$  and extend  $F$  to  $W + \mathbb{R}x$  preserving the inequality  $F \leq e$ . This contradicts maximality of  $W, F$  in  $X$ . Hence,  $W = V$ .

**Theorem 16.2** (Hahn-Banach). Assume  $V$  is a real or complex vector space,  $p$  a seminorm on  $V_0$ ,  $V_0 \subset V$ , and  $f$  a linear functional on  $V_0$  s.t.

$$|f(x)| \leq p(x) \quad \forall x \in V_0.$$

Then  $f$  can be extended to a linear functional  $F$  on  $V$  s.t.  $|F(x)| \leq p(x) \quad \forall x \in V$ .

**Corollary 16.3.** Assume  $V$  is a normed real or complex vector space,  $V_0 \subset V$  and  $f \in V_0^*$ . Then there is a  $F \in V^*$  s.t.

$$F|_{V_0} = f \text{ and } \|F\| = \|f\|.$$

*Proof.* We apply the previous theorem to  $p(x) = \|f\| \cdot \|x\|$ .  $\square$

**Corollary 16.4.** Assume  $V$  is a normed space and  $x \in V, x \neq 0$ . Then there is a  $F \in V^*$  s.t.  $\|F\| = 1$  and  $F(x) = \|x\|$ .

Such an  $F$  is called a *supporting functional* at  $x$ .

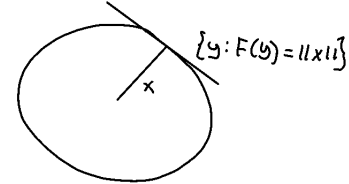


Figure 1: Tangent space?

If  $V$  is a normed vector space, then every  $x \in X$  defines a bounded linear functional on  $V^*$  by

$$V^* \ni f \mapsto f(x).$$

As  $|f(x)| \leq \|f\| \cdot \|x\|$ , this functional has norm  $\leq \|x\|$ . By using a supporting functional at  $x$ , we actually see that we get norm  $\|x\|$ . Thus, we have an isometric embedding  $V \subset V^{**} := (V^*)^*$ . We can therefore see  $V$  as a subspace of  $V^{**}$ .

**Definition 16.5.** A normed space  $V$  is called *reflexive* if  $V^{**} = V$ .

**Remark.** This is stronger than requiring  $V \simeq V^{**}$ .

**Remark.** Every Hilbert space  $\mathcal{H}$  is reflexive. Indeed,  $\mathcal{H}^* = \bar{\mathcal{H}}$ . By Riesz' theorem every bounded linear functional  $f$  on  $\bar{\mathcal{H}}$  has the form

$$f(\bar{x}) = (\bar{x}, \bar{y}) = (y, x),$$

for some  $y \in \mathcal{H}$ , which exactly means that  $f = y$  in  $\mathcal{H}^{**}$ .

As we will see later, the spaces  $\mathcal{L}^p(X, d\mu)$ , with  $\mu$   $\sigma$ -finite and  $1 < p < \infty$ , are reflexive. The spaces  $\mathcal{L}'(X, d\mu)$  and  $\mathcal{L}^\infty(X, \mu)$  are usually not reflexive.

Assume  $(X, \mathcal{B}, \mu)$  is a measure space. Are there other measures on  $(X, \mathcal{B})$ ?

**Example 17.1.** Take a measurable function  $f : X \rightarrow [0, +\infty]$  and define

$$\nu(A) = \int_A f d\mu \text{ for } A \in \mathcal{B}.$$

This is a measure by the monotone convergence theorem. We write  $d\nu = f d\mu$ .

**Proposition 17.2.** Assume  $(X, \mathcal{B})$  is a measurable space,  $\mu$  and  $\nu$  are  $\sigma$ -finite measures on  $(X, \mathcal{B})$ . Then there exist  $N \in \mathcal{B}$  and a measurable  $f : X \rightarrow [0, +\infty]$  s.t.  $\mu(N) = 0$  and  $\nu(A) = \nu(A \cap N) + \int_A f d\mu \forall A \in \mathcal{B}$ .

*Proof.* Assume first that the measure  $\nu, \mu$  are finite. Consider the measure  $\eta = \mu + \nu$  and define a linear functional  $\rho$  on  $L^2(x, d\eta)$  by

$$\rho(g) = \int_X g d\nu.$$

It is well-defined and bounded, since

$$\begin{aligned} |\rho(g)| &\leq \underbrace{\int_X |\rho| d\nu}_{\|g\|_1} \leq \frac{1}{\sqrt{\eta(X)}} \left( \int_X |\rho|^2 d\nu \right)^{\frac{1}{2}} \nu(X)^{\frac{1}{2}} \\ &\leq \nu(X)^{\frac{1}{2}} \left( \int_X |\rho|^2 d\eta \right)^{\frac{1}{2}}. \end{aligned}$$

By Riesz' theorem there exists  $h \in L^2(x, d\eta)$  s.t.  $\rho(g) = \int_X g h d\eta$  for all  $g \in L^2(x, d\eta)$ .

For  $\rho = \mathbb{1}_A$  we get

$$\nu(A) = \int_A h d\eta \forall A \in \mathcal{B}.$$

In particular,  $h \geq 0$  ( $\eta$ -a.e.). As  $\nu(A) \leq \eta(A)$ , we also have  $h \leq 1$  ( $\eta$ -a.e.). From now on we view  $h$  as a function on  $X$  and assume  $0 \leq h \leq 1$ .

For  $g = \mathbb{1}_A$  ( $A \in \mathcal{B}$ ) we have

$$\int_X g d\nu = \rho(g) = \int_X g h d\eta = \int_X g h d\mu + \int_X g h d\nu,$$

hence

$$\int_X g(1-h) d\nu = \int_X g h d\mu. \quad (26)$$

By extending eq. 26 to positive simple functions and then using the monotone convergence theorem, we conclude that eq. 26 holds for all measurable  $g : X \rightarrow [0, +\infty]$ .

We now let

$$N = \{x : h(x) = 1\} \text{ and } f = \frac{h}{1-h} \mathbb{1}_{N^c}.$$

Letting  $g = \mathbb{1}_N$  in eq. 26 we get  $0 = \mu(N)$ . For  $A \in \mathcal{B}$ , letting  $g = \frac{\mathbb{1}_{A \cap N^c}}{1-h}$  in eq. 26 we get

$$\nu(A \cap N^c) = \int_X \frac{\mathbb{1}_A \mathbb{1}_{N^c}}{1-h} h d\mu = \int_X \mathbb{1}_A f d\mu = \int_A f d\mu.$$

Thus

$$\nu(A) = \nu(A \cap N) + \nu(A \cap N^c) = \nu(A \cap N) + \int_A f d\mu.$$

This finishes the proof for finite  $\mu$  and  $\nu$ .

If  $\mu$  and  $\nu$  are  $\sigma$ -finite, we can write  $X$  as disjoint unions  $X = \bigcup_{n \in \mathbb{N}} X_n = \bigcup_{m \in \mathbb{N}} Y_m$  with  $X_n, Y_m \in \mathcal{B}, \mu(X_n) < \infty, \nu(Y_m) < \infty$ . Applying the first part of the proof to  $X_n \cap Y_m$ , we find  $N_{nm} \in \mathcal{B}, N_{nm} \subset X_n \cap Y_m$ , and measurable  $f_{nm} : X_n \cap Y_m \rightarrow [0, +\infty]$  s.t.  $\mu(N_{nm}) = 0$  and

$$\nu(A \cap X_n \cap Y_m) = \nu(A \cap N_{nm}) + \int_{A \cap X_n \cap Y_m} f_{nm} d\mu.$$

We then put  $N = \bigcup_{n,m \in \mathbb{N}} N_{nm}$  and define  $f : X \rightarrow [0, +\infty]$  by letting  $f = f_{nm}$  on  $X_n \cap Y_m$ .  $\square$

When can we discard the term  $\nu(A \cap N)$ ?

**Definition 17.3.** Given measure  $\mu$  and  $\nu$  on  $X, \mathcal{B}$ , we say that  $\nu$  is *absolutely continuous* with respect to  $\mu$  and write  $\nu \ll \mu$ , if  $\nu(A) = 0$  whenever  $A \in \mathcal{B}, \mu(A) = 0$ .

**Lemma 17.4.** Assume  $\mu$  and  $\nu$  are measures on  $(X, \mathcal{B})$ ,  $\nu(X) < \infty$ . Then  $\nu \ll \mu$  iff  $\forall \epsilon > 0 \exists \delta > 0$  s.t. if  $A \in \mathcal{B}, \mu(A) < \delta$ , then  $\nu(A) < \epsilon$ .

*Proof.* " $\Rightarrow$ ": obvious. " $\Leftarrow$ ": Assume this is not true. Then, there is a  $\epsilon > 0$  s.t.  $\forall \delta > 0$  we can find  $A \in \mathcal{B}$  satisfying  $\mu(A) < \delta, \nu(A) \geq \epsilon$ . Let  $A_n$  be such a set  $A$  for  $\delta = 1/2^n$ . Put  $A = \bigcap_{n \in \mathbb{N}} \bigcup_{k=n} A_k$ . Then

$$\begin{aligned} \mu(A) &\leq \lim_{n \rightarrow \infty} \mu \left( \bigcup_{k=n} A_k \right) \leq \lim_{n \rightarrow \infty} \sum_{k=n} \mu(A_k) \\ &\leq \lim_{n \rightarrow \infty} \sum_{k=n} \frac{1}{2^k} = \lim_{n \rightarrow \infty} \frac{1}{2^{n-1}} = 0. \end{aligned}$$

As  $\nu(X) < \infty$ , we also have

$$\nu(A) = \lim_{n \rightarrow \infty} \nu \left( \bigcup_{k=n} A_k \right) \geq \epsilon.$$

This contradicts the assumption  $\nu \ll \mu$ .  $\square$

**Remark.** The result is not true for infinite  $\nu$ .

**Theorem 17.5** (Radon-Nikodym). Assume  $\mu$  and  $\nu$  are  $\sigma$ -finite measures on a measurable space  $(X, \mathcal{B})$ ,  $\nu \ll \mu$ . Then there is a measurable function  $f : X \rightarrow [0, +\infty)$  s.t.  $d\nu = f d\mu$  (that is,  $\nu(A) = \int_A f d\mu$ ). If  $\tilde{f}$  is another function with the same properties, then  $f = \tilde{f}$   $\mu$ -a.e.

The function is called the Radon-Nikodym derivative at  $\nu$  w.r.t.  $\mu$  and is denoted by  $\frac{d\nu}{d\mu}$ .

*Proof.* By the proposition above, we can find  $N \in \mathcal{B}$  and  $f : X \rightarrow [0, +\infty)$  s.t.  $\mu(N) = 0$  and

$$\nu(A) = \nu(A \cap N) + \int_A f d\mu.$$

As  $\nu(A \cap N) = 0$  by the assumption  $\nu \ll \mu$ , this proves the existence of  $f$ .

Assume we have another  $\tilde{f}$ . Then

$$\nu(A) = \int_A f d\mu = \int_A \tilde{f} d\mu \forall A \in \mathcal{B}.$$

If  $B \in \mathcal{B}, \nu(B) < \infty$ , then consider

$$A_1 = \left\{ x \in \mathcal{B} : f(x) > \tilde{f}(x) \right\}, A_2 = \left\{ x \in \mathcal{B} : f(x) < \tilde{f}(x) \right\}.$$

Then

$$\int_{A_1} (f - \tilde{f}) d\mu = 0 \text{ and } \int_{A_2} (\tilde{f} - f) d\mu = 0,$$

hence  $\mu(A_1) = \mu(A_2) = 0$ . Therefore,  $f = \tilde{f}$   $\mu$ -a.e. on  $\mathcal{B}$ . As  $\nu$  is  $\sigma$ -finite, we have  $X = \cup_{n \in \mathbb{N}} B_n$ ,  $\nu(B_n) < \infty$ . Then  $f = \tilde{f}$   $\mu$ -a.e. on  $B_n$  for all  $n$ , hence  $f = \tilde{f}$   $\mu$ -a.e. on  $X$ .  $\square$

**Example 17.6.** Consider a real-valued function  $f \in C'[a, b]$  s.t.  $f'(t) > 0 \forall t \in [a, b]$ . Let  $c = f(a), d = f(b)$ . We know that for every Riemann integrable function  $g$  on  $[c, d]$  we have

$$\int_c^d g(f) dt = \int_a^b g(f(t)) f'(t) dt.$$

Equivalently,

$$\int_c^d g \circ g^{-1} dt = \int_a^b g f'(t) dt. \quad (27)$$

Denote by  $\lambda_{[a,b]}, \lambda_{[c,d]}$  the Lebesgue measure restricted to the Borel subsets of  $[a, b]$  and  $[c, d]$ , respectively. Then eq. 27 implies that

$$d((f^{-1})_* \lambda_{[c,d]}) = f' d\lambda_{[a,b]},$$

since the integration of  $g = \mathbb{1}_{[\alpha, \beta]}$  gives the same results for any interval  $[\alpha, \beta] \subset [a, b]$  and since a finite Borel measure on  $[a, b]$  is

determined by its values on such intervals. Thus,  $(f^{-1})_* \lambda_{[c,d]} < \lambda_{[a,b]}$  and

$$\frac{d((f^{-1})_* \lambda_{[c,d]})}{d\lambda_{[a,b]}} = f'.$$

## Tips'n Tricks

- Assume we can write  $X$  as a finite union:  $X = \cup_{n \in I} A_n$ ,  $i = 1, \dots, N$ . Then

$$\int f d\mu = \int_X f d\mu = \int_{A_1} f d\mu + \int_{A_2} f d\mu + \dots + \int_{A_N} f d\mu.$$

## Questions

- In problem 26.18 we are supposed to show that  $Y_n \perp Y_m = 0$ , i.e. that  $\langle y_n, y_m \rangle = 0$ ,  $n \neq m$ . I get ...

$$\langle y_n, y \rangle \subset \int_{A_n^c} |y_n|^2 |y_m|^2 d\mu,$$

and I want to argue that this is zero since  $\int_{A_m^c} |y_m|^2 d\mu = 0$ , but I don't see how. The solutions are not clear, and I think perhaps my setup is wrong. I am assuming  $\langle f, g \rangle = \int_X f \bar{g} d\mu$ , i.e. from  $L^2$ , but perhaps it is rather  $\langle f, g \rangle = \int_{A_m^c \cup A_n^c} f \bar{g} d\mu$  or something?

## References

- [Schilling(2017)] Schilling, R. 2017, Measures, Integrals and Martingales, Measures, Integrals and Martingales (Cambridge University Press). <https://books.google.no/books?id=sdAoDwAAQBAJ>
- [Teschl(2010)] Teschl, G. 2010, Topics in Linear and Nonlinear Functional Analysis (Universität Wien). <https://www.mat.univie.ac.at/~gerald/ftp/book-fa/fa.pdf>