

MATH602  
Real Analysis II

Pingbang Hu

October 6, 2022

### **Abstract**

This is a graduate level functional analysis taught by [Joseph Conlon](#). The prerequisites include linear algebra, complex analysis and also [real analysis](#). We'll use Peter Lax[[Lax02](#)] and Reed-Simon[[RS80](#)] as textbooks.

This course is taken in Fall 2022, and the date on the covering page is the last updated time.

# Contents

<b>1</b>	<b>Banach and Hilbert Spaces</b>	<b>2</b>
1.1	Linear Space . . . . .	2
1.2	Quotient Space . . . . .	3
1.3	Normed Spaces . . . . .	3
1.4	Banach Spaces . . . . .	6
1.5	Inner Product Spaces . . . . .	9
1.6	Hilbert Spaces . . . . .	11
1.7	Fourier Series . . . . .	13
<b>2</b>	<b>Bounded Linear Operators</b>	<b>18</b>
2.1	Bounded Linear Functionals . . . . .	18
2.2	Representation Theorems . . . . .	20
2.3	Hahn-Banach Theorem . . . . .	23
2.4	Second Dual Space . . . . .	25
2.5	Reflexive Spaces . . . . .	25
2.6	Separation of Convex Sets . . . . .	26
2.7	Bounded Linear Operators . . . . .	29
<b>3</b>	<b>Main Principles of Functional Analysis</b>	<b>36</b>
3.1	Open Mapping Theorem . . . . .	36
<b>A</b>	<b>Additional Proofs</b>	<b>38</b>

# Chapter 1

## Banach and Hilbert Spaces

### Lecture 1: Introduction

We first briefly review different kinds of vector spaces.

30 Aug. 14:30

#### 1.1 Linear Space

**Definition 1.1.1** (Linear vector space). A *linear vector space*  $E$  over a field  $\mathbb{F}$  is a set with operations of addition and multiplication (by a scalar) such that it's closed under operations, and also the addition and scalar multiplication obey

- (a)  $u + v = v + u$  for  $u, v \in E$
- (b)  $u + (v + w) = (u + v) + w$  for  $u, v, w \in E$
- (c)  $\exists 0 \in E$  such that  $0 + u = u + 0 = u$  for  $u \in E$
- (d)  $\forall u \in E, \exists -u \in E$  such that  $u + (-u) = 0$
- (e)  $\lambda(u + v) = \lambda u + \lambda v$  for  $u, v \in E, \lambda \in \mathbb{F}$
- (f)  $(\lambda + \mu)u = \lambda u + \mu u$  for  $u \in E, \lambda, \mu \in \mathbb{F}$
- (g)  $\lambda(\mu u) = (\lambda\mu)u$  for  $u \in E, \lambda, \mu \in \mathbb{F}$

**Remark.** If  $v, w \in E, \lambda, \mu \in \mathbb{R}$  (or  $\mathbb{C}$ ), then  $\lambda v + \mu w \in E$ .

**Notation** (Real and complex vector space). If  $E$  is over  $\mathbb{F} = \mathbb{C}$ , we usually call  $E$  a *complex vector space*; if  $\mathbb{F} = \mathbb{R}$ , we say  $E$  is a *real vector space*.

**Example.**  $\mathbb{R}^n$  an  $n$  dimensional real linear vector space,  $\mathbb{C}^n$  an  $n$  dimensional complex linear vector space.

We concentrate on  $\infty$  dimensional linear vector space.

**Example.** Let  $K$  is a compact Hausdorff space, then

$$E = \{f: K \rightarrow \mathbb{R} \mid f(\cdot) \text{ is continuous}\}.$$

We then see that  $E$  is a  $\infty$  dimensional real linear vector space.

## 1.2 Quotient Space

Observe that a **linear vector space** can have many subspaces. Say  $E$  is a **linear vector space**, and  $E_1 \subset E$  where  $E_1$  is a proper subspace, i.e.,  $E_1 \neq E$ .

**Definition 1.2.1 (Quotient Space).** The *quotient space*  $E / E_1$  of two **linear vector spaces**  $E, E_1$  such that  $E_1 \subseteq E$  is the set of equivalence classes of vectors in  $E$  where equivalence is given by  $x \sim y$  if  $x - y \in E_1$ . Additionally, denote  $[x]$  as the equivalence class of  $x \in E$ , i.e.,  $[x] = x + E_1$ .

**Remark.** A **quotient space**  $E / E_1$  is a **linear vector space**

**Proof.** Since if  $x_1 + x_2 \in E$ ,  $[x_1] + [x_2] = [x_1 + x_2]$ , and also,  $\lambda[x] = [\lambda x]$  for  $\lambda \in \mathbb{R}$  or  $\mathbb{C}$ , i.e.,  $v, w \in E / E_1$ ,  $\lambda, \mu \in \mathbb{R}$  or  $\mathbb{C}$  implies  $\lambda v + \mu w \in E$ .  $\ast$

Turns out that the way of defining dimensions for finite dimensional **vector spaces** doesn't work here: since we may encounter something like  $\frac{\infty}{\infty}$ . Hence, we introduce **Definition 1.2.2**.

**Definition 1.2.2 (Codimension).** If  $E / E_1$  has finite dimension, then the dimension of  $E / E_1$  is called the *codimension* of  $E_1$  in  $E$ , denoted as  $\text{codim}(E_1)$ .

**Example.** There exists the case that  $\dim(E) = \infty$ ,  $\dim(E_1) < \infty$  where  $\dim(E / E_1) < \infty$ .

**Proof.** Let  $E = \{f: K \rightarrow \mathbb{R} \mid f(\cdot) \text{ continuous}\}$ , and  $E_1 = \{f \in E: f(k_1) = 0\}$  where  $k_1 \in K$  is fixed. We see that the dimension of  $E / E_1$  is exactly 1 since  $E / E_1$  is the set of constant functions.  $\ast$

**Theorem 1.2.1.** If  $E$  is finite dimensional, then  $\text{codim}(E_1) + \dim(E_1) = \dim(E)$

**Definition 1.2.3 (Linear operator).** A map  $T: E \rightarrow F$  between 2 **linear spaces** is a *linear operator* if it preserves the properties of addition and multiplication by a scalar, i.e., for  $v, w \in E$  and  $\lambda, \mu \in \mathbb{R}$  or  $\mathbb{C}$ ,

$$T(\lambda v + \mu w) = \lambda T(v) + \mu T(w).$$

**Definition.** Given a **linear operator**  $T: E \rightarrow F$  we have the following.

**Definition 1.2.4 (Kernel).** The *kernel* of  $T$  is the subspace  $\ker(T) = \{x \in E \mid Tx = 0\}$ .

**Definition 1.2.5 (Image).** The *image* of  $T$  is the subspace  $\text{Im}(T) = \{Tx \in F \mid x \in E\}$ .

## 1.3 Normed Spaces

We review some basic notions.

**Definition 1.3.1 (Norm).** Let  $E$  be a **linear vector space**. A *norm*  $\|\cdot\|: E \rightarrow \mathbb{R}$  on  $E$  is a function from  $E$  to  $\mathbb{R}$  with the properties:

- (a)  $\|x\| \geq 0$  and  $\|x\| = 0 \Leftrightarrow x = 0$ .
- (b)  $\|\lambda x\| = |\lambda| \|x\|$ ,  $\lambda \in \mathbb{R}$  or  $\mathbb{C}$ .
- (c)  $\|x + y\| \leq \|x\| + \|y\|$ .

**Notation** (Dilation). We say that the second condition is the *dilation* property.

**Definition 1.3.2** (Normed vector space). A linear vector space  $E$  equipped with a norm  $\|\cdot\|$  is called a *normed vector space*.

**Remark** (Induced metric space). A normed vector space  $E$  induces a *metric space* with metric  $d(x, y) = \|x - y\|$ , where the metric has properties

- (a)  $d(x, y) \geq 0$ . Also,  $d(x, x) = 0$  and  $d(x, y)$  implies  $x = y$ .
- (b)  $d(x, y) = d(y, x)$ .
- (c)  $d(x, z) \leq d(x, y) + d(y, z)$ .

**Example** (Bounded sequences  $\ell_\infty$ ). Let  $\ell_\infty$  be the space of bounded sequences  $x = (x_1, x_2, \dots)$  with  $x_i \in \mathbb{R}$  for  $i = 1, 2, \dots$ . Then we define  $\|x\| = \|x\|_\infty = \sup_{i \geq 1} |x_i|$ .

**Example** (Absolutely summable sequences  $\ell_1$ ). Let  $\ell_1$  be the space of absolutely summable sequences  $x = (x_1, x_2, \dots)$  and  $\sum_{i=1}^\infty |x_i| < \infty$ . Then we define  $\|x\| = \|x\|_1 = \sum_{i=1}^\infty |x_i| < \infty$ .

**Example** (Continuous functions  $C(k)$ ). The space  $C(k)$  of continuous functions  $f: K \rightarrow \mathbb{R}$  where  $K$  is compact Hausdorff. Then we define  $\|f\| = \|f\|_\infty = \sup_{x \in K} |f(x)|$ .

### 1.3.1 Geometry of Normed Spaces

**Definition 1.3.3** (Ball). A (closed) *ball* centered at a point  $x_0 \in E$  with radius  $r > 0$  is the set  $B(x_0, r) = \{x \in E \mid \|x - x_0\| \leq r\}$ .

**Definition 1.3.4** (Sphere). The *sphere* centered at  $x_0$  with radius  $r > 0$  is the set  $S(x_0, r) = \{x \in E \mid \|x - x_0\| = r\}$ .

**Remark.** We see that  $S(x_0, r)$  is the **boundary** of  $B(x_0, r)$ , i.e.,  $S(x_0, r) = \partial B(x_0, r)$ .

**Note** (Infinite dimensional geometry). We know that in finite dimensional, all norms are equivalent, which is not true for infinite dimensional vector spaces. This has something to do with the geometry of balls.

Explicitly, balls can have different geometries depending on the properties of the norms. We see that a  $\|\cdot\|_\infty$  can have multiple supporting hyperplane at the corner, while for a  $\|\cdot\|_2$  can have only one at each point.

Also, unit balls for  $\|\cdot\|_1$  is also a **square**, where we have

$$B(0, 1) = \{x = (x_1, x_2, \dots) \mid -1 < y_\epsilon < 1 \forall \epsilon\}$$

such that  $y_\epsilon = \sum_{i=1}^\infty \epsilon_i x_i$ ,  $\epsilon_i = \pm 1$  and  $\epsilon = (\epsilon_1, \epsilon_2, \dots)$ .

We see that different norms give different geometry, but they have important common features, most notably, convexity properties.

**Definition 1.3.5** (Convex set). Given  $E$  a linear vector space, a set  $K \subset E$  is *convex* if for  $x, y \in K$  and  $0 \leq \lambda \leq 1$ ,

$$\lambda x + (1 - \lambda)y \in K.$$

**Definition 1.3.6** (Convex function). Given  $E$  a linear vector space, a function  $f: E \rightarrow \mathbb{R}$  is called *convex* if for  $x, y \in E$  and  $0 \leq \lambda \leq 1$ ,

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

**Remark.** If  $f: E \rightarrow \mathbb{R}$  is a convex function, then for any  $M \in \mathbb{R}$  the set  $\{x \in E \mid f(x) \leq M\}$  is convex.

The upshot is that norms are convex, and the unit balls are convex as well.

## Lecture 2: Banach Spaces and Completion

Let's first see a proposition.

01 Sep. 14:30

**Proposition 1.3.1.** Let  $(E, \|\cdot\|)$  be a normed linear space, then the norm is convex and continuous.

**Proof.** Let  $f: E \rightarrow \mathbb{R}$  be  $f(x) = \|x\|$ . Then  $f(x) - f(y) = \|x\| - \|y\| \leq \|x - y\|$ , which implies  $|f(x) - f(y)| \leq \|x - y\|$  for  $x, y \in E$ , i.e.,  $f$  is Lipschitz continuous hence continuous. For convexity, let  $0 < \lambda < 1$ , we have

$$f(\lambda x + (1 - \lambda)y) = \|\lambda x + (1 - \lambda)y\| \leq \|\lambda x\| + \|(1 - \lambda)y\| = \lambda \|x\| + (1 - \lambda)\|y\| = \lambda f(x) + (1 - \lambda)f(y).$$

■

**Note.** Note that  $f(\cdot)$  is continuous implies the closed ball

$$B(x_0, r) = \{x \in E \mid \|x - x_0\| \leq r\} = \{x \in E \mid f(x - x_0) \leq r\}$$

is closed in topology of  $E$ . Also,  $f(\cdot)$  is convex implies  $B(x_0, r)$  is convex.

**Remark.** If  $f: E \rightarrow \mathbb{R}$  is convex, then the sets  $\{x \in E \mid f(x) \leq M\}$  is also convex. However, it's possible to have non-convex functions  $f$  such that all sets  $\{x \in E \mid f(x) \leq M\}$  are convex.

**Proof.** Take  $f(x) = |x|^p$  for  $x \in \mathbb{R}$  and  $p > 0$ . We see that  $f$  is convex if  $p > 1$ , and non-convex if  $p < 1$ . However, the sets  $\{x \in \mathbb{R} \mid f(x) \leq M\}$  are all convex since it's independent of  $p$ . ⊗

**Lemma 1.3.1.** Suppose  $x \mapsto \|x\|$  satisfies

- (a)  $\|x\| \geq 0$  and  $\|x\| = 0 \Leftrightarrow x = 0$ .
- (b)  $\|\lambda x\| = |\lambda| \|x\|$ ,  $\lambda \in \mathbb{R}$  or  $\mathbb{C}$ .
- (c) The unit ball  $B(0, 1)$  is convex.

Then  $f(x) = \|x\|$  satisfies the triangle inequality  $\|x + y\| \leq \|x\| + \|y\|$ .

**Proof.** We see that if the third condition is true, then for  $u, v \in B(0, 1)$  and  $0 < \lambda < 1$ , we have  $\lambda u + (1 - \lambda)v \in B(0, 1)$ . Let  $x, y \in E$ , and

$$\lambda = \frac{\|x\|}{\|x\| + \|y\|} \Rightarrow 1 - \lambda = \frac{\|y\|}{\|x\| + \|y\|}.$$

By letting  $u = x / \|x\|$ ,  $v = y / \|y\|$  we see that

$$\lambda u + (1 - \lambda)v = \frac{\|x\|}{\|x\| + \|y\|} \frac{x}{\|x\|} + \frac{\|y\|}{\|x\| + \|y\|} \frac{y}{\|y\|} \in B(0, 1) \Rightarrow \left\| \frac{x}{\|x\| + \|y\|} + \frac{y}{\|x\| + \|y\|} \right\| \leq 1.$$

From the second condition, it follows that  $\|x + y\| \leq \|x\| + \|y\|$ , which is the triangle inequality. ■

**Remark.** If  $x \mapsto \|x\|$  satisfies the first two conditions and is **convex**, then it satisfies the triangle inequality.

**Proof.** Since  $\frac{1}{2}\|x+y\| = \left\|\frac{x}{2} + \frac{y}{2}\right\| \leq \frac{1}{2}\|x\| + \frac{1}{2}\|y\|$ . \*

Now, given a **quotient space**  $E/E_1$ , the question is can we try to define a **norm**?

**Problem 1.3.1.** On  $E/E_1$ , is  $\|[x]\| := \inf_{y \in E_1} \|x+y\|$  a **norm**?

**Answer.** We see that if  $x \in \overline{E_1} \setminus E_1$ , then  $\|[x]\| = 0$  but  $0 \neq [x] \in E/E_1$ . \*

**Note.** Notice the difference from finite dimensional situation. All finite dimensional spaces  $E_1$  are closed but not in general if  $E_1$  has  $\infty$  dimensions.

**Example.** Let  $\ell_1(\mathbb{R})$  be the sequence of  $x_n$  for  $n \geq 1$  in  $\mathbb{R}$  such that  $\sum_{i=1}^{\infty} |x_i| \leq \infty$ . Define

$$\|x\|_1 := \sum_{i=1}^{\infty} |x_i|,$$

and let  $E_1$  be all sequences with finite number of the  $x_n$  are nonzero. We see that  $\overline{E_1} = \ell_1(\mathbb{R})$  is infinite dimensional.

**Proposition 1.3.2.** Let  $(E, \|\cdot\|)$  be a **normed space** and  $E_1 \subseteq E$ ,  $E_1$  is closed. Then

$$\|\cdot\| : E/E_1 \rightarrow \mathbb{R}, \quad \|[x]\| = \inf_{y \in E_1} \|x+y\|$$

is a **norm** on  $E/E_1$ .

**Proof.** If  $\|[x]\| = 0$ , then  $\inf_{y \in E_1} \|x-y\| = 0$ , which implies  $x \in E_1$  since  $E_1$  is closed, so  $[x] = 0$ . Also, since

$$\|\lambda[x]\| = \inf_{y \in E_1} \|\lambda x + y\| = \inf_{z \in E_1} \|\lambda x + \lambda z\| = |\lambda| \inf_{z \in E_1} \|x + z\| = |\lambda| \|[x]\|,$$

the dilation property is satisfied. Finally, for triangle inequality, we have

$$\|[x] + [y]\| = \inf_{x_1, y_1 \in E_1} \|x + y + x_1 + y_1\| \leq \inf_{x_1 \in E_1} \|x + x_1\| + \inf_{y_1 \in E_1} \|y + y_1\| = \|[x]\| + \|[y]\|.$$

■

**Remark.** This shows that the only obstacle for this kind of **norm** being an actual **norm** is the closeness of  $E_1$ .

## 1.4 Banach Spaces

**Definition 1.4.1 (Banach space).** A **linear normed space** is a *Banach space* if it's complete, i.e., every Cauchy sequence converges.

This implies that given a **Banach space**  $(E, \|\cdot\|)$ , if  $\{x_n \in E : n \geq 1\}$  is a sequence with property such that  $\lim_{m \rightarrow \infty} \sup_{n \geq m} \|x_n - x_m\| = 0$ , then  $\exists x_\infty \in E$  such that  $\lim_{n \rightarrow \infty} \|x_n - x_m\| = 0$  as well.

**Example.** The spaces  $\ell_1$ ,  $\ell_\infty$  and  $C(K)$  are **Banach spaces**.



We want to give a different criterion for showing  $(E, \|\cdot\|)$  is **Banach**. Let  $E$  be a **linear normed space** and  $\{x_\ell \mid \ell \geq 1\}$  a sequence in  $E$ .

### 1.4.1 Completion of Normed Space

We now show an important theorem which characterizes completeness in terms of convergence of series rather than sequences. We first see the definition.

**Definition 1.4.2 (Absolutely summable).** A sequence is *absolutely summable* if  $\sum_{i=1}^{\infty} \|x_i\| < \infty$ .

Then, we have the following.

**Theorem 1.4.1 (Criterion for completeness).** A **normed space**  $(E, \|\cdot\|)$  is a **Banach space** if and only if every **absolutely summable** series in  $E$  converges.

**Proof.** We need to prove two directions.

$(\Rightarrow)$  Suppose  $E$  is a **Banach space** and  $\{x_k \mid k \geq 1\}$  an **absolutely summable** series. Set  $s_n = \sum_{k=1}^n x_k$ ,  $n \geq 1$ , we want to show  $s_n$  is Cauchy, and if this is the case, completeness of  $E$  implies  $\exists s_\infty$  and  $\lim_{n \rightarrow \infty} \|s_n - s_\infty\| = 0$ . Let  $n > m$ , we see that

$$\|s_n - s_m\| = \left\| \sum_{k=m+1}^n x_k \right\| \leq \sum_{k=m+1}^n \|x_k\| \leq \sum_{k=m+1}^{\infty} \|x_k\|.$$

Observe that  $\lim_{m \rightarrow \infty} \sum_{k=m+1}^{\infty} \|x_k\| = 0$ , we see that the sequence  $\{s_n\}$  is Cauchy.

$(\Leftarrow)$  Conversely, suppose  $E$  is **not** complete. Then there exists a Cauchy sequence  $\{x_n \mid n \geq 1\}$  which does not converge. Furthermore, no subsequence of  $\{x_n \mid n \geq 1\}$  converges.<sup>a</sup> We now construct an **absolutely summable** series which does not converge.

Define  $n(1) \geq 1$  such that  $\|x_n - x_{n(1)}\| \leq \frac{1}{2}$  if  $n \geq n(1)$ , similarly, let  $n(2) > n(1)$  be such that  $\|x_n - x_{n(2)}\| \leq \frac{1}{2^2}$  if  $n \geq n(2)$ . In all, we have  $n(1) < n(2) < n(3) < \dots$  such that  $\|x_n - x_{n(k)}\| \leq \frac{1}{2^k}$  if  $n \geq n(k)$ . Define  $w_j := x_{n(j+1)} - x_{n(j)}$  for  $j = 1, 2, \dots$ . We see that

$$x_{n(m)} = x_{n(1)} + \sum_{j=1}^m w_j$$

for  $m = 1, 2, \dots$ , and  $\{x_{n(m)}\}$  does not converge, hence so does the series  $\sum_{j=1}^{\infty} w_j$ . However,  $\sum_{j=1}^{\infty} \|w_j\| \leq \sum_{j=1}^{\infty} \frac{1}{2^j} = 1$ , which implies  $\{w_j\}$  is **absolutely summable**. ■

<sup>a</sup>Otherwise, the whole sequence converges by the fact that it's Cauchy.

**Theorem 1.4.2 (Completion).** Suppose  $E$  is a **normed space**. Then there exists a **Banach space**  $\hat{E}$  called *the completion* of  $E$  with the following properties:

- (a) There exists a linear map  $\iota: E \rightarrow \hat{E}$  such that  $\|\iota x\| = \|x\|$ .<sup>a</sup>
- (b)  $\text{Im}(\iota)$  is dense in  $\hat{E}$ , and  $\hat{E}$  is the smallest **Banach space** containing image of  $E$ .

<sup>a</sup>This is called an *isometric embedding* of  $E$  into  $\hat{E}$ .

## Lecture 3: Banach, Inner Product Spaces

Notice that  $\ell_1$  and  $\ell_\infty$  are **Banach**, and we want to generalize to  $\ell_p$  with  $1 < p < \infty$ . For  $x = \{x_n, n \geq 1\} \in \ell_p$  and if  $\sum_{n=1}^{\infty} |x_n|^p < \infty$ , for  $\|x\|_p = (\sum_{n=1}^{\infty} |x_n|^p)^{1/p}$ , we want to show that  $x \rightarrow \|x\|_p$

06 Sep. 14:30

satisfies properties of a **norm**. The first two properties of a **norm** is easy check. As for triangle inequality, we have the following.

**Lemma 1.4.1** (Minkowski inequality). Let  $1 \leq p < \infty$ , for  $x, y \in \ell_p$ ,

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

**Proof.** Recall that from [Lemma 1.3.1](#), we only need to show that  $B(0, 1)$  is **convex**, where

$$B(0, 1) = \left\{ x = \{x_n : n \geq 1\} \mid f(x) = \sum_{n=1}^{\infty} |x_n|^p \leq 1 \right\}.$$

But  $f(x)$  is **convex** since  $x \mapsto |x|^p$ ,  $x \in \mathbb{R}$  is **convex** if  $p \geq 1$ , we're done. Hence,  $\|x + y\|_p \leq \|x\|_p + \|y\|_p$ , i.e.,

$$\left( \sum_{j=1}^{\infty} |x_j + y_j|^p \right)^{1/p} \leq \left( \sum_{j=1}^{\infty} |x_j|^p \right)^{1/p} \left( \sum_{j=1}^{\infty} |y_j|^p \right)^{1/p}.$$

■

**Lemma 1.4.2** (Hölder's inequality). Let  $1 < p < \infty$ , for  $x \in \ell_p$ ,  $y \in \ell_q$ , we have

$$\|x \cdot y\|_1 \leq \|x\|_p \|y\|_q$$

where  $1/p + 1/q = 1$ .

**Proof.** Note first that we can assume without loss of generality,  $\sum_{j=1}^{\infty} |x_j|^p = \sum_{j=1}^{\infty} |y_j|^q = 1$ . Then, result follows from the **Young's inequality**,

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

for  $x, y > 0$ ,  $x, y \in \mathbb{R}$ .

**Remark** (Legendre transform and the inequality). **Young's inequality** is a special case of the inequality

$$xy \leq f(x) + \mathcal{L}f(y)$$

where  $\mathcal{L}f(\cdot)$  is the **Legendre transform** of  $f(\cdot)$ , i.e.,  $\mathcal{L}f(y) = \sup_x [xy - f(x)]$ .

If  $f$  is **convex**, then the function  $xy \mapsto xy - f(x)$  is concave so has unique maximum. And  $\mathcal{L}f(\cdot)$  always **convex** even if  $f(\cdot)$  is not. In particular, if  $f(x) = x^p/p$ , then  $\mathcal{L}f(y) = y^q/q$ .

■

**Note.** **Minkowski inequality** is usually proved via the **Hölder's inequality**. To show this, since

$$\sum_{j=1}^{\infty} |x_j + y_j|^p \leq \sum_{j=1}^{\infty} |x_j| |x_j + y_j|^{p-1} + \sum_{j=1}^{\infty} |y_j| |x_j + y_j|^{p-1}.$$

Then **Holder inequality** implies

$$\sum_{j=1}^{\infty} |x_j| |x_j + y_j|^{p-1} \leq \left( \sum_{j=1}^{\infty} |x_j|^p \right)^{1/p} \left( \sum_{j=1}^{\infty} |x_j + y_j|^{(p-1)q} \right)^{1/q},$$

and we're done.<sup>a</sup>

<sup>a</sup>Note that  $(p-1)q = p$ .

The above argument applies to more general spaces of  $p$  integrable functions. Let  $(\Omega, \Sigma, \mu)$  be a measure space and  $L_p(\Omega, \Sigma, \mu)$  where all  $\Sigma$  measure functions  $f: \Omega \rightarrow \mathbb{R}$  (or  $\mathbb{C}$ ) such that  $\int_{\Omega} |f|^p d\mu < \infty$ . Then,  $L_p(\Omega, \Sigma, \mu)$  is a **normed space** with **norm**

$$\|f\|_p = \left( \int_{\Omega} |f|^p d\mu \right)^{1/p}.$$

It's more tricky to show that  $L^p$  is a **Banach space**, but it's indeed still the case.

**Theorem 1.4.3 (Riesz-Fisher).** The space  $L_p(\Omega, \Sigma, \mu)$  is a **Banach space** for  $1 \leq p < \infty$ .

**Proof.** Toward using **Theorem 1.4.1**, let  $\{f_n: n \geq 1\}$  be an **absolutely summable** sequence in  $L_p$ . Then the **norm** satisfies

$$\left\| \sum_{k=1}^N f_k \right\|_p \stackrel{!}{\leq} \sum_{k=1}^N \|f_k\|_p \leq C < \infty \Rightarrow \int_{\Omega} \left| \sum_{k=1}^N f_k \right|^p d\mu \leq C^p.$$

- Assume all  $f_k$  are non-negative. From **monotone convergence theorem**, we have

$$\lim_{N \rightarrow \infty} \int_{\Omega} \left( \sum_{k=1}^N f_k \right)^p d\mu = \int_{\Omega} \left( \sum_{k=1}^{\infty} f_k \right)^p d\mu \leq C^p.$$

Hence,  $g = \sum_{k=1}^{\infty} f_k \in L_p$ . We now want to show that  $\sum_{k=1}^N f_k \rightarrow g$  in  $L_p$ . Set  $r_n = \sum_{k=n+1}^{\infty} f_k$  where  $r_n$  is a decreasing sequence where  $r_n \rightarrow 0$  a.e. and also

$$\int_{\Omega} r_1^p d\mu < \infty.$$

This means that  $\lim_{n \rightarrow \infty} \|r_n\|_p = 0$  by **dominate convergence theorem**.

- For arbitrary  $f_k: \Omega \rightarrow \mathbb{R}$ , write  $f_k = f_k^+ + f_k^-$  where  $f_k^+ = \sup(f_k, 0)$  and  $f_k^- = \inf(f_k, 0)$ . The sequence  $\{f_k^+: k \geq 1\}$  are **absolutely summable**, and we just proceed as before. Similarly, if  $f_k: \Omega \rightarrow \mathbb{C}$ , we have the similar result.

■

## 1.5 Inner Product Spaces

**Definition 1.5.1 (Inner product).** Let  $E$  be a **linear space** over  $\mathbb{C}$ . An *inner product*  $\langle \cdot, \cdot \rangle: E \times E \rightarrow \mathbb{C}$  is a function which has the following properties:

- $\langle x, x \rangle \geq 0$  and  $\langle x, x \rangle = 0$  if and only if  $x = 0$ .
- $\langle ax + by, z \rangle = a \langle x, z \rangle + b \langle y, z \rangle$  for  $a, b \in \mathbb{C}$ .
- $\langle x, y \rangle = \overline{\langle y, x \rangle}$ .

**Remark (Real inner product).** We can also define **inner products** of spaces over  $\mathbb{R}$  with no extra conjugation in the last property.

**Definition 1.5.2 (Inner product space).** An *inner product space* is a **linear space**  $E$  with an **inner product**  $\langle \cdot, \cdot \rangle: E \times E \rightarrow \mathbb{C}$ .

**Definition 1.5.3 (Orthogonal).** Given a linear space  $E$ ,  $x, y \in E$  are *orthogonal* if  $\langle x, y \rangle = 0$ , denote as  $x \perp y$ .

**Theorem 1.5.1 (Cauchy-Schwarz inequality).** Let  $x, y \in E$  and an inner product  $\langle \cdot, \cdot \rangle$ , then

$$|\langle x, y \rangle| \leq \langle x, x \rangle^{\frac{1}{2}} \langle y, y \rangle^{\frac{1}{2}}.$$

**Proof.** Define  $Q(t)$  by  $Q(t) = \langle x + ty, x + ty \rangle = \langle y, y \rangle t^2 + 2t \operatorname{Re} \langle x, y \rangle + \langle x, x \rangle$  if  $t \in \mathbb{R}$ . Then we see that  $Q(t) \geq 0$  with  $t \in \mathbb{R}$ , by looking at the discriminant, we have  $(\operatorname{Re} \langle x, y \rangle)^2 \leq \langle x, x \rangle \langle y, y \rangle$ . Finally, the result follows by choosing  $\theta \in \mathbb{R}$  such that  $\langle x, y \rangle = \operatorname{Re} \langle x e^{i\theta}, y \rangle$ , we then see that

$$|\langle x, y \rangle| = |\operatorname{Re} \langle x e^{i\theta}, y \rangle| = |\operatorname{Re} \langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \langle y, y \rangle},$$

proving the result. ■

**Corollary 1.5.1.** The function  $x \mapsto \|x\| := \langle x, x \rangle^{\frac{1}{2}}$  is a norm on  $E$ .

**Proof.** The triangle inequality is a consequence of Theorem 1.5.1 such that

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \stackrel{!}{\leq} \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

■

**Remark (Pythagorean theorem).** The calculation in Corollary 1.5.1 clearly implies *Pythagorean theorem*: if  $x \perp y$ , then  $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ .

**Example.** The space  $\ell_2$  of square summable sequences  $x = (x_1, x_2, \dots)$  and  $y = (y_1, y_2, \dots)$ ,

$$\langle x, y \rangle := \sum_{j=1}^{\infty} x_j \bar{y}_j$$

defines an inner product.

**Example (Canonical inner product on  $L_2$ ).** The space  $L_2(\Omega, \Sigma, \mu)$  of square integrable functions  $f, g$ ,

$$\langle f, g \rangle = \int_{\Omega} f(x) \bar{g}(x) \, d\mu(x)$$

defines an inner product. Furthermore,  $\|f\|_2 = \langle f, f \rangle^{1/2}$ .

**Proof.** The only non-trivial fact to prove is that  $\langle f, g \rangle$  is finite, i.e.,  $f\bar{g}$  is integrable. Firstly,  $f^2, \bar{f}^2$  and  $(f + g)^2$  are all integrable since  $f, \bar{g}$  and  $f + \bar{g}$  are all in  $L_2$ , hence  $f\bar{g}$  is also integrable. ⊗

**Example.** The space of  $m \times n$  matrices  $A = (a_{ij})$ ,  $1 \leq i \leq m, 1 \leq j \leq n$ . Then

$$\langle A, B \rangle = \operatorname{tr}(AB^*)$$

defines an inner product, where  $B^*$  is the **Hermitian adjoint** of  $B$ , i.e., for  $B = (b_{ij})$ , then  $B^* = (b_{ij}^*)$  for  $b_{ij}^* = \bar{b}_{ji}$ .

**Remark** (Hilbert-Schmidt (Frobenius) norm). Specifically, the norm corresponding to this inner product is

$$\|A\|_{\text{HS}} := \left( \sum_{i,j} |a_{ij}|^2 \right)^{1/2},$$

which is known as the *Hilbert-Schmidt* or *Frobenius* norm.

**Note** (Angle). Recall that in Euclidean space  $\mathbb{R}^n$ , the inner product can be computed by the formula

$$\langle x, y \rangle = \|x\| \|y\| \cos \theta(x, y)$$

where  $\theta(x, y)$  denotes the angle between  $x$  and  $y$ . We see that we can similarly define the angle between  $x, y$  in an inner product space by

$$\cos \theta(x, y) := \frac{\langle x, y \rangle}{\|x\| \|y\|} \in [-1, 1]$$

where the range is ensured by Theorem 1.5.1, so it's well-defined. Though this concept is rarely used anyway.

Other geometry result also follows in an inner product space: since inner product can be expressed in terms of the norm. This is because both parallelogram law and polarization identity hold.

**Lemma 1.5.1** (Parallelogram law). Given  $E$  an inner product space, we have

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

**Proof.** Recall that  $\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + 2\operatorname{Re} \langle x, y \rangle + \|y\|^2$  and similarly,  $\|x - y\|^2 = \|x\|^2 - 2\operatorname{Re} \langle x, y \rangle + \|y\|^2$ , hence the result follows. ■

**Lemma 1.5.2** (Polarization identity). Given  $E$  an inner product space, we have

$$\langle x, y \rangle = \frac{1}{4} \left\{ \|x + y\|^2 - \|x - y\|^2 + i \|x + iy\|^2 - i \|x - iy\|^2 \right\}$$

**Proof.** The proof is just to expand the right-hand side in terms of inner product. ■

**Remark.** Polarization identity shows that the function  $x \mapsto \|x\|^2$  determines the inner product.

## Lecture 4: Orthogonality and Projection

### 1.6 Hilbert Spaces

08 Sep. 14:30

**Definition 1.6.1** (Hilbert space). A complete inner product space is called a *Hilbert space*.

**Example.** We have seen that  $\ell_2$  and  $L^2(\Omega, \Sigma, \mu)$  are complete, hence are Hilbert space.

#### 1.6.1 Orthogonality

We'll soon see that the key notion in Hilbert space theory is orthogonality.

**Definition 1.6.2** (Orthogonal complement). Let  $A \subseteq \mathcal{H}$  where  $\mathcal{H}$  is a Hilbert space. Then the

orthogonal complement  $A^\perp$  of  $A$  is

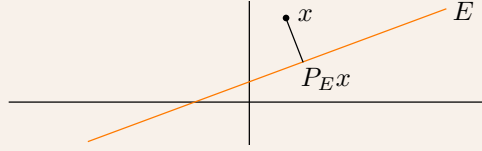
$$A^\perp := \{x \in \mathcal{H} : \langle x, y \rangle = 0 \text{ for } y \in A\}.$$

**Remark.**  $A^\perp$  is also a Hilbert space, in particular, closed and  $A^\perp \cap A \subseteq \{0\}$ .

**Proof.** Since  $A^\perp$  is closed linear subspace of  $\mathcal{H}$ , where the closure follows from the continuity of the function  $x \mapsto \langle x, y \rangle$  for  $x \in \mathcal{H}$  by looking at the inverse image of  $\{0\}$ . Finally, for  $x \in A^\perp \cap A$ ,  $\langle x, x \rangle = 0$  implies  $x = 0$ .  $\circledast$

**Theorem 1.6.1 (Orthogonality principle).** Assume  $E \subseteq \mathcal{H}$  is a closed linear subspace of the Hilbert space  $\mathcal{H}$  and  $x \in \mathcal{H}$ . Then we have the following.

- (a) Then there exists a unique closest point  $y = P_E x \in E$  to  $x$ , i.e.,  $\|x - P_E x\| = \inf_{y' \in E} \|x - y'\|$ .
- (b) The point  $y = P_E x \in E$  is the unique vector such that  $x - y \in E^\perp$ .



**Proof.** Note that the function  $y' \mapsto \|x - y'\|$  for  $y' \in E$  is convex. We expect a minimizer  $y'$ .

- (a) Let  $y_n \in E$  for  $n = 1, 2, \dots$  be a minimizing sequence, i.e.,

$$\lim_{n \rightarrow \infty} \|x - y_n\| = \inf_{y' \in E} \|x - y'\| =: d.$$

From parallelogram law, we have

$$\|y_n - y_m\|^2 + 4\|x - (y_n + y_m)/2\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2.$$

As  $n, m \rightarrow \infty$ , the right-hand side goes to  $4d^2$ . But since  $\frac{1}{2}(y_n + y_m) \in E$ , we have  $\|x - \frac{1}{2}(y_n + y_m)\| \geq d$ , so

$$\lim_{m \rightarrow \infty} \sup_{m \geq n} \|y_n - y_m\|^2 = 0,$$

which further implies  $\{y_n\}$  is a Cauchy sequence. As  $\mathcal{H}$  is complete, we see that  $y_n \rightarrow y_\infty \in E$ , with  $\|x - y_\infty\| = d$ .

Now, with the fact that  $E$  is closed, we set  $y_\infty = P_E x$  where  $y_\infty$  is unique since if  $\|x - y_\infty\| = \|x - y'_\infty\| = d$ , again by the parallelogram law where we now plug in  $y_\infty$  and  $y'_\infty$  instead of  $y_n$  and  $y_m$  as above, we see that  $\|y_\infty - y'_\infty\| = 0$ . In all,  $P_E x \in E$  is uniquely defined.

- (b) We now show  $P_E x$  is the unique vector  $y \in E$  such that  $x - y \perp E$ , i.e.,  $x - y \in E^\perp$ . Let  $y' \in E$  and let  $Q(t)$  be the quadratic

$$Q(t) := \langle x - P_E x + ty', x - P_E x + ty' \rangle = \|x - P_E x + ty'\|^2.$$

Since  $t \mapsto Q(t)$  has a **strict** minimum at  $t = 0$ , which implies  $Q'(0) = 0$ , i.e.,  $\operatorname{Re} \langle x - P_E x, y' \rangle = 0$  for all  $y' \in E$ , which further implies  $\langle x - P_E x, y' \rangle = 0$  for all  $y' \in E$ . This shows that  $x - P_E x \in E^\perp$ .

Finally, we need to show  $P_E x \in E$  is the unique vector such  $x - P_E x \in E^\perp$ . This can be seen from  $Q(t) = \|x - P_E x\|^2 + t^2 \|y'\|^2$  for any  $y' \in E$ .

■

**Note.** To show this exists, we typically need

1. Compactness properties
2. Non-degeneracy properties for uniqueness

Here by using [parallelogram law](#), we don't need compactness.

**Remark.** [Theorem 1.6.1](#) shows that the minimizer for the function  $y' \mapsto \|x - y'\|$  for  $y' \in E$  is characterized by the orthogonality condition, i.e.,  $x - y \perp E$  for some  $y \in E$ .

**Definition 1.6.3 (Orthogonal projection).** Let  $\mathcal{H}$  be a [Hilbert space](#) and let  $E \subseteq \mathcal{H}$  be a closed subspace. The *orthogonal projection operator*  $P_E: \mathcal{H} \rightarrow E$  is given by  $x \mapsto P_E x$  where  $P_E x$  is defined uniquely via  $x - P_E x \in E^\perp$ .

**Definition 1.6.4 (Bounded linear map).** Given a mapping  $A: \mathcal{B} \rightarrow \mathcal{B}$  on a [Banach space](#)  $\mathcal{B}$ , we say it's a *bounded linear map* if it's [bounded](#) and [linear](#).

**Definition 1.6.5 (Linear map).** The operator  $A$  is *linear* if for  $x, y \in \mathcal{B}$ ,  $a, b \in \mathbb{C}$ ,

$$A(ax + by) = aA(x) + bA(y).$$

**Definition 1.6.6 (Bounded map).** The operator  $A$  is *bounded* if

$$\|A\| := \sup_{\|x\|=1} \|Ax\| < \infty.$$

**Remark.** Note that  $\|Ax\| \leq \|A\| \|x\|$  for  $x \in \mathcal{B}$ .

We see that  $P_E x$  is a [bounded linear operator](#)  $P_E: \mathcal{H} \rightarrow E$  with the properties  $P_E^2 = P_E$  and  $\|x\|^2 = \|P_E x\|^2 + \|(I - P_E)x\|^2$  since  $(I - P_E)x \perp P_E x$ . The latter property shows that

$$\|P_E\| \leq 1, \quad \|(I - P_E)\| \leq 1,$$

and fact,  $\|P_E\| = \|I - P_E\| = 1$ . Also,  $I - P_E$  is also an [orthogonal projection](#) onto  $E^\perp$ .

## 1.7 Fourier Series

[Hilbert space](#) gives a geometric framework for studying [Fourier series](#). The classical Fourier analysis studies situations where a function  $f: [-\pi, \pi] \rightarrow \mathbb{C}$  can be expanded as [Fourier series](#)

$$f(t) = \sum_{k=-\infty}^{\infty} \hat{f}(k) \frac{1}{\sqrt{2\pi}} e^{ikt}$$

with the Fourier coefficients

$$\hat{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt.$$

In order to make Fourier analysis rigorous, we have to understand what functions  $f$  can be written as [Fourier series](#), and in what sense the [Fourier series](#) converges. To do so, it's of great advantage to depart from this specific situation and carry out Fourier analysis in an abstract [Hilbert space](#). Let  $f(t)$  be a vector in the function space  $L_2[-\pi, \pi]$ , and the exponential functions  $e^{-ikt}$  will form a set of [orthogonal](#) vectors in this space. Then, [Fourier series](#) will become an orthogonal decomposition of a vector  $f$  w.r.t. an [orthogonal system](#) of coordinates.

### 1.7.1 Orthogonal Systems

We first give the definition.

**Definition 1.7.1** (Orthogonal system). A sequence  $\{x_k : k \geq 1\}$  of non-zero vectors in a Hilbert space  $\mathcal{H}$  is *orthogonal* if  $\langle x_k, x_\ell \rangle = 0$  for all  $\ell \neq k$ .

**Definition 1.7.2** (Orthonormal system). An orthogonal system is called an *orthonormal system* if in addition, we have  $\|x_k\| = 1$  for all  $k$ .

**Remark** (Equivalence definition of orthonormal system).  $\{x_k : k \geq 1\}$  is *orthonormal* if  $\langle x_k, x_\ell \rangle = \delta_{k,\ell}$  where  $\delta$  is the Kronecker delta.

We first see an immediate generation given the remark.

**Theorem 1.7.1** (Pythagorean theorem). Let  $\{x_k\}_k$  be an orthogonal system in a Hilbert space  $\mathcal{H}$ . Then for every  $n \in \mathbb{N}$ ,

$$\left\| \sum_{k=1}^n x_k \right\|^2 = \sum_{k=1}^n \|x_k\|^2$$

**Proof.** From orthogonality,

$$\left\langle \sum_{k=1}^n x_k, \sum_{k=1}^n x_k \right\rangle = \sum_{k,j=1}^n \langle x_k, x_j \rangle = \sum_{k=1}^n \langle x_k, x_k \rangle,$$

proving the result ■

We now see some examples.

**Example** (Canonical basis of  $\ell_2$ ). In the space  $\ell_2$ ,  $x_k = (0, 0, \dots, 1, 0, \dots, 0) \in \ell_2$  for  $k = 1, 2, \dots$  is an orthonormal system in  $\ell_2$ .

**Example** (Fourier basis in  $L_2$ ). In the space  $L_2[-\pi, \pi]$ , consider the exponentials

$$e_k(t) = \frac{1}{\sqrt{2\pi}} e^{ikt}$$

for  $t \in [-\pi, \pi]$ . The set  $\{e_k\}_{k=-\infty}^{\infty}$  is an orthonormal-system in  $L_2[-\pi, \pi]$ .

### 1.7.2 Fourier Series

We can further generalize Fourier series to any Hilbert space by letting  $\{x_k : k \geq 1\}$  be an orthonormal set in  $\mathcal{H}$  as follows.

**Definition.** Consider an orthonormal-system  $\{x_k\}_{k=1}^{\infty}$  in a Hilbert space  $\mathcal{H}$  and a vector  $x \in \mathcal{H}$ .

**Definition 1.7.3** (Fourier series). The *Fourier series* of  $x$  w.r.t.  $\{x_k\}$  is the formal series

$$\sum_{k=1}^{\infty} \langle x, x_k \rangle x_k.$$

**Definition 1.7.4** (Fourier coefficient). The coefficient  $\langle x, x_k \rangle$  in the Fourier series are called *Fourier coefficients* of  $x$ .



To understand the convergence of **Fourier series**, we first focus on the finite case and study the partial sums of **Fourier series**. For  $n = 1, 2, \dots$ , we define  $S_n: \mathcal{H} \rightarrow E_n$  such that

$$S_n(x) = \sum_{k=1}^n \langle x, x_k \rangle x_k$$

for  $x \in \mathcal{H}$  where  $E_n = \text{span}\{x_1, \dots, x_n\}$ . We see that  $S_n$  is a **linear operator** and  $S_n = P_{E_n}$  is the **orthogonal projection** onto  $E_n$  since  $\langle x - S_n(x), x_k \rangle = 0$  for  $k = 1, \dots, n$  and  $S_n(x) \in E_n, x - S_n(x) \perp E_n$ .

**Theorem 1.7.2** (Bessel's inequality). Let  $\{x_k\}_k$  be an **orthogonal system** in a **Hilbert space**  $\mathcal{H}$ . then for every  $x \in \mathcal{H}$ ,

$$\sum_k |\langle x, x_k \rangle|^2 \leq \|x\|^2.$$

**Proof.** To estimate the size of  $S_n(x)$ , since  $x - S_n(x)$ , by **Theorem 1.7.1**,

$$\|S_n(x)\|^2 + \|x - S_n(x)\|^2 = \|x\|^2 \Rightarrow \|S_n(x)\|^2 \leq \|x\|^2.$$

On the other hand, again by **Theorem 1.7.1** and **orthogonality**,

$$\|S_n(x)\|^2 = \sum_{k=1}^n \|\langle x, x_k \rangle x_k\|^2 = \sum_{k=1}^n |\langle x, x_k \rangle|^2.$$

We see that by combining these two inequalities and let  $n \rightarrow \infty$ , we have the result.  $\blacksquare$

**Remark.** In particular, we see that  $\|S_n(x)\|^2 = \sum_{k=1}^n |\langle x, x_k \rangle|^2$ , with  $S_n = P_{E_n}$  and  $\|P_{E_n}x\|^2 \leq \|x\|^2$ , we have

$$\sum_{k=1}^n |\langle x, x_k \rangle|^2 \leq \|x\|^2$$

for  $x \in \mathcal{H}$ .

**Theorem 1.7.3.** Let  $\{x_k\}_k$  be an **orthonormal system** in a **Hilbert space**  $\mathcal{H}$ . Then the corresponding **Fourier series**  $S_n(x) = \sum_{k=1}^n \langle x, x_k \rangle x_k$  converges, i.e.,  $\lim_{n \rightarrow \infty} S_n(x) = S_\infty(x)$  exists for  $x \in \mathcal{H}$ . Furthermore,  $S_n = P_{E_n}$  for every  $n$  where  $E_n$  is the space spanned by  $\{x_i\}_{i=1}^n$ .<sup>a</sup>

<sup>a</sup>This includes  $n = \infty$ , where  $E_\infty$  is the **closure** of the space spanned by  $\{x_i\}_i$ .

**Proof.** We show that the sequence  $S_n(x)$  for  $n = 1, 2, \dots$  is Cauchy. This is because

$$\|S_n(x) - S_m(x)\|^2 = \sum_{k=m+1}^n |\langle x, x_k \rangle|^2,$$

and **Bessel's inequality** implies  $\sum_{k=1}^\infty |\langle x, x_k \rangle|^2 \leq \|x\|^2$ . Hence, for any  $\epsilon > 0$ , there exists  $m(\epsilon)$  such that

$$\sum_{k=m(\epsilon)+1}^\infty |\langle x, x_k \rangle|^2 < \epsilon,$$

which implies  $\|S_n(x) - S_m(x)\|^2 < \epsilon$  if  $n > m(\epsilon)$ , hence  $\{S_n(x) : n \geq 1\}$  is Cauchy, and  $\lim_{n \rightarrow \infty} S_n(x) = S_\infty(x) \in \mathcal{H}$ . Also,  $S_\infty = P_{E_\infty}$  where  $E_\infty$  is the closure of the **linear space** generated by the sequence  $\{x_k : k \geq 1\}$ .  $\blacksquare$

**Remark.** Note that the closeness of  $E_\infty$  makes sense since the self-dual of a set's **orthogonal complement** is itself if it's closed in the first place.

## Lecture 5: Abstract Fourier Series

### 1.7.3 Orthonormal Bases

13 Sep. 14:30

Follows from the discussion, we note that the [Fourier series](#) of  $x$  needs not converge to  $x$ , but we can still compute the point where it converges. However, we can now identify an extra condition so that the [Fourier series](#) of every vector  $x$  converges to  $x$ .

**Definition 1.7.5 (Complete system).** A system of vector  $\{x_k\}_k$  in [Hilbert space](#)  $\mathcal{H}$  is *complete* if the space spanned by  $\{x_k\}_k$  is **dense** in  $\mathcal{H}$ , i.e.,  $\text{span}(\{x_k\}_k) = \mathcal{H}$ .

**Definition 1.7.6 (Orthonormal basis).** A [complete orthonormal system](#) in a [Hilbert space](#)  $\mathcal{H}$  is called an *orthonormal basis* of  $\mathcal{H}$ .

**Theorem 1.7.4 (Fourier expansions).** Let  $\{x_k\}_k$  be an [orthonormal basis](#) of a [Hilbert space](#)  $\mathcal{H}$ . Then every vector  $x \in \mathcal{H}$  can be expanded in its [Fourier series](#)

$$x = \sum_k \langle x, x_k \rangle x_k.$$

This is sometimes called [Fourier inversion formula](#).

**Proof.** If an [orthogonal set](#)  $\{x_k\}_k$  is [complete](#), then  $E_\infty = \mathcal{H}$ ,  $P_{E_\infty} = I$ . This implies  $x = \sum_{k=1}^\infty \langle x, x_k \rangle x_k$  for  $x \in \mathcal{H}$ . ■

**Remark (Parseval's identity).** From [Theorem 1.7.4](#), we have  $\|x\|^2 = \|P_{E_n}x\|^2 + \|(I - P_{E_n})x\|^2$ . By letting  $n \rightarrow \infty$ , we have

$$\|x\|^2 = \lim_{n \rightarrow \infty} \|P_{E_n}x\|^2 = \lim_{n \rightarrow \infty} \sum_{k=1}^n |\langle x, x_k \rangle|^2 = \sum_{k=1}^\infty |\langle x, x_k \rangle|^2.$$

### 1.7.4 Gram-Schmidt Orthogonalization

Suppose  $x_1, x_2, \dots \in \mathcal{H}$  is a set of vectors and  $E_n = \text{span}(\{x_1, \dots, x_n\})$ . Then we can find an [orthonormal set](#)  $\{y_k \in \mathcal{H} : k \geq 1\}$  such that  $E_n = \text{span}(\{y_1, y_2, \dots, y_{m(n)}\})$  where  $m(n) \leq n$ .

Firstly, set  $y_1 = x_1 / \|x_1\|$ , and

$$y_n = \frac{(I - P_{E_{n-1}})x_n}{\|(I - P_{E_{n-1}})x_n\|}$$

if  $x_n \notin E_{n-1}$ , i.e.,  $E_{n-1}$  is properly contained in  $E_n$ .

**Remark.** Proving [completeness](#) of a set of vectors  $\{x_k : k \geq 1\}$  in  $\mathcal{H}$  can be **non-trivial**.

**Example (Haar basis).** We consider the *Haar basis* for  $L^2([0, 1])$ . Let  $h : (0, 1) \rightarrow \mathbb{R}$  where

$$h(t) = \begin{cases} 1, & \text{if } 0 < t < 1/2; \\ -1, & \text{if } 1/2 < t < 1. \end{cases}$$

Extend  $h(\cdot)$  by zero outside  $(0, 1)$ , we get  $h : \mathbb{R} \rightarrow \mathbb{R}$ ,  $h(t) = 0$  if  $t \notin (0, 1)$ , otherwise it's the same as above. The function  $t \mapsto h(2^k t)$  has support in interval  $0 < t < 2^{-k}$ . Move the support to interval  $\ell 2^{-k} < t < (\ell + 1)2^{-k}$  by translation. Set

$$h_{k,\ell}(t) = h(2^k t - \ell), \quad \ell = 0, 1, \dots, 2^k - 1.$$

The constant function plus functions  $h_{k,\ell}$ ,  $k = 0, 1, 2, \dots$ ,  $0 \leq \ell \leq 2^k - 1$  are a [complete](#)

orthogonal set for  $\mathcal{H} = L^2([0, 1])$ .

**Proof.** The span of the Haar functions includes characteristics functions  $\chi_F$  for all dyadic intervals  $[2^{-k}\ell, 2^{-k}(\ell + 1)]$  for  $\ell = 0, 1, \dots, 2^k - 1$ ,  $k = 0, 1, \dots$ . If the set is **not complete**, then there exists  $f \in L^2([0, 1])$  such that

$$\int_F f \, dt = 0$$

for all dyadic intervals  $F$ . Since we can approximate any measurable set  $E \subseteq (0, 1)$  by a union of dyadic intervals.

**Intuition.** An easy way to see this is to consider

$$\left\{ F \in \mathcal{B}: \int_F f \, dt = 0 \right\},$$

which is the Borel subalgebra of  $\mathcal{B}$ , which indeed is a Borel algebra on  $(0, 1)$ . Then observe that dyadic intervals generate all open intervals.

Hence, we see that  $\int_F f \, dt = 0$  for all measurable  $F \subseteq (0, 1)$ . Let  $F = \{t \in (0, 1): f(t) > 0\}$ , if  $m(F) > 0$ , then

$$\int_F f \, dt > 0.$$

Hence, a contradiction, so  $m(F) = 0$ . \*

**Example (Fourier basis).** Consider the Fourier basis  $e_k(t) = \frac{1}{\sqrt{2\pi}} e^{ikt}$  for  $k \in \mathbb{Z}$ ,  $-\pi < t < \pi$ . This is **complete** in  $L^2([-\pi, \pi])$ .

**Proof.** We use **Stone-Weierstrass theorem** and apply it to Fourier basis. All  $e_k(\cdot)$  are in  $C[-\pi, \pi]$ , i.e., continuous functions  $f: [-\pi, \pi] \rightarrow \mathbb{C}$ . We know that  $C([-\pi, \pi])$  is a **Banach space** with supremum norm  $\|f\| := \sup_{t \in [-\pi, \pi]} |f(t)|$ . Stone-Weierstrass theorem implies density of the space spanned by  $e_k(\cdot)$ ,  $k \in \mathbb{Z}$  in  $C([-\pi, \pi])$ , hence the completeness in  $L^2([-\pi, \pi])$  follows from the density of continuous functions in  $L^2([-\pi, \pi])$ . \*

**Proposition 1.7.1.** Let  $\{x_k\}_k$  be a linear independent system in a **Hilbert space**  $\mathcal{H}$ . Then the system  $\{y_k\}_k$  obtained by Gram-Schmidt orthogonalization of  $\{x_k\}_k$  is an **orthonormal system** in  $\mathcal{H}$ , and

$$\text{span}(\{y_k\}_{k=1}^n) = \text{span}(\{x_k\}_{k=1}^n)$$

for all  $n \in \mathbb{N}$ .

**Proof.** The system  $\{y_k\}_k$  is **orthonormal** by construction, and we obviously have the inclusion  $\text{span}(\{y_k\}_k) \subseteq \text{span}(\{x_k\}_k)$ . Furthermore, since the dimensions of these subspaces both equal  $n$  by construction, so they're indeed equal. ■

### 1.7.5 Existence of Orthogonal Bases

We see that from **Proposition 1.7.1**, we'll obtain that every **Hilbert space** that is not *too large* has an **orthonormal basis**. We call this **Hilbert space separable**.

**Definition 1.7.7 (Separable).** A metric space is *separable* if it contains a countable dense subset.

**Remark (Banach space).** For **Banach space**, **separability** follows from finding a countable set of vectors  $\{x_k\}_k$  such that the span of  $\{x_k\}_k$  is dense in  $E$ .

## Chapter 2

# Bounded Linear Operators

In this chapter we study certain transformations of [Banach spaces](#). Because these spaces are linear, the appropriate transformations to study will be linear operators. Furthermore, since [Banach spaces](#) carry topology, it is most appropriate to study continuous transformations, i.e. continuous linear operators. They are also called bounded linear operators for the reasons that will become clear shortly.

### 2.1 Bounded Linear Functionals

#### 2.1.1 Continuity and Boundedness

**Definition.** Let  $E$  be a [linear space](#) over  $\mathbb{R}$  or  $\mathbb{C}$ .

**Definition 2.1.1** (Linear functional). A *linear functional* on  $E$  is a linear operator  $f: E \rightarrow \mathbb{R}$  or  $\mathbb{C}$  such that

$$f(ax + by) = af(x) + bf(y)$$

for  $x, y \in E$ ,  $a, b \in \mathbb{R}$  or  $\mathbb{C}$ .

**Definition 2.1.2** (Bounded linear functional). A [linear functional](#)  $f(\cdot)$  is *bounded* if

$$\|f\| := \sup_{\|x\|=1} |f(x)| < \infty.$$

Clearly, the boundedness of  $f(\cdot)$  implies  $|f(x - y)| \leq \|f\| \|x - y\|$  for  $x, y \in E$ . Hence,  $f(\cdot)$  is continuous and in fact Lipschitz continuous.

**Remark.** Conversely, if a [linear functional](#) is continuous then it is bounded.

**Proof.** Suppose  $f(\cdot)$  is not bounded, then there exists a sequence  $x_n \in E$  such that  $|f(x_n)| \geq n \|x_n\|$  for  $n = 1, 2, \dots$ . By linearity,

$$\left| f\left(\frac{x_n}{n \|x_n\|}\right) \right| \geq 1, \quad n = 1, 2, \dots$$

But we know  $\lim_{n \rightarrow \infty} \frac{x_n}{n \|x_n\|} = 0$  and  $f(0) = 0$ , hence  $f(\cdot)$  is not continuous at 0. ⊗

#### 2.1.2 Dual Spaces and Hyperplanes

**Definition 2.1.3** (Dual space). Let  $E$  be a [normed space](#). The space of all [bounded linear functionals](#)  $f(\cdot)$  on  $E$  is known as the *dual space*  $E^*$  of  $E$ .

The [dual space](#) is also a [normed space](#) with [norm](#)  $\|f\| := \sup_{\|x\|=1} |f(x)|$ , which is in fact a [Banach space](#). And it is a [Banach space](#) even if the original  $E$  is not. This definition implies  $|f(x)| \leq \|f\| \|x\|$  for  $x \in E$ ,

$f \in E^*$ . Also,  $\|f\|$  is the smallest number in this inequality that makes it valid for all  $x \in X$ .

**Definition 2.1.4 (Hyperplane).** Let  $E$  be a linear space and  $H \subseteq E$  is a subspace. Say  $H$  is a *hyperplane* if  $\text{codim}(H) = 1$ , i.e.,  $\dim(E/H) = 1$ .

The goal is to make an equivalence between bounded linear functionals on  $E$  and closed hyperplanes in  $E$ .

**Problem 2.1.1.** Does there exist a non-closed hyperplane?

**Answer.** We know that this is not the case in finite dimension. And this question is analogous to asking *does there exist a subset  $F \subseteq \mathbb{R}$  which is not Lebesgue measurable?* The answer to this is yes in both cases. However, construction uses **axiom of choice**. \*

Turns out that there is a canonical correspondence between the linear functionals and the hyperplanes in  $E$ . This is clarified in Proposition 2.1.1.

**Proposition 2.1.1 (Linear functionals and hyperplanes).** Let  $E$  be a linear space.

- (a) For every linear functional on  $E$ ,  $\ker(f)$  is a hyperplane in  $E$ . If  $E$  is a Banach space, and  $f(\cdot)$  is bounded, then  $\ker(f) = H$  is closed.
- (b) If  $f, g \neq 0$  are linear functionals on  $E$  such that  $\ker(f) = \ker(g)$ , then  $f = ag$  for some  $a \neq 0$ .
- (c) For every hyperplane  $H \subseteq E$ , there exists a linear functional  $f \neq 0$  on  $E$  such that  $\ker(f) = H$ . If  $E$  is a Banach space, and  $\ker(f) = H$  is closed, then  $f(\cdot)$  is bounded.

## Lecture 6: Riesz Representation Theorem

Let's first see the proof of Proposition 2.1.1.

15 Sep. 14:30

**Proof of Proposition 2.1.1.** We prove them in order.

- (a) Let  $x, y \notin \ker(f)$ , then  $f(x), f(y) \neq 0$ , meaning that there exists a scalar  $\lambda \neq 0$  such that  $f(x) = \lambda f(y)$ , i.e.,  $x - \lambda y \in \ker(f)$ . Hence, if  $[x], [y] \in E/\ker(f)$ ,  $[x] = \lambda[y]$ , implying  $\dim(E/\ker(f)) = 1$ . Now, if  $f$  is bounded, then  $f$  is continuous, so  $\ker(f) = f^{-1}(\{0\})$  is closed.
- (b) Consider the induced functionals  $\tilde{f}, \tilde{g}: E/H \rightarrow \mathbb{R}$  or  $\mathbb{C}$  where  $H = \ker(f) = \ker(g)$ . This implies
 
$$\dim(E/H) = 1 \Rightarrow \tilde{f} = a\tilde{g} \text{ for some } a \neq 0 \Rightarrow f = ag.$$
- (c) Assume  $\dim(E/H) = 1$ , so  $E/H = \{a[x_0] : a \in \mathbb{C} \text{ (or } \mathbb{R})\}$  for some  $x_0 \in E$ . Then, for any  $x \in E$ ,  $[x] = a(x)[x_0]$  for some  $a(x) \in \mathbb{C}$  or  $\mathbb{R}$ . Define  $f(x) := a(x)$ , we see that  $f$  is linear and  $\ker(f) = H$ . Now, if  $E$  is a Banach space and  $H$  is closed with  $\dim(E/H) = 1$ . Recall that  $E/H$  is also a Banach space with norm  $\|[x]\| = \inf_{y \in H} \|x + y\|$  for  $x \in E$ .<sup>a</sup> Let  $\tilde{f}$  be a linear functional on  $E/H$ . Since  $\dim(E/H)$  is finite,  $\tilde{f}$  is continuous, implying  $|\tilde{f}([x])| \leq A\|[x]\|$  for all  $x \in E$  for some scalar  $A$ . Finally, we define  $f(x) = \tilde{f}([x])$  for  $x \in E$ , then  $\ker(f) = H$  and  $|f(x)| \leq A\|[x]\| \leq A\|x\|$ .

■

<sup>a</sup>We see now why we need the closure: otherwise we'll get a non-zero function with norm 0.

## 2.2 Representation Theorems

In concrete Banach spaces, the bounded linear functionals usually have a specific and useful form. Generally speaking, all **linear functionals** on function spaces (such as  $L_p$  and  $C(K)$ ) act by integration of the function (with respect to some weight or measure). Similarly, all **linear functionals** on sequence spaces (such as  $\ell_p$ ) act by summation with weights.

We now start by characterizing **bounded linear functionals** on a **Hilbert space**  $\mathcal{H}$ .

**Theorem 2.2.1 (Riesz representation theorem).** Let  $\mathcal{H}$  be a **Hilbert space**. Then we have the following.

- (a) For every  $y \in \mathcal{H}$ , then function  $f(x) = \langle x, y \rangle$  for  $x \in \mathcal{H}$  is a **bounded linear functional** on  $\mathcal{H}$ .
- (b) If  $f: \mathcal{H} \rightarrow \mathbb{C}$  or  $\mathbb{R}$  is a **bounded linear functional** on  $\mathcal{H}$ , then there exists  $y \in \mathcal{H}$  such that  $f(x) = \langle x, y \rangle$  for all  $x \in \mathcal{H}$ . Hence, the **dual**  $\mathcal{H}^*$  of  $\mathcal{H}$  is isometric to  $\mathcal{H}$ .

**Proof.** We prove this in order.

- (a)  $f(x) = \langle x, y \rangle$  is clearly a **linear functional**. Boundedness follows from **Cauchy-Schwarz inequality** such that

$$|f(x)| = |\langle x, y \rangle| \leq \|x\| \|y\|,$$

and we can achieve  $\|f\| = \|y\|$  by setting  $x = y / \|y\|$ .

**Note.** Note that there exists  $x_f$  such that  $\|x_f\| = 1$  since  $\|f\| = \sup_{\|x\|=1} |f(x)| = f(x_f)$ , i.e., the supremum is achieved, although we're working on an infinite dimensional space. This property does not always hold for **bounded linear functionals** on **Banach space** since the unit ball can be not compact. But this holds for **Hilbert space**.

- (b) Let  $f: \mathcal{H} \rightarrow \mathbb{C}$  or  $\mathbb{R}$  be a **bounded linear functional** on  $\mathcal{H}$ . Let  $H = \ker(f)$ , which is closed from **Proposition 1.7.1**. Let  $H^\perp$  be the **orthogonal complement** of  $H$ , i.e.,  $\mathcal{H} = H \oplus H^\perp$ . Then  $\dim(\mathcal{H} / H) = 1 \Rightarrow \dim(H^\perp) = 1$ . Choose  $y' \in H^\perp$  such that  $g(x) = \langle x, y' \rangle$ , which is in  $\mathcal{H}^*$  from (i). Furthermore, we see that  $\ker(g) = \ker(f)$ , so from **Proposition 1.7.1**,  $f$  and  $g$  are equal up to a constant  $\lambda \in \mathbb{C}$  or  $\mathbb{R}$ , i.e.,  $f = \lambda g$ . It follows that

$$f(x) = \lambda g(x) = \lambda \langle x, y' \rangle = \langle x, \lambda y' \rangle =: \langle x, y \rangle$$

for  $y := \lambda y'$ , hence we're done.<sup>a</sup>

■

<sup>a</sup>We can even show that  $y$  here is unique.

In a concise form, **Riesz representation theorem** can be realized as  $\mathcal{H}^* = \mathcal{H}$ . Given a **Hilbert space**  $\mathcal{H}$ , **Riesz representation theorem** identifies the **dual space**  $\mathcal{H}^*$ , which can be used to show **Radon-Nikodym theorem**.

**Theorem 2.2.2 (Radon-Nikodym theorem).** Let  $\mu, \nu$  be two finite measures such that  $\nu \ll \mu$ , i.e.,  $\nu$  is absolutely continuous w.r.t.  $\mu$ .<sup>a</sup> Then there exists  $g \geq 0$  such that  $g$  is  $\mu$ -integrable and

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

for  $A$  measurable.

<sup>a</sup>This means  $\mu(A) = 0 \Rightarrow \nu(A) = 0$ .

**Proof.** Consider the **linear functional**  $F: L^2(\mu) \rightarrow \mathbb{R}$  or  $\mathbb{C}$  such that

$$F(f) = \int_\Omega f \, d\mu.$$

Then we have  $\|F(f)\| \leq \|f\|_2 \sqrt{\mu(\Omega)}$ , i.e.,  $F$  is also a **bounded linear functional** on  $L^2(\mu + \nu)$ , hence by [Theorem 2.2.1](#), there exists  $h \in L^2(\mu + \nu)$  such that

$$F(f) = \int_{\Omega} fh \, d(\mu + \nu)$$

for  $f \in L^2(\mu + \nu)$ , i.e.,

$$\int_{\Omega} f \, d\mu = \int_{\Omega} fh \, d\mu + \int_{\Omega} fh \, d\nu \quad (2.1)$$

if  $f \in L^2(\mu + \nu)$ . This further implies

$$\int_{\Omega} fh \, d\nu = \int_{\Omega} f[1 - h] \, d\mu \quad (2.2)$$

for  $f \in L^2(\mu + \nu)$ .

**Claim.** Such  $h$  satisfies  $0 < h \leq 1$   $\mu$ -a.e., moreover,  $(\mu + \nu)$ -a.e.

**Proof.** We first note that  $\mu(A) = 0 \Leftrightarrow \mu(A) + \nu(A) = 0$ . Let  $A = \{h \leq 0\}$ ,  $f = \mathbb{1}_A$  be the characteristic function on  $A$ . Then [Equation 2.1](#) implies

$$\int_A h \, (d\mu + d\nu) \leq 0 \Rightarrow \mu(A) = 0 \Rightarrow h > 0 \, \mu \text{ a.e.}$$

But since  $g$  is a positive function, so we also need  $h \leq 1$ . Again, set  $B = \{h > 1\}$ ,  $f = \mathbb{1}_B$ . Then [Equation 2.1](#) implies

$$\mu(B) = \int_B h \, (d\mu + d\nu) > \mu(B)$$

unless  $\mu(B) = 0$ . ⊗

Now, by using **monotone convergence theorem**, we conclude<sup>a</sup> that [Equation 2.2](#) holds for all  $f \geq 0$ ,  $f \in L^2(\mu + \nu)$ .<sup>b</sup> Finally, let  $A \subseteq \Omega$  measurable and  $hf = \chi_A$ , from [Equation 2.2](#),

$$\nu(A) = \int_A \frac{1 - h}{h} \, d\mu.$$

By letting  $g := 1 - h/h \Rightarrow g = d\nu/d\mu$ , we're done. ■

<sup>a</sup>Consider  $f_n(t) := \min(f(t), n)$  and let  $n \rightarrow \infty$ .

<sup>b</sup>Both sides could be  $\infty$ .

**Notation** (Radon-Nikodym derivative).  $g$  in [Theorem 2.2.2](#) is referred to as the *Radon-Nikodym derivative* where  $g := d\nu/d\mu$ .

**Note (Uniqueness).** The uniqueness of Radon-Nikodym derivatives can be shown via

$$\int_A g \, d\mu = 0$$

for all  $\mu$ -measurable  $A$ , i.e.,  $g = 0$   $\mu$ -a.e.

Another useful application of [Theorem 2.2.1](#) is to characterize  $L_p$  and  $\ell_p$  spaces and their dual  $L_p^*$  and  $\ell_p^*$ . We first see the following.

**Remark.** Consider spaces  $L^p(\Omega, \mu)$  for  $1 \leq p \leq \infty$ ,  $L^q(\Omega, \Sigma, \mu) \subseteq (L^p(\Omega, \Sigma, \mu))^*$  where  $1/p + 1/q = 1$ .

**Proof.** The easy part is that  $g \in L^q$  induces a bounded linear functional on  $L^p$  by setting

$$F(f) = \int_{\Omega} f g \, d\mu.$$

By Hölder's inequality,  $|F(f)| \leq \|f\|_p \|g\|_q$ , hence  $\|F\| \leq \|g\|_q$ . To show the equality and  $\sup_{\|f\|_p} |F(f)|$  is attained for  $1 < p < \infty$ , we choose  $f = g^{q-1} \operatorname{sgn}(g)$  since

$$F(f) = \int_{\Omega} |g|^q \, d\mu = \|g\|_q^q,$$

and from  $1/p + 1/q = 1 \Rightarrow q - 1 = q/p$ , we have

$$\|f\|_p^p = \int_{\Omega} |f|^p \, d\mu = \int_{\Omega} |g|^q \, d\mu = \|g\|_q^q \Rightarrow \|f\|_p = \|g\|_q^{q/p} = \|g\|_q^{q-1}.$$

This implies

$$F(f) = \int_{\Omega} |g|^q \, d\mu \Rightarrow \|g\|_q^q = \|g\|_q \|f\|_p.$$

**Note.** We see that  $\sup_{\|f\|_p=1} |F(f)|$  is attained by taking  $f = \operatorname{sgn}(g)$ .

⊗

In particular, we have the following.

**Theorem 2.2.3** ( $L_p^* = L_q$ ). Consider the space  $L_p = L_p(\Omega, \Sigma, \mu)$  with finite measure or  $\sigma$ -finite measure  $\mu$ . Then for  $1 \leq p < \infty$  and the conjugate exponent  $q$  of  $p$ .

(a) For every weight function  $g \in L_q$ , integration with weight

$$\int_{\Omega} f g \, d\mu$$

for  $f \in L_p$  is a bounded linear functional on  $L_p$ , and its norm is  $\|G\| = \|g\|_q$ .

(b) Conversely, every bounded linear functional  $G \in L_p^*$  can be represented as integration with weight for some unique weight function  $g \in L_q$ . Moreover,  $\|G\| = \|g\|_q$ .

## Lecture 7: Hahn-Banach Theorem

**Remark.** When  $p = 1$ , the supremum is not attained necessarily. Consider  $g \in L_{\infty}$ ,  $F(f) := \int f g \, d\mu$  is dual of  $L_1$ . If  $g(\cdot)$  is continuous on  $\mathbb{R}$  with unique maximum, then the supremum  $\sup_{\|f\|_1} |F(f)|$  is not attained. In all, for  $1 \leq p \leq \infty$ ,  $L_q$  contained in the dual of  $L_p$ . If  $1 < p \leq \infty$ , then  $\sup_{\|f\|_p=1} |F(f)|$  is attained. For  $p = 1$ , the supremum is not necessarily attained.

20 Sep. 14:30

Now, we're ready to prove Theorem 2.2.3.

**Proof of Theorem 2.2.3.** To show that the dual of  $L_p^*$  is  $L_q$  if  $1 \leq p < \infty$  where  $1/p + 1/q = 1$ , we use Theorem 2.2.2. Suppose  $E = L_p(\Omega, \Sigma, \mu)$  with  $1 \leq p < \infty$  and  $f \in E^*$ . Just consider finite measure space, i.e.,  $\mu(\Omega) < \infty$ . We define a measure  $\nu$  on  $\Sigma$  by  $\nu(A) := F(\chi_A)$  for  $A \in \Sigma$ , where  $\chi_A$  is the characteristic function of  $A$ . We see that

$$\mu(A) = 0 \Rightarrow \nu(A) = 0 \Rightarrow \nu \ll \mu,$$

and Theorem 2.2.2 implies

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



for some  $g = \frac{d\nu}{d\mu} \in L_1(\Omega, \Sigma, \mu)$ . Note that  $g$  may not be in  $L_q$  since  $q > 1$ . Hence,  $F(f) = \int_{\Omega} fg \, d\mu$  for all simple function  $f$  assuming  $g \geq 0$ . Set  $f = g^{q-1}$  with the fact that  $\|F(f)\| \leq \|F\|_p \|f\|_p$ . Recall that  $q - 1 = q/p$ , hence

$$\int g^q \, d\mu \leq \|F\|_p \left( \int g^q \, d\mu \right)^{1/p} \Rightarrow \|g\|_q^q \leq \|g\|_q^q \leq \|F\|_p \|g\|_q^{q/p} = \|F\|_p \|g\|_q^{q-1},$$

hence  $\|g\|_q \leq \|F\|_p$ .

**Note.** We assume  $g \geq 0$  is because  $\nu$  is a sign measure, then if we have a bounded variation function, we can just break it into  $\nu^+ + \nu^-$ .

■

**Remark.**  $L_1$  is a subset of  $L_{\infty}^*$  but not equal to it. If  $F: L_{\infty}(\mu) \rightarrow \mathbb{C}$  is bounded linear functional, then if  $\Omega = K$  is a compact Hausdorff space,  $F$  induces a bounded linear functional on  $C(K)$ , i.e., the space of continuous functions on  $K$ . We see that  $C(K) \subseteq L_{\infty}(K, \Sigma, \mu)$  where  $\Sigma$  is the Borel algebra on  $K$ .

**Theorem 2.2.4.** Let  $E = C(K)$  be the space of continuous functions on compact Hausdorff space  $K$ . Then we have the following.

- (a) For every Borel regular signed measure on  $K$ , the functional  $F(f) = \int_K f \, d\mu$  is a bounded linear functional on  $K$ .
- (b) Every bounded linear functional on  $C(K)$  can be expressed as  $F(f) = \int_K f \, d\mu$  for some measure  $\mu$ , and  $\|F\| = |\mu|(K)$ , i.e.,  $TV(K)$ .

## 2.3 Hahn-Banach Theorem

**Hahn-Banach theorem** allows one to extend continuous linear functionals  $f$  from a subspace to the whole normed space, while preserving the continuity of  $f$ . **Hahn-Banach theorem** is a major tool in functional analysis. Together with its variants and consequences, this result has applications in various areas of mathematics, computer science, economics and engineering.

**Theorem 2.3.1 (Hahn-Banach theorem).** Let  $E_0$  be a subspace of a Banach space  $E$ . Then every  $f_0: E_0 \rightarrow \mathbb{R}$  or  $\mathbb{C}$  has a continuous extension  $f: E \rightarrow \mathbb{R}$  or  $\mathbb{C}$  such that  $\|f\| = \|f_0\|$ .

Before proving this, let's first see some implications.

**Theorem 2.3.2 (Supporting functional).** Let  $E$  be a Banach space. For every  $x \in E$ , there exists  $f \in E^*$  such that  $\|f\| = 1$ ,  $f(x) = \|x\|$ . i.e.,  $\sup_{\|y\|=1} |f(y)|$  attained at  $y = x$ .

**Proof.** Consider dimension 1 space  $E_0 = \text{span}(x) = \{tx, t \in \mathbb{R} \text{ or } \mathbb{C}\}$ . Define  $f_0: E_0 \rightarrow \mathbb{R}$  or  $\mathbb{C}$  such that  $f_0(tx) = t\|x\|$ . We see that  $\|f_0\| = 1$ , and **Theorem 2.3.1** implies there exists  $f \in E^*$  with  $\|f\| = 1$ . We see that  $f(x) = \|x\|$  explicitly attain the norm and  $\|\cdot\|$  is clearly a continuous extension of  $\|\cdot\|_{E_0} = f_0$  as required. ■

**Remark (Geometric interpretation).** Let  $B$  be a unit ball  $\{x \in E: \|x\| \leq 1\}$  in a real Banach space  $E$ . Choose  $x_0 \in \partial B$  such that  $\|x_0\| = 1$ . Then there exists  $f \in E^*$ ,  $\|f\| = 1$ ,  $f(x) = \|x\|$ . Let  $H = \ker(f) + x_0$  where  $H$  intersects  $B$  at  $x_0$ , we see that  $H$  divides  $E$  into 2 disjoint subsets, while  $B$  lies in one of which.

**Proof.** Since  $x \in B$  and  $\|x\| < 1$  implies  $|f(x)| \leq \|x\| < 1$ , we have  $f(x) < 1$ , i.e.,  $B \subseteq \{x: f(x) < 1\}$

and  $E = \{x: f(x) < 1\} \cup H \cup \{x: f(x) > 1\}$ . ⊗

**Note.** Notice that we don't have uniqueness (as we don't have it in [Theorem 2.3.1](#)) since a unit ball in  $L^\infty$  has corner, which will give multiple [hyperplanes](#)...

## Lecture 8: Proof of Hahn-Banach Theorem and Duality

We now see the proof of [Hahn-banach theorem](#).

20 Sep. 14:30

**Proof of Theorem 2.3.1.** We assume  $E$  is separable, otherwise we need [transfinite induction](#).

**Note.** Separability allows us to extend  $f_0$  one dimension at a time.

Let  $\{x_n: n \geq 1\}$  have the property that its span is dense in  $E$ . Now, if we can extend  $f_0$  such that  $E_0 \rightarrow E_0 + \{x_1\} \rightarrow E_0 + \{x_1, x_2\} \rightarrow \dots \rightarrow E_0 + \text{span}(\{x_n: n \geq 1\})$ , then we can have  $\|f\| = \|f_0\|$ , with the final space is dense in  $E$ , we can extend  $f$  to  $E$  by continuity.

To extend  $f$  by 1 dimension, i.e.,  $E \rightarrow E + \{x_1\}$ . Note that extension is determined by a single number  $\gamma = f(x_1)$  since  $f$  is a linear functional. Firstly, we want that  $\|f\| = \|f_0\|$  such that the linear functional  $f_0: E_0 \rightarrow \mathbb{R}$  extends to  $f: D_0 + \{x_1\} \rightarrow \mathbb{R}$ , i.e., we want

$$|f_0(x_0) + \lambda\gamma| \leq \|x_0 + \lambda x_1\|$$

for  $x_0 \in E_0$ ,  $\lambda \in \mathbb{R}$ . By dividing the inequality by  $\lambda \neq 0$ , it's sufficient to find  $\gamma$  such that  $|f_0(x_0) + \gamma| \leq \|x_0 + x_1\|$ ,  $x_0 \in E_0$ .

Suppose  $f_0$  is a real-valued function, we need

$$-\|x_0 + x_1\| \leq f_0(x_0) + \gamma \leq \|x_0 + x_1\|$$

for all  $x_0 \in E_0$ . Such a  $\gamma$  exists, provides  $\|x_0 + x_1\| - f_0(x_0) \geq -\|x'_0 + x_1\| - f_0(x'_0)$  for all  $x_0, x'_0 \in E_0$ . Furthermore, this is equivalent to write

$$f_0(x_0 - x'_0) \leq \|x_0 + x_1\| + \|x'_0 + x_1\|$$

for all  $x_0, x'_0 \in E_0$ , i.e.,  $f_0(x_0 - x'_0) \leq \|x_0 + x_1\| + \|-x_1 - x'_0\|$  for  $x_0, x'_0 \in E_0$ . Recall that  $\|f_0\| = 1$ , we have

$$f_0(x_0 - x'_0) \leq \|x_0 - x'_0\| \leq \|x_0 + x_1\| + \|-x_1 - x'_0\|.$$

For complex valued  $f$ , consider  $f: E \rightarrow \mathbb{C}$  be a linear functional over  $\mathbb{C}$  and let  $g(x) = \text{Re } f(x)$ . Then  $g: E \rightarrow \mathbb{R}$  is a real-valued linear functional. We see that  $f(x) = g(x) - ig(ix)$  for all  $x \in E$ .<sup>a</sup> Conversely, if  $g: E \rightarrow \mathbb{R}$  is a real linear functional on Banach space  $E$  over  $\mathbb{C}$ , then  $f: E \rightarrow \mathbb{C}$  defined by  $f(x) = g(x) - ig(ix)$ ,  $x \in E$  is a complex linear functional on  $E$ .

But we need to be a bit careful since when we extend  $f_0: E_0 \rightarrow \mathbb{C}$ , we're extending 2 real dimensions since for  $g_0 = \text{Re } f_0$ , we need to do  $E_0 \rightarrow E_0 + \{x_1\} \rightarrow E_0 + \{x_1, ix_1\}$ . Again, define  $f(\cdot) = g(\cdot) - ig(i\cdot)$ , we want to show  $|f| = \|f_0\|$ . We use the fact that for  $x \in E_0 + \{\lambda x_0: \lambda \in \mathbb{C}\}$ ,

$$e^{i\theta} f(x) = f(xe^{i\theta})$$

for  $\theta \in \mathbb{R}$ . Choose  $\theta$  such that  $f(xe^{i\theta}) = g(xe^{i\theta})$ , and since we already have  $|g(xe^{i\theta})| \leq \|f_0\| \|xe^{i\theta}\|$ , we see that  $|f(x)| \leq \|f_0\| \|x\|$  for  $x \in E_0 + \{\lambda x_1: \lambda \in \mathbb{C}\}$ . ■

<sup>a</sup>Since  $f(ix) = if(x)$ , hence  $g(ix) = -\text{Im } f(x)$ .

Before we end this section, we see some corollaries of [Hahn-Banach theorem](#). From [Theorem 2.3.2](#), we see that for every vector  $x$ , we indeed attain its [norm](#) on some [functional](#)  $f \in E^*$ , i.e., their supporting [functional](#). But recall that the [norm](#) of a [functional](#)  $f \in E^*$  is defined as

$$\|f\| := \sup_{x \neq 0} \frac{|f(x)|}{\|x\|},$$

and in general,  $f$  will not attain its [norm](#) on some vector  $x$ . Above observation leads to the following.

**Corollary 2.3.1.** For every vector  $x$  in a **normed space**  $E$ ,

$$\|x\| = \max_{f \neq 0} \frac{|f(x)|}{\|f\|}$$

where the maximum is taken over all non-zero **functionals**  $f \in E^*$ .

**Hahn-Banach theorem** implies that there are enough **bounded linear functionals**  $f \in E^*$  on every space  $E$ . One manifestation of this is the following.

**Corollary 2.3.2 (Separation of points).** For every two vectors  $x_1 \neq x_2$  in a **normed space**  $E$ , there exists a **functional**  $f \in E^*$  such that  $f(x_1) \neq f(x_2)$ .

**Proof.** The **supporting functional**  $f \in E^*$  of the vector  $x = x_1 - x_2$  must satisfy

$$f(x_1 - x_2) = \|x_1 - x_2\| \neq 0,$$

as required. ■

## 2.4 Second Dual Space

Let  $E$  be a **normed space**, then the **functionals**  $f^*$  are designed to act on vectors  $x \in E$  via

$$f: x \mapsto f(x).$$

But indeed, we can instead say that *vectors*  $x \in E$  act on **functionals**  $f \in E^*$  via

$$x: f \mapsto f(x).$$

Thus, a vector  $x \in E$  can itself be considered as a function from  $E^*$  to  $\mathbb{R}$ . Furthermore, this function  $x$  is clearly linear, so we may consider  $x$  as a **linear functional** on  $E^*$ . Also, the inequality

$$|f(x)| \leq \|x\| \|f\|$$

shows that this **functional** is bounded, so  $x \in E^{**}$ . We may instead write  $x$  as  $x^{**}$  for clarity. Note that the **norm** of  $x^{**}$  as a **functional** is  $\|x^{**}\|_{E^{**}} \leq \|x\|$  since

$$\|x^{**}\| = \sup_{\substack{\|f\|=1 \\ f \in E^*}} |x^{**}(f)| = \sup_{\substack{\|f\|=1 \\ f \in E^*}} |f(x)| \leq \|x\|,$$

implying that  $\|x^{**}\| \leq \|x\|$  for all  $x \in E$ . But from **supporting functional**  $f \in E^*$  of  $x$ , we actually have

$$\|x^{**}\| = \|x\|,$$

i.e., we have a *canonical embedding* of  $E$  into  $E^{**}$ . The above discussion leads to **Theorem 2.4.1**.

**Theorem 2.4.1 (Second dual space).** Let  $E$  be a **normed space**. Then  $E$  can be considered as a **linear subspace** of  $E^{**}$ . For this, a vector  $x \in E$  is considered as a **bounded linear functional** on  $E^*$  via the action

$$x: f \mapsto f(x), \quad f \in E^*.$$

## 2.5 Reflexive Spaces

**Definition 2.5.1 (Reflexive space).** A **normed space**  $E$  is called *reflexive space* if  $E \rightarrow E^{**}$ .<sup>a</sup>

<sup>a</sup>Under the canonical embedding.

**Example.**  $L_p$  spaces for  $1 < p < \infty$  are [reflexive spaces](#).

**Proof.** We know that  $L_p^* = L_q$  where  $1 \leq p < \infty$  for  $q$  being the conjugate index of  $p$ . ⊗

**Example.**  $L_p$  spaces for  $p = 1$  or  $\infty$  are not [reflexive spaces](#)

**Proposition 2.5.1.** Let  $E$  be a [reflexive space](#), then every [linear functional](#)  $f \in E^*$  attains its [norm](#) on  $E$ .

**Proof.** By [reflexivity](#), the [supporting functional](#) of  $f$  is a vector  $x \in E^{**} = E$ , thus  $\|x\| = 1$  and  $f(x) = \|f\|$ , as required. ■

**Remark (James' theorem).** The converse of [Proposition 2.5.1](#) is also true, i.e., if every [functional](#)  $f \in E^*$  on a [Banach space](#)  $E$  attains its [norm](#), then  $E$  is [reflexive](#). This is the so-called *James' theorem*.

## Lecture 9: Hahn-Banach Theorem for Sublinear Functions

From [Proposition 2.5.1](#), we see that to show a [Banach space](#)  $E$  is not [reflexive](#), it's sufficient to find  $f \in E^*$  such that  $\sup_{\|x\|=1} |f(x)|$  is not attained. 27 Sep. 14:30

**Example.** Let  $C([0, 1])$  be the space of continuous functions  $g: [0, 1] \rightarrow \mathbb{C}$  with  $\|g\| := \sup_{0 \leq t \leq 1} |g(t)|$ . Then for  $f \in E^*$ ,

$$f(g) = \int_0^1 h(x)g(x) \, dx$$

for

$$h(x) = \begin{cases} -1, & \text{if } 0 < x < \frac{1}{2}; \\ 1, & \text{if } \frac{1}{2} < x < 1. \end{cases}$$

Then we have  $\|f\| = 1 = \sup_{\|g\|=1} |f(g)|$ , but the supremum is not attained since  $g$  needs to be continuous.

## 2.6 Separation of Convex Sets

In this section, we can extend [supporting functional theorem](#) such that we now have it for arbitrary [convex sets](#) other than the unit ball. We see that [supporting functional theorem](#) depends on [Hahn-Banach theorem](#), so we should first generalize [Hahn-Banach theorem](#).

### 2.6.1 Sublinear Functions

By looking into the proof of [Hahn-Banach theorem](#), we see that we only used positive homogeneity and triangle inequality of the axiom of [norm](#), which suggests we define the following.

**Definition 2.6.1 (Sublinear).** Let  $E$  be a [linear vector space](#). a function  $\|\cdot\| : E \rightarrow [0, \infty)$  is *sublinear* if it satisfies

- (a)  $\|\lambda x\| = \lambda \|x\|$  for  $\lambda \in \mathbb{R}^+$ ,  $x \in E$ .
- (b)  $\|x + y\| \leq \|x\| + \|y\|$ ,  $x, y \in E$ .

**Remark (Differences from norm).** Note that for a [sublinear](#) function to be a [norm](#), we need

- (a)  $\|-x\| = \|x\|$ ,  $x \in E$
- (b)  $\|x\| = 0 \Rightarrow x = 0$ .

**Theorem 2.6.1** (Hahn-Banach theorem for sublinear functions). Let  $E_0$  be a subspace of a linear vector space over  $\mathbb{R}$ . Let  $\|\cdot\|$  be a sublinear functional on  $E$ , and  $f_0: E_0 \rightarrow \mathbb{R}$  be a linear functional on  $E_0$  satisfying  $f_0(x) \leq \|x\|$  for  $x \in E_0$ . Then  $f_0$  admits an extension  $f$  to  $E$  such that  $f(x) \leq \|x\|$  for  $x \in E$ .

**Proof.** The idea is the same from Theorem 2.3.1. ■

## 2.6.2 Geometric Properties of Sublinear Functions

We see that by considering sublinear functionals instead of norms offers us more flexibility in geometric applications. In particular, sublinear functionals arise as Minkowski functionals of convex sets.

**Definition 2.6.2** (Absorbing). A subset  $K$  of a linear vector space is *absorbing* if

$$E = \bigcup_{t \geq 0} tK$$

where  $tK := \{tk : k \in K\}$ .

**Definition 2.6.3** (Minkowski functional). Let  $K$  be an absorbing convex subset of a linear vector space  $E$  such that  $0 \in K$ . Then the *Minkowski functional*  $\|\cdot\|_K$  is defined as

$$\|x\|_K := \inf \{t > 0 : x/t \in K\}.$$

**Proposition 2.6.1.** Let  $K$  be an absorbing convex subset of a linear vector space  $E$  such that  $0 \in K$ . Then Minkowski functional  $\|x\|_K$  is a sublinear functional on  $E$ . Conversely, let  $\|\cdot\|$  be a sublinear functional on a linear vector space  $E$ , then the sub-level set

$$K = \{x \in E : \|x\| \leq 1\}$$

is an absorbing convex set, and  $0 \in K$ .

**Proof.** To prove the forward direction, the main observation is that since  $0 \in K$  and  $K$  is convex, then  $x \in K \Rightarrow tx \in K$  if  $0 \leq t < 1$ . To show dilation, for  $\lambda > 0$ ,

$$\|\lambda x\| = \inf \left\{ t > 0 : x \in \frac{t}{\lambda} K \right\} = \lambda \inf \{s > 0 : x \in sK\} = \lambda \|x\|.$$

To show triangle inequality, suppose  $x \in tK$ ,  $y \in sK$ , then  $x = tk_1$ ,  $y = sk_2$  for some  $k_1, k_2 \in K$ . We then have

$$x + y = (t + s) \left( \frac{t}{t+s} k_1 + \frac{s}{t+s} k_2 \right) = (t + s)k$$

for some  $k \in K$  since  $K$  is convex, hence  $x + y \in (t + s)K$ , we then have  $\|x + y\| \leq \|x\| + \|y\|$ .

Now, if  $\|\cdot\|$  is sublinear, then  $K = \{x \in E : \|x\| \leq 1\}$  is absorbing, convex and  $0 \in K$ .<sup>a</sup> ■

<sup>a</sup> $0 \in K$  since  $\|0\| = 0$ , while the convexity comes from the triangle inequality.

**Remark.** If  $K \neq -K$ , then  $\exists x \in E$  with  $\|x\| \neq \|-x\|$ . If  $K = E$ , then  $\|\cdot\| \equiv 0$ .

## 2.6.3 Separation of Convex Sets

Hahn-Banach theorem has some remarkable geometric implications, which are grouped together under the name of *separation theorems*. Under mild topological requirements, these results guarantee that two

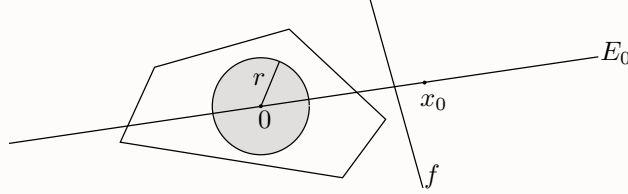
convex sets  $A, B$  can always be separated by a [hyperplane](#).

**Theorem 2.6.2** (Separation of a point from a convex set). Let  $K$  be an open convex subset of a normed space  $E$  and  $x_0 \notin K$ . Then there exists a continuous linear functional  $f: E \rightarrow \mathbb{R}$  with  $f \neq 0$  and  $f(x) < f(x_0)$  for  $x \in K$ .

**Proof.** By translation, we can assume without loss of generality that  $0 \in K$ . Since  $K$  is open, it is [absorbing](#). Now, let  $\|\cdot\|_K$  be the [Minkowski functional](#), then

$$\|x\|_K \leq \frac{1}{r} \|x\|$$

for  $x \in E$  if  $B(0, r) \subseteq K$ .



Proceed as in [Theorem 2.3.2](#) for unit ball, we define  $f_0$  on  $\text{span}(\{x_0\})$  by

$$f_0(tx_0) = t \|x_0\|_K$$

for  $t \in \mathbb{R}$ . Then if  $E_0 = \{\lambda x_0 : \lambda \in \mathbb{R}\}$ ,  $f_0(x) \leq \|x\|_K$  for  $x \in E_0$  (i.e.,  $\|\cdot\|_K$  dominates  $f_0$ ) since for  $t \geq 0$ ,

$$f_0(tx_0) = t \|x_0\|_K = \|tx_0\|_K;$$

while for  $t \leq 0$ ,

$$f_0(tx_0) = t \|x_0\|_K \leq 0 \leq \|tx_0\|_K.$$

Then from [Theorem 2.3.1](#), we can extend  $f_0$  to  $f: E \rightarrow \mathbb{R}$  such that

$$f(x) \leq \|x\|_K \leq \frac{1}{r} \|x\|$$

for  $x \in E$ , hence  $f \in E^*$ . For separation, we see that if  $x \in K$  (hence in  $E$ ),

$$f(x) \leq \|x\|_K \leq 1 \leq \|x_0\|_K = f_0(x_0) = f(x_0),$$

hence  $f(x) \leq f(x_0)$ . To get a strict separation, since  $K$  is open, so  $x + tv \in K$  for  $x \in K$  and some  $t > 0$  and all  $v$  with  $\|v\| = 1$ . Hence, for all  $t = t_x > 0$ , we have

$$f(x + tv) \leq f(x_0) \Rightarrow f(x) + t \sup_{\|v\|=1} f(v) \leq f(x_0).$$

With the fact that  $f \neq 0$ , so  $\|f\| = \sup_{\|v\|=1} f(v) \neq 0$ , we conclude that

$$f(x) < f(x_0).$$

■

**Theorem 2.6.3** (Separation of convex sets). Let  $A, B$  be disjoint [convex subsets](#) of a [Banach space](#)  $E$ .

- (a) If  $A$  is open, then there  $\exists f: E \rightarrow \mathbb{R}$  such that  $f(a) < f(b)$  for  $a \in A, b \in B$ .
- (b) If  $A, B$  are closed and  $B$  is compact, then there  $\exists f: E \rightarrow \mathbb{R}$  such that  $\sup_{a \in A} f(a) < \inf_{b \in B} f(b)$ .

**Proof.** We have the following.

- (a) Let  $K = A - B = \{a - b : a \in A, b \in B\}$ , we then see that  $K$  is open, [convex](#) and  $0 \notin K$ . By

**Theorem 2.6.2**, there exists  $f \in E^*$  such that

$$f(a - b) < f(0) = 0$$

for  $a \in A, b \in B$ , hence  $f(a) < f(b)$  for  $a \in A, b \in B$ .

(b) Let  $A$  be closed,  $B$  be compact. Then we have

$$d(A, B) = \inf \{\|x - y\| : x \in A, y \in B\} = r > 0.$$

Define  $A_\delta := \{x \in E : d(x, A) < \delta\}$  where  $A_\delta$  is open. By setting  $\delta := r/2$ , we have  $A_\delta \cap B = \emptyset$ . From (a), we see that there exists  $f \in E^*$  such that  $f(x) < f(y)$  for  $x \in A_\delta, y \in B$ . Then  $a \in A$  implies  $a + \delta/2v \in A_\delta$  if  $\|v\| = 1$ , hence

$$f(a + \delta/2v) < f(b)$$

for  $b \in B$ . So

$$f(a) + \frac{\delta}{2}f(v) < f(b)$$

for  $b \in B, \|v\| = 1$ . Take the supremum over  $\|v\| = 1$ , we have  $\sup_{\|v\|=1} |f(v)| = \delta > 0$ , implying  $f(a) < f(b) - \delta, a \in A, b \in B$ . Finally, we have

$$\sup_{a \in A} f(a) < \inf_{b \in B} f(b).$$

■

## Lecture 10: Adjoint Operators and Ergodic Theorem

Before ending this section, we have this final characterization of **convex sets**: they're intersections of **half-spaces**! 29 Sep. 14:30

**Definition 2.6.4 (Half-space)**. A *half-space*  $H \subseteq E$  has the form of

$$H = \{x \in E : f(x) \leq \lambda\}$$

for  $f \in E^*$ , i.e., it is what lies on one side of a **hyperplane**.

**Corollary 2.6.1**. Let  $K \subseteq E$  be closed **convex set**. Then  $K$  is the intersection of all **half-spaces** containing  $K$ .

**Proof**. Firstly,  $K$  is trivially contained in the intersection of the **half-spaces** that contain  $K$ . Denote such an intersection as  $S$ , then we have  $K \subseteq S$ . On the other hand, to show  $K \supseteq S$ , if  $x_0 \notin K$ , we show that there's a **half-space** contains  $K$  but not  $x_0$ , hence  $x_0 \notin S$  too, i.e.,  $S \subseteq K$ .

From **Theorem 2.6.3** with  $A = K$  and  $B = \{x_0\}$ , there exists  $f \in E^*$  such that  $\lambda := \sup_{k \in K} f(k) < f(x_0)$ . We then see that the **half-space**  $\{x \in E : f(x) \leq \lambda\}$  contains  $K$  but not  $x_0$ . ■

## 2.7 Bounded Linear Operators

Turns out that we can generalize the notion of **linear functionals**  $f: E \rightarrow \mathbb{R}$  or  $\mathbb{C}$  by further abstracting out the domain by another **Banach space**.

As one can imagine, several results for **linear operators** will be generalizations of those we have already seen for **linear functionals**, but there'll be important differences though. For example, a natural extension of **Hahn-Banach theorem** fails for **linear operators**.

Firstly, same as before, the **operator norm** is defined as follows.

**Definition 2.7.1** (Operator norm). Given an operator  $T$ , its *operator norm* is defined as

$$\|T\| := \sup_{\|x\|=1} \|Tx\|.$$

**Remark.** The *operator norm* is a *norm* on *bounded linear operators*.

### 2.7.1 Continuity and Boundedness

Then, we have the following.

**Definition.** Let  $X, Y$  be two *Banach spaces*.

**Definition 2.7.2** (Linear operator). A map  $T: X \rightarrow Y$  is called a *linear operator* if

$$T(ax + by) = aTx + bTy$$

for  $a, b \in \mathbb{R}$  or  $\mathbb{C}$ ,  $x, y \in X$ .

**Definition 2.7.3** (Bounded linear operator). A *linear operator*  $T$  is *bounded* if  $\|T\| < \infty$ .

**Remark** (Bounded operator). We can also talk about boundedness of a(n) (nonlinear) operator  $T$  just the same as requiring  $\|T\| < \infty$ .

As before, given [Definition 2.7.1](#), we always have

$$\|Tx\| \leq \|T\| \|x\|$$

for a *linear operator*  $T: X \rightarrow Y$ ,  $x \in X$ .

**Definition 2.7.4** (Lipschitz). The operator  $T$  is called *Lipschitz* if

$$\|Tx - Ty\| \leq \|T\| \|x - y\|$$

for  $x, y \in E$ .

**Remark** (Continuity and Boundedness). Same as *linear functionals*, the continuity and boundedness of *linear operators* are equivalent.

### 2.7.2 Space of Operators

Let  $X$  and  $Y$  be *normed space*, and let  $\mathcal{L}(X, Y)$  be the space of *bounded linear operators*  $T: X \rightarrow Y$ , then  $\mathcal{L}(X, Y)$  is a *Banach space* under the *norm*  $T \rightarrow \|T\|$ .

**Example.** The *dual space* of  $E$  is just  $E^* = \mathcal{L}(E, \mathbb{R})$ .

**Remark.** In particular, we have

- (a)  $\|T\| = 0 \Leftrightarrow T = 0$ .
- (b)  $\|\lambda T\| = |\lambda| \|T\|$  for  $\lambda \in \mathbb{R}$  or  $\mathbb{C}$ ,  $T \in \mathcal{L}(X, Y)$ .
- (c)  $\|T + S\| \leq \|T\| + \|S\|$ ,  $T, S \in \mathcal{L}(X, Y)$ .



$$(d) \|TS\| \leq \|T\| \|S\|, T, S \in \mathcal{L}(X, Y).$$

### 2.7.3 Adjoint Operators

The concept of **adjoint operators** is a generalization of matrix transpose in linear algebra. Recall that if  $A = (a_{ij})$  is an  $n \times n$  matrix with complex entries, then the Hermitian transpose of  $A$  is an  $n \times n$  matrix  $A^* = (\overline{a_{ij}})$ . The transpose thus satisfies the identity

$$\langle A^*x, y \rangle = \langle x, Ay \rangle$$

for  $x, y \in \mathbb{C}^n$ . We now extend this to **linear operators**.

**Definition 2.7.5 (Adjoint operator).** Let  $T \in \mathcal{L}(X, Y)$ , the *adjoint*  $T^* \in \mathcal{L}(Y^*, X^*)$  of  $T$  is defined as

$$T^*f: X \rightarrow \mathbb{R} \text{ or } \mathbb{C}$$

for  $f \in Y^*$ , and  $T^*f(x) = f(Tx)$  for  $x \in X$ .

We should note that  $T^*$  is indeed a **bounded linear operator** since

$$|T^*f(x)| = |f(Tx)| \leq \|f\| \|Tx\| \leq \|f\| \|T\| \|x\|$$

for  $x \in X$ , hence  $T^*f$  is a **linear functional** where

$$\|T^*f\| = \sup_{\|x\|=1} |T^*f(x)| \leq \sup_{\|x\|=1} \|f\| \|T\| \|x\| = \|f\| \|T\|,$$

hence,  $T^*f \in X^*$  and  $\|T^*f\| \leq \|T\| \|f\|$ . So, we have  $T^*: Y^* \rightarrow X^*$  with  $T^*$  being a **linear operator** and  $T^*$  is **bounded** with

$$\|T^*\| \leq \|T\|.$$

In fact, we can achieve equality, which is shown in **Proposition 2.7.1**.

**Proposition 2.7.1.** For every  $T \in \mathcal{L}(X, Y)$ , the **adjoint**  $T^*$  is in  $\mathcal{L}(Y^*, X^*)$  with  $\|T^*\| = \|T\|$ .

**Proof.** Since

$$\begin{aligned} \|T^*\| &= \sup_{\|f\|_{Y^*}=1} \|T^*f\|_{X^*} = \sup_{\|f\|_{Y^*}=1} \sup_{\|x\|_X=1} |T^*f(x)| \\ &= \sup_{\|f\|_{Y^*}=1} \sup_{\|x\|_X=1} |f(Tx)| = \sup_{\|x\|_X=1} \sup_{\|f\|_{Y^*}=1} |f(Tx)|. \end{aligned}$$

By choosing  $f$  to be a **supporting functional** of  $Tx$ ,  $\sup_{\|f\|_{Y^*}=1} |f(Tx)| = \|Tx\|_{Y^*}$ , hence

$$\|T^*\| = \sup_{\|x\|_X=1} \|Tx\|_{Y^*} = \|T\|.$$

■

Let's look at some properties of **adjoint operators**. Let  $T, S \in \mathcal{L}(X, Y) \Rightarrow T^*, S^* \in \mathcal{L}(Y^*, X^*)$ , then

- (a)  $(aT + bS)^* = aT^* + bS^*$ ,  $a, b \in \mathbb{R} \text{ or } \mathbb{C}$ . Also,  $(aT)^*f(x) = f(aTx) = af(Tx) = aT^*f(x)$ .
- (b)  $(ST)^* = T^*S^*$ . This implies that if  $T \in \mathcal{L}(X, X)$  is invertible, then  $T^* \in \mathcal{L}(X^*, X^*)$  is invertible and  $(T^*)^{-1} = (T^{-1})^*$ .

**Remark (Adjoint operators on Hilbert spaces).** Specialize to **Hilbert space**  $\mathcal{H}$ , then by **Riesz representation theorem**,  $\mathcal{H}^* \equiv \mathcal{H}$ , i.e.,  $f \in \mathcal{H}^* \Leftrightarrow \exists y \in \mathcal{H}$  such that  $f(x) = \langle x, y \rangle$  for  $x \in \mathcal{H}$ . Let  $T \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ , and  $T^* \in \mathcal{L}(\mathcal{H}^*, \mathcal{H}^*)$  with  $T^*f(x) = f(Tx) = \langle Tx, y \rangle$  for  $x, y \in \mathcal{H}$ ,  $f \in \mathcal{H}^*$ .

Write  $T^*f(x) = \langle x, T^*y \rangle$ , which defined  $T^*y: \mathcal{H} \rightarrow \mathcal{H}$ , hence  $\langle Tx, y \rangle = \langle x, T^*y \rangle$  for  $x, y \in \mathcal{H}$ . Clearly,  $T^*$  is a **bounded linear operator** on  $\mathcal{H}$ , i.e.,  $T^* \in \mathcal{L}(\mathcal{H}^*, \mathcal{H}^*)$  since

$$\|T^*\| = \sup_{\|y\|=1} \|T^*y\| = \sup_{\|y\|=\|x\|=1} \langle x, T^*y \rangle = \sup_{\|y\|=\|x\|=1} \langle Tx, y \rangle = \|T\|$$

just like [Proposition 2.7.1](#). We see that  $T^* \in \mathcal{L}(\mathcal{H}^*, \mathcal{H}^*) \Rightarrow T^* \in \mathcal{L}(\mathcal{H}, \mathcal{H})$  via [Riesz representation](#). Note that if  $T^* \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ ,

$$(aT)^* = \bar{a}T^*$$

for  $a \in \mathbb{C}$ .

Just as with [Hilbert space](#), we have a generalized notion of [orthogonality](#), which we call [annihilator](#).

**Definition 2.7.6 (Annihilator).** Let  $A \subseteq X$  where  $X$  is a [Banach space](#), then the *annihilator*  $A^\perp$  of  $A$  is a subset of  $X^*$  defined as

$$A^\perp := \{f \in X^* : f(x) = 0, x \in A\}.$$

**Note.**  $A^\perp$  is a closed linear subspace of  $X^*$ .

**Proposition 2.7.2.** Given two [Banach spaces](#)  $X$  and  $Y$ , let  $T \in \mathcal{L}(X, Y)$  and  $T^* \in \mathcal{L}(Y^*, X^*)$ . Then  $(\text{Im } T)^\perp, \ker(T^*) \subseteq Y^*$  satisfy

$$(\text{Im } T)^\perp = \ker(T^*).$$

**Proof.** Since  $f \in (\text{Im } T)^\perp \Leftrightarrow f(Tx) = 0$  for all  $x \in X$ , i.e.,  $T^*f(x) = 0 \Leftrightarrow T^*f = 0 \Leftrightarrow f \in \ker(T^*)$ , proving the result. ■

**Corollary 2.7.1.** Let  $\mathcal{H}$  be a [Hilbert space](#), and  $T \in \mathcal{L}(\mathcal{H}, \mathcal{H})$ . Then the orthogonal decomposition holds, i.e.,

$$\mathcal{H} = \overline{\text{Im } T} \oplus \ker(T^*).$$

**Proof.** By [Proposition 2.7.2](#),  $\ker(T^*) = (\text{Im } T)^\perp$ . And since  $\mathcal{H}$  is [Hilbert space](#),  $\overline{\text{Im } T} = \text{Im } T$  from the fact that if  $E \subseteq \mathcal{H}$ ,  $(E^\perp)^\perp = \overline{E}$ , hence  $(\overline{\text{Im } T})^\perp = \ker T^*$ . Just by a simple application of [Theorem 1.6.1](#), the proof is complete. ■

## 2.7.4 Ergodic Theory

We now see an application on ergodic theorems. Ergodic theorems allow one to compute space averages as time averages. Given a probability space  $(\Omega, \mathcal{F}, P)$  with  $P(\Omega) = 1$ , let  $T: \Omega \rightarrow \Omega$  be a measurable map, i.e.,  $T^{-1}A \in \mathcal{F}$  if  $A \in \mathcal{F}$ . Then, we define the following.

**Definition 2.7.7 (Measure-preserving).** Let  $(\Omega, \mathcal{F}, P)$  be a probability space. A transformation  $T: \Omega \rightarrow \Omega$  is called *measure-preserving* if

$$P(T^{-1}A) = P(A)$$

for  $A \in \mathcal{F}$ , where  $T^{-1}A = \{\omega \in \Omega : T\omega \in A\}$ .

Let's first see some examples which illustrate the so-called *time and space averages*. We start with simple dynamical systems corresponding to rotation.

**Example (Rotation).** Let  $\Omega = [0, 1]$ ,  $P$  be the Lebesgue measure and  $\mathcal{F}$  be Borel sets. Given  $\lambda \in \mathbb{R}$ , define

$$T\omega = \omega + \lambda \bmod 1.$$

This is equivalent to rotation on the unit circle through an angle  $2\pi\lambda$ , and we see that  $T$  is [measure-preserving](#) and one-to-one, and  $T^{-1}$  exists.

**Example (Shift Operator).** Let  $\Omega = [0, 1]$ ,  $P$  be the Lebesgue measure and  $\mathcal{F}$  be Borel sets. Now, let

$$T\omega = 2\omega \bmod 1.$$

Then we see that  $T$  is just the shift operator on the binary representation, i.e., given  $\omega = \sum_{j=1}^{\infty} \frac{a_j}{2^j}$  for  $a_j = 0$  or  $1$ , then

$$T\omega = \sum_{j=1}^{\infty} \frac{a_{j+1}}{2^j}.$$

Now, let the *dyadic interval*  $I_{n,k}$  be defined as

$$I_{n,k} := \left[ \frac{k-1}{2^n}, \frac{k}{2^n} \right]$$

for  $1 \leq k \leq 2^n$ , we have  $T^{-1}I_{n,k} = I_{n+1,k} \cup I_{n+1,k+2^n}$ , hence  $P(T^{-1}I_{n,k}) = P(I_{n,k})$  for all dyadic intervals  $I_{n,k}$ . This implies

$$P(T^{-1}O) = P(O)$$

for all  $O \in \mathcal{F}$ , hence  $T$  is **measure-preserving**, but not one-to-one. In fact,  $T$  is a two-to-one mapping. The action of  $T$  is  $[0, 1/2] \xrightarrow{T} [0, 1]$ ,  $[1/2, 1] \xrightarrow{T} [0, 1]$ . We see that  $T$  doubles the length of a dyadic interval. To summarize,

- $T$  is **measure-preserving** since it is two-to-one.
- $T$  is an expanding map, which is called hyperbolic.

## Lecture 11: Ergodic Theorem and Open Mapping

Now, we're ready to discuss ergodic theorem formally. Suppose  $T: \Omega \rightarrow \Omega$  is **measure-preserving**, we can associate operator  $U$  on  $L_2(\Omega)$  by defining  $Uf(\omega) = f(T\omega)$  for  $f \in L_2(\Omega)$  and  $\omega \in \Omega$ . Notice that 4 Oct. 14:30

$$\int_{\Omega} f(T\omega) d\mu(\omega) = \int_{\Omega} f(\omega) d\mu(\omega)$$

for all  $f \in L_1(\Omega)$ ,<sup>1</sup> so for  $\varphi \in L_2(\Omega)$ ,  $U\varphi(\omega) = \varphi(T\omega)$  and since

$$\langle U\varphi, U\psi \rangle = \int_{\Omega} \varphi(T\omega)\psi(T\omega) d\mu(\omega) = \int_{\Omega} \varphi(\omega)\psi(\omega) d\mu(\omega) = \langle \varphi, \psi \rangle$$

for  $\varphi, \psi \in L_2(\Omega)$ , we see that  $U$  is a **bounded linear operator** on  $\mathcal{H} = L_2(\Omega)$  with  $\|U\| = 1$ ,  $\|U\varphi\| = \|\varphi\|$ ,  $\varphi \in \mathcal{H}$ . In addition, for  $\varphi, \psi \in \mathcal{H}$ ,  $\langle U\varphi, U\psi \rangle = \langle \varphi, \psi \rangle$  implies  $\langle U^*U\varphi, \psi \rangle = \langle \varphi, \psi \rangle$ , which further implies  $U^*U = I$ , so  $U$  is one-to-one. Let's first see one more definition before we proceed.

**Definition 2.7.8 (Unitary operator).** A *unitary operator* is a **bounded linear operator**  $U: \mathcal{H} \rightarrow \mathcal{H}$  on a **Hilbert space**  $\mathcal{H}$  such that  $U$  is surjective and for all  $x, y \in \mathcal{H}$ ,

$$\langle Ux, Uy \rangle_{\mathcal{H}} = \langle x, y \rangle_{\mathcal{H}}.$$

Notice that  $U$  is not necessarily onto. However, if  $U$  is indeed onto, then  $UU^* = U^*U = I$ , implying that  $U$  is a **unitary operator** on  $\mathcal{H}$  and invertible.

**Note.**  $U$  is invertible if and only if  $T$  is one-to-one.

**Proof.** Since  $U$  just need to be onto for  $U$  being invertible, with  $U^*\varphi(\omega) = \varphi(T^{-1}\omega)$  for  $\omega \in \Omega$ , if  $T$  is one-to-one then  $T^{-1}$  is onto, implying  $U^*$  is onto, so is  $U$ . ⊛

**Remark.**  $T: \Omega \rightarrow \Omega$  is one-to-one implies  $T$  is almost onto.

**Proof.** Let  $A$  be a set such that  $T(\Omega) \subset A$ , and hence  $T^{-1}A = \Omega$  so  $P(T^{-1}A) = P(\Omega) = 1$ , implying that  $P(A) = 1$ , hence  $P(\Omega \setminus A) = 0$ . ⊛

<sup>1</sup>This is true by letting  $f = \mathbb{1}_A$  and then extend to  $L_1(\Omega)$ .

In the case  $T$  is not invertible (e.g. a 2-1 mapping), one might expect a similar formula for  $U^*$ . In the [shift operator](#) example,  $T_1: [0, 1/2] \rightarrow [0, 1]$ ,  $T_2: [1/2, 1] \rightarrow [0, 1]$ , and  $T_1, T_2$  are invertible, we have

$$U^* \varphi(\omega) = \frac{1}{2} (\varphi(T_1^{-1}\omega) + \varphi(T_2^{-1}\omega)).$$

**Definition 2.7.9 (Ergodic transformation).** A one-to-one, [measure-preserving](#) transformation  $T$  is *ergodic* if the only functions  $f \in L_2(\Omega, \mathcal{F}, P)$  which satisfy  $f(T\omega) = f(\omega)$  for almost all  $\omega \in \Omega$  are the constant functions.

**Remark (Eigenfunction).** Phrasing differently, a [measure-preserving](#) mapping  $T: \Omega \rightarrow \Omega$  is *ergodic* if and only if the only eigenfunction  $\varphi \in L_2(\Omega)$  of the corresponding operator  $U$  is the constant function, i.e.  $U\varphi = \varphi$  implying  $\varphi$  is a constant.

**Lemma 2.7.1.** A [measure-preserving](#) mapping  $T: \Omega \rightarrow \Omega$  is *ergodic* if and only if invariant sets of  $T$  have probability 0 or 1, i.e. if  $A \in \mathcal{F}$  satisfies

$$P((A - T^{-1}A) \cup (T^{-1}A - A)) = 0,$$

then  $P(A) = 0$  or  $P(A) = 1$ .

**Proof.** Assume  $T$  is not *ergodic*, then there exists  $\varphi \in L_2(\Omega)$  such that  $U\varphi = \varphi$ . Hence, we can find  $a, b \in \mathbb{R}$ ,  $a < b$  such that  $A = \{\omega \in \Omega : a < \varphi(\omega) < b\}$  has  $0 < P(A) < 1$ . However,

$$T^{-1}A = \{\omega : T\omega \in A\} = \{\omega : a < \varphi(T\omega) < b\} = \{\omega : a < \varphi(\omega) < b\} = A,$$

and thus  $A$  is invariant.

Conversely, suppose  $A \in \mathcal{F}$ , we have  $A = T^{-1}A$  up to measure-zero sets and  $0 < P(A) < 1$ , then  $\varphi = \mathbb{1}_A$  satisfies  $U\varphi = \varphi \in L_2(\Omega)$  with the fact that  $\varphi$  is not constant., proving the result. ■

**Proposition 2.7.3.** Suppose  $T: \Omega \rightarrow \Omega$  is [measure-preserving](#) and  $\varphi \in L_2(\Omega)$ ,  $\mathbb{E}[\varphi] = 0$ , then

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \varphi(T^n \cdot) \rightarrow 0$$

in  $L_2(\Omega)$ .

**Proof.** Note it suffices to assume  $\mathbb{E}[\varphi] = 0$ . We want to show

$$\lim_{N \rightarrow \infty} \frac{1}{N} [I + U + U^2 + \dots + U^{N-1}] \varphi(\cdot) = 0$$

in  $L_2(\Omega)$ . If  $\varphi$  is orthogonal to the constant function. Since  $\mathbb{E}[\varphi] = 0$ , then  $\langle \varphi, 1 \rangle = 0$ . Define a *derivative* operator on  $L_2(\Omega)$  such that

$$D\varphi = (U - I)\varphi = \varphi(T\cdot) - \varphi(\cdot).$$

Use Fundamental Theorem of Calculus argument,

$$[I + U + U^2 + \dots + U^{N-1}]D\varphi = (U^N - I)\varphi.$$

Hence,

$$\left\| \frac{I + U + U^2 + \dots + U^{N-1}}{N} \varphi \right\| \leq \frac{2 \|\varphi\|}{N}$$

if  $\varphi = D\psi$ . In that case  $\lim_{N \rightarrow \infty}$  is zero, i.e. if  $\varphi \in \text{Im}\{D\} \subset \mathcal{H} = L_2(\Omega)$ , then finished. Note also that

$$\left\| \frac{I + U + U^2 + \dots + U^{N-1}}{N} \right\| \leq 1$$

since  $\|U\| = 1$ . Hence, converge to zero if  $\varphi \in \overline{\text{Im}\{D\}}$ .

$$\varphi \in \overline{\text{Im}\{D\}} \Rightarrow \exists \varphi_\epsilon \in \text{Im}\{D\}, \|\varphi_\epsilon - \varphi\| < \epsilon,$$

which implies  $\left\| \frac{I+U+\dots+U^{N-1}}{N}(\varphi_\epsilon - \varphi) \right\| < \epsilon$ .

Recall  $\overline{\text{Im}\{D\}} \oplus \ker D^* = \mathcal{H} = L_2(\Omega)$ . It suffices to show  $\ker D^*$  is spanned by constant functions.

Note  $T$  ergodic implies  $\ker D$  is spanned by constants, we have  $D\varphi = 0 \Leftrightarrow U\varphi = \varphi$ , and

$$(D^*\varphi = 0 \Leftrightarrow U^*\varphi = 0) \Rightarrow (\langle \varphi, U^*\varphi \rangle = \langle \varphi, \varphi \rangle).$$

Therefore,

$$\begin{aligned} \langle U\varphi, \varphi \rangle &= \langle \varphi, \varphi \rangle \\ \int \varphi(T\omega)\varphi(\omega) dP(\omega) &= \int \varphi(\omega)^2 d\omega \\ &= \int \varphi(T\omega)^2 d\omega, \end{aligned}$$

which implies

$$\frac{1}{2} \int [\varphi(T\omega)^2 + \varphi(\omega)^2] dP(\omega) - \int \varphi(T\omega)\varphi(\omega) dP(\omega) = 0.$$

i.e.  $\frac{1}{2} \int [\varphi(T\omega) - \varphi(\omega)]^2 dP(\omega) = 0$ , which means

$$\varphi(T\omega) = \varphi(\omega), \omega \in \Omega.$$

i.e.  $\varphi \equiv \text{constant}$  by ergodicity. ■

**Theorem 2.7.1** (von Neumann ergodic theorem). Suppose  $T: \Omega \rightarrow \Omega$  is [measure-preserving](#), then for any  $\varphi \in L_2(\Omega)$ , one has

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} \varphi(T^n \cdot) = \int_{\Omega} \varphi(\omega) dP(\omega).$$

**Remark.** Convergence is in the  $L_2(\Omega)$  sense, i.e. mean square.

## Chapter 3

# Main Principles of Functional Analysis

### 3.1 Open Mapping Theorem

Suppose  $T: X \rightarrow Y$  is a **bounded linear operator** on **Banach spaces**, and  $T$  is injective and surjective, i.e.  $T^{-1}: Y \rightarrow X$  exists. In this section, we'll see that the open mapping theorem implies  $T^{-1}$  is a **bounded operator**. The main argument relies on **Baire category theorem**.

**Definition 3.1.1** (Nowhere dense). A set  $S$  in a metric space  $M$  is *nowhere dense* if its closure  $\overline{S}$  has empty interior.

**Example** (Cantor set). The **Cantor set** is a **nowhere dense** set.

# Appendix

## Appendix A

# Additional Proofs



# Bibliography

- [Lax02] P.D. Lax. *Functional Analysis*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2002. ISBN: 9780471556046. URL: <https://books.google.com/books?id=18VqDwAAQBAJ>.
- [RS80] M. Reed and B. Simon. *I: Functional Analysis*. Methods of Modern Mathematical Physics. Elsevier Science, 1980. ISBN: 9780125850506. URL: <https://books.google.com/books?id=hInvAAAAMAAJ>.