

# Functional Analysis Class Notes

Hao Fan [hao2718281828@gmail.com](mailto:hao2718281828@gmail.com)

2022 Autumn Semester, 1-16 Weeks

Version: 2022-11-23 13:35:10+08:00

This note is taken for the Functional Analysis course, lectured by Professor Yong Jiao. This note contains my personal thoughts, so all errors in this notes should be mine.

Here are some conventions:

- $\mathbb{R}, \mathbb{Q}, \mathbb{C}$  are fields learnt in Mathematical Analysis. The mappings  $\text{Re}, \text{Im}: \mathbb{C} \rightarrow \mathbb{R}$  are taking real part and imaginary part of a complex number respectively.  $\mathbb{K}$  is one of  $\mathbb{R}$  and  $\mathbb{C}$ , usually used to state different cases conveniently.  $\mathbb{N}$  is the set of **positive** integers. For all  $n \in \mathbb{N}$ , the set  $\{k \in \mathbb{N}: 1 \leq k \leq n\}$  is denoted by  $[n]$ .
- $\mathbb{K}^{n \times n}$  means the matrix space containing all  $n \times n$  matrices.
- $\forall, \exists$  and  $\exists!$  means “for all, there is and there is unique” respectively.
- Formula  $A := B$  means  $A$  is defined as  $B$ . For example,

$$\mathbb{C} := \mathbb{R}[x] / (1 + x^2)$$

means  $\mathbb{C}$  is defined as the quotient ring.

- For each set  $A$ , the identity map is  $\text{id}_A: A \rightarrow A, a \mapsto a$ . It's also used in the case that  $A \subseteq X$ , then  $\text{id}_A: A \rightarrow X, a \mapsto a$ .
- For a mapping  $f: A \rightarrow B$ , we write  $A = \text{dom}(f), B = \text{cod}(f)$ . For two arbitrary sets  $X, Y$ , the set  ${}^Y X$  is the set containing all mappings from  $X$  to  $Y$ .
- For  $a, b \in \mathbb{R}$ , define minimum function

$$\wedge: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (a, b) \mapsto \frac{a + b - |a - b|}{2},$$

and maximum function

$$\vee: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (a, b) \mapsto \frac{a + b + |a - b|}{2}.$$

- Subtraction of sets  $A, B$  is  $A \setminus B := \{x \in A : x \notin B\}$ .
- For a set  $A$ ,  $\mathcal{P}(A)$  means the power set of  $A$ .

For a set  $A$ , and its subset  $B$ , the mapping  $\chi_B$  is defined by

$$\chi_B: A \rightarrow \{0, 1\}, x \mapsto \begin{cases} 1 & x \in B; \\ 0 & x \notin B. \end{cases}$$

There is a bijection between  $\mathcal{P}(A)$  and  ${}^A 2$ , where 2 is an arbitrary set with exactly 2 elements.

- A sequence in  $X$  is a map  $x: \mathbb{N} \rightarrow X, n \mapsto x_n$  (in other words,  $x \in {}^{\mathbb{N}}X$ ), and  $x: \mathbb{N} \rightarrow X$  is usually denoted as  $(x_n)_{n \in \mathbb{N}} \subseteq X$ .

If  $X$  is a topological space, the definition of limit is just the definition of a net in a topological space.

Furthermore, limit of a double indexed sequence  $(x_{m,n})_{m,n \in \mathbb{N}}$  is defined as the limit for the product directed set  $\mathbb{N} \times \mathbb{N}$ .

- For proposition  $p, q$ , we use  $p \wedge q$  to mean the proposition “ $p$  and  $q$ ”,  $\wedge$  has the truth table as follows.

$p$	$q$	$p \wedge q$
1	1	1
1	0	0
0	1	0
0	0	0

Similarly, we define  $p \vee q$ .

- Addition of mappings whose codomains are the same linear space is defined pointwisely. That is: let  $f, g: X \rightarrow V$  be given, we define the function

$$f + g: X \rightarrow V, x \mapsto f(x) + g(x).$$

- For  $f: X \rightarrow \mathbb{K}$  and  $k \in \mathbb{K}$ , we define that function  $f + k$  by  $x \mapsto f(x) + k$ . That is, respect  $k$  as a constant function  $x \mapsto k$ .
- $\lim_n$  means  $\lim_{n \rightarrow \infty}$  for short.

- Given a linear map  $f: X \rightarrow Y$  where  $X, Y$  are linear spaces. Then

$$\ker f := f^{-1}(0) = \{x \in X : f(x) = 0\}.$$

- We say a diagram commutes, if all the morphisms (and their possible compositions) with the same domain and same codomain coincide.
- If an arrow is unique/injective/surjective, we denote the arrow by  $\dashrightarrow / \hookrightarrow / \twoheadrightarrow$  respectively.
- The Kronecker symbol on a set is defined as

$$\delta: X \times X \rightarrow \{0, 1\}, (x, y) \mapsto \delta_y^x := \begin{cases} 1, & x = y; \\ 0, & x \neq y. \end{cases}$$

And  $\delta_y^x$  is also denoted by  $\delta_{x,y}$ .

- The sign function is defined as follows

$$\text{sign}: \mathbb{K} \rightarrow \mathbb{K}, z \mapsto \begin{cases} 0, & z = 0; \\ \bar{z}/|z|, & z \neq 0. \end{cases}$$

- The rounding down function is

$$\lfloor \rfloor : \mathbb{R} \rightarrow \mathbb{Z}: x \mapsto \sup\{k \in \mathbb{Z}: k \leq x\}.$$

- Let  $(\Omega, \mathcal{F})$  be a measurable space. We say a function  $f: \Omega \rightarrow \mathbb{R}$  is measurable, if the preimage of Borel subsets of  $\mathbb{K}$  under  $f$  is  $\mathcal{F}$ -measurable. That is, assume  $\mathbb{K}$  is equipped with Borel  $\sigma$ -algebra.
- The symbol  $\square$  means that a proof or solution ends.
- Somewhere you can see `color different`, that is reminding you to think about what here should be. (Just like  $1 + 1 = 2$ .)

# Contents

<b>0</b>	<b>Introduction</b>	<b>7</b>
<b>1</b>	<b>Week 1</b>	<b>9</b>
1.1	Week 1, Lecture 1 . . . . .	9
1.1.1	Linear Normed Spaces . . . . .	9
1.2	Week 1, Lecture 2 . . . . .	13
1.2.1	Lebesgue integrable function spaces . . . . .	13
<b>2</b>	<b>Week 2</b>	<b>18</b>
2.1	Week 2, Lecture 1 . . . . .	18
2.1.1	Quotient Spaces . . . . .	18
2.2	Week 2, Lecture 2 . . . . .	22
2.2.1	Metric Spaces . . . . .	22
<b>3</b>	<b>Week 3</b>	<b>28</b>
3.1	Week 3, Lecture 1 . . . . .	28
3.1.1	Banach Space . . . . .	29
3.2	Week 3, Lecture 2 . . . . .	35
3.2.1	Completion . . . . .	36
<b>4</b>	<b>Week 4</b>	<b>40</b>
4.1	Week 4, Lecture 1 . . . . .	40
4.1.1	Exercise course . . . . .	40
4.2	Week 4, Lecture 2 . . . . .	41
4.2.1	Banach Fixed-point Theorem . . . . .	43
<b>5</b>	<b>Week 5</b>	<b>45</b>
5.1	Week 5, Lecture 1 . . . . .	45
5.1.1	Bounded Linear Operators/Maps . . . . .	45
5.2	Week 5, Lecture 2 . . . . .	50
5.2.1	Some exercises . . . . .	50
<b>6</b>	<b>Week 6</b>	<b>54</b>
6.1	Week 6, Lecture 1 . . . . .	54
6.1.1	Compactness, Relative Compactness and Total Boundedness . . . . .	54
6.2	Week 6, Lecture 2 . . . . .	59
6.2.1	Finite Dimensional Linear Normed Spaces . . . . .	59

<b>7</b>	<b>Week 7</b>	<b>64</b>
7.1	Week 7, Lecture 1 . . . . .	64
7.1.1	Construct more Linear Normed Spaces . . . . .	64
7.1.2	Unbounded Linear Functional . . . . .	66
7.2	Week 7, Lecture 2 . . . . .	68
7.2.1	Theorems about Banach Spaces . . . . .	68
7.2.2	Baire Category Theorem . . . . .	70
<b>8</b>	<b>Week 8</b>	<b>74</b>
8.1	Week 8, Lecture 1 . . . . .	74
8.1.1	Application of Banach-Steinhaus Theorem . . . . .	74
8.2	Week 8, Lecture 2 . . . . .	84
8.2.1	Open Mapping Theorem (general version) . . . . .	84
8.2.2	Closed Graph Theorem . . . . .	86
<b>9</b>	<b>Week 9</b>	<b>89</b>
9.1	Week 9, Lecture 1 . . . . .	89
9.1.1	Hahn-Banach Theorem . . . . .	89
9.2	Week 9, Lecture 2 . . . . .	94
9.2.1	Review . . . . .	94
<b>10</b>	<b>Week 10</b>	<b>98</b>
10.1	Week 10, Lecture 1 . . . . .	98
10.1.1	Something about Dual Space . . . . .	100
10.2	Week 10, Lecture 2 . . . . .	104
10.2.1	More Dual Spaces . . . . .	104
10.2.2	Weak Convergence . . . . .	106
<b>11</b>	<b>Week 11</b>	<b>113</b>
11.1	Week 11, Lecture 1 . . . . .	113
11.1.1	More about Weak Convergence and Weak-star Convergence . . . . .	113
11.1.2	Conjugate Operators . . . . .	115
11.2	Week 11, Lecture 2 . . . . .	119
11.2.1	Compact Operators and Finite-rank Operators . . . . .	119
<b>12</b>	<b>Week 12</b>	<b>125</b>
12.1	Week 12, Lecture 1 . . . . .	125
12.1.1	Inner Product and Inner Product Space . . . . .	125
12.1.2	Hilbert Space . . . . .	128
12.2	Week 12, Lecture 2 . . . . .	128
12.2.1	Orthogonality . . . . .	128

<b>13 Week 13</b>	<b>132</b>
13.1 Week 13, Lecture 1 . . . . .	132
13.1.1 Orthonormal Basis . . . . .	133
<b>A Hamel Basis</b>	<b>135</b>
<b>B Banach Functor</b>	<b>137</b>
<b>C Completion is unique up to a unique isomorphism</b>	<b>141</b>
<b>D Riesz Representation Theorem and so on</b>	<b>143</b>
D.1 Riesz Representation Theorem . . . . .	143
D.2 Application . . . . .	146

# 0 Introduction

Is there anything...?

## Syllabus

This lecture note contains topics as follows:

- Foundations:
  - Linear Normed space;
  - Bounded Linear Map;
  - Banach Space and Completion.
- Important Theorems:
  - Baire Category Theorem;
  - Banach-Steinhaus Theorem;
  - Open Mapping Theorem;
  - Closed Graph Theorem.
- Topics about Duality:
  - Dual Space;
  - Natural Embedding;
  - Weak/Weak\* Convergence.
- Compact/Finite-Rank Operator:
- Hilbert Space:
  - Definition and Examples;
  - Orthogonality and related topics:
    - \* Pythagoras Theorem;
    - \* Bessel's Inequality;
    - \* Gram-Schmidt process;
    - \* Complete Orthonormal Basis.
  - Projection:
    - \* what
    - \* what
  - Riesz Representation Theorem (on Hilbert space):

- Applications:
  - \* Sesquilinear functional and Representation Theorem;
  - \* Hilbert Adjoint Operator;
  - \* Self-Adjoint Operator;
  - \* Lax-Milgram Theorem and PDE.



# 1 Week 1

## 1.1 Week 1, Lecture 1

We begin from **Banach Space** and **Metric Space**. Before the definition of **Banach space**, we should recall the definition of Vector spaces(or Linear Spaces). Given a set  $X$ , a vector space is a triple  $(X, +, \cdot)$  where  $+: X \times X \rightarrow X$  is called the addition on  $X$ , and  $\cdot: \mathbb{K} \times X \rightarrow X$  is called scalar-multiplication on  $X$ , satisfying 8 axioms.

### Recall

An isomorphism between vector space means a bijection that keeps the linear structure, that is  $\varphi: X \rightarrow Y$  satisfies:  $\forall k, l \in \mathbb{K}, \forall x, x' \in X$  we have  $\varphi(kx + lx') = k\varphi(x) + l\varphi(x')$ . Isomorphism in categories should be in mind:

Category	Grp	$\text{Lin}_{\mathbb{K}}$	Top
Isomorphism	Group isomorphism	$\mathbb{K}$ -Linear isomorphism	Homeomorphism

### 1.1.1 Linear Normed Spaces

**Definition** (Linear Normed Space). Let  $X$  be a linear space. Define a map  $\| \cdot \|: X \rightarrow \mathbb{R}_{\geq 0}$  satisfying:

- (i)  $\|x\| = 0 (\in \mathbb{K}) \iff x = 0 (\in X)$ ;
- (ii)  $\|kx\| = |k| \cdot \|x\| (\forall k \in \mathbb{K}, x \in X)$ ;
- (iii)  $\|x + y\| \leq \|x\| + \|y\| (\forall x, y \in X)$ .

Then  $\| \cdot \|$  is called a **norm** over  $X$ , and  $(X, \| \cdot \|)$  is called a **linear normed space**.

**Remark.** There is some similar weaker definitions:

- If (only) (i) is not satisfied, we call  $\| \cdot \|$  a **semi-norm**.
- If (only) (iii) becomes  $\|x + y\| \leq C(\|x\| + \|y\|)$  for some  $C \in \mathbb{R}_{>1}$ , we call  $\| \cdot \|$  a **quasi-norm**.

Equivalently, we can change the codomain of  $\| \cdot \|$  to  $\mathbb{R}$  and (i) to

$$(\forall x \in X, \|x\| \geq 0) \wedge (\|x\| = 0 \iff x = 0).$$

**Example 1** (Euclidean Spaces).  $(\mathbb{R}^n, \|\cdot\|)$  is a linear normed space, whose norm is defined as follow:

$$\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}, x = (x_1, \dots, x_n) \mapsto \left( \sum_{j=1}^n x_j^2 \right)^{1/2} (= d(x, 0)).$$

Triangle inequality for this norm comes to be the particular triangle inequality for the metric, which can be shown by Cauchy-Schwarz inequality for real numbers.

**Example 2** (Continuous Functions Spaces).  $(C([a, b], \mathbb{K}), \max_{[a, b]} |\cdot|)$  is a linear normed space. Recall the definition of  $C([a, b], \mathbb{K})$  the family of continuous function from  $[a, b]$  to  $\mathbb{K}$ . whose norm is defined as follow:

$$\max_{[a, b]} |\cdot|: (C([a, b], \mathbb{K}) \rightarrow [0, \infty), f \mapsto \max_{x \in [a, b]} |f(x)|.$$

Recall why  $C([a, b], \mathbb{K})$  is a vector space. What is needed to show is just “addition of continuous functions is continuous”, and there is lots of ways to do this, see remark. Notice that  $[a, b]$  is compact and so is  $f([a, b])$ , guaranteeing the existence of  $\max_{x \in [a, b]} |f(x)|$ . Compatibility with multiplication and triangle inequality is trivial.

**Remark.** We have many methods for proving “addition of continuous functions is continuous”. They give the same result with different standpoints. Suppose  $f, g \in C([a, b], \mathbb{K})$

1. By the definition of continuity: We prove pointwisely: Fix  $x \in [a, b]$ .  $\forall \varepsilon > 0$ , we can find  $\delta_1, \delta_2 > 0$  such that  $\forall y: 0 < |y - x| < \delta_1, |f(y) - f(x)| < \varepsilon/2$  and  $\forall y: 0 < |y - x| < \delta_2, |g(y) - g(x)| < \varepsilon/2$ . Therefore, let  $\delta := \delta_1 \wedge \delta_2$  and we have  $\forall y: 0 < |y - x| < \delta$ ,

$$\begin{aligned} |(f+g)(y) - (f+g)(x)| &= |f(y) + g(y) - f(x) - g(x)| \\ &\leq |f(y) - f(x)| + |g(y) - g(x)| \\ &< \varepsilon/2 + \varepsilon/2 \\ &= \varepsilon. \end{aligned}$$

Therefore,  $f+g$  is continuous at  $x$ .

2. By sequence: We prove pointwisely: Fix  $x \in [a, b]$ . Suppose sequence  $(x_n)_{n \in \mathbb{N}} \subseteq [a, b]$  converges to  $x$ , then:

$$\begin{aligned} \lim_{n \rightarrow \infty} (f+g)(x_n) &= \lim_{n \rightarrow \infty} (f(x_n) + g(x_n)) \\ &= \lim_{n \rightarrow \infty} f(x_n) + \lim_{n \rightarrow \infty} g(x_n) \\ &= f(x) + g(x) \\ &= (f+g)(x). \end{aligned}$$

Therefore,  $f + g$  is continuous at  $x$ .

3. By the topological definition ( $\mathbb{K} = \mathbb{R}$  case): an observation :

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

which should be prove by  $A \subseteq B \wedge B \subseteq A \implies A = B$ . Right hand side is union of intersection of two open sets, and similarly for  $(f + g)^{-1}(-\infty, t)$ . We're done.

4. By the continuity of addition ( $\mathbb{K} = \mathbb{R}$  case): We decompose  $f + g$  as following communicative diagrams

$$\begin{array}{ccc} [a, b] & \xrightarrow{\langle f, g \rangle} & \mathbb{R} \times \mathbb{R} \\ & \searrow f+g & \downarrow + \\ & & \mathbb{R} \end{array} \quad \begin{array}{ccc} x & \xrightarrow{\quad} & (f(x), g(x)) \\ & \searrow & \downarrow \\ & & f(x) + g(x) \end{array}$$

The right diagram explains what the functions in the left diagram mean. By the property of product topology and continuity of  $f$  and  $g$ , we know  $\langle f, g \rangle$  is continuous. Continuity of  $+$ :  $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  is trivial. Therefore  $f + g = + \circ (\langle f, g \rangle)$  is continuous.

To get rid of the assumption  $\mathbb{K} = \mathbb{R}$ , use the fact that  $f: X \rightarrow \mathbb{C}$  is continuous if and only if both  $\text{Re}(f), \text{Im}(f)$  are continuous.

**Example 3** ( $p$ -summable sequence spaces). Given  $p \in [1, \infty]$  we define  $(\ell_p, \| \cdot \|_p)$ , where

$$\ell_p := \{ (a_n)_{n \in \mathbb{N}} : \sum_{n \geq 1} |a_n|^p < \infty \}, \text{ (for } p < \infty \text{)}$$

$$\ell_\infty := \{ (a_n)_{n \in \mathbb{N}} : \sup_{n \in \mathbb{N}} |a_n| < \infty \}.$$

And norms are

$$\|a\|_p := \left( \sum_{n \geq 1} |a_n|^p \right)^{1/p}, \quad \text{(for } p < \infty \text{)}$$

$$\|a\|_\infty := \sup_{n \in \mathbb{N}} |a_n|. \quad \text{(Here } a \text{ means } (a_n)_{n \in \mathbb{N}} \text{)}$$

**Proposition 1.1.**  $(\ell_\infty, \| \cdot \|_\infty)$  is a normed space.

**Proof.** Clearly  $\ell_\infty$  is a vector space. Now we prove  $\| \cdot \|_\infty$  is a norm.

1.  $\|a\|_\infty \geq 0$  and  $\|a\|_\infty = 0 \iff a = 0$ :  $\|a\|_\infty \geq 0$  is trivial. Suppose  $\|a\|_\infty = 0$ , that is  $\sup_{n \in \mathbb{N}} |a_n| = 0$ . By definition of supremum,  $|a_n| \leq 0 (\forall n \in \mathbb{N})$ . Therefore,  $a = 0$ .
2.  $\forall k \in \mathbb{K}$ , by property of absolute value we know  $\|ka\|_\infty = |k| \|a\|_\infty$ .
3. Let  $a, b \in \ell_\infty$  and  $M_a = \|a\|_\infty, M_b = \|b\|_\infty$ . Now from definition of supremum

$$\forall n \in \mathbb{N} : |a_n + b_n| \leq |a_n| + |b_n| \leq M_a + M_b$$

Again using definition of supremum, we get  $\|a + b\|_\infty \leq M_a + M_b$ , which was what we wanted.  $\square$

**Theorem 1.2** (Minkowski's Inequality). For each measure space  $(\Omega, \mathcal{F}, \mu)$  and  $f, g \in \mathcal{L}_p (1 \leq p \leq \infty)$ , we have

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

**Remark.** In general, the inequality  $\|f + g\|_p \leq \|f\|_p + \|g\|_p (p \geq 1)$  is called the Minkowski's inequality.

**Example 4.**  $\ell_\infty$  has linear subspaces:  $c_0 \subseteq c \subseteq \ell_\infty$ , where

$$\begin{aligned} c &:= \{(x_n)_{n \in \mathbb{N}} \in \mathbb{R}^\mathbb{N} : (x_n)_{n \in \mathbb{N}} \text{ is a convergent sequence}\}, \\ c_0 &:= \{(x_n)_{n \in \mathbb{N}} \in \mathbb{R}^\mathbb{N} : (x_n)_{n \in \mathbb{N}} \text{ is a convergent sequence with limit } 0\}. \end{aligned}$$

## 1.2 Week 1, Lecture 2

### 1.2.1 Lebesgue integrable function spaces

Recall the left problem: Minkowski's inequality, which makes  $(\ell_p, \|\cdot\|_p)$  a normed space. Now, we need a lemma.

**Lemma 1.3** (Hölder's Inequality). Let  $a \in \ell^p, b \in \ell^q$  for  $p \in (1, \infty)$  and  $q \in (1, \infty)$  satisfy  $1/p + 1/q = 1$ , we have:

$$\|ab\|_1 \leq \|a\|_p \|b\|_q, \quad (1)$$

**Remark.**  $q = p/(p-1)$  is also called the dual index of  $p$ , usually denoted by  $p'$ .

**Remark.** Before start of the proof, we have a look at (1). Recall what we have learned in mathematical analysis, and have a problem in mind: is there anything similar? That is Cauchy-Schwarz Inequality, since they coincide when  $p = q = 2$ . Now we have a direct goal.

**Aim.** Prove (1) by imitating the proof of Cauchy-Schwarz Inequality.

Now, recall all the proofs of Cauchy-Schwarz Inequality you know and think: Which would be useful in this case? [5] Lagrange's Identity, Schwarz's argument (inner product  $\langle x + ty, x + ty \rangle \geq 0$ ), or just  $2xy \leq x^2 + y^2$ ? When  $p \neq 2$ , Schwarz's argument is a nonstarter since there is no quadratic polynomial in sight. Similarly, the absence of a quadratic form means that one is unlikely to find an effective analog of Lagrange's identity.

This brings us to our most robust proof of Cauchy-Schwarz Inequality, the one that starts with the so-called "humble bound,"

$$xy \leq \frac{x^2}{2} + \frac{y^2}{2}, \forall x, y \in \mathbb{R}. \quad (2)$$

(2) proves Cauchy's inequality as follows.

**Proof of Cauchy's inequality from (2).** Without loss of generality, suppose that  $\sum_{n \geq 1} a_n^2 = A^2 \neq 0$  and  $\sum_{n \geq 1} b_n^2 = B^2 \neq 0$ . Let

$$a'_j = a_j/A, b'_j = b_j/B, \forall j \in \mathbb{N}.$$

Notice that  $\sum_{n \geq 1} a'_n = \sum_{n \geq 1} b'_n = 1$ . Now (2) implies

$$\sum_{n \geq 1} a_n b_n \leq \sum_{n \geq 1} (a_n^2 + b_n^2)/2 = \sum_{n \geq 1} a_n^2/2 + \sum_{n \geq 1} b_n^2/2 = 1.$$

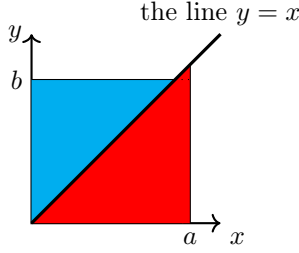


Figure 1: Area meaning of (2)

And, in terms of  $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ , and multiply  $AB$  on both sides, we have

$$\sum_{n \geq 1} a_n b_n \leq \left( \sum_{n \geq 1} a_n^2 \right)^{1/2} \left( \sum_{n \geq 1} b_n^2 \right)^{1/2}. \quad \square$$

This bound may now remind us that the general AM-GM inequality

$$x^p y^q \leq \frac{x}{p} + \frac{y}{q} \quad \text{for all } x, y \geq 0 \text{ and } q = p'(p, q > 1). \quad (3)$$

(3) is the perfect analog of the “humble bound”(2).

**Proof of (2).** There is many ways to to this, see[5]. We choose the way by area of regions. Consider the region under the function  $x \mapsto x$ :

$$\begin{aligned} A &:= \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq x \leq a\}, \\ B &:= \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq y \leq b\}. \end{aligned}$$

Then Figure 1 shows that  $m(A) + m(B) \geq m([0, a] \times [0, b])$ , where  $m$  denotes the Lebesgue measure on  $\mathbb{R}^2$ .  $\square$

Now, by imitating the proof of (2), we need to get the  $x^p/p$  as area of some region under a function, so consider the function  $x \mapsto x^{p-1}$ .

**Proof of (3).**

It's easy to verify that

$$m(A) = \int_{[0, a]} f \, dm, m(B) = b^{\frac{p}{p-1}} - \int_{[0, b^{p/(p-1)}]} f \, dm,$$

where  $m$  is the Lebesgue measure on  $\mathbb{R}$ . By simple calculation, we have  $m(A) = a^p/p, m(B) = b^q/q$ . Notice that  $A \cup B$  contains  $[0, a] \times [0, b]$ , we're done.  $\square$

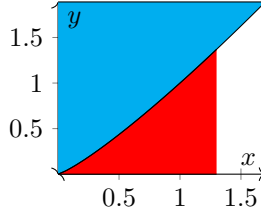


Figure 2: Area meaning of (3)

**Proof of (1).** Without loss of generality, suppose  $\|a + b\|_p \neq 0$ . And suppose  $a \neq 0 (\in \ell_p), b \neq 0 (\in \ell_q)$ . As what we do in the proof of Cauchy's inequality, let

$$a'_j = a_j / \|a\|_p, b'_j = b_j / \|b\|_p, \forall j \in \mathbb{N}.$$

Notice that  $\|a'\|_p = \|b'\|_q = 1$ . Now, apply (3) to  $|a_j b_j|$ , we have

$$\sum_{n \geq 1} |a'_n b'_n| \leq \sum_{n \geq 1} |a'_n|^p / p + \sum_{n \geq 1} |b'_n|^q / q = 1/p + 1/q = 1,$$

which implies

$$\|ab\|_1 \leq \|a\|_p \|b\|_p. \quad \square$$

**Proof of Minkowski's inequality.**

$$\begin{aligned} \|x + y\|_p^p &= \sum_{n \geq 1} |(x + y)_n|^p \\ &= \sum_{n \geq 1} |x_n + y_n|^{p-1} |x_n + y_n| \\ &\leq \sum_{n \geq 1} |x_n + y_n|^{p-1} (|x_n| + |y_n|) \text{ (Triangle inequality on } \mathbb{R}) \\ &= \sum_{n \geq 1} |x_n + y_n|^{p-1} |x_n| + \sum_{n \geq 1} |x_n + y_n|^{p-1} |y_n| \\ &= \|(x + y)^{p-1} x\|_1 + \|(x + y)^{p-1} y\|_1 \text{ (def of norm)} \\ &\leq \|(x + y)^{p-1}\|_q \|x\|_p + \|(x + y)^{p-1}\|_q \|y\|_p \text{ (see (1)) (*)} \\ &= \|(x + y)\|_p^{p/q} (\|x\|_p + \|y\|_p) \text{ ((p-1)q = p)}, \end{aligned}$$

and divide  $\|x + y\|_p^{p/q} (\neq 0)$  from both sides, getting

$$\|x + y\|_p^{p-p/q} \leq \|x\|_p + \|y\|_p.$$

We're done, since  $p - p/q = 1$ . □

To summarize what we have done, we need the language of measure.

**Definition** ( $\sigma$ -algebra). A  **$\sigma$ -algebra** on a set  $\Omega$  is a subset  $\mathcal{F}$ , satisfying:

1.  $\Omega, \emptyset \in \mathcal{F}$ ;
2.  $A \in \mathcal{F} \implies \Omega \setminus A \in \mathcal{F}$ ;
3.  $(A_n)_{n \in \mathbb{N}} \subseteq \mathcal{F} \implies \bigcup_{n \geq 1} A_n \in \mathcal{F}$ .

**Definition** (Measurable Space). A **measurable space** is a double  $(\Omega, \mathcal{F})$  where  $\Omega$  is an arbitrary set and  $\mathcal{F}$  is a  $\sigma$ -algebra over  $\Omega$ . Elements of  $\mathcal{F}$  are called **measurable sets** of  $(\Omega, \mathcal{F})$ .

**Definition** (Measure, Measure space). A **measure** is a  $\sigma$ -additive function from  $\mathcal{F}$  to  $[0, \infty]$ . A triple  $(\Omega, \mathcal{F}, \mu)$  is called a **measure space**, if  $(\Omega, \mathcal{F})$  is a measurable space and  $\mu$  is a measure.

**Definition** (Integral with respect to measure). Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space. We have a glance at “how to define integral with respect to measure”. For the detail, see [2].

**Step 1:** Define **integral**  $\int$  for measurable simple nonnegative function:

$$\sum_{k=1}^n a_k \chi_{A_k} \longmapsto \sum_{k=1}^n a_k \mu(A_k).$$

**Step 2:** Define **integral**  $\int$  for measurable nonnegative function:

$$f \longmapsto \sup \left\{ \int \varphi : \varphi \leq f, \varphi \text{ is nonnegative simple function} \right\}.$$

**Step 3:** Define **integral**  $\int$  for measurable function:

$$f \longmapsto \int f^+ d\mu - \int f^- d\mu,$$

$$\text{where } f^+ = f \chi_{f^{-1}[0, \infty)}, f^- = -f \chi_{f^{-1}(-\infty, 0]}.$$

**Definition** ( $p$ -integrable space). Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space, then the  $p$ -integrable space over  $\mathcal{L}_p(\Omega, \mathcal{F}, \mu)$  is defined as

$$\mathcal{L}_p(\Omega, \mathcal{F}, \mu) := \left\{ f \in \mathbb{K}^\Omega : f \text{ is measurable and } \int |f|^p d\mu < \infty \right\}.$$



**Fact.** The proof of Minkowski's inequality over  $\ell_p$  actually proved the Minkowski's inequality of **every**  $p$ -integrable space  $\mathcal{L}_p(\Omega, \mathcal{F}, \mu)$ .

To understand this fact, we should have another way to illustrate  $\sum$ . That is,  $\sum$  is a kind of integral.

**Definition** (Counting Measure). Given a measurable space  $(\Omega, \mathcal{F})$ . Define  $\mu: \mathcal{F} \rightarrow [0, \infty]$ ,  $A \mapsto \sharp A$ . Where  $\sharp A = \infty$  if  $A$  is an infinite set, and  $\sharp A = n$  if  $A$  has exactly  $n$  elements.  $\mu$  is called the **counting measure** over  $(\Omega, \mathcal{F})$ .

**Remark.** It can be shown that, [1] for real sequence  $(a_n)_{n \in \mathbb{N}}$  (equivalent to a function  $a: \mathbb{N} \rightarrow \mathbb{R}$ ), we have

$$\sum_{n \geq 1} a_n = \int a \, d\mu.$$

That's why we can respect  $\sum$  as  $\int$ . And hence, the fact above is just regard  $\sum$  as integral with respect to counting measure, and the proof works for arbitrary measure space.

**Remark.** We can also prove Minkowski's inequality of  $L_p$  by using the  $L_{p'}$ . Since

$$\|f\|_p = \sup \left\{ \left| \int f g \, d\mu \right| : g \in L_{p'}(\Omega, \mathcal{F}, \mu), \|g\|_{p'} \leq 1 \right\}.$$

## 2 Week 2

### 2.1 Week 2, Lecture 1

#### 2.1.1 Quotient Spaces

Let  $X$  be a vector space with a linear subspace  $X_0$ , denoted as  $X_0 \hookrightarrow X$ .

**Definition** (Coset).  $\forall x \in X$ , the coset of  $x$  (with respect to  $X_0$ ), denoted as  $[x]$  or  $x + X_0$  is defined as

$$[x] = x + X_0 := \{x + y : y \in X_0\}.$$

**Definition** (Quotient Space).  $X/X_0 := \{[x] : x \in X\}$ , called the quotient space of  $X$  (with respect to  $X_0$ ).

We want  $X/X_0$  to be a vector space, so we define operations as follows:

$$\begin{aligned}\oplus : X/X_0 \times X/X_0 &\rightarrow X/X_0, ([x], [y]) \mapsto [x + y]; \\ \odot : \mathbb{K} \times X/X_0 &\rightarrow X/X_0, ([x], k) \mapsto [kx].\end{aligned}$$

Where  $[x + y]$  means the addition (and take the coset), and the  $[kx]$  means the scalar multiplication of  $X$  (and take the coset). You should verify that the operations are well defined. For simplicity, we write  $+$ ,  $\cdot$  instead of  $\oplus, \odot$ .

**Claim.**  $(X/X_0, +, \cdot)$  is a vector space.

**Question 2.1.** Think this questions:

1. Clearly, the zero element in  $X/X_0$  is  $[0]$ . But,  $[0] = ?$ ;
2. If  $[x] \neq [y]$ , what is  $[x] \cap [y]$  ?
3. Show that  $x \in [y] \iff x - y \in X_0$ .

Answers are as follows:

1.  $[0] = X_0$ , from definition of coset.
2.  $\emptyset$ . Since (3) implies  $[x] \cap [y] \neq \emptyset$  means  $\exists z : z - x, z - y \in X_0$ , therefore  $x - y = (z - y) - (z - x) \in X_0$  since  $X_0$  is a linear subspace. Now,  $\forall a \in [x]$ , from  $a = x + w$  ( $w \in X_0$ ), we have  $a = y + (w + (x - y))$  and  $(w + (x - y)) \in X_0$  so  $a \in [y]$ . Above all,  $[x] \subseteq [y]$ . It is the same to know  $[y] \subseteq [x]$ .

3. Since

$$\begin{aligned} x \in [y] &\iff x = y + z \text{ for some } z \in X_0 \\ &\iff x - y = z \text{ (} = 0 + z \text{)} \text{ for some } z \in X_0 \\ &\iff x - y \in [0] = X_0. \end{aligned}$$

Let's see a simple example:

**Example 5.** From Example 4,  $c_0 \hookrightarrow c \hookrightarrow \ell_\infty$ . And we introduce a new notion:

**Definition** (Codimension). Suppose  $X$  a vector space and  $X_0 \hookrightarrow X$ . Then the codimension of  $X_0$ , is  $\text{codim}_X X_0 := \dim X/X_0$ . Also denoted by just  $\text{codim}(X_0)$  if there is no confusion.

**Claim.**  $\text{codim}_c c_0 = 1$ .

**Proof.** Let  $(1_n)_{n \in \mathbb{N}}$  be the sequence with all elements 1. We want to show that  $\{(1_n)_{n \in \mathbb{N}}\}$  is a basis of  $c/c_0$ . Let  $(x_n)_{n \in \mathbb{N}} \in c$ , and suppose  $\lim_n x_n = x \in \mathbb{K}$ . We have  $[(x_n)_{n \in \mathbb{N}}] = [x(1_n)_{n \in \mathbb{N}}]$ , since  $x(1_n)_{n \in \mathbb{N}}$  is just the sequence with all elements  $x$ , and clearly  $\lim_n (x_n - x) = 0 \implies (x_n)_{n \in \mathbb{N}} - x(1_n)_{n \in \mathbb{N}} \in c_0$ . That is,  $[(x_n)_{n \in \mathbb{N}}] = [x(1_n)_{n \in \mathbb{N}}] = x[(1_n)_{n \in \mathbb{N}}]$ . We're done.  $\square$

**Remark.** There is an isomorphism from  $c/c_0$  to  $\mathbb{K}$ , that is  $[(x_n)_{n \in \mathbb{N}}] \mapsto \lim_n x_n$ .

**Example 6.** Consider  $X = \mathbb{R}^2$ ,  $X_0 \hookrightarrow X$  with  $\dim X_0 = 1$ . It is easy to see that  $\forall x \in \mathbb{R}$ , the coset containing  $x$  is just translating  $X_0$  such that  $0(\in X_0)$  is translated to  $x$ . And

$$X/X_0 = \{X_0\} \cup \{\text{all lines that are parallel to } X_0\}.$$

Now we want to define a norm on  $X/X_0$ . An intuitive norm is the distance between  $X_0$  and the coset.

**Definition** (Norm on  $X/X_0$ ). Define

$$\| \cdot \| : X/X_0 \rightarrow \mathbb{R}_{\geq 0}, [x] \mapsto \inf_{y \in X_0} \|x - y\|.$$

The norm in green color is the usual norm in  $\mathbb{R}^2$ , see Example 1.

We should verify that  $\| \cdot \|$  is actually a norm. That is

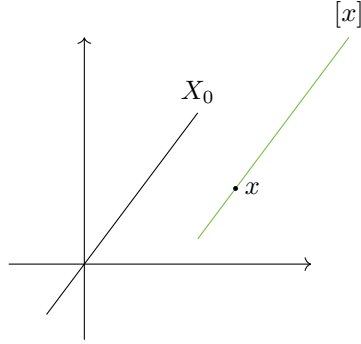


Figure 3:  $X, X_0$  and points of  $X/X_0$

**Question 2.2.** Verify that :

1.  $\forall [x] \in X/X_0 : \|[x]\| \geq 0$  and  $\|x\| = 0 \iff x = X_0$ ;
2.  $\forall [x] \in X/X_0 : \|k[x]\| = |k| \cdot \|x\|$ ;
3.  $\|[x] + [y]\| \leq \|[x]\| + \|[y]\|$ .

**Proof.** For 1: Only needed is to show that  $\|x\| = 0 \iff x = X_0$ . Here we use a Theorem (in the below remark) and a trivial fact:

**Fact.**  $X_0$  is a closed subset of  $X$ .

Now suppose  $[x] \in X/X_0$  satisfying  $\|[x]\| = 0$ . By definition, we have  $\inf_{y \in X_0} \|x - y\| = 0$ . From the definition of infimum :  $\forall n \in \mathbb{N} \exists y_n \in X_0$  such that  $\|x - y_n\| < 1/n$ , therefore we have a sequence  $(y_n)_{n \in \mathbb{N}} \subseteq X_0$  converging to  $x$ . From the theorem below, we know  $x \in X_0$ , so  $[x] = X_0$  as we wanted.

2: It holds naturally when  $k = 0$ . If  $k \neq 0$ , it just follows from property of norm and  $k^{-1}X_0 = X_0$ .

3: Intuitively , we have

$$\begin{aligned}
 \|[x] + [y]\| &= \|[x + y]\| \\
 &= \inf_{z \in X_0} \|x + y - 2z\| \\
 &\leq \inf_{z \in X_0} (\|x - z\| + \|y - z\|) \text{ (triangle inequality of norm)} \\
 &\leq \inf_{z \in X_0} \|x - z\| + \inf_{z \in X_0} \|y - z\| \\
 &= \|[x]\| + \|[y]\|.
 \end{aligned} \tag{4}$$

So easy, isn't it? However, look at the  $\leq$ , this inequality is non-trivial and we should prove. By simple application of definition of infimum, we find: the inequality is **reversed**! But (4) can be corrected:  $\forall \varepsilon > 0$ ,  $\exists z_\varepsilon \in X_0, w_\varepsilon \in X_0$  such that

$$\begin{aligned}\|x - z_\varepsilon\| &< \inf_{z \in X_0} \|x - z\| + \varepsilon/2 = \|x\| + \varepsilon/2, \\ \|y - w_\varepsilon\| &< \inf_{z \in X_0} \|y - z\| + \varepsilon/2 = \|y\| + \varepsilon/2.\end{aligned}$$

Therefore we have

$$\inf_{z \in X_0} (\|x - z\| + \|y - z\|) \leq \|x\| + \|y\| + \varepsilon.$$

Since  $\varepsilon$  is arbitrary, we know  $\inf_{z \in X_0} (\|x - z\| + \|y - z\|) \leq \|x\| + \|y\|$  and then  $\|x + y\| \leq \|x\| + \|y\|$ .

However, this is wrong again. Since  $z_\varepsilon$  may not coincide with  $w_\varepsilon$ . To fix this, write

$$\|[x + y]\| = \inf_{z, w \in X_0} \|x + y - (z + w)\|. \quad (5)$$

By (5), and  $\|x + y - (z + w)\| \leq \|x - z\| + \|y - w\|$ , we use the definition of  $\inf$  for  $\inf_{z \in X_0} \|x - z\|, \inf_{w \in X_0} \|y - w\|$ . We can find  $z_\varepsilon, w_\varepsilon$  as above and get  $\|[x + y]\| \leq \|[x]\| + \|[y]\| + \varepsilon$ , we're done.

Above all,  $\| \cdot \|$  is actually a norm.  $\square$

**Remark.** We define the topology of linear normed space as follows:

**Definition** (Topology of linear normed space). Let  $(X, \| \cdot \|)$  be a linear normed space. Then there is a natural metric on  $X$ , that is  $d: X \times X \rightarrow \mathbb{R}_{\geq 0}, (x, y) \mapsto \|x - y\|$ . The topology induced by this metric is called the (usual) topology of  $(X, \| \cdot \|)$ .

Now we have a topology of  $X$ , and we have a result characterizing the closed subsets of  $X$ .

**Theorem 2.1.** Given a linear normed space  $X$  with  $X_0 \hookrightarrow X$ . Then,  $X$  is closed **if and only if** for all  $(x_n)_{n \in \mathbb{N}} \subseteq X_0$  such that  $\lim_n x_n = x \in X$ , we have  $x \in X_0$ .

**Remark.** A quotient semi-norm in  $X/X_0$  is a norm if and only if  $X_0$  is closed.

## 2.2 Week 2, Lecture 2

### 2.2.1 Metric Spaces

**Definition** (Metric, Metric Spaces). Let  $X$  be a set.  $d: X \times X \rightarrow \mathbb{R}$  is called a metric, if  $d$  satisfies:

1.  $\forall x, y \in X: d(x, y) \geq 0$  and  $d(x, y) = 0 \iff x = y$ .
2.  $\forall x, y \in X: d(x, y) = d(y, x)$ .
3.  $\forall x, y, z \in X: d(x, y) + d(y, z) \geq d(x, z)$ .

The ordered pair  $(X, d)$  is called a metric space.

**Remark.** Every metric space has a topology, we will discuss this later.

**Remark.** Let's compare normed spaces and metric spaces: normed space need linear structures but metric spaces don't need. A normed space  $(X, \|\cdot\|)$  is naturally a metric space by the metric induced by norm  $d: X \times X \rightarrow \mathbb{R}, (x, y) \mapsto \|x - y\|$ .

**Remark.** Let  $X$  be an arbitrary set, we can define a metric on  $X$  by the Kronecker symbol  $\delta$ .

**Example 7.**  $(\mathbb{R}^n, d)$  is a metric space, where

$$d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, ((x_1, \dots, x_n), (y_1, \dots, y_n)) \mapsto \left( \sum_{j=1}^n (x_j - y_j)^2 \right)^{1/2}.$$

**Example 8.**  $(\mathbb{R}^{\mathbb{N}}, d)$  is a metric space, where

$$d: \mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}, ((x_n)_{n \in \mathbb{N}}, (y_n)_{n \in \mathbb{N}}) \mapsto \sum_{j \geq 1} \frac{1}{2^j} \frac{|x_j - y_j|}{1 + |x_j - y_j|}.$$

$d$  is well-defined, since the series can be dominated by  $\sum_{j=1}^{\infty} 1/2^j$ . To verify the triangle inequality, we use the monotone function  $f: [0, \infty) \rightarrow [0, 1], x \mapsto x/(1+x)$ . So,  $|x_j - y_j| + |y_j - z_j| \geq |x_j - z_j|$  implies

$$\frac{|x_j - y_j| + |y_j - z_j|}{1 + |x_j - y_j| + |y_j - z_j|} \geq \frac{|x_j - z_j|}{1 + |x_j - z_j|},$$

and clearly the left-hand side is no more than  $f(|x_j - y_j|) + f(|y_j - z_j|)$ . Sum for  $j \in \mathbb{N}$  and we're done.

**Example 9.** Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space with  $\mu(\Omega) < \infty$ . Let  $\mathcal{L}_0(\Omega)$  be the space of all  $\mathcal{F}$ -measurable functions from  $\Omega$  to  $\mathbb{K}$ , written  $\mathcal{L}_0$  for short. Define

$$\mathcal{Z} := \{f \in \mathcal{L}_0(\Omega) : f(x) = 0 \text{ for } \mu\text{-almost every } x \in \Omega\},$$

the (linear) subspace containing all functions equal 0  $\mu$ -almost everywhere. Now consider the quotient space  $\mathcal{L}_0/\mathcal{Z}$ . We define

$$\begin{aligned} d: \mathcal{L}_0/\mathcal{Z} \times \mathcal{L}_0/\mathcal{Z} &\longrightarrow \mathbb{R} \\ (f + \mathcal{Z}, g + \mathcal{Z}) &\longmapsto \int_{\Omega} \frac{|f - g|}{1 + |f - g|} d\mu. \end{aligned} \quad (6)$$

Integrand on the right-hand side can be dominated by  $1_{\Omega}(=1)$ , hence the integral is finite. The definition of  $d$  involves the selection of representative element, so we should verify that  $d$  is well-defined. Suppose  $f + \mathcal{Z} = f' + \mathcal{Z}, g = g' + \mathcal{Z}$ , and suppose  $f, g$  is finite everywhere, then

$$\begin{aligned} \exists A_1 : \mu(A_1) = 0 \quad \forall x \in A_1^c \quad f(x) &= f'(x); \\ \exists A_2 : \mu(A_2) = 0 \quad \forall x \in A_2^c \quad g(x) &= g'(x). \end{aligned} \quad (7)$$

Then  $f(x) - g(x) = f'(x) - g'(x)$  for all  $x \in (A_1 \cup A_2)^c$  and  $\mu(A_1 \cup A_2) = 0$ . Therefore  $f - g = f' - g'$   $\mu$ -almost everywhere, and hence  $\frac{|f - g|}{1 + |f - g|} = \frac{|f' - g'|}{1 + |f' - g'|}$   $\mu$ -almost everywhere, implying that their integration coincide. Above all,  $d(f + \mathcal{Z}, g + \mathcal{Z}) = d(f' + \mathcal{Z}, g' + \mathcal{Z})$  whenever  $f - f' \in \mathcal{Z}, g - g' \in \mathcal{Z}$ .

Proof of “ $d$  is a metric” is the same as the previous example.

**Example 10.** These are all metric spaces, since they are linear normed spaces:  $\ell_p, c_0, c, C([a, b], \mathbb{K}), L_p, \mathbb{R}^n$ .

**Definition** (Convergence in metric space). Let  $(X, d)$  be a metric space. A sequence in  $X$ , say  $(x_n)_{n \in \mathbb{N}} \subseteq X$ . We say  $(x_n)_{n \in \mathbb{N}}$  is convergent to  $x \in X$ , if  $\lim_n d(x_n, x) = 0$  (limit of real sequence).  $(x_n)_{n \in \mathbb{N}}$  is convergent to  $x$  is usually denoted by  $(x_n)_{n \in \mathbb{N}} \xrightarrow{d} x$  or  $(x_n)_{n \in \mathbb{N}} \rightarrow x$  if there is no ambiguity.

**Example 11.** Suppose  $X$  is an arbitrary set.  $(X, \delta)$  is a metric space, where  $\delta$  means the Kronecker symbol. Then

$$(x_n)_{n \in \mathbb{N}} \rightarrow x \iff \exists N \in \mathbb{N} \quad \forall n \geq N \quad x_n = x.$$

**Example 12.** Consider  $((C[a, b], \mathbb{K}), d)$ , where

$$d: (C[a, b], \mathbb{K}) \times (C[a, b], \mathbb{K}) \rightarrow \mathbb{R}, (f, g) \mapsto \max_{[a, b]} |f - g|.$$

Then  $(f_n)_{n \in \mathbb{N}} \xrightarrow{d} f \iff (f_n)_{n \in \mathbb{N}}$  converge to  $f$  uniformly, as we learned in Mathematical Analysis.

**Example 13.** Recall  $(L_0/\mathcal{Z}, d)$ ,  $(f + \mathcal{Z})_{n \in \mathbb{N}} \xrightarrow{d} f + \mathcal{Z} \iff (f_n)_{n \in \mathbb{N}} \xrightarrow{\mu} f$ .

**Proof.** Necessity:  $(f + \mathcal{Z})_{n \in \mathbb{N}} \xrightarrow{d} f + \mathcal{Z}$  means

$$\lim_n \int_{\Omega} \frac{|f_n - f|}{1 + |f_n - f|} d\mu = 0.$$

Given  $\sigma > 0$ . Define a set  $E_n^\sigma := \{x \in \Omega : |f_n(x) - f(x)| > \sigma\}$ , we need to show  $\lim_n \mu(E_n^\sigma) = 0$ . By Chebyshev's inequality:

$$\begin{aligned} \mu(E_n^\sigma) &= \mu\{x \in \Omega : |f_n(x) - f(x)| > \sigma\} \\ &= \mu\left\{x \in \Omega : \frac{|f_n(x) - f(x)|}{1 + |f_n(x) - f(x)|} > \frac{\sigma}{1 + \sigma}\right\} \\ &\leq \frac{1 + \sigma}{\sigma} \int_{E_n^\sigma} \frac{|f_n(x) - f(x)|}{1 + |f_n(x) - f(x)|} d\mu \\ &\leq \frac{1 + \sigma}{\sigma} \int_{\Omega} \frac{|f_n(x) - f(x)|}{1 + |f_n(x) - f(x)|} d\mu \\ &= \frac{1 + \sigma}{\sigma} d(f_n + \mathcal{Z}, f + \mathcal{Z}). \end{aligned}$$

$\lim_n d(f_n + \mathcal{Z}, f + \mathcal{Z}) = 0$  implies  $\lim_n \mu(E_n^\sigma) = 0$ , that is  $f_n \xrightarrow{\mu} f$ .

Sufficiency: Given  $\sigma \in (0, 1)$ , we know:

$$\left\{x \in \Omega : \frac{|f_n - f|}{1 + |f_n - f|} > \sigma\right\} = \{x \in \Omega : |f_n - f| > \frac{\sigma}{1 - \sigma}\}.$$

This implies that  $\frac{|f_n - f|}{1 + |f_n - f|} \xrightarrow{\mu} 0$ .

Now, from the dominated convergence theorem ( $1_\Omega$  being the dominated function), we have:

$$\begin{aligned} \lim_n d(f_n + \mathcal{Z}, f + \mathcal{Z}) &= \lim_n \int_{\Omega} \frac{|f_n - f|}{1 + |f_n - f|} d\mu \\ &= \int_{\Omega} \lim_n \frac{|f_n - f|}{1 + |f_n - f|} d\mu \\ &= 0. \end{aligned} \quad \square$$

**Remark.** I didn't get this solution for sufficiency at the class. So, it is meaningful to have a look after class.



## Topology of metric spaces

**Definition** (Topology of metric space). The topology of a metric space  $(X, d)$  is generated by the base

$$\mathcal{B} = \{B(x, r) : x \in X, r \in (0, \infty)\},$$

where  $B(x, r) := \{y \in X : d(y, x) < r\}$ .

**Remark.** Now we can define these things for metric spaces:

- Interior points of a set.
- Interior of sets.
- Limit points of a set.
- Derived sets.
- Closure.
- Isolated point.
- Boundary.

**Fact.** For a metric space  $(X, d)$ :

1. A set  $G$  is open  $\iff \forall x \in G \exists r > 0 \ B(x, r) \subseteq G$ .

**Proof.** Sufficiency is trivial. For necessity, since each open set is union of bases, then  $x \in G$  must lie in a open ball contained in  $G$ , and we can find some  $r > 0$  such that  $B(x, r)$  is contained in the open ball.  $\square$

2. Intersection of open sets may not be open. For example,

$$\bigcap_{n \in \mathbb{N}} (-1/n, 1/n) = \{0\}.$$

**Definition** (Continuity for Met). Let  $(X, d_X), (Y, d_Y)$  be two metric spaces. We say  $f: X \rightarrow Y$  is continuous at  $x \in X$ , if  $\forall \varepsilon > 0 \exists r > 0$  such that  $f(B(x, r)) \subseteq B(f(x), \varepsilon)$  (two balls are in  $X$  and  $Y$  respectively).  $f$  is continuous if  $f$  is continuous at every  $x \in X$ .

**Theorem** (Continuity's equivalent conditions). Let  $(X, d_X), (Y, d_Y)$  be two metric spaces. A map  $f: X \rightarrow Y$  is continuous at  $x$  if and only if  $\forall (x_n)_{n \in \mathbb{N}} \subseteq X (\lim_n x_n = x \implies \lim_n f(x_n) = f(x))$ .

**Proof.** Suppose  $f$  is continuous at  $x$  and  $(x_n)_{n \in \mathbb{N}} \rightarrow x$ .  $\forall \varepsilon > 0$ , by continuity of  $f$  at  $x$ ,  $\exists r > 0$  such that  $f(B(x, r)) \subseteq B(f(x), \varepsilon)$ . For this  $r > 0$ , by convergence of  $(x_n)_{n \in \mathbb{N}}$ ,  $\exists N \in \mathbb{N}$  such that  $\forall n > N$   $x_n \in B(x, r)$  and hence  $\forall n > N$   $f(x_n) \in B(f(x), \varepsilon)$ . Therefore,  $\lim_n f(x_n) = f(x)$ .

Suppose  $\forall (x_n)_{n \in \mathbb{N}} \subseteq X (\lim_n x_n = x \implies \lim_n f(x_n) = f(x))$ . If  $f$  is not continuous at  $x$ , by definition of continuity,

$$\exists \varepsilon_0 > 0 \forall \delta > 0 \exists y \in B(x, \delta) f(y) \notin B(f(x), \varepsilon_0).$$

In particular, take  $\delta_n = 1/n$ . Then there is  $y_n \in B(x, 1/n)$  and  $f(y_n) \notin B(f(x), \varepsilon_0)$ . Now we have a sequence  $(y_n)_{n \in \mathbb{N}}$  converge to  $x$  but  $\lim_n f(y_n) \neq f(x)$ , contradiction. Therefore,  $f$  must be continuous at  $x$ .  $\square$

**Definition** (Continuity for  $\mathbf{Top}$ ). Let  $(X, \mathcal{T})$ ,  $(Y, \mathcal{U})$  be two topological spaces. We say  $f: X \rightarrow Y$  is continuous if  $\forall O \in \mathcal{U} f^{-1}(O) \in \mathcal{T}$ .

**Theorem** (Equivalence of definitions of continuity).  $f: (X, d) \rightarrow (Y, d)$  is continuous if and only if  $f: (X, \mathcal{T}_{d_X}) \rightarrow (Y, \mathcal{T}_{d_Y})$  is continuous.

**Remark.** Here we mean  $f: (X, d) \rightarrow (Y, d)$  is continuous, if it satisfies the definition of continuous maps between metric spaces. And “ $f: (X, \mathcal{T}_{d_X}) \rightarrow (Y, \mathcal{T}_{d_Y})$  is continuous” means it satisfies the definition of continuous maps between topological spaces.

**Proof.** Suppose  $f: (X, d) \rightarrow (Y, d)$  is continuous. Since  $(Y, \mathcal{T}_{d_Y})$  has the topology base

$$\mathcal{B}_Y = \{B(y, r) : y \in Y, r \in (0, \infty)\},$$

it suffices to show that  $\forall B(y, r) \in \mathcal{B}_Y$  we have  $f^{-1}(B(y, r)) \in \mathcal{T}_{d_X}$ . Suppose  $f^{-1}(B(y, r)) \neq \emptyset$ , else it's automatically open. Since  $f(x_1) \in B(y, r)$ ,  $\exists r_1 > 0$  such that  $B(f(x_1), r_1) \subseteq B(y, r)$ . Using the continuity of  $f$  at  $x_1$ ,  $\exists \delta > 0$  such that  $f(B(x_1, \delta)) \subseteq B(f(x_1), r_1) \subseteq B(y, r)$ . Therefore  $B(x_1, \delta) \subseteq f^{-1}(B(y, r))$ . This means  $f^{-1}(B(y, r))$  contains a neighbourhood for each point of itself, and hence  $f^{-1}(B(y, r))$  is open.

Suppose  $f: (X, \mathcal{T}_{d_X}) \rightarrow (Y, \mathcal{T}_{d_Y})$  is continuous. Then  $\forall x \in X$ ,  $f^{-1}(B(f(x), r))$  is open for all  $r > 0$ .  $x \in f^{-1}(B(f(x), r))$  and  $f^{-1}(B(f(x), r))$  is union of sets like  $B(x_0, \delta_0)$ , so we can suppose  $x \in B(x_0, \delta_0)$  for some  $x_0 \in X, \delta_0 > 0$ . Now choose  $\delta > 0$  such that  $B(x, \delta) \subseteq B(x_0, \delta_0)$  and we have

$$f(B(x, \delta)) \subseteq f(B(x_0, \delta_0)) \subseteq f(f^{-1}(B(f(x), r))) \subseteq B(f(x), r).$$

We're done.



## 3 Week 3

### 3.1 Week 3, Lecture 1

#### Recall

Every linear normed space  $(X, \|\cdot\|)$  has a metric (induced by its norm)  $d: X \times X \rightarrow \mathbb{R}, (x, y) \mapsto \|x - y\|$ . This is surely a metric, ensured by the properties of norm. However, a metric space  $(X, d)$  need not to be a linear normed space, since it is possible that  $X$  has no linear structure.

Now, suppose  $(X, d)$  a metric space, where  $X$  is a linear space. We have a question: is there some norm  $\|\cdot\|$  such that  $d$  is induced from  $\|\cdot\|$ ? If there is a norm that we want, it is clear that  $\|\cdot\|: X \rightarrow \mathbb{R}, x \mapsto \|x\| := d(x, 0)$ . We want  $\|\cdot\|$  is a norm, so it should satisfy:

1.  $\|x\| \geq 0$  and  $\|x\| = 0 \iff x = 0$ . This holds, since  $d$  is a metric.
2.  $\forall k \in \mathbb{K}, x \in X, d(kx, 0) = |k|d(x, 0)$ . This should be satisfied.
3.  $d(x, 0) + d(y, 0) \geq d(x + y, 0)$  as the triangle inequality.

Moreover,  $d$  should satisfy  $d(x + z, y + z) = d(x, y)$ , since  $(x + z) - (y + z) = x - y$ . In fact, the following conditions ensure that  $d$  is induced by a norm:

**Condition 1.**  $d(kx, 0) = |k|d(x, 0)$ .

**Condition 2.**  $d$  is translation-invariant, that is  $d(x + z, y + z) = d(x, y)$ .

Suppose  $d$  satisfies condition 1 and condition 2, then it is enough to show that  $\|\cdot\|$  satisfies the triangle inequality.

#### Proof.

$$\begin{aligned}\|x + y\| &= d(x + y, 0) \\ &= d(x + y, -y + y) \\ &= d(x, -y) && \text{(condition 2)} \\ &\leq d(x, 0) + d(0, -y) && \text{(triangle inequality of } d) \\ &= d(x, 0) + d(-y, 0) && \text{(} d \text{ is symmetric)} \\ &= d(x, 0) + d(y, 0) && \text{(condition 1)} \\ &= \|x\| + \|y\|.\end{aligned}$$

We're done. □

Here comes an important notion of functional analysis.

### 3.1.1 Banach Space

**Definition** (Banach Space). A **complete** linear normed space  $(X, \| \cdot \|)$  is called a **Banach Space**.

Here the word “complete” should be defined.

**Definition** (Completeness). A metric space  $(X, d)$  is complete if every Cauchy sequence in  $X$  converges.

**Definition** (Cauchy sequence). Let  $(X, d)$  be a metric space. A sequence  $(x_n)_{n \in \mathbb{N}} \subseteq X$  is said to be a Cauchy sequence, if

$$\lim_{m, n} \|x_m - x_n\| = 0.$$

**Remark.** Here  $\{\|x_m - x_n\|\}_{m, n \in \mathbb{N}}$  is a double index real sequence, and “the double index limit is 0” should be interpreted as

$$\forall \varepsilon > 0 \exists M \in \mathbb{N} \forall n \in \mathbb{N} (\forall m > M \forall n > N \|x_m - x_n\| - 0 < \varepsilon).$$

**Warning.** Convergent sequence must be Cauchy sequence ( from definition ), while Cauchy sequence may not converge ( as the following examples ).

**Example 14.** Let  $d: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (x, y) \mapsto |x - y|$  be the normal metric on  $\mathbb{R}$ . Consider  $(\mathbb{Q}, d|_{\mathbb{Q} \times \mathbb{Q}})$ . This is not a complete metric space, since  $\mathbb{Q}$  is dense in  $\mathbb{R}$  and for arbitrary  $x \in \mathbb{R}$  we can find a sequence  $(x_n)_{n \in \mathbb{N}}$  that converges to  $x$  in  $\mathbb{R}$ . Consider  $x \in \mathbb{R} \setminus \mathbb{Q}$  and we get a sequence in  $\mathbb{Q}$ , that is Cauchy in  $(\mathbb{Q}, d|_{\mathbb{Q} \times \mathbb{Q}})$  and doesn’t converge to any  $r \in \mathbb{Q}$ .

**Example 15.** Consider  $(C[0, 1], \| \cdot \|_{L_1})$ , where  $\| \cdot \|_{L_1}$  means the norm

$$\| \cdot \|_{L_1}: C[0, 1] \rightarrow \mathbb{R}, f \mapsto \int_{[0, 1]} |f| \, dm.$$

This is a norm, since  $\|f\|_{L_1} = 0 \iff |f| = 0 \text{ m-a.e.}$  and continuity of  $f$  ensures  $f = 0$ . Other conditions for norm is trivial. And this is a incomplete normed vector space, since  $C[0, 1]$  is dense ( with respect to the norm  $\| \cdot \|_{L_1}$  ) in  $L_1$ .

From now on,  $C_p[a, b]$  means  $(C[0, 1], \| \cdot \|_{L_p})$ .

**Remark.** The completion ( which will be defined the next class ) of  $C_p[a, b], 1 \leq p < \infty$  is  $L_p[a, b]$ , since  $C[a, b]$  is dense in  $L_1[a, b]$

**Example 16.** Let  $P[a, b] := \{\text{Polynomial functions defined on } [a, b]\}$ , then the linear normed space  $(P[a, b], \max_{[a, b]} | \cdot |)$  is incomplete. Since

$\exists f \in C[a, b]$  such that  $f$  is not a polynomial, such as  $f = \exp|_{[a, b]}$ . Suppose  $\exp: \mathbb{R} \rightarrow \mathbb{R}$  is defined as the power series for convenience. Then by **Weierstrass Approximation Theorem**, for each fixed  $\varepsilon > 0$ , there is some  $p \in P[a, b]$  such that  $\max_{[a, b]} p - f < \varepsilon$ .

In fact, for  $f = \exp|_{[a, b]}$ , it is enough to take

$$p_n: [a, b] \rightarrow \mathbb{R}, x \mapsto \sum_{j=1}^n \frac{x^j}{j!}.$$

By the result in power series theory, we know  $p_n \xrightarrow{\max_{[a, b]} | \cdot |} f$ .

Now we compare two normed spaces sharing the underlying set  $C[a, b]$ .  $C[a, b]$  means the normed space  $(C[a, b], \max_{[a, b]} | \cdot |)$  somewhere. And we will prove the completeness of  $C[a, b]$ .

Normed space	$C[a, b]$	$C_p[a, b]$
Underlying set	$C[a, b]$	$C[a, b]$
Norm	$\max_{[a, b]}   \cdot  $	$\  \cdot \ _p$
Completeness	complete	incomplete

**Proof of completeness.** Let  $(f_n)_{n \in \mathbb{N}} \subseteq C[a, b]$  be a Cauchy sequence. That is

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall m, n \geq N \max_{[a, b]} |f_m - f_n| < \varepsilon.$$

Therefore, given any  $x \in [a, b]$  we have

$$|f_m(x) - f_n(x)| \leq \max_{[a, b]} |f_m - f_n| < \varepsilon.$$

That is the sequence  $(f_n(x))_{n \in \mathbb{N}}$  is a Cauchy sequence. By the completeness of  $\mathbb{R}$ ,  $(f_n(x))_{n \in \mathbb{N}}$  converge. Then we can define a function

$$f: [a, b] \rightarrow \mathbb{R}, x \mapsto \lim_n f_n(x).$$

$\lim_n f_n(x)$  is surely a real number, as explained above. And we have two claims.

**Claim.**  $f_n \xrightarrow{\max_{[a, b]} | \cdot |} f$ .

$\forall n > N$ , we have

$$\max_{[a, b]} |f_m - f_n| < \varepsilon.$$

It's equivalent to

$$|f_m(x) - f_n(x)| < \varepsilon (\forall x \in [a, b]),$$

and let  $m \rightarrow \infty$ , using the continuity of  $| \cdot |$  ( to change the order of  $\lim_m$  and  $| \cdot |$  )

$$|f(x) - f_n(x)| < \varepsilon (\forall x \in [a, b]),$$

which is equivalent to

$$\max_{x \in [a, b]} |f - f_n| < \varepsilon.$$

Therefore,  $f_n \xrightarrow{\max_{[a, b]} | \cdot |} f$ .

**Claim.**  $f \in C[a, b]$ .

It suffices to show that  $f$  is uniformly continuous. Given arbitrary  $\varepsilon > 0$ , by the convergence of  $(f_n)_{n \in \mathbb{N}}$

$$\exists N \forall n \geq N \max_{[a, b]} |f_n - f| < \varepsilon/3.$$

Fix this  $N$ , and the continuity ( equivalent to uniform continuity for functions on  $[a, b]$  ) of  $f_N$  ensures that  $\exists \delta > 0$  such that

$$\forall x \forall y (|x - y| < \delta \implies |f_N(x) - f_N(y)| < \varepsilon/3).$$

And  $\forall x \forall y$  such that  $|x - y| < \delta$ , we have

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| \\ &\leq \max_{[a, b]} |f_N - f| + \varepsilon/3 + \max_{[a, b]} |f_N - f| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\ &= \varepsilon. \end{aligned}$$

Thus  $f$  is uniformly continuous. □

**Example 17.** Suppose  $1 \leq p \leq \infty$ . then  $L_p(\Omega, \mathcal{F}, \mu)$  is a Banach space.

**Proof.** First, suppose  $1 \leq p < \infty$ . Here is a proof different from our textbook. Suppose  $(f_n)_{n \in \mathbb{N}}$  is a Cauchy sequence, then  $(f_n)_{n \in \mathbb{N}}$  is Cauchy in measure ( by Chebyshev's Inequality ). By the [lemma](#),  $\exists$  a

subsequence  $(f_{n_j})_{j \in \mathbb{N}}$  such that  $f_{n_j} \rightarrow f$   $\mu$ -a.e.. Therefore, by **Fatou's Lemma**:

$$\begin{aligned} \lim_j \|f_{n_j} - f\|_p^p &= \lim_j \int_{\Omega} |f_{n_j} - f|^p d\mu \\ &\leq \int_{\Omega} \liminf_j |f_{n_j} - f|^p d\mu \quad (\text{Fatou's Lemma}) \\ &= 0. \quad (f_{n_j} \rightarrow f \text{ } \mu\text{-a.e.}) \end{aligned}$$

While the inequality should be reversed. This can be corrected:

$$\begin{aligned} \|f_{n_j} - f\|_p^p &= \int_{\Omega} \lim_n |f_{n_j} - f_n|^p d\mu \\ &\leq \liminf_j \int_{\Omega} |f_{n_j} - f|^p d\mu, \quad (\text{Fatou's Lemma}) \end{aligned}$$

and

$$\begin{aligned} \lim_{n_j} \|f_{n_j} - f\|_p^p &= \lim_{n_j} \int_{\Omega} \lim_n |f_{n_j} - f_n|^p d\mu \\ &\leq \lim_{n_j} \liminf_n \int_{\Omega} |f_{n_j} - f_n|^p d\mu \quad (\text{Fatou's Lemma}) \\ &= 0. \quad (\text{Cauchy sequence}) \end{aligned}$$

So  $f_{n_j} \xrightarrow{\|\cdot\|_{L_p}} f$ . Minkowski's inequality shows

$$\|f_n - f\| \leq \|f_n - f_{n_j}\| + \|f - f_{n_j}\|.$$

Let  $n_j, n \rightarrow \infty$  and use the fact that  $(f_n)_{n \in \mathbb{N}}$  is Cauchy in norm, we have  $f_n \xrightarrow{\|\cdot\|_{L_p}} f$ .

If  $f \in L_p$ , we are done.  $f$  is a  $\mu$ - a.e. limit of  $(f_{n_j})_{j \in \mathbb{N}}$  and hence is measurable. Minkowski's inequality shows

$$\|f\|_p \leq \|f - f_{n_j}\|_p + \|f_{n_j}\|_p.$$

The first term is bounded ( since the real sequence has limit 0), and the second term is finite since  $f_{n_j} \in L_p$ .

Then, suppose  $p = \infty$ . There is  $(A_{m,n})_{m,n \in \mathbb{N}} \in \mathcal{F}$  such that  $\mu(A_{m,n}) = 0 \forall m, n \in \mathbb{N}$  and

$$\forall \omega \in A_{m,n}^c \quad |f_m(\omega) - f_n(\omega)| \leq \|f_n - f_m\|_{\infty}.$$

Clearly for  $A := \cup_{m,n \geq 1} A_{m,n}$ , we have  $\mu(A) = 0$ . And we have

$$\forall \omega \in A^c \quad |f_n(\omega) - f_m(\omega)| \leq \|f_n - f_m\|_{\infty}.$$



Let  $m \rightarrow \infty$

$$\forall \omega \in A^c |f_n(\omega) - f(\omega)| \leq \lim_m \|f_n - f_m\|_\infty,$$

and hence

$$\|f_n - f\|_\infty \leq \lim_m \|f_n - f_m\|_\infty.$$

Let  $n \rightarrow \infty$  and use

$$\lim_n \lim_m \|f_n - f_m\|_\infty = 0.$$

We're done. □

**Lemma.** Let  $(\Omega, \mathcal{F}, \mu)$  be a measure space, and  $(f_n)_{n \in \mathbb{N}} \subseteq \mathcal{L}_0(\Omega)$  is Cauchy in measure, where

$$\mathcal{L}_0(\Omega) := \{f: \Omega \rightarrow (\mathbb{K}, \mathcal{B}(\mathbb{K})) \text{ that is measurable}\}.$$

Then there is a subsequence  $(g_n)_{n \in \mathbb{N}}$  of  $(f_n)_{n \in \mathbb{N}}$  such that  $g_n \rightarrow f$   $\mu$ -a.e.. Here  $f \in \mathcal{L}_0(\Omega)$ .

**Proof.** We can choose a subsequence  $(g_n)_{n \in \mathbb{N}} = (f_{n_j})_{j \in \mathbb{N}}$  such that if  $E_j := |g_j - g_{j+1}|^{-1}[2^{-j}, \infty)$  then  $\mu(E_j) \leq 2^{-j}$ . Because

$$\forall j \in \mathbb{N} \quad \lim_{m, n \rightarrow \infty} |f_m - f_n|_* \mu[2^{-j}, \infty) = 0.$$

And pick  $n_j$  inductively, such that  $n_{j+1} > n_j$  and

$$\mu_* |f_m - f_n|[2^{-j}, \infty) < 2^{-j} \quad \forall m, n \geq n_j.$$

Set  $F_k := \bigcup_{j \geq k} E_j$  then  $\mu(F_k) \leq \sum_{j \geq k} 2^{-j} = 2^{1-k}$ . Continuity from above is allowed! If  $x \notin F_k$ , for  $i \geq j \geq k$  we have

$$|g_i(x) - g_j(x)| \leq \sum_{l=j}^{i-1} |g_{l+1}(x) - g_l(x)| \leq \sum_{l=j}^{i-1} 2^{-l} \leq 2^{1-j},$$

which ensures that  $\forall x \in F_k^c$ ,  $(g_j(x))_{j \in \mathbb{N}}$  is a Cauchy sequence. Let

$$F = \bigcap_{j \geq 1} F_j = \limsup_j E_j,$$

we have  $\mu(F) = \mu(\lim_j F_j) = \lim_j \mu(E_j) = 0$ . □

**Exercise 3.1.** Prove that  $\ell_p$  is complete when  $1 \leq p < \infty$ .

Suppose  $(X, \| \cdot \|)$  is a linear normed space. Let  $(x_n)_{n \in \mathbb{N}} \subseteq X$  satisfies  $\sum_{n \geq 1} \|x_n\| < \infty$ , and we can define the infinite sum for this sequence as

$$\sum_{n \geq 1} x_n := \lim_{n \rightarrow \infty} S_n, \text{ where } S: \mathbb{N} \rightarrow X, j \mapsto \sum_{j=1}^N x_j.$$

**Theorem 3.1.**  $(X, \| \cdot \|)$  is a Banach space if and only if  $\forall (x_n)_{n \in \mathbb{N}} \subseteq X$ ,

$$\sum_{n \geq 1} \|x_n\| < \infty \implies \sum_{n \geq 1} x_n < \infty.$$

Here  $\sum_{n \geq 1} x_n < \infty$  means  $\sum_{n \geq 1} x_n$  exists for short.

**Proof.** Necessity: suppose  $X$  is a Banach space, then  $\sum_{n \geq 1} \|x_n\| < \infty$  implies

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n > N \left( \sum_{j=1}^p \|x_{n+j}\| < \varepsilon (\forall p \in \mathbb{N}) \right),$$

and therefore  $\forall n > N \|S_{n+p} - S_n\| \leq \sum_{j=1}^p \|x_{n+j}\| < \varepsilon$ , this means that  $(S_n)_{n \in \mathbb{N}}$  is a Cauchy sequence.  $X$  is complete, so  $(S_n)_{n \in \mathbb{N}}$  converges. That is  $\sum_{n \geq 1} x_n < \infty$ .

Sufficiency: suppose  $X$  satisfies the condition above. If  $X$  is not complete, then  $\exists (x_n)_{n \in \mathbb{N}} \subseteq X$  that is Cauchy but has no limit in  $X$ . Now, select a subsequence of  $(x_n)_{n \in \mathbb{N}}$ , say  $(x_{n_j})_{j \in \mathbb{N}}$  such that

$$\forall j \in \mathbb{N} \|x_{n_{j+1}} - x_{n_j}\| < 2^{-j}.$$

Define  $y: \mathbb{N} \rightarrow X, j \mapsto x_{n_{j+1}} - x_{n_j}$ , then  $(y_n)_{n \in \mathbb{N}}$  is a Cauchy sequence, satisfying

$$\forall j \in \mathbb{N} \|y_j\| < 2^{-j}.$$

Therefore,  $\sum_{n \geq 1} \|y_j\| < \infty$ . Then  $X$  satisfies the condition, which implies that  $\sum_{n \geq 1} y_n < \infty$ . Equivalently,  $\lim_j x_{n_j}$  exists in  $X$ . While  $(x_n)_{n \in \mathbb{N}}$  is Cauchy, so  $\lim_n x_n = \lim_j x_{n_j}$  exists, that's a contradiction (see how we selected  $(x_n)_{n \in \mathbb{N}}$ ).  $\square$

## 3.2 Week 3, Lecture 2

### Recall

1.  $L_p(\Omega)$  ( $1 \leq p \leq \infty$ ) is complete. The outline of proof for  $p < \infty$  is here:

**Step 1.** Show that if  $(f_n)_{n \in \mathbb{N}}$  is Cauchy (in norm), then  $(f_n)_{n \in \mathbb{N}}$  is Cauchy in measure.

**Step 2.** Show that  $(f_n)_{n \in \mathbb{N}}$  is Cauchy in measure, then  $(f_n)_{n \in \mathbb{N}}$  has a subsequence  $(f_{n_j})_{j \in \mathbb{N}}$  that converges to a measurable function  $f$   $\mu$ -a.e..

**Step 3.** Use Fatou's lemma to show that  $(f_{n_j})_{j \in \mathbb{N}} \xrightarrow{\|\cdot\|_p} f$ .

**Step 4.** Show that  $(f_n)_{n \in \mathbb{N}} \xrightarrow{\|\cdot\|_p} f$  and  $f \in L_p$

2. About quotient space. Given a normed space  $(X, \|\cdot\|)$  and a closed subspace  $X_0 \hookrightarrow X$ . We can define the quotient space

$$X/X_0 := \{[x] = x + X_0 : x \in X\},$$

whose norm is

$$\|[x]\| = \inf_{y \in X_0} \|x - y\| = \inf_{y \in [x]} \|y(-0)\|.$$

The second equality can be verified by change  $y \in [x] \iff y = x + x_0, x_0 \in X_0$ .

3. Norm and semi-norm ( $p, p(x) = 0 \not\iff x = 0$ ). Let  $X$  be a linear semi-normed space, with the semi-norm  $p$ . A familiar linear semi-normed is  $\mathcal{L}_p(1 \leq p \leq \infty)$ . Let  $X_0 := \{x \in X : p(x) = 0\} \hookrightarrow X$ .

**Claim.**  $X_0$  is closed subspace of  $X$  (so,  $X/X_0$  is allowed, see [this remark](#)).

**Proof.**  $X_0$  is a linear subspace, since  $p$  is a semi-norm.

$p$  is a continuous map, since the triangle inequality holds. Then  $N = p^{-1}(0)$  must be closed.  $\square$

Now, [the remark](#) ensures that  $\|\cdot\|: X/X_0, [x] \mapsto p(x)$  is a norm on  $X/X_0$ .

**Proof.** It should be verified that  $p$  is well-defined (though this should have been proved in [the remark](#)). Suppose  $[x] = [y]$ , that is  $[x - y] = [y - x] = [0]$ . Since  $p$  is a semi-norm, we have the triangle inequality

$$p(x) + p(y - x) \geq p(y), p(y) + p(x - y) \geq p(x),$$

and  $[x - y] = [y - x] = 0 \implies p(x - y) = p(y - x) = 0$ , that is  $p(x) = p(y)$ . Thus,  $[x] \mapsto p(x)$  is well-defined. And

$$(1) \quad \|[x]\| = 0 \iff p(x) = 0 \iff x \in X_0 = [0] \iff [x] = [0] (\in X/X_0).$$

$$(2) \quad \|k[x]\| = \|[kx]\| = p(kx) = |k|p(x) = |k|\|x\|.$$

$$(3) \quad \|[x] + [y]\| = \|[x + y]\| = p(x + y) \leq p(x) + p(y) = \|[x]\| + \|[y]\|.$$

Above all,  $\| \cdot \|$  is a norm on  $[X]$ .  $\square$

### 3.2.1 Completion

In this class,  $X$  is a linear normed space, unless otherwise specified.

**Definition** (Isometry). Suppose  $X, Y$  are two linear normed spaces. We say  $X$  is isometric with  $Y$ , if there is a linear surjection  $T: X \rightarrow Y$  such that

$$\|Tx\| = \|x\| (\forall x \in X),$$

or equivalently  $\| \cdot \|_Y \circ T = \| \cdot \|_X$ .

**Remark.** Isometry is automatically injective, since  $Tx = 0 \iff \|Tx\| = \|x\| = 0 \iff x = 0$ . That is  $\ker T = \{0\}$ . Therefore,  $T$  is automatically injective and hence bijective as we want.

**Definition** (Density). Let  $(X, \| \cdot \|)$  be a linear normed space and  $X_0 \hookrightarrow X$ .  $X_0$  is said to be dense in  $X$ , if  $\overline{X_0} = X$ .

**Question 3.1.** How to verify  $\overline{X_0} = X$  ?

$\overline{X_0} = X$ , if

$$\forall x \in X \forall \varepsilon > 0 \exists x_\varepsilon \in X_0 (\|x_\varepsilon - x\| < \varepsilon).$$

And equivalently

$$\forall x \in X \forall n \in \mathbb{N} \exists x_n \in X_0 (\|x_n - x\| < 1/n).$$

That is,  $\exists (x_n)_{n \in \mathbb{N}} \subseteq X_0$  that converges to  $x$ .

**Theorem 3.2** (Existence of Completion). Let  $(X, \|\cdot\|)$  be a linear normed space. There is a Banach space  $(\widehat{X}, \|\cdot\|)$  such that  $X$  is isometric to a dense subspace of  $\widehat{X}$ .

**Remark.** in fact, the completion  $\widehat{X}$  is unique up to an isometry (but the definition of completion should be different).

**Definition** (Completion). A pair  $(\widehat{X}, \iota)$  is called a completion of  $X$ , if  $\iota: X \rightarrow \widehat{X}$  satisfies  $\forall x \in X: \|\iota(x)\| = \|x\|$  and  $\iota(X)$  is dense in  $\widehat{X}$ .

**Proof.** We will construct a completion of  $X$ . Let

$$\mathcal{E} := \{(x_n)_{n \in \mathbb{N}} \subseteq X : (x_n)_{n \in \mathbb{N}} \text{ is a Cauchy sequence}\},$$

and define  $p: \mathcal{E} \rightarrow \mathbb{R}, x(= (x_n)_{n \in \mathbb{N}}) \mapsto \lim_n \|x_n\|$ . Here  $\lim_n \|x_n\|$  exists in  $\mathbb{R}$ , because  $(x_n)_{n \in \mathbb{N}}$  is a Cauchy sequence implies that  $(\|x_n\|)_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\mathbb{R}$ , and  $\mathbb{R}$  is complete. Moreover,  $p$  is a seminorm on  $\mathcal{E}$ . Now define  $N := p^{-1}(0)$ . Then  $N \hookrightarrow \mathcal{E}$  and  $N$  is closed (by the continuity of  $p$ ). Therefore we can consider  $\widehat{X} := \mathcal{E}/N$ , with the norm  $\|\cdot\|: \widehat{X} \rightarrow \mathbb{R}, x + N \mapsto p(x)$ .

Now, we prove this theorem in 3 steps.

**Step 1.**  $X$  is isometric to a subspace of  $\widehat{X}$ . Let  $X_0 := \{[(x)_{n \in \mathbb{N}}] : x \in X\}$  and

$$T: X \rightarrow X_0, x \mapsto [(x)_{n \in \mathbb{N}}] = (x)_{n \in \mathbb{N}} + N,$$

where  $(x)_{n \in \mathbb{N}}$  means the constant sequence  $(x, \dots, x, \dots)$ . That is,  $T(x) = (x, \dots, x, \dots) + N$ . Clearly  $T$  is a linear surjection. We want to show  $T$  is isometric, that is  $\forall x \in X, \|T(x)\| = \|x\|$ . By definition

$$\begin{aligned} \|T(x)\| &= \|[(x)_{n \in \mathbb{N}}]\| && \text{(def of } T) \\ &= p((x)_{n \in \mathbb{N}}) && \text{(def of } \|\cdot\|_{\widehat{X}}) \\ &= \lim_n \|x\| && \text{(def of } p) \\ &= \|x\|. \end{aligned}$$

To sum up,  $T$  is an isometry as we want.

**Step 2.**  $X_0 \hookrightarrow \widehat{X}$  is dense. As discussed above, it suffices to show that  $\forall [x] = (x_1, \dots, x_n, \dots) + N \in \widehat{X}$ , there is a sequence in  $X_0$  converge to  $X$ . Let

$$[x]^{(m)}: \mathbb{N} \rightarrow [(x_m)_{n \in \mathbb{N}}] = (x_m, \dots, x_m, \dots) + N,$$

and we prove that the sequence  $([x^{(m)}])_{m \in \mathbb{N}}$  is convergent to  $[x]$ .

$$\begin{aligned}
& \lim_m \left\| [x]^{(m)} - [x] \right\| \\
&= \lim_m \left\| (x_m - x_1, \dots, x_m - x_n, \dots) + N \right\| && \text{(def of } \pm \text{)} \\
&= \lim_m p((x_m - x_n)_{n \in \mathbb{N}}) && \text{(def of } \parallel \parallel \text{)} \\
&= \lim_m \lim_n \|x_m - x_n\| && \text{(def of } p \text{)} \\
&= 0. && \text{(see remark)}
\end{aligned}$$

**Step 3.**  $\hat{X}$  is a Banach space. That is  $\hat{X}$  is complete. Let  $([x]^{(n)})_{n \in \mathbb{N}}$  be a Cauchy sequence in  $\hat{X}$ . By the density of  $X_0 = TX$ , we have a sequence  $(y_n)_{n \in \mathbb{N}} \subseteq X$  such that

$$\forall n \in \mathbb{N} \quad \left\| T(y_n) - [x]^{(n)} \right\| \leq 1/n.$$

**Claim.**  $(y_n)_{n \in \mathbb{N}}$  is a Cauchy sequence.

$$\begin{aligned}
& \|y_m - y_n\| \\
&= \|T(y_m) - T(y_n)\| \\
&\leq \|T(y_m) - [x]^{(m)}\| + \|[x]^{(m)} - [x]^{(n)}\| + \|T(y_n) - [x]^{(n)}\| \\
&\leq 1/m + \|[x]^{(m)} - [x]^{(n)}\| + 1/n.
\end{aligned}$$

Apply  $\limsup_{m,n}$  on both sides and we have

$$\limsup_{m,n} \|y_m - y_n\| \leq 0.$$

Therefore,  $(y_n)_{n \in \mathbb{N}}$  is Cauchy, and  $(y_n)_{n \in \mathbb{N}} \in \mathcal{E}$ . Now we show that  $([x]^{(n)})_{n \in \mathbb{N}} \rightarrow [y] = (y_1, \dots, y_n, \dots) + N$ . By definition of  $\parallel \parallel_{\hat{X}}$

$$\begin{aligned}
\|[x]^m - [y]\| &\leq \|[x]^m - T(y_m) + T(y_m) - [y]\| \\
&\leq \|[x]^m - T(y_m)\| + \|T(y_m) - [y]\| \\
&\leq 1/m + p((y_n - y_m)_{n \in \mathbb{N}}) \\
&= 1/m + \lim_n \|y_n - y_m\|,
\end{aligned}$$

and let  $m \rightarrow \infty$ , we have

$$\limsup_m \|[x]^m - [y]\| \leq \limsup_m 1/m + \limsup_m \lim_n \|y_n - y_m\|.$$

The second limit must be 0, since  $\lim_m \lim_n \|y_n - y_m\| = 0$  (see remark) .

□

**Remark.** Here we explain why  $\lim_m \lim_n \|x_m - x_n\| = 0$ . We may want to write: suppose  $\lim_n x_n = x$ , then

$$\lim_m \lim_n \|x_m - x_n\| = \lim_m \|x_m - x\| = 0,$$

where the first equality is using the continuity of  $\|\cdot\|$  and the second equality follows from the definition of  $\lim_n x_n = x$ . Everything makes sense, except  $\lim_n x_n = x$ . Notice that  $(x_n)$  is a sequence in  $X$  and none said that  $X$  is complete.

So, why  $\lim_m \lim_n \|x_m - x_n\| = 0$  holds? It suffices to show that we have

$$\lim_m \lim_n d(x_m, x_n) = \lim_{m,n} d(x_m, x_n) = 0.$$

whenever  $(x_n)_{n \in \mathbb{N}}$  is Cauchy. See <https://math.stackexchange.com/a/633595/1061247>.

## 4 Week 4

### 4.1 Week 4, Lecture 1

#### Recall

No recall today.

#### 4.1.1 Exercise course

We have only 3 exercises this course.

**Question 4.1.** Let  $(X, \| \cdot \|)$  be a linear normed space,  $X_0 \hookrightarrow X$ . If  $X$  is complete and  $X_0$  is closed then  $X_0$  is complete.

**Question 4.2.** Let  $(X, d)$  be a metric space.  $T: X \rightarrow X$  such that  $\exists \lambda \in (0, 1)$

$$d(T(x), T(y)) \leq \lambda d(x, y), \forall x, y \in X.$$

Prove that  $\exists! x_0 \in X$  such that  $Tx_0 = x_0$ .

**Remark.** This result doesn't hold when  $\lambda = 1$ . To see this, consider

$$(X, d) = ([0, \infty), d), T: X \rightarrow X, x \mapsto \sqrt{1 + x^2}.$$

And completeness is necessary too, consider  $(X, d) = ((0, \infty), d)$  and  $T: X \rightarrow X, x \mapsto x/2$ . Other examples can be found.

**Question 4.3.** Let  $(X, \| \cdot \|)$  be a linear normed space. Then  $X$  is a Banach space if and only if for each closed decreasing non-empty subsets sequence  $(A_n)_{n \in \mathbb{N}}$ ,  $\bigcap_{n \geq 1} A_n$  is a singleton set whenever we have  $\lim_n \text{diam}(A_n) = 0$ .

There are answers in the next section.



## 4.2 Week 4, Lecture 2

### Recall

For all l.n.s  $(X, \|\cdot\|)$ , there is a Banach space  $\widehat{X}$  such that  $X \cong X_0 \hookrightarrow \widehat{X}$ , where  $X_0$  is a dense subspace of  $\widehat{X}$ . It's ok to say  $X = X_0 \hookrightarrow \widehat{X}$ , and hence  $\overline{X} = \widehat{X}$ . The proof has 3 steps: construction of  $\widehat{X}$ , embedding  $X$  to  $\widehat{X}$  and showing the completeness.

**Remark.** In the final exam and Phd qualifying exam, stating this theorem and its proof is common.

### Review of exercise class

Here are the proofs of the questions of the exercise class.

**Proof of the first.** Suppose  $(x_n)_{n \in \mathbb{N}} \subseteq X_0$  is a Cauchy sequence in  $X_0$ , then  $(x_n)_{n \in \mathbb{N}}$  is Cauchy in  $X$ .  $X$  is complete so  $\exists x \in X$  such that  $(x_n)_{n \in \mathbb{N}} \rightarrow x$ . Now,  $X_0$  is closed and hence  $x \in X_0$ . Thus,  $(x_n)_{n \in \mathbb{N}} \rightarrow x \in X_0$ . That is every Cauchy sequence in  $X_0$  is convergent to some point  $x \in X_0$ , which is equivalent to  $X_0$ 's completeness.  $\square$

**Proof of the second.** Let  $a$  be an arbitrary point in  $X$ . Define a sequence inductively:

$$(x_n)_{n \in \mathbb{N}}: \mathbb{N} \mapsto X, n \mapsto x_n := \begin{cases} a, & n = 1; \\ T(x_{n-1}), & n \geq 2. \end{cases}$$

Then  $(x_n)_{n \in \mathbb{N}}$  is Cauchy, because for all  $n \geq 2$

$$d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1})) \leq \lambda(x_n, x_{n-1}).$$

By induction, we have  $d(x_{n+1}, x_n) \leq \lambda^{n-1}d(x_2, x_1)$ , and hence

$$\sum_{n \geq 1} d(x_{n+1}, x_n) \leq \sum_{n \geq 1} \lambda^{n-1}d(x_2, x_1) = \frac{1}{1-\lambda}d(x_2, x_1) < \infty.$$

Therefore, the sequence  $(S_n)_{n \in \mathbb{N}}$  is Cauchy, where

$$S: \mathbb{N} \rightarrow \mathbb{R}, n \mapsto S_n := \sum_{j=1}^n d(x_j, x_{j+1}).$$

The triangle inequality implies that

$$\forall m, n > 1 (S_{m \vee n} - S_{m \wedge n-1} \geq d(x_m, x_n)),$$

which ensures that  $(x_n)_{n \in \mathbb{N}}$  is Cauchy ( let  $S_0 = 0$  and then the inequality above always holds ). By the completeness of  $X$ ,  $\exists! x_0 \in X$  such that  $(x_n)_{n \in \mathbb{N}} \rightarrow x$ . Now, the continuity ( from  $d(T(x), T(y)) \leq \lambda d(x, y)$  ) of  $T$  implies

$$T(x_0) = \lim_n T(x_n) = \lim_n x_{n+1} = x_0.$$

This proves the existence. Suppose there is  $y \in X$  such that  $T(y) = y$ , then

$$d(y, x_0) = d(T(y), T(x_0)) \leq \lambda d(y, x_0).$$

$\lambda < 1$  implies that  $d(y, x_0) = 0$ . Equivalently,  $x_0 = y$ . This proves the uniqueness.  $\square$

**Proof of the third.** I think this proof is similar to the proof of [4, Chapter 5, Thm 2].

Necessity: suppose  $X$  is a Banach space. Given a closed decreasing non-empty subsets sequence  $(A_n)_{n \in \mathbb{N}}$ , choose  $x_n \in A_n$  for each  $n \in \mathbb{N}$ . This is possible since  $\forall n \in \mathbb{N} A_n \neq \emptyset$ . Since  $(A_n)_{n \in \mathbb{N}}$  is decreasing, we have

$$\forall m, n \in \mathbb{N} (x_m \in A_{m \wedge n}, x_n \in A_{m \wedge n}),$$

and hence

$$d(x_m, x_n) \leq \text{diam } A_{m \wedge n} \rightarrow 0 (m, n \rightarrow \infty).$$

Therefore,  $(x_n)_{n \in \mathbb{N}}$  is a Cauchy sequence. Then the completeness of  $X$  ensures that  $\exists a \in X$  such that  $(x_n)_{n \in \mathbb{N}} \rightarrow a$ .  $\forall n \in \mathbb{N}$ , since  $A_n$  is closed and  $x_j \in A_n$  for all except for finite  $j \in \mathbb{N}$ , we have  $a \in A_n$ . Therefore,  $a \in \bigcap_{n \geq 1} A_n$ . Clearly  $\bigcap_{n \geq 1} A_n$  can't have more than 1 elements. If so,  $\exists y \in A_n \forall n \in \mathbb{N}$  and hence  $\text{diam}(A_n) \geq d(x, y) \geq 0$ . That's a contradiction.

Sufficiency: suppose  $X$  satisfies the condition above. Let  $(x_n)_{n \in \mathbb{N}}$  be a Cauchy sequence in  $X$ , define  $(A_n)_{n \in \mathbb{N}}$  as follows

$$\forall n \in \mathbb{N}, A_n := \{x_m \in X : m \geq n\}.$$

Then  $(\overline{A_n})_{n \in \mathbb{N}}$  satisfies the condition for set sequence: clearly  $(\overline{A_n})_{n \in \mathbb{N}}$  is decreasing, and  $\text{diam}(\overline{A_n}) = \text{diam}(A) \rightarrow 0$  since  $(x_n)_{n \in \mathbb{N}}$  is Cauchy. The reason of  $\text{diam}(\overline{A_n}) = \text{diam}(A_n)$  is written in remark. Therefore,  $\exists! a \in \bigcap_{n \geq 1} A_n$ . Now, it suffices to show that  $(x_n)_{n \in \mathbb{N}} \rightarrow a$ . This follows from

$$d(x_n, a) \leq \text{diam}(\overline{A_n}) \rightarrow 0 (n \rightarrow \infty). \quad \square$$

**Remark.**  $\forall n \in \mathbb{N}$ , we want to show that  $\text{diam}(\overline{A_n}) = \text{diam}(A_n)$ . Since  $n$  is fixed, we can omit the index. Given  $A \subseteq X$  and  $\varepsilon > 0$ ,  $\forall x, y \in \overline{A}$ , there is  $x_\varepsilon, y_\varepsilon \in A$  such that

$$\|x - x_\varepsilon\| < \varepsilon/2, \|y - y_\varepsilon\| < \varepsilon/2.$$

Therefore

$$\|x - y\| \leq \|x - x_\varepsilon\| + \|x_\varepsilon - y_\varepsilon\| + \|y_\varepsilon - y\| \leq \|x_\varepsilon - y_\varepsilon\| + \varepsilon,$$

and use  $\|x_\varepsilon - y_\varepsilon\| \leq \text{diam}(A)$ ,

$$\|x - y\| \leq \text{diam}(A) + \varepsilon.$$

Since  $x, y \in \overline{A}$  are arbitrary, we have

$$\text{diam}(\overline{A}) \leq \text{diam}(A) + \varepsilon.$$

And  $\varepsilon$  is arbitrary, so

$$\text{diam}(\overline{A}) \leq \text{diam}(A).$$

The reversed inequality is trivial.

#### 4.2.1 Banach Fixed-point Theorem

Here we introduce a classical result about Banach spaces.

**Definition** (Contraction mapping). Given a metric space  $(X, d)$ . Then a mapping  $T: X \rightarrow X$  is called a contraction if  $\exists \lambda \in (0, 1)$  such that  $d(T(x), T(y)) \leq \lambda d(x, y)$ .

**Remark.** Every linear normed space  $(X, \|\cdot\|)$  has the natural metric  $d(x, y) = \|x - y\|$  and hence a contraction on  $(X, \|\cdot\|)$  means  $T: X \rightarrow X$ , such that  $\exists \lambda \in (0, 1), \forall x, y \in X$

$$\|T(x) - T(y)\| \leq \lambda \|x - y\|.$$

The  $\neq$  above means that  $T$  may not be a linear map.

It is easy to verify that each contraction is continuous.

**Theorem 4.1** (Banach fixed-point theorem). Suppose  $(X, d)$  is a complete metric space and  $T$  is a contraction on  $X$ . Then  $\exists! x_0 \in X$  such that  $Tx_0 = x_0$ .

**Proof.** See the [the proof of the second question](#). □

Let's have some applications. Suppose  $X$  is a Banach space and  $U: X \rightarrow X$ . We want to solve the equation  $U(x) = y$ .

**Proof.** To use 4.2.1, we should rewrite the equation  $U(x) = y$  as  $T(x) = x$  for some  $T$ .

$$U(x) = y \iff U(x) - y = 0 \iff U(x) + x - y = x,$$

thus consider  $T: X \rightarrow X, x \mapsto U(x) + x - y$ . And

$$\|T(u) - T(v)\| = \|U(u) + u - y - (U(v) + v - y)\|.$$

If it's verified that  $T$  is a contraction, then 4.2.1 (Banach Fixed-point Theorem) implies that  $T$  has a unique fixed-point, i.e.  $U(x) = y$  has a unique solution.  $\square$

**Example 18.**  $X$  is a Banach space, on which  $U$  is a contraction. Prove that  $U(x) = x + y$  has a unique solution.

**Proof.** We want solve  $U(x) - x = y$ , i.e.  $(U - \text{id})(x) - y = 0$ . So the discussion above tells us that we should consider  $T = U - \text{id} + \text{id} - y = U - y$ . Let  $x_1 \in X$  be an arbitrary point. Define  $x_{n+1} = T(x_n) = U(x_n) - y$  for all  $n \in \mathbb{N}$ . Then  $T$  is a contraction since

$$\|T(a) - T(b)\| = \|U(a) - U(b)\|,$$

and  $U$  is a contraction. Then use Theorem 4.2.1 (Banach Fixed-point Theorem) and we're done.  $\square$

## 5 Week 5

### 5.1 Week 5, Lecture 1

In this part,  $X, Y$  are supposed to be two linear normed spaces  $(X, \| \cdot \|)$ ,  $(Y, \| \cdot \|)$ .

#### Recall

A map  $T: X \rightarrow Y$  is said to be continuous, if

$$\forall x \in X \forall (x_n)_{n \in \mathbb{N}} \xrightarrow{\| \cdot \|_X} x, (Tx_n)_{n \in \mathbb{N}} \xrightarrow{\| \cdot \|_Y} Tx.$$

#### 5.1.1 Bounded Linear Operators/Maps

Here is the definition of Bounded linear operators/maps

**Definition** (Bounded linear operators/maps).  $T: X \rightarrow Y$  is said to be bounded, if  $\exists C > 0$  such that  $\| \cdot \|_Y \circ T \leq \| \cdot \|_X$ , equivalently  $\|Tx\|_Y \leq \|x\|_X, \forall x \in X$ . The set of all bounded linear operators from  $X$  to  $Y$  is denoted as  $\mathcal{B}(X, Y)$ . If  $Y = X$ ,  $\mathcal{B}(X, X)$  is also written as  $\mathcal{B}(X)$ .

**Remark.**  $\exists C > 0 : \|Tx\|_Y \leq \|x\|_X, \forall x \in X$  is **not** equivalent to  $\forall x \in X \exists C > 0 : \|Tx\|_Y \leq \|x\|_X$ .

**Remark.** Usually we don't distinguish map and operator, but a functional should be distinguished ( see the definition of [Bounded linear functional](#) ).

It is easy to verify: a bounded map is continuous. Then it's natural to consider the inverse proposition. To do this, we define bounded sets.

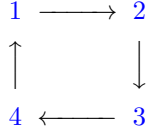
**Definition** (Bounded set). Suppose  $A \subseteq X$ . If  $\exists M > 0$  such that  $\sup_{x \in A} \|x\| \leq M$ , then  $A$  is said to be bounded.

**Remark.**  $T$  is a bounded map  $\iff T$  maps bounded sets to bounded sets.

**Proposition 5.1.** The following statements are equivalent.

1.  $T$  is continuous;
2.  $T$  is continuous at some point  $x_0 \in X$ ;
3.  $T$  is continuous at 0;
4.  $T$  is bounded.

**Proof.** We prove in the following order



1 $\rightarrow$ 2: is done automatically.

2 $\rightarrow$ 3: Suppose  $T$  is continuous at  $x_0$ , then  $\forall (x_n)_{n \in \mathbb{N}} \rightarrow x_0$  we have  $(Tx_n)_{n \in \mathbb{N}} \rightarrow Tx_0$ . Now  $\forall (y_n)_{n \in \mathbb{N}} \rightarrow 0$ , we have  $(y_n + x_0)_{n \in \mathbb{N}} \rightarrow x_0$  since

$$\|(y_n + x_0) - x_0\| = \|y_n\| \rightarrow 0 (n \rightarrow \infty).$$

Thus,  $T(y_n + x_0) \rightarrow T(x_0)$  by  $T$ 's continuity at  $x_0$  and hence

$$\|T(y_n) - 0\| = \|T(y_n + x_0) - T(x_0)\| \rightarrow 0.$$

Therefore,  $(Ty_n)_{n \in \mathbb{N}} \rightarrow 0$  as we wanted.

3 $\rightarrow$ 4: Given  $T$  that is continuous at 0. If  $T$  isn't bounded, then there is a bounded subset of  $X$ , denoted by  $A$ , such that  $TA$  is unbounded. **Replace**  $A$  with  $\bigcup_{0 \leq t \leq 1} tA$ , still denoted by  $A$ . By the definition of unboundedness:

$$\forall n \in \mathbb{N} \exists x_n \in A : \|Tx_n\| > n.$$

Now we want a sequence  $(y_n)_{n \in \mathbb{N}} \subseteq A$  satisfying  $(\|Ty_n\|)_{n \in \mathbb{N}}$  is unbounded. Take  $y_n = x_n / \sqrt{n}$ , and we're done. Since  $\{y_n : n \in \mathbb{N}\}$  is a bounded subset of  $A$  whose image under  $T$  is unbounded. That's a contradiction.

4 $\rightarrow$ 1:  $T$  is bounded, then  $T$  is uniformly continuous.  $\square$

**Remark.** There is another proof of 3  $\rightarrow$  4, see the textbook.

Now, we have a set and it's naturally to consider it's linear structure and topology. There is a natural linear structure on  $\mathcal{B}(X, Y)$  as follows

$$\begin{aligned}
 + : \mathcal{B}(X, Y) \times \mathcal{B}(X, Y) &\rightarrow \mathcal{B}(X, Y) \\
 (S, T) &\mapsto S + T := (x \mapsto S(x) + T(x)),
 \end{aligned}$$

and

$$\begin{aligned}
 \cdot : \mathcal{B}(X, Y) \times \mathbb{K} &\rightarrow \mathcal{B}(X, Y) \\
 (S, k) &\mapsto k \cdot S := (x \mapsto k \cdot S(x)).
 \end{aligned}$$

**Definition** (Operator norm). The operator norm on  $\mathcal{B}(X, Y)$  is defined as follows

$$\| \cdot \|: \mathcal{B}(X, Y) \rightarrow \mathbb{R}_{\geq 0}, T \mapsto \sup_{\|x\| \leq 1} \|Tx\|.$$

It's easy to verify that the operator norm is a norm on  $\mathcal{B}(X, Y)$ .

**Remark.**  $(\mathcal{B}(X, Y), \| \cdot \|)$  is a linear normed space.

**Remark.** Equivalent definitions:

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|.$$

**Proof.** Since

$$\begin{aligned} \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} &= \sup_{x \neq 0} \left\| \frac{1}{\|x\|} Tx \right\| \\ &= \sup_{x \neq 0} \left\| T \left( \frac{x}{\|x\|} \right) \right\| \\ &= \sup_{\|x\|=1} \|Tx\| \\ &= \sup_{\|x\| \leq 1} \|Tx\| \\ &= \sup_{0 < \|y\| \leq \delta} \frac{1}{\delta} \|Ty\| \quad (y = \delta x) \\ &\leq \sup_{0 < \|y\| \leq \delta} \frac{1}{\|y\|} \|Ty\| \\ &= \sup_{y \neq 0} \frac{\|Ty\|}{\|y\|}. \end{aligned}$$

And

$$\sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{y \neq 0} \frac{\|Ty\|}{\|y\|},$$

which ensures that the  $\leq$  above can be replaced with  $=$ . □

**Definition** (Bounded linear functional). An element of  $\mathcal{B}(X, \mathbb{K})$  is called a linear functional on  $X$ .  $\mathcal{B}(X, \mathbb{K})$  is also called the dual space of  $X$ , denoted by  $X^*$ .

**Remark.** Discontinuous linear functionals exist (but only when  $X$  is infinite dimensional. See [this post](#)).

**Example 19.** Fix  $a = (a_n)_{n \in \mathbb{N}} \in \ell_1$ . Define

$$T: c_0 \rightarrow \ell_1, x = (x_n)_{n \in \mathbb{N}} \mapsto a \cdot x = (a_n x_n)_{n \in \mathbb{N}}. \quad (8)$$

Show that:

- 1)  $T$  is bounded;
- 2)  $\|T\| = \|a\|_1$ .

**Proof.**

- 1) Recall that  $c_0 \hookrightarrow \ell_\infty$  is equipped with the norm  $\|\cdot\|_\infty = \sup_{\mathbb{N}} |\cdot|$ .  
 $\forall x \in c_0$ , we have

$$\begin{aligned} \|Tx\|_1 &= \|a \cdot x\|_1 \\ &= \sum_{n \geq 1} |a_n x_n| \\ &\leq \sum_{n \geq 1} |a_n| \|x\|_\infty \\ &= \|a\|_1 \|x\|_\infty. \end{aligned}$$

Thus pick  $C = \|a\|_1$ , we have  $\|Tx\|_{\ell_1} \leq C \|x\|_\infty$ . This means  $T \in \mathcal{B}(c_0, \ell_1)$ .

- 2) We have proved  $\|T\| \leq \|a\|_1$ . Thus it suffices to show the reversed inequality. From the definition of  $\|\cdot\|_1$

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N : \sum_{j=1}^n |a_j| > \|a\|_1 - \varepsilon.$$

In particular

$$\sum_{j=1}^N |a_j| > \|a\|_1 - \varepsilon.$$

Now consider

$$c_0 \ni x_N := (\underbrace{1, \dots, 1}_{N \text{ terms}}, 0, 0, \dots),$$

whose image under  $T$  is

$$\|Tx_N\| = \sum_{j=1}^N |a_j| > \|a\|_1 - \varepsilon.$$



$\|x\|_\infty = 1$  ensures that

$$\|T\| \geq \|Tx_N\|_{\ell_1} > \|a\|_1 - \varepsilon.$$

$\varepsilon > 0$  is arbitrary, therefore  $\|T\| \geq \|a\|_1$ .

□

**Remark.** In fact,  $c_0^* \cong \ell_1$ . Here  $\cong$  means “isometrically isomorphic”.

Here is a left exercise:

**Exercise 5.1.** Consider  $X = C[0, 1]$ , with the norm  $x \mapsto \max_{[0,1]} |x|$ . Define the linear functional

$$f: X \rightarrow \mathbb{K}, x \mapsto \int_0^{1/2} x \, dm - \int_{1/2}^1 x \, dm.$$

Here  $m$  is the Lebesgue measure on  $\mathbb{R}$ . Show that:

- 1)  $f$  is a bounded linear functional ( i.e.  $f \in (C[0, 1])^*$  );
- 2)  $\|f\| = 1$ .

## 5.2 Week 5, Lecture 2

Here is a remark for Exercise 5.1. We want to find  $x \in C[0, 1]$ ,  $\|x\| = 1$ ,  $|f(x)| = 1$ , i.e.

$$\int_0^{1/2} x \, dm = 1/2, \int_{1/2}^1 x \, dm = -1/2.$$

But this is impossible, by  $\max_{[0,1]} |x| = 1$  and the continuity of  $x$ . Now, consider the approximation of  $x$ :  $\forall \varepsilon \in (0, 1/2)$ , let

$$x_\varepsilon: [0, 1] \rightarrow \mathbb{R}, t \mapsto \begin{cases} 1, & t \in [0, 1/2 - \varepsilon] \\ l(t), & t \in (1/2 - \varepsilon, 1/2 + \varepsilon) \\ -1, & [1/2 + \varepsilon, 1] \end{cases},$$

where  $l$  is the unique affine function determined by

$$l(1/2 - \varepsilon) = 1, l(1/2 + \varepsilon) = -1.$$

Since  $|f(x_\varepsilon)| = 1 - \varepsilon$  and  $|x_\varepsilon| = 1$ , we have  $\|f\| \geq 1 - \varepsilon$ . Therefore,  $\|f\| \geq 1$ .

### 5.2.1 Some exercises

Here are exercises for this class.

**Exercise 5.2.** Given a measure space  $(\Omega, \mathcal{F}, \mu)$  and  $\alpha \in L_1(\Omega)$ . Let

$$T_\alpha: L_\infty(\Omega) \rightarrow L_1(\Omega), x \mapsto \alpha \cdot x,$$

where  $\alpha \cdot x$  means pointwise product. Try to find  $\|T_\alpha\|$ .

**Solution:** It's natural to guess that  $\|T_\alpha\| = \|\alpha\|_1$ . Hölder's inequality implies that

$$\|T_\alpha(x)\|_1 = \|\alpha \cdot x\|_1 \leq \|\alpha\|_1 \|x\|_\infty.$$

Thus,  $\|T_\alpha\| \leq \|\alpha\|_1$ . On the other hand,

$$L_\infty(\Omega) \ni x: \Omega \rightarrow \mathbb{K}, \omega \mapsto 1$$

then  $\|x\|_\infty = 1$ , and  $T_\alpha(x) = \alpha$ , hence  $\|T_\alpha\| \geq \|\alpha\|_1$ .  $\square$

**Remark.** We have proved this for  $\mu$  being the counting measure, see 19.

**Fact.** A matrix (with respect to the normal base)  $T \in \mathbb{K}^{n \times n}$  considered as a linear map  $T: \mathbb{K}^n \rightarrow \mathbb{K}^n, x \mapsto Tx$  is bounded.

**Proof.** Since  $\mathbb{K}$  is equipped with the norm  $\| \cdot \|: x \mapsto (\sum_{j=1}^n |x_j|^2)^{1/2}$  that is not very convenient. It can be proved that  $\| \cdot \|_\infty \leq \| \cdot \| \leq \sqrt{n} \| \cdot \|_\infty$ . So it suffices to show that  $T: (\mathbb{K}^n, \| \cdot \|_\infty) \rightarrow (\mathbb{K}^n, \| \cdot \|_\infty)$  is continuous. Suppose  $T = (a_{i,j})_{n \times n}$ . Now  $\forall x = (x_1, \dots, x_n)^t \in \mathbb{K}^n$

$$\begin{aligned}
\|Tx\|_\infty &= \left\| \left( \sum_{j=1}^n a_{1,j}x_1, \dots, \sum_{j=1}^n a_{n,j}x_n \right)^t \right\|_\infty \\
&\leq \sum_{k=1}^n \left\| \left( \sum_{j=1}^n a_{1,j}x_1\delta_{k,j}, \dots, \sum_{j=1}^n a_{n,j}x_n\delta_{k,j} \right)^t \right\|_\infty \\
&\leq \sum_{k=1}^n \sum_{j=1}^n \left\| \left( a_{1,j}x_1\delta_{k,j}, \dots, a_{n,j}x_n\delta_{k,j} \right)^t \right\|_\infty \quad (9) \\
&= \sum_{k=1}^n \sum_{j=1}^n |a_{k,j}| \cdot |x_k| \\
&\leq \left( \sum_{k=1}^n \sum_{j=1}^n |a_{k,j}| \right) \|x\|_\infty.
\end{aligned}$$

Thus, let  $C := \sum_{j=1}^n \sum_{k=1}^n |a_{k,j}|$  and we have proved  $\| \cdot \|_\infty \circ T \leq C \| \cdot \|_\infty$ , i.e.  $T$  is bounded.  $\square$

**Claim.** Each finite dimensional linear normed space  $X$  is linear homeomorphic to  $\mathbb{K}^n$ .

**Proof.** Suppose  $\mathbb{K}$  is equipped with  $\| \cdot \|_\infty$  and  $\{\alpha_1, \dots, \alpha_n\}$  is a base of  $X$ . Thus there is a map

$$\varphi: \mathbb{K}^n \rightarrow X, (x_1, \dots, x_n)^t \mapsto \sum_{j=1}^n x_j \alpha_j,$$

which is a bijection from definition of base. And  $\varphi$  is bounded, since

$$\begin{aligned}
\|\varphi(x_1, \dots, x_n)\|_X &\leq \sum_{j=1}^n |x_j| \|\alpha_j\|_X \\
&\leq \left( \sum_{j=1}^n \|\alpha_j\|_X \right) \| (x_1, \dots, x_n) \|_\infty. \quad (10)
\end{aligned}$$

Let  $C := \sum_{j=1}^n \|\alpha_j\|_X$ , then  $\| \cdot \|_\infty \circ \varphi \leq C \| \cdot \|_\infty$  and thus  $\varphi$  is bounded.

Now we prove that  $\Phi := \varphi^{-1}$  is bounded. Given  $(x_1, \dots, x_n) \in \mathbb{K}^n$  such that

$$\left\| \sum_{j=1}^n x_j \alpha_j \right\|_X \leq 1,$$

i.e. an element in the unit ball of  $X$ . We prove that  $\Phi(\sum_{j=1}^n x_j \alpha_j) = (x_1, \dots, x_n)$  lies in some ball of  $\mathbb{K}^n$ .  $\{\alpha_1, \dots, \alpha_n\}$  is a base for  $X$ , thus  $\alpha_j \neq 0 (\forall j)$  and let  $\delta = \min_{1 \leq j \leq n} \|\alpha_j\| > 0$ . Now

$$1 \geq \left\| \sum_{j=1}^n x_j \alpha_j \right\|_X \geq \sum_{j=1}^n |x_j| \|\alpha_j\| \geq \delta \sum_{j=1}^n |x_j| \geq \delta \|(x_1, \dots, x_n)^t\|_\infty.$$

Therefore  $\|(x_1, \dots, x_n)^t\|_\infty \leq 1/\delta$ , i.e.  $\left\| \Phi(\sum_{j=1}^n x_j \alpha_j) \right\|_\infty \leq 1/\delta$ . This means that  $\Phi$  is bounded.  $\square$

**Exercise 5.3.** Given a measure space  $(\Omega, \mathcal{F}, \mu)$  and  $\alpha \in L_\infty(\Omega)$ . Let  $p$  be a real number fixed in  $(1, \infty)$ . Define

$$T_\alpha: L_p(\Omega) \rightarrow L_p(\Omega), x \mapsto \alpha \cdot x.$$

Try to find  $\|T_\alpha\|$ .

**Proof.** Let  $T$  denotes  $T_\alpha$  for short. First,  $\|T\| \leq \|\alpha\|_\infty$ : since  $|\alpha(\omega)| \leq \|\alpha\|_\infty$  for a.e.  $\omega \in \Omega$ , and

$$\|Tx\|_p = \left( \int_\Omega |\alpha|^p |x|^p d\mu \right)^{1/p} \leq \|\alpha\|_\infty \left( \int_\Omega |x|^p d\mu \right)^{1/p} = \|\alpha\|_\infty \|x\|_p.$$

The reversed inequality needs a condition: “ Suppose  $L_p \neq \{0\}$ . Then  $\forall A \in \mathcal{F}$  such that  $\mu(A) = \infty$ ,  $\exists A_0 \subseteq A$  such that  $0 < \mu(A_0) < \infty$ ”. Now,  $\forall \varepsilon > 0$ , consider the set  $E_\varepsilon := \{\omega \in \Omega : |\alpha(\omega)| > \|\alpha\|_\infty - \varepsilon\}$ .

**Case 1:**  $\mu(E_{\varepsilon_1}) < \infty$  for some  $\varepsilon_1 > 0$ . Since  $0 < a < b$  implies  $E_a \subseteq E_b$ , by considering  $\varepsilon < \varepsilon_1$  we have  $\mu(E_\varepsilon) < \infty$ . Then  $\chi_{E_\varepsilon} \in L_p$ . And hence

$$\|T\| \geq \frac{\|T\chi_{E_\varepsilon}\|}{\|\chi_{E_\varepsilon}\|} \geq \frac{(\|\alpha\|_\infty - \varepsilon) \left( \int_{E_\varepsilon} \chi_{E_\varepsilon}^p d\mu \right)^{1/p}}{\|\chi_{E_\varepsilon}\|} = \|\alpha\|_\infty - \varepsilon.$$

Since  $\varepsilon \in (0, \varepsilon_1)$  is arbitrary, we have  $\|T\| \leq \|\alpha\|_\infty$ .

**Case 2:**  $\mu(E_\varepsilon) = \infty$  for all  $\varepsilon > 0$ . If  $\exists A_\varepsilon \subseteq E_\varepsilon$  such that  $0 < \mu(A_\varepsilon) < \infty$  and hence  $\chi_{A_\varepsilon} \in L_p$ , then

$$\|T\| \geq \frac{\|T\chi_{A_\varepsilon}\|}{\|\chi_{A_\varepsilon}\|} \geq \frac{(\|\alpha\|_\infty - \varepsilon) \left( \int_{A_\varepsilon} \chi_{A_\varepsilon}^p d\mu \right)^{1/p}}{\|\chi_{A_\varepsilon}\|} = \|\alpha\|_\infty - \varepsilon.$$

Since  $\varepsilon > 0$  is arbitrary, we have  $\|T\| \leq \|\alpha\|_\infty$ .

For the case that there is some  $\varepsilon > 0$  such that  $\mu(E_\varepsilon) = \infty$  and  $\mu(A) \in \{\infty, 0\}$  for all  $A \subseteq E_\varepsilon$ , we can't prove that  $\|T\| \geq \|\alpha\|_\infty$  and there is a example such that  $\|T\| \neq \|\alpha\|_\infty$  in this case.

**Example 20.** Consider the measure space  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu)$  where  $\mu$  is defined as the unique measure such that

$$\nu(\{1\}) = \infty, \nu(A) = \text{card}(A) (\forall 1 \notin A),$$

where  $\text{card}(A)$  is the number of elements of the set  $A$  when  $A$  is finite, and  $\infty$  when  $A$  is infinite. Now the function  $\alpha = \chi_{\{1\}} \in L_\infty(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu)$  and  $\forall f \in L_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu)$  we have  $f(1) = 0$ . Therefore

$$T_\alpha: L_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu) \rightarrow L_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu), f \mapsto \alpha \cdot f$$

is just a zero operator and hence  $\|T_\alpha\| = 0 \neq \|\alpha\|_\infty$ .

Therefore, the operator

$$T: L_\infty(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu) \rightarrow \mathcal{B}(L_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu))$$

has a nontrivial kernel  $\ker T \ni \alpha \neq 0$ . □

## 6 Week 6

### 6.1 Week 6, Lecture 1

#### 6.1.1 Compactness, Relative Compactness and Total Boundedness

**Definition** (Open Cover). Given a topological space  $(X, \mathcal{T})$ .  $A \subseteq X$  is said to have an open cover  $(O_i)_{i \in I}$  if

$$A \subseteq \bigcup_{i \in I} O_i.$$

**Definition** (Compact). A topological space  $(X, \mathcal{T})$  is said to be compact, if each open cover of  $X$  has a finite subcover.

**Remark.** Compactness is topological invariant.

**Definition** (Relative Compactness). Let  $(X, \mathcal{T})$  be a topological space. A subset  $F$  of  $(X, \mathcal{T})$  is said to be relatively compact, if its closure  $\overline{F}$  is compact.

**Definition** (Sequential Compactness). Let  $(X, \mathcal{T})$  be a topological space. A subset  $F$  of  $(X, \mathcal{T})$  is said to be sequentially compact, if every sequence  $(x_n)_{n \in \mathbb{N}} \subseteq A$  there is a subsequence  $(x_{n_k})_{k \in \mathbb{N}} \subseteq (x_n)_{n \in \mathbb{N}}$  such that  $(x_{n_k})_{k \in \mathbb{N}} \rightarrow x \in A$ .

**Definition** ( $\varepsilon$ -net). Let  $(X, d)$  be a metric space.  $E \subseteq X$  is called an  $\varepsilon$ -net of  $A$ , if  $A \subseteq \bigcup_{x \in E} B(x, \varepsilon)$ .

**Definition** (Total Boundedness). Let  $(X, d)$  be a metric space.  $A \subseteq X$  is said to be totally bounded, if  $\forall \varepsilon > 0$  there is a finite  $\varepsilon$ -net of  $A$ .

**Remark.** This is not a topological invariant (since it needs a metric), but is invariable under bi-Lipschitz mappings.

Now, we will compare the following notions in metric space: compact sets, relatively compact sets and totally bounded sets.

**Theorem 6.1.** Let  $(X, d)$  be a metric space and  $A \subseteq X$ . The following statements are equivalent:

1.  $A$  is compact.
2.  $A$  is sequentially compact.

**Proof.** 1  $\Rightarrow$  2: Suppose  $A$  is compact while not sequentially compact. Then  $\exists (x_n)_{n \in \mathbb{N}} \subseteq A$  such that  $\forall a \in A$ ,  $a$  is not a limit point of  $(x_n)_{n \in \mathbb{N}}$ . Thus

$$\forall a \in A \exists \varepsilon_a > 0 (\exists N_a \in \mathbb{N} \forall n \geq N_a d(x_{N_a}, a) > \varepsilon_a).$$

Now we have an open cover of  $A$ ,  $\{B(a, \varepsilon_a) : a \in A\}$ . Since  $A$  is compact, there is  $a_1, \dots, a_m \in A$  such that

$$A \subseteq \bigcup_{k=1}^m B(a_k, \varepsilon_{a_k}),$$

Let  $N := N_{a_1} \vee \dots \vee N_{a_m}$  then  $x_N \notin B(a_k, \varepsilon_{a_k}) \forall 1 \leq k \leq m$ . But  $x_N \in A = \bigcup_{k=1}^m B(a_k, \varepsilon_{a_k})$ . That's a contradiction.

2  $\Rightarrow$  1: Let  $(O_i)_{i \in I}$  be an open covering of  $A$ . First, we prove that  $\exists \lambda > 0$  such that  $\forall 0 < r < \lambda \forall x \in A, B(x, r) \subseteq O_i$  for some  $i \in I$  ( This constant  $\lambda$  is called an Lebesgue number of the open covering  $(O_i)_{i \in I}$ ).

If there is no Lebesgue number for  $(O_i)_{i \in I}$ , then  $\forall n \in \mathbb{N} \exists x_n \in A$  such that  $B(x_n, 1/n)$  is not contained in any element of  $(O_i)_{i \in I}$ . Therefore we have a sequence  $(x_n)_{n \in \mathbb{N}}$ . 2 ensures that  $(x_n)_{n \in \mathbb{N}}$  has a convergent subsequence  $(x_{n_k})_{k \in \mathbb{N}}$  with its limit  $x_0$ . Notice that  $x_0 \in O_{i_0}$  for some  $i_0 \in I$  and  $O_{i_0}$  is open, so  $\exists r > 0$  such that  $B(x_0, r) \subseteq O_{i_0}$ . From the definition of convergence,  $\exists K$  such that  $\forall k \geq K d(x_{n_k}, x_0) < r/2$ . WLOG, suppose  $n_K > 2/r$ . Now,  $\forall y \in B(x_{n_K}, 1/n_K)$ , we have

$$d(y, x_0) \leq d(y, x_{n_K}) + d(x_{n_K}, x_0) < \frac{1}{n_K} + \frac{r}{2} < r.$$

This means  $B(x_{n_K}, 1/n_K) \subseteq B(x_0, r)$ . Since  $B(x_0, r) \subseteq O_{i_0}$ , we get  $B(x_{n_K}, 1/n_K) \subseteq O_{i_0}$ . That's a contradiction with the selection of  $(x_n)_{n \in \mathbb{N}}$ . Therefore, there is a Lebesgue number.

Let  $\lambda$  be a Lebesgue number, whose existence is proved above. Then  $A$  has an open cover  $\{B(x, \lambda/2) : x \in A\}$ . Take arbitrary  $x_1 \in A$ . If  $A \subseteq B(x_1, \lambda/2)$  we're done. Else, it's possible to take  $x_2 \in A \setminus B(x_1, \lambda/2)$ . Similarly we can take  $x_3, \dots, x_n, \dots$  if possible. This process must end in finite steps, i.e. we can only get a finite sequence as above. If we get a infinite sequence  $(x_n)_{n \in \mathbb{N}}$  as above, then

$$d(x_m, x_n) \geq \frac{\lambda}{2}, \forall m \neq n.$$

That's a contradiction since  $A$  is supposed to be sequentially compact. Suppose we get a sequence having only  $m$  terms and then

$$A \subseteq \bigcup_{k=1}^m B\left(x_k, \frac{\lambda}{2}\right).$$

Recall the selection of  $\lambda, x_k$  ensures that  $B(x_k, \lambda/2)$  lies in an element of  $(O_i)_{i \in I}$  for each  $k$ . Therefore  $(O_i)_{i \in I}$  has a finite subcover.  $\square$

**Theorem 6.2.** Let  $(X, d)$  be a metric space and  $A \subseteq X$ . The following statements are equivalent:

1.  $A$  is relatively compact.
2.  $\forall (x_n)_{n \in \mathbb{N}} \subseteq A, \exists (x_{n_k})_{k \in \mathbb{N}} \subseteq (x_n)_{n \in \mathbb{N}}$  such that  $(x_{n_k})_{k \in \mathbb{N}} \xrightarrow{d} x \in X$ .

**Remark.** Notice that  $(x_{n_k})_{k \in \mathbb{N}} \xrightarrow{d} x \in X$  but not  $(x_{n_k})_{k \in \mathbb{N}} \xrightarrow{d} x \in A$ .

**Proof.** We use Theorem 6.1 to prove this theorem.

**1  $\Rightarrow$  2:** Suppose 1 holds, then  $\overline{A}$  is compact, Theorem 6.1 implies  $\overline{A}$  is sequentially compact and hence 2 holds.

**2  $\Rightarrow$  1:** Suppose 2 holds, then clearly  $x \in \overline{A}$ . Now we want to prove that  $\overline{A}$  is compact. 1 means that it suffices to show that  $\overline{A}$  is sequentially compact. Given an arbitrary sequence  $(x_n)_{n \in \mathbb{N}} \subseteq \overline{A}$ , we want to show that there is a subsequence  $x_{n_k} \rightarrow x$  for some  $x \in X$ . Since  $(x_n)_{n \in \mathbb{N}} \subseteq \overline{A}$  doesn't mean that  $(x_n)_{n \in \mathbb{N}} \subseteq A$ , we should find a sequence  $(y_n)_{n \in \mathbb{N}} \subseteq A$  such that  $x_{n_k} \rightarrow x$  whenever  $y_{n_k} \rightarrow x$ . By the property of closure: we can define  $(y_n)_{n \in \mathbb{N}} \subseteq A$  such that

$$\forall n \in \mathbb{N}, y_n := \begin{cases} x_n, & x_n \in A; \\ x'_n, & x_n \notin A, x'_n \in A, d(x'_n, x_n) < 1/n. \end{cases}$$

Now 2 implies that  $\exists (y_{n_k})_{k \in \mathbb{N}}$  such that  $(y_{n_k})_{k \in \mathbb{N}} \rightarrow x \in X$ , and hence  $(x_{n_k})_{k \in \mathbb{N}} \rightarrow x \in X$  as we want.  $\square$

**Theorem 6.3.** Let  $(X, d)$  be a metric space and  $A \subseteq X$ . The following statements are equivalent:

1.  $A$  is totally bounded.
2.  $\forall (x_n)_{n \in \mathbb{N}} \subseteq A, \exists (x_{n_k})_{k \in \mathbb{N}} \subseteq (x_n)_{n \in \mathbb{N}}$  such that  $(x_{n_k})_{k \in \mathbb{N}}$  is a Cauchy sequence.

**Proof.** **1  $\Rightarrow$  2:** proof given by our professor is omitted here and should be found in your notes. And the “another proof” is not very different from this.

“Another proof” of **1  $\Rightarrow$  2:** suppose  $A$  is totally bounded. Given an arbitrary sequence  $(x_n)_{n \in \mathbb{N}} \subseteq X$ . WLOG, suppose  $(x_n)_{n \in \mathbb{N}}$  has infinite distinct terms, else we're done.  $\forall \varepsilon > 0$  there is a finite  $\varepsilon$ -net of  $A$ . Thus for each  $k \in \mathbb{N}$  there is a finite  $\varepsilon$ -net  $F_k$  for  $A$ . Let  $J_0 = \mathbb{N}$



and define  $J_0 \supseteq J_1 \supseteq J_2 \supseteq \cdots$  inductively as follows. Suppose  $J_k$  is defined. Since  $F_{k+1}$  is finite and  $J_k$  is infinite, for each  $n \in J_k$  there is an element  $p_{k+1} \in F_{k+1}$  such that the ball  $B(p_{k+1}, 1/(k+1))$  contains infinite elements of  $\{x_n : n \in J_k\}$ . Let

$$J_{k+1} := \{n \in J_k : d(x_n, p_{k+1}) < 1/(k+1)\}.$$

Now, let  $n_1 \in J_1$  be an arbitrary element. And inductively select  $n_{k+1} \in J_{k+1}$  such that  $n_{k+1} > n_k$ . We have defined a subsequence  $(x_{n_k})_{k \in \mathbb{N}}$ .  $\forall \varepsilon > 0 \exists N \in \mathbb{N}$  such that  $2/N < \varepsilon$  and hence  $\forall j, k \geq N$  we have  $d(x_{n_j}, p_N) < 1/n_j < 1/N$ ,  $d(x_{n_k}, p_N) < 1/N$ . Therefore

$$d(x_{n_k}, x_{n_j}) \leq d(x_{n_k}, p_N) + d(x_{n_j}, p_N) < 1/N + 1/N < \varepsilon$$

by the triangle inequality. Now  $(x_{n_k})_{k \in \mathbb{N}}$  is a subsequence of  $(x_n)_{n \in \mathbb{N}}$  that is Cauchy.

**2  $\Rightarrow$  1:** Suppose  $A$  satisfies **1**. If  $A$  isn't totally bounded, then  $\exists \varepsilon_0 > 0$  such that  $A$  has no finite  $\varepsilon_0$ -net. Thus pick an arbitrary point  $x_1 \in \mathbb{N}$  and  $X \setminus B(x_1, \varepsilon_0) \neq \emptyset$  (since  $A$  has no finite  $\varepsilon_0$ -net). Pick an arbitrary point  $x_2 \in X \setminus B(x_1, \varepsilon_0)$  and pick  $x_3$  similarly. We have defined a sequence  $(x_n)_{n \in \mathbb{N}}$  inductively, satisfying

$$d(x_m, x_n) \geq \varepsilon_0 (\forall m \neq n),$$

which implies that  $(x_n)_{n \in \mathbb{N}}$  has no Cauchy subsequence. That's a contradiction. Therefore,  $A$  must be totally bounded.  $\square$

**Corollary 6.4.** Let  $(X, d)$  be a metric space and  $A \subseteq X$ . Then

1.  $A$  is compact  $\implies A$  is relatively compact  $\implies A$  is totally bounded.
2.  $A$  is compact  $\implies A$  is closed and bounded.
3. Suppose  $A$  is closed. Then  $A$  is compact  $\iff A$  is relatively compact.
4. Suppose  $X$  is complete. Then  $A$  is relatively compact  $\iff A$  is totally bounded.
5.  $X$  is compact  $\iff X$  is complete and totally bounded.
6.  $X = \mathbb{K}^n$ , then  $A$  is bounded  $\iff A$  is totally bounded  $\iff A$  is relatively compact.

**Proof.** We imply some results from the **point set topology** course.

1:  $(X, d)$  is a metric space and hence a Hausdorff space. Compact sets in Hausdorff space is closed. Therefore  $A = \overline{A}$  and  $\overline{A}$  is compact, i.e.  $A$  is relatively compact. The definition of compactness ensures that  $A$  is totally bounded.

2:  $A$  is closed as talked above. To see that  $A$  is bounded, consider an arbitrary point  $x_0 \in X$  and the open covering

$$\{B(x, r) : r > 0\}. \quad (11)$$

(11) is an open cover of  $A$ . Compactness of  $A$  means that there is a finite subcover of (11), which ensures that  $A$  is bounded.

3: Since  $\overline{A} = A$ .

4:  $A$  is totally bounded if and only if for all  $(x_n)_{n \in \mathbb{N}} \subseteq A$ ,  $(x_n)_{n \in \mathbb{N}}$  has a Cauchy subsequence i.e. a convergent subsequence. Therefore,  $A$  is totally bounded if and only if  $\overline{A}$  is sequentially compact i.e.  $\overline{A}$  is compact.

5: Necessity follows from 1 and 6.3. Apply 4 for sufficiency.

6: Heine-Borel theorem [4, Chapter 5, Thm 14] implies this.  $\square$

**Remark.** The inverse proposition of 2 is **incorrect**. Consider  $(\mathbb{R}, d_1)$  where  $d_1$  is defined as

$$d_1 : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, (x, y) \mapsto |\phi(x) - \phi(y)|,$$

where

$$\phi : \mathbb{R} \rightarrow (-1, 1), x \mapsto \frac{x}{1 + |x|}.$$

Then  $(\mathbb{R}, d_1)$  has a closed and bounded subset that is not compact:  $\mathbb{R}$ , itself. But clearly  $(\mathbb{R}, d_1)$  has the same topology as the usual topological space  $\mathbb{R}$ . Therefore  $(\mathbb{R}, d_1)$  is not compact since the open covering  $\{(n, -n) : n \in \mathbb{N}\}$  has no finite subcover.

In fact,  $(\mathbb{R}, d_1) \cong \mathbb{R}$ . Here  $\cong$  means there is a homeomorphism. Thus “boundedness” is not topological invariant.

## 6.2 Week 6, Lecture 2

### Recall

Let  $X, Y$  be two linear normed spaces.

- $T: X \rightarrow Y$  is said to be bounded/continuous, if  $\exists C > 0$  such that  $\| \|_Y \circ T \leq C \| \|_X$  (i.e.  $\|T\| \leq C$ ).
- $X \cong Y$  means that  $X$  is isometric to  $Y$ , i.e.  $\exists T: X \rightarrow Y$  such that  $T$  is linear, surjective and satisfies  $\| \|_Y \circ T = \| \|_X$ .

**Definition** (Isomorphism).  $X$  is **isomorphic** to  $Y$ , if there is a linear surjection  $T$  and  $C_1, C_2 > 0$  such that

$$C_1 \| \|_X \leq \| \|_Y \leq C_2 \| \|_X,$$

and this  $T$  is called an **isomorphism** from  $X$  to  $Y$ .  $X$  is isomorphic to  $Y$  is denoted by  $X \simeq Y$ .

**Remark.** In the category  $\text{Vect}_{\mathbb{K}}$ , an isomorphism is a linear bijection and vice versa. In the category  $\text{Nor}$ :  $\text{Ob}(\text{Nor})$  are normed spaces and  $\text{Mor}(\text{Nor})$  are bounded linear maps. An isomorphism in  $\text{Nor}$  is a linear homeomorphism. In the category  $\text{Nor}_1$ :  $\text{Ob}(\text{Nor}_1)$  are normed spaces and  $\text{Mor}(\text{Nor}_1)$  are contraction operators. An isomorphism in  $\text{Nor}_1$  is an isometry. In this notes,  $X \cong Y$  means that  $X$  is isometric to  $Y$  and  $X \simeq Y$  means that  $X$  is isomorphic to  $Y$ .

### 6.2.1 Finite Dimensional Linear Normed Spaces

**Definition** (Equivalent norms). Let  $(X, \| \|_1), (X, \| \|_2) \in \text{Ob}(\text{Nor})$ . We say  $\| \|_1$  is equivalent to  $\| \|_2$ , if  $\exists a, b > 0$  such that

$$a \| \|_2 \leq \| \|_1 \leq b \| \|_2.$$

**Remark.**  $\sim$  is an equivalent relation between norms on  $X$ , as you should verify.

See the definition of **Isomorphism** and we get  $\| \|_1 \sim \| \|_2$  if and only if

$$\text{id}: (X, \| \|_2) \rightarrow (X, \| \|_1), x \mapsto x$$

is an isomorphism.

**Example 21.** Consider  $(\mathbb{R}^n, \| \|_2)$  and  $(\mathbb{R}^n, \| \|_\infty)$ . Clearly

$$\| \|_\infty \leq \| \|_2 \sqrt{n} \| \|_\infty,$$

and hence  $\| \|_2 \sim \| \|_\infty$ .

**Theorem 6.5** (Classification of Finite Dimensional Spaces). Let  $X \in \text{Ob}(\text{Nor})$  with  $\dim(X) = n < \infty$ , then  $X \simeq \mathbb{K}^n$ .

**Proof.** WLOG, suppose  $\mathbb{K}^n$  is equipped with  $\|\cdot\|_\infty$ . Consider

$$\varphi: \mathbb{K}^n \rightarrow X, (x_1, \dots, x_n) \mapsto \sum_{j=1}^n x_j \alpha_j,$$

where  $\{\alpha_1, \dots, \alpha_n\}$  is a base of  $X$ .  $\varphi$  is proved to be continuous because

$$\begin{aligned} \|\varphi(x_1, \dots, x_n)\|_X &\leq \sum_{j=1}^n |x_j| \|\alpha_j\| \\ &\leq \sum_{j=1}^n \|(x_1, \dots, x_n)\|_\infty \|\alpha_j\| \\ &= \left( \sum_{j=1}^n \|\alpha_j\| \right) \|(x_1, \dots, x_n)\|_\infty. \end{aligned}$$

Then let

$$\Phi: \mathbb{K}^n \rightarrow \mathbb{R}, (x_1, \dots, x_n) \mapsto \|\varphi(x_1, \dots, x_n)\|_X.$$

Now  $\Phi = \|\cdot\|_X \circ \varphi$  is continuous. Hence  $\Phi$  obtains a minimal value on  $S = \{x \in \mathbb{K}^n : \|x\|_\infty = 1\}$ . Suppose  $\delta = \min \Phi|_S$  (such  $\delta$  exists, since  $S$  is compact). Then  $\delta > 0$  since  $\|\cdot\|_X$  is a norm and  $0 \notin S$ . Now we have  $\forall 0 \neq (x_1, \dots, x_n) \in \mathbb{K}^n$ ,

$$\begin{aligned} \Phi(x_1, \dots, x_n) &= \|(x_1, \dots, x_n)\|_\infty \Phi\left(\frac{(x_1, \dots, x_n)}{\|(x_1, \dots, x_n)\|_\infty}\right) \\ &\geq \delta \|(x_1, \dots, x_n)\|_\infty, \end{aligned}$$

i.e.

$$\|\varphi(x_1, \dots, x_n)\|_X \geq \delta \|(x_1, \dots, x_n)\|_\infty. \quad (12)$$

(12) holds for  $\forall (x_1, \dots, x_n) \in \mathbb{K}^n$  and means that  $\varphi^{-1}$  is continuous. Above all,  $\varphi$  is a linear homeomorphism, i.e. an isomorphism.  $\square$

**Remark.** Consider  $\min \Phi|_S$  is natural, just like

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

**Corollary 6.6.** Let  $(X, \|\cdot\|) \in \text{Ob}(\text{Nor})$ .

1)  $\dim X = n$  implies that  $X$  is complete.

- 2)  $X$  is an arbitrary linear normed space and  $X_0 \hookrightarrow X$  such that  $\dim(X_0) < \infty$ . Then  $X_0$  is closed.
- 3)  $\dim(X) < \infty$  implies that  $\mathcal{L}(X) = \mathcal{B}(X)$ .

Theorem 6.5 implies that: if  $\dim(X) < \infty$ , then  $A \subseteq X$  is compact if and only if  $A$  is closed and bounded. But it is not true for some (all, in fact, see Theorem 6.8) infinite dimensional normed spaces.

**Example 22** (A closed bounded set that is not compact). Consider  $\ell_2$  and its base  $\{e_n : n \in \mathbb{N}\}$ , where

$$e_n := (\underbrace{0, \dots, 0}_{n-1 \text{ terms}}, 1, 0, \dots), \forall n \in \mathbb{N}.$$

**Proof.**  $B := \{e_n : n \in \mathbb{N}\}$  is what we want.

- $B$  is closed: consider an arbitrary convergent sequence  $(x_n)_{n \in \mathbb{N}} \subseteq B$ , then there is some  $m \in \mathbb{N}$  such that  $x_n = e_m$  for all but finite many  $n \in \mathbb{N}$ , because  $\|e_m - e_n\| = \sqrt{2}\delta_n^m$ . Thus  $(x_n)_{n \in \mathbb{N}} \rightarrow e_m \in B$ .
- $B$  is bounded: since  $\text{diam}(B) = \sqrt{2}$ .
- $B$  is not compact: since  $(e_n)_{n \in \mathbb{N}} \subseteq B$  is a sequence having no convergent subsequence. Thus  $B$  is not sequentially compact and hence not compact.  $\square$

**Lemma 6.7** (Riesz). Let  $X$  be a linear normed space and  $X_0 \hookrightarrow X$ ,  $X_0 \neq X$  is a closed subspace. Then

$$\forall \varepsilon \in (0, 1) \exists x_\varepsilon \in X (\|x_\varepsilon\| = 1 \wedge d(x_\varepsilon, X_0) > \varepsilon.)$$

**Proof.** Taking arbitrary  $x' \in X \setminus X_0$ , then  $d(x', X_0) > 0$ . Let  $d = d(x', X_0)$ , now  $d/\varepsilon > d$  and hence

$$\exists \bar{x} \in X_0 \|\bar{x} - x'\| < d/\varepsilon.$$

Taking  $x_\varepsilon := \frac{\bar{x} - x'}{\|\bar{x} - x'\|}$ , then  $\|x_\varepsilon\| = 1$  and  $\forall x \in X_0$

$$\begin{aligned} \|x_\varepsilon - x\| &= \left\| \frac{\bar{x} - x' - \|\bar{x} - x'\|x}{\|\bar{x} - x'\|} \right\| \\ &= \frac{1}{\|\bar{x} - x'\|} \|\bar{x} - x' - \|\bar{x} - x'\|x\| \\ &> \varepsilon. \end{aligned}$$

The last inequality comes from  $\|\bar{x} - x'\| < d/\varepsilon$  and  $\|y - x'\| \geq d$ , where  $y = \bar{x} - \|\bar{x} - x'\|x \in X_0$ .  $\square$

**Theorem 6.8.** Let  $X$  be a linear normed space and  $\overline{B(0,1)}$  is its closed unit ball. The following statements are equivalent:

1.  $X$  is finite dimensional.
2.  $\partial B(0,1)$  is compact.
3.  $\overline{B(0,1)}$  is compact.
4.  $\forall A \subseteq X$ ,  $A$  is closed and bounded if and only if  $A$  is compact.

**Proof.** We want to show that

$$\begin{array}{ccc} 1 & \Longleftarrow & 2 \\ \Downarrow & & \Uparrow \\ 4 & \Longrightarrow & 3 \end{array}$$

We get  $1 \implies 4$  from Theorem 6.5,  $4 \implies 3$  is trivial and  $3 \implies 2$  since a closed subset of a compact set is compact.

It suffices to prove that  $2 \implies 1$ . Consider proof by contradiction. Suppose  $\dim(X) = \infty$ . Let  $\forall x_1 \in X$  such that  $x_1 \neq 0$ . Consider the closed linear subspace  $\text{span}\{x_1\}$  (this is a closed linear subspace, see the third corollary of Theorem 6.5). From Lemma 6.7, there is  $x_2 \in X \setminus \text{span}\{x_1\}$  such that  $\|x_2\| = 1$  and  $d(x_2, \text{span}\{x_1\}) > 1/2$ . Then consider the closed linear subspace  $\text{span}\{x_1, x_2\}$  (that is closed by the same reason as  $\text{span}\{x_1\}$ ),  $\text{span}\{x_1, x_2\} \neq X$  and Lemma 6.7 implies that there is  $x_3 \in X \setminus \text{span}\{x_1, x_2\}$  such that  $\|x_3\| = 1$  and  $d(x_3, \text{span}\{x_1, x_2\}) > 1/2$ . Thus, We can define a sequence  $(x_n)_{n \in \mathbb{N}} \subseteq \partial B(0,1)$  inductively such that

$$\forall m \neq n, d(x_m, x_n) > 1/2.$$

Therefore,  $\partial B(0,1)$  is not sequentially compact and hence not compact.

Above all,  $X$  is infinite dimensional implies that  $\partial B(0,1)$  is not compact. Thus  $2 \implies 1$ .  $\square$

## Summary

We have proved that

1.  $\dim X < \infty \implies X \simeq \mathbb{K}^n$  and hence:
  - (a) The space  $X$  is complete;
  - (b) Every finite dimensional subspace of an arbitrary linear normed space is closed;

- (c) Two spaces coincide:  $\mathcal{L}(X) = \mathcal{B}(X)$ .
2. [Riesz's Lemma](#)  $\implies$  [Theorem 6.8](#) which gives equivalent descriptions of finite dimensions.

## 7 Week 7

### 7.1 Week 7, Lecture 1

#### 7.1.1 Construct more Linear Normed Spaces

Let  $(X_i, \|\cdot\|_{X_i}) \in \text{Ob}(\text{Nor}), 1 \leq i \leq n$ . Define

$$\times_{i=1}^n X_i := \prod_{i=1}^n X_i$$

with operations

$$k(x_1, \dots, x_n) + l(y_1, \dots, y_n) = (kx_1 + ly_1, \dots, kx_n + ly_n).$$

$\forall p \in [1, \infty]$ , define a norm on  $X = \times_{i=1}^n X_i$

$$\|\cdot\|_X : X \rightarrow \mathbb{R}, (x_1, \dots, x_n) \mapsto \left( \sum_{i=1}^n \|x_i\|_{X_i}^p \right)^{1/p}.$$

At the case of  $p = \infty$ ,  $\|x\|$  should be interpreted like  $\|\cdot\|_\infty$ . To see that  $\|\cdot\|$  is a norm, it suffices to show that

1. it's positive definite;
2. it's homogeneous;
3. triangle inequality holds. And this follows from the [Minkowski's Inequality](#) for  $\ell_p$ , since

$$\begin{aligned} \|x + y\| &= \left( \sum_{i=1}^n \|x_i + y_i\|^p \right)^{1/p} \\ &\leq \left( \sum_{i=1}^n (\|x_i\| + \|y_i\|)^p \right)^{1/p} \\ &\leq \left( \sum_{i=1}^n \|x_i\|^p \right)^{1/p} + \left( \sum_{i=1}^n \|y_i\|^p \right)^{1/p}. \end{aligned}$$

The first inequality comes from the triangle inequality of  $\|\cdot\|_X$  and the second inequality comes from the [Minkowski's Inequality](#):

$$\begin{aligned} &\|(\|x_1\|, \dots, \|x_n\|, 0, \dots)\|_{\ell_p} + \|(\|y_1\|, \dots, \|y_n\|, 0, \dots)\|_{\ell_p} \\ &\geq \|(\|x_1\| + \|y_1\|, \dots, \|x_n\| + \|y_n\|, 0, \dots)\|_{\ell_p}. \end{aligned}$$

Now, we have some questions



**Question 7.1.**

- 1) Let  $p_i: X \rightarrow X_i, (x_1, \dots, x_n) \mapsto x_i$  be the projection to the  $i$ -coordinate. Show that  $p_i$  is continuous and  $\|p_i\| = 1$ .
- 2)  $X = \times_{i=1}^n X_i$  is complete  $\iff$  all  $X_i$  is complete.

**Proof.**

- 1) On the one hand:  $\forall x = (x_1, \dots, x_n) \in X$ , we have  $\|p_i(x)\| = \|x_i\|$  and

$$\|x\|_X = \left( \sum_{i=1}^n \|x_i\|_{X_i}^p \right)^{1/p} \geq \|x_i\|_{X_i}.$$

Thus  $\|p_i\| \leq 1$ . On the other hand: taking  $x = (0, \dots, x_i, \dots, 0) \in X$  with  $x_i \neq 0 \in X_i$ , we get  $\|p_i(x)\|_{X_i} = \|x_i\|_{X_i} = \|x\|_X$  which implies  $\|p_i\| \geq 1$ .

- 2) **Sufficiency:** taking an arbitrary Cauchy sequence

$$(x_m)_{m \in \mathbb{N}} = ((x_m^{(1)}, \dots, x_m^{(n)}))_{m \in \mathbb{N}}$$

in  $X$ . Then

$$\max_{1 \leq i \leq n} \|x_p^{(i)} - x_q^{(i)}\|_{X_i} \leq \|x_p - x_q\| \rightarrow 0 (p, q \rightarrow \infty),$$

which means that  $(x_m^{(i)})_{m \in \mathbb{N}}$  is Cauchy in  $X_i$  and hence converges to some  $y^{(i)} \in X_i$ . Then

$$\lim_m x_m = (y^{(1)}, \dots, y^{(n)}) =: y \in X$$

Because

$$\lim_m \|x_m - y\|_X = \lim_m \left( \sum_{i=1}^n \|x_m^{(i)} - y^{(i)}\|_{X_i}^p \right)^{1/p} = 0.$$

Therefore  $X$  is complete.

Necessity:  $\forall 1 \leq i \leq n$ ,  $X_i$  is isometric to  $E_i \hookrightarrow X$ , where  $E_i := \{(0, \dots, x_i, 0, \dots) : x_i \in X_i\}$ . The isometry is

$$\iota_i: X_i \rightarrow E_i, x \mapsto (0, \dots, x, 0, \dots).$$

$E_i$  is closed, since

$$E_i = \bigcap_{j \neq i} p_j^{-1}(0)$$

is a finite intersection of closed sets. Thus  $E_i$  is complete since  $X$  is complete. And now,  $X_i$  is isometric to a Banach space. We're done.  $\square$

### 7.1.2 Unbounded Linear Functional

This proposition gives a description of unbounded linear functional.

**Lemma 7.1.** Let  $X \in \text{Ob}(\text{Nor})$  and  $f \in \mathcal{L}(X, \mathbb{K})$  is a unbounded linear functional. Then

$$f(B(0, r)) = \mathbb{K}, \forall r > 0.$$

**Proof.** Given an arbitrary  $\alpha \in \mathbb{K}$ ,  $\alpha \neq 0$  there is  $x' \in B(0, r)$  such that  $|f(x')| \geq |\alpha|$  (else,  $f$  maps a bounded set to a bounded set and hence  $f$  is bounded). Taking  $x = \frac{\alpha}{f(x')}x'$ , we're done since

$$f(x) = f\left(\frac{\alpha}{f(x')}x'\right) = \frac{\alpha}{f(x')}f(x') = \alpha$$

and  $x \in B(0, r)$  since

$$\|x\| = \frac{|\alpha|}{|f(x')|}\|x'\| \leq \|x'\| < r. \quad \square$$

By lemma 7.1, we have

**Proposition 7.2.** Suppose  $f \in X^*$  and  $f \neq 0$ . The following statements are equivalent:

- 1)  $f$  is continuous;
- 2)  $\ker f$  is closed.

**Proof.**

- 1)  $\implies$  2)  $\{0\}$  is closed in  $\mathbb{K}$  and 2) follows from the topological definition of continuity.
- 2)  $\implies$  1)  $\ker f$  is closed and hence is not dense in  $X$  since  $f \neq 0$ . Therefore

$$\exists x_0 \in X \exists r > 0 (B(x_0, r) \cap \ker f = \emptyset.) \quad (13)$$

You can check (13) by denying the proposition “ $\ker f$  is dense in  $X$ ”. If  $f$  is not continuous, then Lemma 7.1 ensures that  $f(B(0, r)) = \mathbb{K}$ . Thus,  $\exists y \in B(0, r)$  such that  $f(y) = -f(x_0)$ . And now

$$f(y + x_0) = f(y) + f(x_0) = 0,$$

i.e.  $y + x_0 \in \ker f$  while  $y + x_0 \in B(x_0, r)$ . This is a contradiction since  $B(x_0, r) \cap \ker f = \emptyset$ .

□

**Exercise 7.1.** Determine which of the following sets are closed

1)  $M := \{x \in \ell_2 : \sum_{n \geq 1} x_n / \sqrt{n} = 0\};$

2)  $M := \{x \in \ell_2 : \sum_{n \geq 1} x_n / n = 0\}.$

**Solution.** In fact, **2)** is simpler.

- 1) We will not prove this, because  $f: \ell_2 \rightarrow \mathbb{K}$  is not well-defined. There is an element in  $\ell_2$ :

$$x: \mathbb{N} \rightarrow \mathbb{K}, n \mapsto \begin{cases} 0 & n = 1 \\ \frac{1}{\sqrt{n} \log n} & n \geq 2 \end{cases}$$

such that  $f(x) \notin \mathbb{K}$ .

The set in **1)** can be proved to be not closed by the theory of **Hilbert Space**. See [this post](#).

- 2) Let

$$f: \ell_2 \rightarrow \mathbb{K}, x \mapsto \sum_{n=1}^{\infty} \frac{1}{n} x_n. \quad (14)$$

Clearly  $f$  is well-defined. Furthermore,  $\forall x \in \ell_2$ , we have

$$|f(x)| \leq \sum_{n \geq 1} \frac{1}{n} |x_n| \leq \left( \sum_{n \geq 1} 1/n^2 \right)^{1/2} \|x\|_2$$

And hence  $M = f^{-1}(0)$  is closed. □

**Remark.** We have  $\|f\| = \pi/\sqrt{6}$  by taking

$$\ell_2 \ni x = (1, 1/2, \dots, 1/n, \dots),$$

since  $\|x\|_{\ell_2} = \pi/\sqrt{6}$ .

**Remark.** In fact,

$$\ell_p^* \cong \ell_q,$$

where  $p \in [1, \infty)$  and  $q = p/(p-1)$ .

## 7.2 Week 7, Lecture 2

### 7.2.1 Theorems about Banach Spaces

Here are some topics of this lecture:

1. Open mapping theorem, see Theorem 7.3;
2. Banach-Steinhaus Theorem, see Theorem 7.6;
3. Hahn-Banach Theorem, see Theorem 9.1 and Theorem 9.2.

To state Theorem 7.3 better, we need a topological notion:

**Definition** (Open mapping). Let  $(X, \mathcal{T}), (Y, \mathcal{T})$  be two topological spaces and  $f: X \rightarrow Y$  be an arbitrary map (not continuous possibly).  $f$  is said to be an open mapping, if  $\forall O \in \mathcal{T}_X, f(O) \in \mathcal{T}_Y$ .

And then we have

**Theorem 7.3** (Open mapping theorem). Let  $X, Y$  be two Banach spaces and  $T \in \mathcal{B}(X, Y)$ . If  $T$  is surjective then  $T$  is open.

Proof of Theorem 7.3 is delayed to next (maybe) course.

**Theorem 7.4** (Boundedness of inverse mapping). Let  $X, Y$  be two Banach spaces and  $T \in \mathcal{B}(X, Y)$ . If  $T$  is bijective, then  $T^{-1} \in \mathcal{B}(Y, X)$ .

**Proof.** Theorem 7.3 implies that  $T$  is an open mapping and equivalently  $T^{-1}$  is continuous.  $\square$

Theorem 7.4 implies

**Corollary 7.5.** Let  $(X, \|\cdot\|_1), (X, \|\cdot\|_2)$  be two Banach spaces. If  $\exists C > 0$  such that  $\|\cdot\|_1 \leq C\|\cdot\|_2$ , then  $\|\cdot\|_1 \sim \|\cdot\|_2$ .

**Proof.** Consider  $\text{id}_X: (X, \|\cdot\|_2) \rightarrow (X, \|\cdot\|_1)$ .  $\|\cdot\|_1 \leq C\|\cdot\|_2$  implies that  $\text{id}_X$  is continuous. Apply Theorem 7.4 to  $\text{id}_X$  and we get that  $\text{id}_X^{-1}$  is bounded.  $\square$

**Theorem 7.6** (Banach-Steinhaus). Let  $(X, \|\cdot\|_X)$  be a Banach space,  $(Y, \|\cdot\|_Y) \in \text{Ob}(\text{Nor})$  and  $\{T_\lambda\}_{\lambda \in \Gamma} \subseteq \mathcal{B}(X, Y)$ . If

$$\forall x \in X, \sup_{\lambda \in \Gamma} \|T_\lambda x\|_Y < \infty,$$

then  $\sup_{\lambda \in \Gamma} \|T_\lambda\| < \infty$ .

The other name of this theorem is “the uniform boundedness principle”.

**Proof.** Here is a proof using the [Corollary](#) above.

Let  $\| \cdot \|_I$  be a new norm on  $X$ , defined as

$$\| \cdot \|_I : X \rightarrow \mathbb{R}_{\geq 0}, x \mapsto \|x\|_X + \sup_{\lambda \in \Gamma} \|T_\lambda x\|_Y.$$

It's easy to verify that  $\| \cdot \|_I$  is actually a norm. Clearly  $\text{id}_X : (X, \| \cdot \|_I) \rightarrow (X, \| \cdot \|_X)$  is continuous. If  $(X, \| \cdot \|_I)$  is a Banach space, then [Corollary](#) can be applied and we're done. Now, taking an arbitrary Cauchy sequence  $(x_n)_{n \in \mathbb{N}} \subseteq (X, \| \cdot \|_I)$ , i.e.

$$\lim_{m,n} \|x_n - x_m\|_I = 0. \quad (15)$$

And (15) is equivalent to

$$\lim_{m,n} \|x_n - x_m\|_X = 0, \quad (16)$$

$$\lim_{m,n} \sup_{\lambda \in \Gamma} \|T_\lambda x_n - T_\lambda x_m\|_Y = 0. \quad (17)$$

Since  $(X, \| \cdot \|_X)$  is a Banach space, (16) implies that  $(x_n)_{n \in \mathbb{N}} \xrightarrow{\| \cdot \|_X} x \in X$ . Now we prove that  $(x_n)_{n \in \mathbb{N}} \xrightarrow{\| \cdot \|_I} x \in X$  and it suffices to show that  $\lim_n \sup_{\lambda \in \Gamma} \|T_\lambda x_n - T_\lambda x\|_Y = 0$ . And this proof is similar to [the proof](#) of the completeness of  $C[a, b]$ .

To see this, from the definition of limit of double indexed sequence:

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} : \forall m, n > N, \sup_{\lambda \in \Gamma} \|T_\lambda x_n - T_\lambda x_m\|_Y < \varepsilon.$$

The definition of sup implies that

$$\forall m, n > N, \|T_\lambda x_n - T_\lambda x_m\|_Y < \varepsilon (\forall \lambda \in \Gamma).$$

Let  $m \rightarrow \infty$ , the continuity of  $\| \cdot \|_Y$  and  $T_\lambda$  (for each  $\lambda \in \Gamma$ ) implying that

$$\forall n > N, \|T_\lambda x_n - T_\lambda x\|_Y \leq \varepsilon (\forall \lambda \in \Gamma).$$

Therefore,

$$\forall n > N, \sup_{\lambda \in \Gamma} \|T_\lambda x_n - T_\lambda x\|_Y \leq \varepsilon.$$

Equivalently,

$$\lim_n \sup_{\lambda \in \Gamma} \|T_\lambda x_n - T_\lambda x\|_Y = 0,$$

which was what we wanted. □

### 7.2.2 Baire Category Theorem

**Definition** ( $G_\delta$ -set,  $F_\sigma$ -set). Let  $(X, \mathcal{T}) \in \text{Ob}(\text{Top})$ .

- A set of the form  $\bigcap_{n=1}^{\infty} G_n$  is called a  $G_\delta$ -set, where  $(G_n)_{n \in \mathbb{N}} \subseteq \mathcal{T}$ .
- A set of the form  $\bigcup_{n=1}^{\infty} F_n$  is called a  $F_\sigma$ -set, where  $(F_n)_{n \in \mathbb{N}} \subseteq \mathcal{T}$ .

**Remark.** Here “G” is German (Gebiet) and “F” is French (Fermé).

**Definition** (First Category Set). Let  $(X, \mathcal{T}) \in \text{Ob}(\text{Top})$ .  $A \subseteq X$  is called a **set of the first category**, if  $A \subseteq B$  for some  $F_\sigma$  set  $B$  with  $\overset{\circ}{B} = \emptyset$ .

**Definition** (Second Category Set). Let  $(X, \mathcal{T}) \in \text{Ob}(\text{Top})$ .  $A \subseteq X$  is called a **set of the second category**, if  $A$  is not of the first category.

**Definition** (Baire Space). Let  $(X, \mathcal{T}) \in \text{Ob}(\text{Top})$ .  $(X, \mathcal{T})$  is called a **Baire space**, if each countable intersection of dense open sets is dense in  $X$ .

Here is an equivalent definition of Baire space

**Definition** (Baire space). Let  $(X, \mathcal{T}) \in \text{Ob}(\text{Top})$ .  $(X, \mathcal{T})$  is called a **Baire space**, if each countable union of closed sets with empty interior has empty interior.

And now we can talk about Baire category Theorem.

**Theorem 7.7** (Baire category Theorem). If  $(X, \mathcal{T})$  is a topological space whose topology  $\mathcal{T}$  can be induced by a complete metric, then  $X$  is a Baire space.

**Proof.** Suppose  $(X, d)$  is the metric space whose topology induced by  $d$  is  $\mathcal{T}$ . Let  $(O_n)_{n \in \mathbb{N}}$  be a sequence of dense open sets in  $(X, d)$ . It suffices to show that  $O = \bigcap_{n \geq 1} O_n$  is dense in  $X$ . Taking an arbitrary open set  $\emptyset \neq U \in \mathcal{T}$ , now we show that  $O \cap U \neq \emptyset$ .

Since  $O_1$  is dense in  $X$ , we have  $O_1 \cap U \neq \emptyset$  and thus  $\exists x_1 \in O_1 \cap U$ . Moreover,  $\exists r > 0$  such that  $B(x_1, r) \subseteq O_1 \cap U$  since  $O_1 \cap U \in \mathcal{T}$ . Let  $F_1 := \overline{B(x_1, r/2)}$ . Then  $F_1 \neq \emptyset$ ,  $F_1 \subseteq O_1 \cap U$  and  $\text{diam } F_1 = r =: r_1$ .

Since  $O_2$  is dense in  $X$  and  $F_1 \neq \emptyset$ , we have  $O_2 \cap F_1 \neq \emptyset$  and thus  $\exists x_2 \in O_2 \cap F_1$ . Moreover,  $\exists r_2 > 0 \wedge r_2 < r_1/2$  such that  $B(x_2, r_2) \subseteq O_2 \cap F_1$ . Let  $F_2 := \overline{B(x_2, r_2)}$ . Then  $F_2 \neq \emptyset$ ,  $F_2 \subseteq O_2 \cap U$  and  $\text{diam } F_2 < r_1/2$ .

Analogically, we can define a sequence of decreasing closed sets  $(F_n)_{n \in \mathbb{N}}$  such that  $F_n \neq \emptyset (\forall n \in \mathbb{N})$ ,  $F_n \subseteq O_n \cap U (\forall n \in \mathbb{N})$  and

$\text{diam } F_n \leq 2^{1-n}r (\forall n \in \mathbb{N})$ . Then the third question of [Week 4, Lecture 1](#) implies that  $\exists! x_0 \in X$  such that

$$\{x\} = \bigcap_{n \geq 1} F_n.$$

Therefore,

$$O \cap U = \left( \bigcap_{n \geq 1} O_n \right) \cap U = \bigcap_{n \geq 1} (O_n \cap U) \supseteq \bigcap_{n \geq 1} F_n = \{x\},$$

and hence  $O \cap U \neq \emptyset$ . Since  $U$  is arbitrary,  $O$  is dense in  $X$ . □

Here is some results about Baire space:

**Theorem 7.8.** Let  $X$  be a Baire space. Then

- 1) Each open subset of  $X$  with the subspace topology is a Baire sapce;
- 2) Suppose  $(F_n)_{n \in \mathbb{N}}$  is a sequence of closed subsets of  $X$  with  $X = \bigcup_{n \geq 1} F_n$ , then  $\bigcup_{n \geq 1} \overset{\circ}{F}_n$  is dense in  $X$ .

**Proof.**

- 1) For  $A \subseteq X$ , let  $\text{cl}_A$  means the closure operator with respect to the subspace topology of  $A$ . Similarly,  $\text{int}$  means the interior operator.

Suppose  $\Omega \subseteq X$  is open. Given  $(O_n)_{n \in \mathbb{N}} \subseteq \Omega$  such that  $\text{cl}_\Omega(O_n) = \Omega (\forall n \in \mathbb{N})$ , i.e.  $O_n$  is dense in  $\Omega$  for all  $n \in \mathbb{N}$ . Since  $\text{cl}_\Omega(O_n) = \Omega \cap \overline{O_n}$ , we have  $\overline{O_n} \supseteq \Omega$  and hence  $\overline{O_n} \supseteq \overline{\Omega}$ . Since the closure of union is the union of closure, we have  $O_n \cup (\overline{\Omega})^c$  is dense in  $X$  for all  $n \in \mathbb{N}$ . Therefore,

$$(U_n)_{n \in \mathbb{N}} := \left( O_n \cup (\overline{\Omega})^c \right)_{n \in \mathbb{N}}$$

is a sequence of dense open sets in  $X$ . Now  $X$  is a Baire space, which means  $\bigcap_{n \geq 1} U_n$  is dense in  $X$ . While

$$\bigcap_{n \geq 1} U_n = \left( \bigcap_{n \geq 1} O_n \right) \cup (\overline{\Omega})^c,$$

and hence

$$\overline{\bigcap_{n \geq 1} U_n} = X = \overline{\bigcap_{n \geq 1} O_n \cup (\overline{\Omega})^c}.$$

To prove that  $\text{cl}_\Omega(\bigcap_{n \geq 1} O_n) = \Omega$ , we want to show that  $\overline{(\Omega)^c}^c \subseteq \Omega^c$ . And this holds

$$(\overline{\Omega})^c \subseteq \Omega^c \implies \overline{(\Omega)^c}^c \subseteq \Omega^c.$$

since  $\overline{(\Omega)^c}^c$  is the smallest closed set containing  $(\overline{\Omega})^c$ .

Above all,  $\Omega$  is a Baire space.

- 2) Let  $\Omega \neq \emptyset$  be an arbitrary open set in  $X$ . Then 7.8 implies that  $\Omega$  is a Baire space. And

$$\Omega = \Omega \cap X = \bigcup_{n \geq 1} (\Omega \cap F_n),$$

the definition of **Baire space** ensures that there is some  $n \in \mathbb{N}$  such that  $\text{int}(\Omega \cap F_n) \neq \emptyset$ . Since “the interior of intersection is the intersection of union” and  $\Omega$  is open, we have  $\Omega \cap \overset{\circ}{F}_n \neq \emptyset$ . Therefore

$$\Omega \cap \left( \bigcup_{j \geq 1} \overset{\circ}{F}_j \right) \supseteq \Omega \cap \overset{\circ}{F}_n \neq \emptyset.$$

Then  $\bigcap_{n \geq 1} F_n$  is dense in  $X$  since  $\Omega$  is arbitrary. □

Now we give another proof of Theorem 7.6 by the **Baire category Theorem**.

**Proof of Theorem 7.6.** Let

$$M: X \rightarrow \mathbb{R}, x \mapsto \sup_{\lambda \in \Gamma} \|T_\lambda x\|_Y,$$

which is well-defined by the assumption. For all  $n \in \mathbb{N}$

$$F_n := M^{-1}[0, n] = \bigcap_{\lambda \in \Gamma} (\| \cdot \|_Y \circ T_\lambda)^{-1}[0, n],$$

and  $\| \cdot \|_Y, T_\lambda (\forall \lambda \in \Gamma)$  is continuous. Therefore,  $F_n$  is closed. Now  $X$  is a Banach space (hence a Baire space) and

$$X = \bigcup_{n \geq 1} F_n.$$

Theorem 7.8 shows that there is some  $k \in \mathbb{N}$  such that  $\overset{\circ}{M}_k \neq \emptyset$ . There is  $x_0 \in \overset{\circ}{F}_k$  and  $r > 0$  such that  $B_X(x_0, r) \subseteq \overset{\circ}{F}_k$ . Now  $\forall x \in B_X(x_0, r)$ ,  $x + x_0 \in B_X(x_0, r)$  and hence  $\forall \lambda \in \Gamma$ , we have

$$\|T_\lambda(x)\|_Y \leq \|T_\lambda(x + x_0)\|_Y + \|T_\lambda(x_0)\|_Y \leq k + M(x_0).$$



Thus  $T(B_X(x_0, r)) \subseteq \overline{B_Y(0, k + M(x_0))}$  holds for all  $\lambda \in \Gamma$ , which implies

$$\|T_\lambda\| \leq \frac{k + M(x_0)}{r}, \forall \lambda \in \Gamma.$$

Above all,  $\sup_{\lambda \in \Gamma} \|T_\lambda\| \leq (k + M(x_0))/r < \infty$ . □

**Remark.** To give  $X = \bigcup_{n \geq 1} F_n$ , we need the assumption

$$\forall x \in X, M(x) < \infty.$$

## 8 Week 8

### 8.1 Week 8, Lecture 1

#### Recall

We have proved Theorem 7.8 by [Open mapping theorem](#). We used the corollary 7.5 and proved that  $\text{cod}(\text{id}_X)$  is a Banach space, where

$$\text{id}_X: (X, \| \cdot \|_X) \rightarrow (X, \| \cdot \|_X + \sup_{\lambda \in \Gamma} \| \cdot \|_Y \circ T_\lambda)$$

is continuous.

Moreover, we proved [Baire category Theorem](#) and applied it to prove [Banach-Steinhaus Theorem](#).

#### 8.1.1 Application of [Banach-Steinhaus Theorem](#)

**Definition** (Strong convergence). Let  $X, Y$  be two normed spaces,  $(T_n)_{n \in \mathbb{N}} \subseteq \mathcal{B}(X, Y)$  and  $T \in \mathcal{B}(X, Y)$ . We say that  $(T_n)_{n \in \mathbb{N}}$  converges to  $T$  strongly, if

$$\forall x \in X, (T_n x)_{n \in \mathbb{N}} \xrightarrow{\| \cdot \|_Y} T x,$$

denoted as  $(T_n)_{n \in \mathbb{N}} \xrightarrow{s} T$ .

**Remark.** The relation between  $(T_n)_{n \in \mathbb{N}} \xrightarrow{s} T$  and  $(T_n)_{n \in \mathbb{N}} \xrightarrow{\| \cdot \|_{\mathcal{B}(X, Y)}} T$  is similar to the pointwise convergence and uniform convergence of function sequence.

We use [Banach-Steinhaus](#) to prove the following theorem about strong convergence.

**Theorem 8.1.** Let  $X$  be a linear normed space,  $Y$  be a Banach space, and  $(T_n)_{n \in \mathbb{N}} \subseteq \mathcal{B}(X, Y)$  is a sequence of operators. Suppose

1.  $\sup_{n \in \mathbb{N}} \|T_n\| < \infty$ ;
2.  $\exists G \subseteq X$  such that  $\overline{G} = X$  and  $\forall x \in G, (T_n x)_{n \in \mathbb{N}}$  converges in  $Y$ .

Then there is a  $T \in \mathcal{B}(X, Y)$  with  $\|T\| \leq \liminf_{n \rightarrow \infty} \|T_n\|$  such that

$$T_n \xrightarrow{s} T (n \rightarrow \infty).$$

**Proof.** Let  $M := \sup_{n \in \mathbb{N}} \|T_n\|$ . Since  $G$  is dense in  $X$ ,  $\forall x \in X$  and  $\forall \varepsilon > 0$ , there is  $y \in G$  such that  $\|y - x\| < \varepsilon$ . Then

$$\begin{aligned} \|T_n x - T_m x\| &\leq \|T_n x - T_n y\| + \|T_n y - T_m y\| + \|T_m y - T_m x\| \\ &\leq \|T_n\| \|x - y\| + \|T_n y - T_m y\| + \|T_m\| \|y - x\| \\ &\leq 2M\varepsilon + \|T_n y - T_m y\|. \end{aligned}$$

Let  $m, n \rightarrow \infty$  and use the strong convergence of  $(T_n)_{n \in \mathbb{N}}$ , we find

$$\limsup_{m, n} \|T_n x - T_m x\| \leq 2M\varepsilon.$$

From arbitrariness of  $\varepsilon > 0$ , we get  $\lim_{m, n} \|T_n x - T_m x\| = 0$ , i.e.  $(T_n x)_{n \in \mathbb{N}}$  is a Cauchy sequence in  $Y$  and hence converges to some point in  $Y$ , because  $Y$  is a Banach space. Therefore, we can define

$$T: X \rightarrow Y, x \mapsto \lim_n T_n x,$$

which is linear since both  $T_n (\forall n)$  and  $\lim_n$  is linear (i.e.  $T$  is a composition of two linear maps,  $f_1: X \rightarrow \mathcal{E}, x \mapsto (T_n x)_{n \in \mathbb{N}}$  and  $f_2: \mathcal{E} \rightarrow Y, (y_n)_{n \in \mathbb{N}} \mapsto \lim_n y_n$ , where  $\mathcal{E}$  is the set of all Cauchy sequences in  $Y$ ). Then  $T = f_2 \circ f_1$  is linear.

Now we show that,  $T$  is what we want. For all  $x \in X$ ,

$$\|Tx\| = \left\| \lim_n T_n x \right\| = \lim_n \|T_n x\|,$$

since  $\|\cdot\|$  is continuous. And  $\forall n \in \mathbb{N}, \|T_n x\| \leq \|T_n\| \|x\|$ , take  $\liminf$  on both sides and we get

$$\liminf \|T_n x\| \leq \liminf \|T_n\| \|x\|. \quad (18)$$

And  $(\|T_n x\|)_{n \in \mathbb{N}}$  is a Cauchy sequence in  $\mathbb{R}$ , thus  $\lim \|T_n x\|$  exists in  $\mathbb{R}$ . Then 18 implies

$$\|Tx\| = \lim \|T_n x\| \leq \liminf \|T_n\| \|x\| \quad (\forall x \in X)$$

so

$$\|T\| \leq \liminf_n \|T_n\|,$$

which ensures that  $T \in \mathcal{B}(X, Y)$ . □

If  $X$  is also a Banach space, then the inverse proposition holds. That is:

**Proposition 8.2.** Let  $X, Y$  be two Banach spaces. Suppose there is some  $T \in \mathcal{B}(X, Y)$  such that  $(T_n)_{n \in \mathbb{N}} \xrightarrow{s} T$ , then

1.  $\sup_{n \in \mathbb{N}} \|T_n\| < \infty$ ;
2.  $\exists G \subseteq X$  such that  $G$  is dense in  $X$  and  $\forall x \in G, (T_n x)_{n \in \mathbb{N}}$  converges in  $Y$ .

**Proof.** For all  $x \in X$ ,  $(\|T_n x\|)_{n \in \mathbb{N}}$  is a Cauchy sequence (and hence bounded) in  $\mathbb{R}$  by strong convergence. Then Theorem 7.8 implies that  $\sup_{n \in \mathbb{N}} \|T_n\|$  is finite. Let  $G = X$  then  $\overline{G} = X$  and  $\forall x \in G, (T_n x)_{n \in \mathbb{N}}$  converges in  $Y$ , since  $(T_n)_{n \in \mathbb{N}} \xrightarrow{s} T$ .  $\square$

To state the next theorem better, there is an essential exercise.

**Exercise 8.1.** If  $Y$  is a Banach space and  $X$  is a linear normed space, then  $\mathcal{B}(X, Y)$  is a Banach space. Especially,  $X^*$  is a Banach space.

Proof of this exercise is written in the Appendix B. Note that the exercise is just saying that  $\mathcal{B}(X, Y)$  is complete in the meaning of the metric induced by the norm, then you can see that the next theorem is just saying that  $\mathcal{B}(X, Y)$  is complete in the meaning for “strongly Cauchy sequence converges to some operator strongly”.

**Theorem 8.3.** If  $X, Y$  are Banach spaces, then  $\forall (T_n)_{n \in \mathbb{N}} \subseteq \mathcal{B}(X, Y)$  such that

$$T_n - T_m \xrightarrow{s} 0 (m, n \rightarrow \infty), \quad (19)$$

we have

$$T_n - T \xrightarrow{s} 0 (n \rightarrow \infty)$$

for some  $T \in \mathcal{B}(X, Y)$ .

**Proof.** Since (19) implies that  $\forall x \in X, (T_n x)_{n \in \mathbb{N}}$  is a Cauchy sequence in  $Y$ . Since  $Y$  is a Banach space,  $(T_n x)_{n \in \mathbb{N}}$  converges to some point in  $Y$ . Notice that  $\forall x \in X, (\|T_n x\|)_{n \in \mathbb{N}}$  is bounded in  $\mathbb{R}$ , thus Theorem 7.6 implies that  $\sup_{n \in \mathbb{N}} \|T_n\| < \infty$ , so let  $G = X$ , apply Theorem 8.1 and we’re done.  $\square$

## Inverse of Hölder’s inequality

We have learnt the Hölder’s inequality (especially, for the measure space  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ ):  $\forall p \in [1, \infty], \forall a \in \ell_p, b \in \ell_q$ ,

$$\|ab\|_1 \leq \|a\|_p \|b\|_q$$

holds, where  $q = p'$ . Now we’re going to show that if  $p \in (1, \infty)$ ,  $\forall x \in \ell_p$  we have  $\sum_{n \geq 1} \alpha_n x_n < \infty$ , then  $(\alpha_n)_{n \in \mathbb{N}} \in \ell_q$ , where  $q = p'$ . For convenience, consider  $\mathbb{K} = \mathbb{R}$ .

**Proof.** Let

$$(\forall k \in \mathbb{N}) f_k: \ell_p \rightarrow \mathbb{R}, (x_n)_{n \in \mathbb{N}} \mapsto \sum_{j=1}^k x_j \alpha_j.$$

Then  $f_k$  is linear, and bounded since Hölder's inequality implies

$$\begin{aligned} |f_k x| &\leq \sum_{j=1}^k |x_j \alpha_j| \\ &\leq \left( \sum_{j=1}^k |x_j|^p \right)^{1/p} \left( \sum_{j=1}^k |\alpha_j|^q \right)^{1/q} \\ &\leq \left( \sum_{j=1}^k |\alpha_j|^q \right)^{1/q} \|x\|_p, \end{aligned}$$

i.e.  $\|f_k\| \leq \left( \sum_{j=1}^k |\alpha_j|^q \right)^{1/q}$ . And the reversed inequality holds, to see this, consider the equality condition of Hölder's inequality (and triangle inequality) and hence pick

$$\ell_p \ni x^{(k)}, (x_n^{(k)})_{n \in \mathbb{N}} := (\text{sign}(\alpha_1) |\alpha_1|^{q/p}, \dots, \text{sign}(\alpha_k) |\alpha_k|^{q/p}, 0, \dots).$$

And

$$\left| f_k(x^{(k)}) \right| = \sum_{j=1}^k |\alpha_j|^q, \left\| x^{(k)} \right\|_p = \left( \sum_{j=1}^k |\alpha_j|^q \right)^{1/p},$$

implies

$$\|f_k\| \geq \left( \sum_{j=1}^k |\alpha_j|^q \right)^{1-1/p} = \left( \sum_{j=1}^k |\alpha_j|^q \right)^{1/q}.$$

Above all,  $\|f_k\| = \left( \sum_{j=1}^k |\alpha_j|^q \right)^{1/q}$ .

By assumption, we have  $\forall x \in \ell_p, (f_n(x))_{n \in \mathbb{N}}$  converges, and hence is bounded. Now apply Theorem 7.6, we get  $\sup_{n \in \mathbb{N}} \|f_n\| < \infty$ . While  $(\|f_n\|)_{n \in \mathbb{N}} \subseteq \mathbb{R}$  is a non decreasing sequence, thus

$$\sup_{n \in \mathbb{N}} \|f_n\| = \lim_{n \rightarrow \infty} \|f_n\| = \lim_{n \rightarrow \infty} \left( \sum_{j=1}^n |\alpha_j|^q \right)^{1/q} = \|\alpha\|_q.$$

Therefore,  $\|\alpha\|_q < \infty$ , i.e.  $\alpha \in \ell_q$ . □

**Remark.** We can drop the restriction  $\mathbb{K} = \mathbb{R}$ .

## Fourier series's divergence

First, we introduce some notions for convenience.

**Definition.** Let  $C_{2\pi}$  be the normed space whose underlying set is

$$\{f: \mathbb{R} \rightarrow \mathbb{C} \mid f \text{ is continuous and } 2\pi\text{-periodic}\},$$

with the norm

$$\| \cdot \|_{\infty}: C_{2\pi} \rightarrow \mathbb{R}, f \mapsto \sup_{x \in \mathbb{R}} |f(x)|.$$

**Remark.** The norm  $\max$  is well-defined since  $f$  is  $2\pi$ -periodic implies that

$$\sup_{\mathbb{R}} |f| = \sup_{[0, 2\pi]} |f| = \max_{[0, 2\pi]} |f|.$$

For clarity, here is the definition of period of a real function.

**Definition** (Period, Periodic function). Let  $f$  be a function  $\mathbb{R} \rightarrow \mathbb{R}$ . A number  $T \in \mathbb{R}$  is called a **period** of  $f$ , if

$$f = \tau_T f,$$

where

$$\tau_T f: \mathbb{R} \rightarrow \mathbb{C}, x \mapsto f(x - T).$$

Function  $f$  has a period  $T$  is called a  **$T$ -periodic function**, or a **periodic function** for short.

**Remark.** Let

$$\text{per}(f) := \{T \in \mathbb{R} : \tau_T f = f\},$$

then  $\text{per}(f)$  is a subgroup of the additive group  $\mathbb{R}$ . The structure of  $\text{per}(f)$  has only 3 possibilities:

1.  $\text{per}(f) = \{0\}$ , i.e.  $f$  is not a periodic function.
2.  $\text{per}(f) = T_0 \mathbb{Z} = \{T_0 k : k \in \mathbb{Z}\}$  for some  $T_0 > 0$ . And such  $T_0$  is usually called the fundamental period or the minimum period.
3.  $\text{per}(f)$  is a dense subgroup of  $\mathbb{R}$ , equivalently  $f$  has no fundamental period. For example,  $\text{per}(\chi_{\mathbb{Q}}) = \mathbb{Q}$ .

The Fourier series of a  $2\pi$ -periodic function is defined as follows

**Definition** (Fourier transform, Fourier coefficient). Given  $f \in C_{2\pi}$ , the **Fourier transform** of  $f$  is the sequence defined as follows

$$\widehat{f}: \mathbb{Z} \rightarrow \mathbb{C}, n \mapsto \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-in\theta} d\theta.$$

We use the notation

$$f(x) \sim \sum_{n \in \mathbb{Z}} a_n e^{inx}$$

to mean that  $a_n = \widehat{f}(n)$ ,  $\forall n \in \mathbb{Z}$ . The  $n$ -th term of  $\widehat{f}$ ,  $\widehat{f}(n)$  is called the  $n$ -th **Fourier coefficient** of  $f$ .

And define the partial sum of Fourier series

**Definition.** The  $n$ -th partial sum of  $f$ 's Fourier series, denoted by  $S_n(f)$  is

$$\sum_{k=-n}^n \widehat{f}(k) e^{-ikx} = \sum_{k=-n}^n \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{ik(\theta-x)} d\theta.$$

To understand the partial sum better, we have to notations: convolution and Dirichlet kernel.

**Definition** (Convolution). Given the measure space  $([-\pi, \pi], \mathcal{B}, \mu)$  and two measurable function  $f, g$ , define

$$f * g: [-\pi, \pi] \rightarrow \mathbb{C}, t \mapsto \begin{cases} \int_{-\pi}^{\pi} f(t-x)g(x) d\mu(x), & \text{if integral is finite} \\ 0. & \text{else} \end{cases}$$

Here  $\mu$  is an arbitrary Borel measure on  $(\mathbb{R}, \mathcal{B})$ .

Fubini's Theorem implies that, if  $f, g \in \mathcal{L}_1[-\pi, \pi]$ , then

$$\int_{[-\pi, \pi]} f(t-x)g(x) d\frac{m}{2\pi}(x) < \infty$$

for almost every  $t \in [-\pi, \pi]$ .

**Definition** (Dirichlet kernel). Given  $n \in \mathbb{N}$ , the function

$$D_n: \mathbb{R} \rightarrow \mathbb{C}, x \mapsto \sum_{k=-n}^n e^{ikx}$$

is called the  $n$ -th Dirichlet kernel.

By this definitions, we have a convolution formula for the  $n$ -th partial sum of Fourier series

$$S_n(f)(x) = \int_{-\pi}^{\pi} f(\theta) \sum_{k=-n}^n e^{ik(x-\theta)} d\frac{m}{2\pi}(x) = D_n * f(x),$$

where the convolution is taking integration with respect to the measure  $\frac{m}{2\pi}$ , i.e. Lebesgue measure multiplied by  $1/2\pi$ .

Now, we have the following result.

**Proposition 8.4.** The set  $\{f \in C_{2\pi} : \sup_{n \in \mathbb{N}} |S_n(f)(0)| = \infty\}$  is a dense  $G_\delta$  subset of  $C_{2\pi}$ .

This means that, there are lots of functions in  $C_{2\pi}$  whose Fourier series diverges at 0.

First, we have a overlook about the proof

**Sketch of proof.** We have the following steps:

**Step 1.** Define a linear functional sequence  $(u_n)_{n \in \mathbb{N}}$  as follows

$$u_n : C_{2\pi} \rightarrow \mathbb{C}, f \mapsto S_n(f)(0). \quad (20)$$

Check that  $(u_n)_{n \in \mathbb{N}} \subseteq \mathcal{B}(C_{2\pi}, \mathbb{C})$ .

**Step 2.** Show that  $\sup_{n \in \mathbb{N}} \|u_n\| = \infty$ .

**Step 3.** Apply the following theorem.

**Theorem 8.5** (Principle of concentration of singularity). Let  $X$  be a Banach space,  $Y$  be a linear normed space and  $\{u_i \in \mathcal{B}(X, Y) : i \in I\}$  such that

$$\sup_{i \in I} \|u_i\| = \infty.$$

Then  $\{x \in X : \sup_{i \in I} \|u_i(x)\| = \infty\}$  is a dense  $G_\delta$  set in  $X$ .

□

And now we give the detailed proof.

**Proof.**

**Step 1.** Define  $(u_n)_{n \in \mathbb{N}}$  as (20). Given  $n \in \mathbb{N}$ , we have

$$\begin{aligned} |u_n(f)| &= |S_n(f)(0)| \\ &\leq \int_{-\pi}^{\pi} \left| f(t) \sum_{k=-n}^n e^{ikt} \right| \frac{dm}{2\pi}(t) \\ &\leq \|f\|_\infty \|D_n\|_1 \text{ (Hölder's inequality)} \end{aligned}$$

Thus  $\forall n \in \mathbb{N}, u_n \in \mathcal{B}(C_{2\pi}, \mathbb{C})$  and  $\|u_n\| \leq \|D_n\|_1$ .

The reversed inequality holds. To see this,  $\forall \varepsilon > 0$ , take a finite union of intervals (denote the union by  $I$ ) such that  $m(I) \leq$



$\pi\varepsilon/(2n+1)$ . This is possible since  $D_n$  has only finite zeros in  $[0, 2\pi]$  (consider  $D_n(t) = \frac{\sin((n+1/2)x)}{\sin(x/2)}$ ). Now Define  $f$

$$f: [0, 2\pi] \rightarrow \mathbb{C}, x \mapsto \begin{cases} 1, & x \notin I \wedge D_n(x) > 0; \\ -1, & x \notin I \wedge D_n(x) < 0; \\ l(x), & x \in I. \end{cases}$$

Here  $l$  is the affine mapping on each subinterval of  $I$  such that  $f$  is continuous.

$$\begin{aligned} |u_n(f)| &\geq \left| \int_{[-\pi, \pi] \setminus I} f(t) D_n(t) \frac{dm}{2\pi}(t) \right| - \left| \int_I f(t) D_n(t) \frac{dm}{2\pi}(t) \right| \\ &= \int_{[-\pi, \pi] \setminus I} |D_n(t)| \frac{dm}{2\pi}(t) - \left| \int_I f(t) D_n(t) \frac{dm}{2\pi}(t) \right| \\ &= \int_I |D_n(t)| \frac{dm}{2\pi}(t) - 2 \left| \int_I f(t) D_n(t) \frac{dm}{2\pi}(t) \right| \\ &\geq \|D_n\|_1 - \|f\|_\infty \varepsilon. \end{aligned}$$

The first inequality is just  $|x+y| \geq |x| - |y|$ , and the last inequality follows from

$$\begin{aligned} \left| \int_I f(t) D_n(t) \frac{dm}{2\pi}(t) \right| &\leq \int_I |f \cdot D_n| \frac{dm}{2\pi} \\ &\leq \|f\|_\infty \int_I |D_n| \frac{dm}{2\pi} \\ &\leq \|f\|_\infty (2n+1) \cdot \frac{m}{2\pi}(I) \\ &\leq \|f\|_\infty (2n+1) \cdot \frac{1}{2\pi} \pi\varepsilon. \end{aligned}$$

since  $\|D_n\|_\infty \leq 2n+1$ . From  $\|f\|_\infty = 1$ , we get

$$\|u_n\| \geq \|D_n\|_1 - \varepsilon \quad (\forall \varepsilon > 0),$$

which implies  $\|u_n\| \geq \|D_n\|$ .

Above all,  $\|u_n\| = \|D_n\|$ . Now we show that  $(\|D_n\|)_{n \in \mathbb{N}}$  is un-

bounded. Let  $dx$  denote  $dm(x)$  for short

$$\begin{aligned}
\|D_n\|_1 &= \int_{-\pi}^{\pi} \frac{|\sin(n+1/2)x|}{|\sin(x/2)|} \frac{dx}{2\pi} \\
&\geq \frac{1}{2\pi} \int_0^{\pi} \frac{|\sin(n+1/2)x|}{x/2} dx \\
&= \frac{2}{\pi} \int_0^{\frac{2n+1}{2}\pi} \frac{|\sin x|}{x} dx \\
&> \frac{2}{\pi} \int_0^{\pi} \frac{|\sin x|}{x} dx \\
&= \frac{2}{\pi} \sum_{k=0}^{n-1} \int_0^{\pi} \frac{|\sin x|}{x+k\pi} dx,
\end{aligned}$$

and

$$\int_0^{\pi} \frac{|\sin x|}{x+k\pi} dx \geq \frac{1}{(k+1)\pi} \int_0^{\pi} |\sin x| dx = \frac{2}{(k+1)\pi}.$$

Therefore

$$\|D_n\|_1 \geq \frac{4}{\pi^2} \sum_{k=1}^n \frac{1}{k} \sim \frac{4}{\pi^2} \log n \rightarrow \infty (n \rightarrow \infty).$$

Then, Theorem 8.5 implies the result.

□

Here is the proof of Theorem 8.5.

**Proof of Theorem 8.5.** Let

$$M: X \rightarrow [0, \infty], x \mapsto \sup_{i \in I} \|u_i(x)\|.$$

Define

$$F_n = \{x \in X : M(x) \leq n\} (\forall n \in \mathbb{N}),$$

then  $F_n$  is closed in  $X$  and hence  $\Omega_n := F_n^c$  is open, for all  $n \in \mathbb{N}$ . Now if  $F_m = \emptyset$  for some  $m \in \mathbb{N}$ , then

$$\{x \in X : M(x) = \infty\} = \bigcap_{n \geq 1} \Omega_n,$$

hence  $\{x \in X : M(x) = \infty\}$  is a  $G_\delta$ -set. If  $\{x \in X : M(x) = \infty\}$  is not dense in  $X$ , i.e.  $\bigcap_{n \geq 1} \Omega_n$  is not dense in  $X$ , then there is some

$m \in \mathbb{N}$  such that  $\Omega_m$  is not dense in  $X$ , since  $X$  is a Banach space (hence a Baire space). But  $\overline{\Omega_m} \neq X$  implies that  $F_m^\circ \neq \emptyset$ , thus the proof of Theorem 7.8 works (but we should replace  $\|T_\lambda(x_0)\| \leq M(x_0)$  with  $\|T_\lambda(x_0)\| \leq k$ ), contradiction with  $\sup_{i \in I} \|u_i\| = \infty$ .  $\square$

## 8.2 Week 8, Lecture 2

The aim of this lecture is

**Aim.** Prove the open mapping theorem.

Though we have proved Theorem 7.3, which needs a strong condition:  $T$  is surjective. We will change this restriction weaker.

### Recall

- A set  $B$  is called a set of first category, if  $B \subseteq \overset{\circ}{F}$ , where  $F$  is a  $F_\sigma$  set such that  $\overset{\circ}{F} = \emptyset$ .
- A topological space  $(X, \mathcal{T})$  is said to be a Baire space, if for all open set sequence  $(O_n)_{n \in \mathbb{N}}$  such that  $\overline{O_n} = X (\forall n \in \mathbb{N})$ , we have  $\bigcap_{n \geq 1} O_n = X$ .
  1. A set  $O$  is open and dense in  $X$  if and only if  $O^c$  is closed and  $(O^c)^\circ = \emptyset$ .
  2. Let  $X$  be a topological space. Then  $X$  is a Baire space if and only if for all closed set sequence  $(F_n)_{n \in \mathbb{N}}$  such that  $\overset{\circ}{F_n} = \emptyset$ , we have  $(\bigcup_{n \geq 1} F_n)^\circ = \emptyset$ .

### 8.2.1 Open Mapping Theorem (general version)

**Theorem 8.6** (Open mapping theorem). Let  $X, Y$  be two Banach spaces and  $T \in \mathcal{B}(X, Y)$ . If  $T(X) \hookrightarrow Y$  is a set of 2 category, then

1. there is  $c > 0$  such that  $B_Y \subseteq cT(B_X)$ . Here  $B_X, B_Y$  means the unit ball in  $X, Y$  respectively.
2.  $T$  is an open mapping.

**Proof.** In this theorem, for  $k \in \mathbb{K}$  and  $A, B \subseteq X$ ,

$$kA := \{kx \in X : x \in A\} \quad A + B := \{x + y : x \in A, y \in B\}.$$

Similarly for  $A, B \subseteq Y$ .

From  $X = \bigcup_{n \geq 1} nB_X$ , we have

$$T(X) = \bigcup_{n \geq 1} T(nB_X) \subseteq \bigcup_{n \geq 1} \overline{T(nB_X)},$$

and hence  $T(X) \subseteq \bigcup_{n \geq 1} \overline{T(nB_X)}$ . Thus  $T(X)$  is a  $F_\sigma$ -set in  $Y$ . So  $(\bigcup_{n \geq 1} \overline{T(nB_X)})^\circ \neq \emptyset$  (else,  $\bigcup_{n \geq 1} \overline{T(nB_X)}$  is a  $F_\sigma$ -set containing  $T(X)$ )

while having empty interior, contradiction with  $T(X)$  is not a set of 2 category). While  $Y$  is a Banach space, hence a Baire space, and there is some  $m \in \mathbb{N}$  such that  $(\overline{T(mB_X)})^\circ \neq \emptyset$ . Thus, there is  $y_0 \in (\overline{T(mB_X)})^\circ$  and  $r > 0$  such that

$$y_0 + rB_Y \subseteq (\overline{T(mB_X)})^\circ \subseteq \overline{T(mB_X)}.$$

Then

$$rB_Y \subseteq \overline{T(mB_X)} - y_0 \subseteq \overline{T(mB_X)} - \overline{T(mB_X)} \subseteq \overline{T(2mB_X)}.$$

The second  $\subseteq$  follows from  $y_0 \in \overline{T(mB_X)}$  and the last  $\subseteq$  can be easily checked by taking sequences convergent to each point. Now we have

$$B_Y \subseteq \overline{T(cB_X)},$$

where  $c := 2m/r$ . And we want  $B_Y \subseteq T(cB_X)$ . To see this, given arbitrary  $y_1 \in B(Y) \subseteq \overline{T(cB_X)}$ , there is  $x_1 \in cB_X$  such that

$$\|y_1 - T(x_1)\| < 1/2,$$

from the definition of closure. Then  $y_2 := 2(y_1 - T(x_1)) \in B_Y$ , since  $\|y_1\| < 1$ . Using the definition of closure again, there is  $x_2 \in cB_X$  such that

$$\|y_2 - T(x_2)\| < 1/2.$$

Now we define  $y_3 := 2(y_2 - T(x_2)) \in B_Y$ . And we can define two sequences  $(x_n)_{n \in \mathbb{N}} \subseteq cB_X, (y_n)_{n \in \mathbb{N}} \subseteq B_Y$  inductively, such that

$$\forall n \in \mathbb{N} \begin{cases} \|y_n - T(x_n)\| < 1/2, \\ y_{n+1} = 2(y_n - T(x_n)). \end{cases}$$

Hence,  $\forall n \in \mathbb{N}$

$$\begin{aligned} y_1 &= y_2/2 + T(x_1) \\ &= y_3/2^2 + T(x_2)/2 + T(x_1) \\ &= \dots \\ &= y_{n+1}/2^n + \sum_{j=1}^n T(x_j)/2^j \\ &= y_{n+1}/2^n + T\left(\sum_{j=1}^n x_j/2^j\right), \end{aligned}$$

where  $\|y_{n+1}/2^n\| < 2^{-n}$  since  $y_{n+1} \in B_Y$ ,  $\sum_{j=1}^n x_j/2^j$  is absolutely convergent since

$$\left\| \sum_{j=1}^n x_j/2^j \right\| \leq \sum_{j=1}^n \|x_j/2^j\| < \sum_{j=1}^n \frac{c}{2^j} < c(\forall n \in \mathbb{N}).$$

and is convergent to some  $x_0 \in cB_X$  because  $X$  is a Banach space. Therefore, by the continuity of  $T$

$$\begin{aligned} y_1 &= \lim_{n \rightarrow \infty} y_{n+1}/2^n + T\left(\sum_{j=1}^n x_j/2^j\right) \\ &= \lim_{n \rightarrow \infty} T\left(\sum_{j=1}^n x_j/2^j\right) \\ &= T\left(\lim_{n \rightarrow \infty} \sum_{j=1}^n x_j/2^j\right) \\ &= T(x_0). \end{aligned}$$

Above all,  $B_Y \subseteq T(cB_X) = cT(B_X)$ . Then

$$Y = \bigcup_{n \geq 1} nB_Y \subseteq \bigcup_{n \geq 1} nT(cB_X) = T\left(\bigcup_{n \geq 1} ncB_X\right) = TX,$$

i.e.  $T$  is surjective.

To see that  $T$  is open, it suffices to show that  $T(x + \delta B_X)$  is open in  $Y$  since  $\{x + \delta B_X : x \in X, \delta > 0\}$  is a topology base for  $X$ . While  $T$  is linear, WLOG, it suffices to show  $T(B_X)$  is open. Given  $x \in B_X$ ,  $T(x) \in T(B_X)$ , there is  $r_x > 0$  such that  $x + r_x B_X \subseteq B_X$  and hence

$$T(B_X) \supseteq T(x) + r_x T(B_X) \supseteq T(x) + r_x c^{-1} B_Y,$$

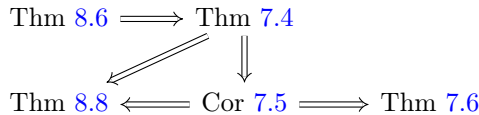
thus  $T(B_X)$  is open. □

**Remark.** This implies Theorem 7.3, since  $T(X) = Y$  and  $Y$  is a Banach space (hence a Baire space) implies  $T(X) = Y$  is of 2 category. Since  $Y$  is an open set that can't be of first category.

### 8.2.2 Closed Graph Theorem

Some results in subsection 7.1 are used here.

This graph tells the relation between theorems.



Now we talk about the Closed Graph Theorem. Here is the natural definition of the graph of an operator.

**Definition** (Graph). Let  $X, Y$  be two sets (allowed to have structures such as topology, norm and so on) and  $T: X \rightarrow Y$  is a map. The graph of  $T$ , denoted by  $G(T)$  is defined as

$$G(T) := \{(x, y) \in X \times Y : y = Tx\} = \{(x, Tx) \in X \times Y : x \in X\}.$$

**Remark.** In this lecture, we assume  $X \times Y$  is a linear normed space, the norm of  $X \times Y$  is

$$\| \cdot \| : X \times Y \rightarrow \mathbb{R}, (x, y) \mapsto \|x\|_X + \|y\|_Y,$$

if nothing else is mentioned. Equivalently, pick  $p = 1$  by default.

Furthermore, we need this notion.

**Definition** (Closed Operator). Suppose  $X, Y$  are two sets,  $T: X \rightarrow Y$  is a map. Then  $T$  is said to be closed, if  $G(T)$  is closed.

Recall that, a sequence  $((x_n, y_n))_{n \in \mathbb{N}} \subseteq X \times Y$  converges to  $(x, y) \in (X \times Y, \| \cdot \|_p)$ , if and only if

$$\lim_n \|(x_n, y_n) - (x, y)\|_p = \lim_n \left( \|x_n - x\|^p + \|y_n - y\|^p \right)^{1/p} = 0,$$

$$\text{i.e. } (x_n)_{n \in \mathbb{N}} \xrightarrow{\| \cdot \|_X} x \wedge (y_n)_{n \in \mathbb{N}} \xrightarrow{\| \cdot \|_Y} y.$$

**Proposition 8.7.** Let  $X, Y \in \text{Ob}(\text{Nor})$  and  $T \in \mathcal{L}(X, Y)$ . Then

1.  $T$  is closed iff  $\forall (x_n)_{n \in \mathbb{N}} \subseteq X, \forall (y_n)_{n \in \mathbb{N}} \subseteq Y, \forall x \in X, \forall y \in Y$  such that  $\lim_n x_n = x, \lim_n T(x_n) = y$ , we have  $T(x) = y$ .
2. If  $T$  is bounded, then  $T$  is closed.

**Proof.** For necessity of 1: suppose  $T$  is closed. Then  $\forall (x_n)_{n \in \mathbb{N}} \subseteq X$  such that  $\lim_n x_n = x \wedge \lim_n T x_n = y$ , we have  $((x_n, T x_n))_{n \in \mathbb{N}} \subseteq G(T)$  converges to  $(x, y) \in X \times Y$ . While  $G(T)$  is closed, we get  $(x, y) \in G(T)$  thus  $y = T x$ .

For sufficiency of 1: suppose  $T$  satisfies the latter condition in 1. Given an arbitrary convergent sequence  $((x_n, y_n))_{n \in \mathbb{N}} \subseteq G(T)$ , i.e.  $(x_n)_{n \in \mathbb{N}} \subseteq X, y_n = T(x_n)$  for all  $n \in \mathbb{N}$  and  $(x_n)_{n \in \mathbb{N}} \rightarrow x \in X$ . Then the latter condition implies  $\lim_n y_n = \lim_n T x_n$ . And continuity of  $T$  implies  $T(x) = \lim_n y_n$ . Thus the limit of  $((x_n, y_n))_{n \in \mathbb{N}}$ , i.e.  $(x, \lim_n y_n) \in X \times Y$  lies in  $G(T)$ . Therefore,  $G(T)$  is closed.

For 2, suppose  $T$  is bounded. Then the continuity of  $T$  implies that  $T$  satisfies the latter condition in 1.  $\square$

Now, the Closed Graph Theorem is

**Theorem 8.8** (Closed Graph). Let  $X, Y \in \text{Ob}(\text{Ban})$ . If  $T: X \rightarrow Y$  is closed, then  $T$  is bounded.

**Proof.** We know that  $X \times Y$  is a Banach space, and hence  $G(T) \hookrightarrow X \times Y$  being a closed subspace of  $X \times Y$  is also a Banach space. Consider the projection mapping

$$\begin{aligned} p: G(T) &\rightarrow X \\ (x, Tx) &\mapsto x, \end{aligned}$$

which is a linear bijection and  $\|p\| \leq 1$  since

$$\forall x \in X, \|p((x, Tx))\|_X = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\|_{X \times Y}.$$

Theorem 7.4 implies  $p^{-1} \in \mathcal{B}(X, G(T))$ . Therefore,

$$\begin{aligned} \forall x \in X, \|Tx\| &\leq \|x\| + \|Tx\| \\ &= \|(x, Tx)\|_{X \times Y} \\ &= \|p^{-1}(x)\|_{X \times Y} \\ &\leq \|p^{-1}\| \|x\|_X, \end{aligned}$$

i.e.  $\|T\| \leq \|p^{-1}\| < \infty$ .

Above all,  $T \in \mathcal{B}(X, Y)$ . □

**Remark.** This can also be proved by Corollary 7.5: since  $(X, \|\cdot\|_X)$  is a Banach space, and

$$(X, \|\cdot\|_X + \|\cdot\|_Y \circ T) \cong (G(T), \|\cdot\|_{X \times Y})$$

is also a Banach space, where the isometry is just  $T: X \rightarrow G(T)$ . Then Corollary 7.5 ensures Theorem 8.8.



## 9 Week 9

### 9.1 Week 9, Lecture 1

In this subsection, I will use the notation  $f \leq g$  lots of times, whose meaning can be found [here](#).

#### Recall

We have proved the relation between Theorems on Banach spaces, see the graph in section [8.2.2](#).

#### 9.1.1 Hahn-Banach Theorem

Today, here is going to prove Theorem [9.1](#) and Theorem [2](#). [Zorn's lemma](#) is needed here, see Appendix [A](#).

There is an important object related to Hahn-Banach Theorem:

**Definition** (Sub-linear functional). Let  $X$  be a **real** linear space. A real-valued function  $p: X \rightarrow \mathbb{R}$  is called a **sub-linear functional**, if

1. For all  $x \in X, \lambda \geq 0$ ,  $p(\lambda x) = \lambda p(x)$  holds;
2. For all  $x, y \in X$ ,  $p(x + y) \leq p(x) + p(y)$  holds.

**Definition** (Linear dual space). For a linear space  $X \in \text{Ob}(\text{Lin}_{\mathbb{K}})$ , the dual space, denoted by  $X^\sharp$ , is

$$X^\sharp := \{f: X \rightarrow \mathbb{K} \text{ that is } \mathbb{K}\text{-linear}\}.$$

The first theorem is irrelevant to topology, considering only linear space.

**Theorem 9.1** (Hahn-Banach). Let  $X$  be a **real** vector space and  $X_0 \hookrightarrow X$  is a subspace. Suppose  $f_0 \in X_0^\sharp$  and  $p: X \rightarrow \mathbb{R}$  is a sub-linear functional such that  $f_0 \leq p|_{X_0}$ . Then there is (at least one)  $f \in X^\sharp$  such that  $(f|_{X_0} = f_0) \wedge (f \leq p)$ .

**Proof.** We will prove this Theorem by [Zorn's lemma](#). Thus we need to construct an partially ordered set whose maximal element is the function we want.

**Step 1:** We construct the partially ordered set. Let

$$\mathcal{F} := \bigcup_{X_0 \hookrightarrow D \hookrightarrow X} \{g \in D^\sharp: g|_{X_0} = f_0 \wedge g \leq p|_D\},$$

where the union is taken over all subspaces  $D$  such that  $X_0 \hookrightarrow D \hookrightarrow X$ . Define an order on  $\mathcal{F}$  as follows

$$g_1 \leq g_2 \iff \text{dom}(g_1) \hookrightarrow \text{dom}(g_2) \wedge g_2|_{\text{dom}(g_1)} = g_1,$$

i.e.  $g_1 \leq g_2$  iff  $g_2$  is an extension of  $g_1$  in the sense above.

**Exercise 9.1.** Check that  $(\mathcal{F}, \leq)$  is a partially ordered set.

**Step 2:** We prove that  $\mathcal{F}$  satisfies the condition of [Zorn's lemma](#). Given an arbitrary linearly ordered subset  $\mathcal{F}_0 \subseteq \mathcal{F}$ . We prove that  $\mathcal{F}_0$  has an upper bound in  $\mathcal{F}$ . Consider the set  $\bigcup_{g \in \mathcal{F}_0} \text{dom}(g) = \text{dom}(\bigcup \mathcal{F}_0)$ . Define the linear structure as follows

$$\begin{aligned} + : \text{dom}(\bigcup \mathcal{F}_0) \times \text{dom}(\bigcup \mathcal{F}_0) &\rightarrow \text{dom}(\bigcup \mathcal{F}_0) \\ (v_1, v_2) &\mapsto v_1 +_{V_1 \cup V_2} v_2, \end{aligned}$$

where  $V_j \in \text{dom}(\bigcup \mathcal{F}_0)$  is a space containing  $v_j$ ,  $V_1 \cup V_2$  is a subspace of  $\text{dom}(\bigcup \mathcal{F}_0)$  since  $\text{dom}(\bigcup \mathcal{F}_0)$  is linearly ordered and  $+_{V_1 \cup V_2}$  is the natural addition of the subspace  $V_1 \cup V_2$  (this is surely a subspace by the linear order). Though it's possible that  $v_1 \in V_1 \cap V'_1, v_2 \in V_2 \cap V'_2$  for some  $V_1, V'_1 \in \mathcal{F}_0$ , the summation  $+$  is well-defined. WLOG, we suppose  $V_1 \hookrightarrow V'_1, V_2 \hookrightarrow V'_2$  by the linear order and then

$$v_1 +_{V_1 \cup V_2} v_2 = v_1 +_{V'_1 \cup V'_2} v_2$$

since  $V_1 \cup V_2 \hookrightarrow V'_1 \cup V'_2$ . And the scalar multiplication is just

$$\cdot : \text{dom}(\bigcup \mathcal{F}_0) \times \mathbb{R} \rightarrow \text{dom}(\bigcup \mathcal{F}_0), (v, k) \mapsto k \cdot_V v,$$

where  $V \in \mathcal{F}_0$  is a subspace of  $X$  containing  $v$ . We can prove that  $\cdot$  is well-defined similarly. Above all,  $\text{dom}(\bigcup \mathcal{F}_0)$  is a vector space of  $X$  and contains  $X_0$ . Now we define a linear functional on  $\text{dom}(\bigcup \mathcal{F}_0)$  as follows

$$h : \text{dom}(\bigcup \mathcal{F}_0) \rightarrow \mathbb{R}, v \mapsto g(v),$$

whenever  $g$  is an element of  $\mathcal{F}_0$  such that  $v \in \text{dom}(g)$ . This is well-defined by the property of  $\mathcal{F}_0$ . And  $h \in \mathcal{F}$  is an upper bound for  $\mathcal{F}_0$ .

**Step 3:** apply [Zorn's lemma](#), thus there is a maximal element in  $\mathcal{F}$  and let  $f$  be the maximal element. We prove that  $f$  is what we want. Equivalently, we prove that

1.  $f \in X^\sharp$ , i.e.  $\text{dom}(f) = X$ ;

2.  $f|_{X_0} = f_0$ ;
3.  $f \leq p$ .

Since  $f \in \mathcal{F}$ , we get  $f|_{X_0} = f_0$  and  $f \leq p|_{\text{dom}(f)}$ . Thus it suffices to show that  $\text{dom}(f) = X$ . Suppose there is an element  $x_0 \in X \setminus \text{dom}(f)$ , then

$$\text{dom}(f) + \mathbb{R}x_0 = \{y + kx_0 : y \in \text{dom}(f) \wedge k \in \mathbb{R}\}$$

is a subspace of  $X$  strictly bigger than  $\text{dom}(f)$ . We prove that  $f$  can be extended to a linear functional  $\tilde{f}$  on  $\text{dom}(f) + \mathbb{R}x_0$  such that  $\tilde{f} \in \mathcal{F}$ , which is a contradiction with  $f$  being a maximal element. In order to define  $\tilde{f}$ , it suffices to check that  $\tilde{f}(x_0)$  can be defined, since

$$\tilde{f}(y + kx_0) = \tilde{f}(y) + k\tilde{f}(x_0) = f(y) + k\tilde{f}(x_0)$$

is determined by  $\tilde{f}(x_0)$ . The only thing restricts the value of  $\tilde{f}(x_0)$  is  $\tilde{f} \leq p|_{\text{dom}(\tilde{f})}$ , i.e.  $\forall y \in \text{dom}(f), \forall k \in \mathbb{R} \setminus \{0\}$  :

$$\begin{cases} \tilde{f}(y + kx_0) \leq p|_{\text{dom}(\tilde{f})}(y + kx_0), & k > 0; \\ \tilde{f}(y + kx_0) \leq p|_{\text{dom}(\tilde{f})}(y + kx_0), & k < 0; \end{cases}$$

i.e.  $\forall y \in \text{dom}(f), \forall k \in \mathbb{R} \setminus \{0\}$  :

$$\begin{cases} \tilde{f}(x_0) \leq p(y + kx_0)/k - \tilde{f}(y)/k, & k > 0; \\ \tilde{f}(x_0) \geq p(y + kx_0)/k - \tilde{f}(y)/k, & k < 0; \end{cases}$$

i.e.  $\forall y, z \in \text{dom}(f), \forall k > 0, k' < 0$  :

$$\begin{cases} \tilde{f}(x_0) \leq p(y/k + x_0) - f(y/k); \\ \tilde{f}(x_0) \geq -p(-z/k' - x_0) - f(z/k'). \end{cases}$$

Here I don't care the case  $k = 0$  since  $\tilde{f}(y) \leq p|_{\text{dom}(\tilde{f})}(y)$  can be deduced from  $f \leq p|_{\text{dom}(f)}$ . Therefore, it suffices to show that  $\forall y, z \in \text{dom}(f), \forall k > 0, k' < 0$  :

$$-p(-z/k' - x_0) - f(z/k') \leq p(y/k + x_0) - f(y/k), \quad (21)$$

And (21) holds. To see this, we can set  $k = -k' = 1$  and then

$$p(z - x_0) + p(y + x_0) \geq p(y + z) \geq f(y + z) = f(y) + f(z).$$

Then (21) implies

$$\sup S^- \leq \inf S^+.$$

where

$$\begin{aligned} S^- &:= \{-p(-z/k' - x_0) - f(z/k') \in \mathbb{R} : z \in \text{dom}(f), k' < 0\}, \\ S^+ &:= \{p(y/k + x_0) - f(y/k) \in \mathbb{R} : y \in \text{dom}(f), k > 0\}. \end{aligned}$$

Then  $\tilde{f}(x_0)$  can be taken as an arbitrary number in the interval

$$[\sup S^-, \inf S^+].$$

We're done. □

Back to linear normed space.

**Theorem 9.2** (Hahn-Banach, general version). Let  $X$  be a linear normed space over the field  $\mathbb{K}$  and  $X_0 \hookrightarrow X$  is a subspace. Suppose  $f \in X_0^*$ , then there is  $f \in X^*$  such that

1.  $f|_{X_0} = f_0$ ;
2.  $\|f\|_{X^*} = \|f_0\|_{X_0^*}$ .

In other words,  $f$  is an extension of  $f_0$  with the same norm.

**Remark.** Before the proof, we should have an observation: a complex vector space can be viewed as a real vector space. For the detail, see Proposition A.2.

**Proof.** To use Theorem 9.1, I will prove the case  $\mathbb{K} = \mathbb{R}$  first, which can be applied for the case  $\mathbb{K} = \mathbb{C}$ .

**Case 1:**  $\mathbb{K} = \mathbb{R}$ . Let  $p$  be the norm defined as follows

$$p: X \rightarrow \mathbb{R}, x \mapsto \|f_0\|_{X_0^*} \|x\|.$$

Then  $p$  is a semi-norm such that  $f_0 \leq p|_{X_0}$ . Thus, Theorem 9.1 implies that there is a function  $f$ , an extension of  $f_0$  that satisfies  $f \leq p$ . For all  $x \in X$

$$\begin{aligned} f(x) &\leq p(x), \\ f(x) &= -f(-x) \geq -p(x). \end{aligned}$$

Thus  $|f(x)| \leq p(x)$  and  $\|f\|_{X^*} \leq \|f_0\|_{X_0^*}$ , hence  $f \in X^*$ . The reversed inequality holds since  $f|_{X_0} = f_0$ .

**Case 2:**  $\mathbb{K} = \mathbb{C}$ . It can be shown that  $\forall h \in \mathcal{B}(X, \mathbb{C})$ ,  $h$  is uniquely determined by  $\text{Re} \circ h = \text{Re}(h) \in \mathcal{B}(X, \mathbb{R})$ . For all  $x \in X$ ,

$$h(ix) = i(\text{Re } h(x) + i \text{Im } h(x)) = -\text{Im } h(x) + i \text{Re } h(x),$$

take real parts for both sides and get  $-\text{Im } h(x) = \text{Re } h(ix)$ . Thus

$$\forall x \in X : h(x) = \text{Re } h(x) - i \text{Re } h(ix).$$

Now, view  $X$  as a real vector space and suppose  $\forall x \in X, f_0(x) = g_0(x) - ig_0(ix)$ , where  $g_0 \in \mathcal{B}(X, \mathbb{R})$ . Define

$$p: X \rightarrow \mathbb{R}, x \mapsto \|g_0\|_{\mathcal{B}(X, \mathbb{R})} \|x\|.$$

Then apply the result in 9.1.1 and we get  $\exists g \in \mathcal{B}(X, \mathbb{R})$  such that

$$(g|_{X_0} = g_0) \wedge (g \leq p) \wedge (\|g\|_{\mathcal{B}(X, \mathbb{R})} = \|g_0\|_{\mathcal{B}(X_0, \mathbb{R})}).$$

Then  $f: X \rightarrow \mathbb{C}, x \mapsto g(x) - ig(ix)$  satisfies

$$(f|_{X_0} = f_0) \wedge (\|f\|_{X^*} = \|f_0\|_{X_0^*}).$$

The first is trivial and the second is true if  $\|f\|_{X^*} \leq \|f_0\|_{X_0^*}$ , equivalently,  $\forall x \in X : |f(x)| \leq p(x)$ . To see this, let

$$\theta: X \rightarrow \mathbb{C}, x \mapsto (\text{sign} \circ f)(x).$$

Notice that  $|\cdot| \circ \theta: X \rightarrow \mathbb{R}$  is a constant function. Then  $\forall x \in X$

$$\begin{aligned} |f(x)| &= f(x) \cdot \theta(x) \\ &= f(\theta(x) \cdot x) \\ &= g(\theta(x) \cdot x) && (f(y) \in \mathbb{R} \iff f(y) = g(y)) \\ &\leq \|g\|_{\mathcal{B}(X, \mathbb{R})} \cdot \|\theta(x) \cdot x\|_X \\ &= \|g\|_{\mathcal{B}(X, \mathbb{R})} \cdot |\theta(x)| \cdot \|x\|_X \\ &= \|g\|_{\mathcal{B}(X, \mathbb{R})} \cdot \|x\|_X. \end{aligned}$$

And

$$\|g\|_{\mathcal{B}(X, \mathbb{R})} = \|g_0\|_{\mathcal{B}(X_0, \mathbb{R})} \leq \|f_0\|_{X_0^*}.$$

Thereby,  $\forall x \in X : |f(x)| \leq \|f_0\|_{X_0^*} \|x\|$ , i.e.  $\|f\|_{X^*} \leq \|f_0\|_{X_0^*}$ .  $\square$

Theorem 9.2 is of great importance in the theory of “dual space of Banach space”, which can be seen later.

## 9.2 Week 9, Lecture 2

### Recall

We have studied

1. Theorem 7.6 (Resonance Theorem/ Uniformly bounded principle):

Let  $X$  be a Banach space and  $Y$  be a linear normed space. Suppose  $\{T_\lambda\}_{\lambda \in \Lambda} \subseteq \mathcal{B}(X, Y)$  satisfies:  $\forall x \in X \exists M_x > 0$  such that  $\sup_{\lambda \in \Lambda} \|T_\lambda x\|_Y < M_x$ . Then there is  $M > 0$  such that  $\sup_{\lambda \in \Lambda} \|T_\lambda\| < M$ .

2. Theorem 7.3 (Open mapping Theorem):

(a) Let  $X, Y$  be two Banach spaces and  $T \in \mathcal{B}(X, Y)$  is a surjection. Then  $T$  is an open mapping.

(b) Theorem 8.6.

3. Theorem 8.8 (Closed graph Theorem):

For a mapping  $T: X \rightarrow Y$ , the graph of  $T$  is  $G(T) := \{(x, Tx) : x \in X\} \subset X \times Y$ . A mapping  $T$  is said to be closed if  $G(T)$  is closed.

Let  $X, Y$  be two Banach spaces and  $T \in \mathcal{L}(X, Y)$ . Then  $T$  is a closed operator implies  $T \in \mathcal{B}(X, Y)$ .

4. Theorem 9.1 and Theorem 9.2 (Hahn-Banach Theorem):

**Remark.** This is one of the most important theorems for functional analysis.

Here is an exercise that explains the name “Resonance Theorem”.

**Exercise 9.2.** Let  $X$  be a Banach space and  $Y$  be a linear normed space. Suppose  $\{T_\lambda\}_{\lambda \in \Lambda} \subseteq \mathcal{B}(X, Y)$  satisfies  $\sup_{\lambda \in \Lambda} \|T_\lambda\| = \infty$ . Show that  $\exists x_0 \in X$  such that  $\sup_{\lambda \in \Lambda} \|T_\lambda x_0\|_Y = \infty$ .

### 9.2.1 Review

Recall the definition of semi-norm

**Definition** (Semi-norm). Let  $X$  be a linear normed space. A function  $p: X \rightarrow \mathbb{R}$  is said to be a semi-norm, if it satisfies:

1. For all  $x \in X$ ,  $p(x) \geq 0$ ;

2. For all  $x, y \in X$ ,  $p(x + y) \leq p(x) + p(y)$ ;
3. For all  $x \in X, k \in \mathbb{K}$ ,  $p(kx) = |k|p(x)$ .

Here is another way to state “semi-norm”.

**Definition** (Sub-additive). Let  $X$  be a linear normed space. A function  $f: X \rightarrow \mathbb{R}$  is said to be sub-additive if

$$\forall x, y \in X: f(x) + f(y) \leq f(x + y).$$

**Definition** (Positive-homogeneity). Let  $X$  be a linear normed space. A function  $f: X \rightarrow \mathbb{R}$  is said to be positive-homogeneous if

$$\forall x \in X, \alpha \in [0, \infty): f(\alpha x) = \alpha f(x).$$

Thus, a function  $p: X \rightarrow \mathbb{R}$  is a semi-norm if and only if ( $p$  is sub-additive  $\wedge$   $p$  is positive-homogeneous). Let

$$X'_+ := \{f: X \rightarrow \mathbb{R} \mid f \text{ is sub-additive} \wedge \text{homogeneous}\}.$$

**Theorem 9.3.** Let  $X$  be a linear normed space over  $\mathbb{R}$  and  $p \in X'_+$ . Suppose  $X_0 \hookrightarrow X$  and  $f \in \mathcal{L}(X, \mathbb{R})$ . Then

1.  $\exists f \in \mathcal{L}(X, \mathbb{R})$  such that  $f|_{X_0} = f$ ;
2. If  $f_0 \leq p|_{X_0}$ , then  $f \leq p$ .

1 is equivalent to the following commutative diagram

$$\begin{array}{ccc} X_0 & \xrightarrow{f_0} & \mathbb{R} \\ & \searrow \iota & \nearrow \exists f \\ & X & \end{array}$$

**Remark.** We’ve prove this theorem, see 9.1. Why we need [Zorn’s lemma](#) here? Suppose  $X_0 \neq X$  and  $x_0 \in X \setminus X_0$ . Let  $M := \text{span}(\{x_0\} \cup X_0) \hookrightarrow X$ . Notice that  $\dim(M \setminus X_0) = 1$ . Thus, by Mathematical Induction, we can prove the case  $\dim(X \setminus X_0) < \infty$ . To get rid of the assumption  $\dim(X \setminus X_0) < \infty$ , we need “Transfinite induction” in some sense, which is relevant to the Axiom of Choice (equivalent to [Zorn’s lemma](#)). And how do we apply [Zorn’s lemma](#)? Recall how do we define the partially ordered set  $(\mathcal{F}, \leq)$ .

**Theorem 9.4.** Let  $X$  be a linear normed space over  $\mathbb{K}$ ,  $X_0 \hookrightarrow X$  and  $f_0 \in \mathcal{B}(X_0, \mathbb{K})$ . Then

1. there exists  $f \in \mathcal{B}(X, \mathbb{K})$  such that  $f|_{X_0} = f_0$ , i.e. the following diagram commutes

$$\begin{array}{ccc} X_0 & \xrightarrow{f_0} & \mathbb{K} \\ & \searrow \iota & \nearrow \exists f \\ & X & \end{array}$$

2. their norms coincide:  $\|f\|_{X^*} = \|f_0\|_{X_0^*}$ .

**Remark.** Proof of this theorem ( $\mathbb{K} = \mathbb{C}$  case) needs an observation:  $\forall f \in \mathcal{L}(X, \mathbb{C})$ ,  $f$  is uniquely determined by  $\operatorname{Re} f$  since

$$\forall x \in X, f(x) = f_1(x) - i f_1(ix),$$

where  $f_1 = \operatorname{Re}(f)$ .

Here is some corollaries of Theorem 9.4.

**Corollary 9.5.** Let  $X$  be a linear normed space over  $\mathbb{K}$ ,  $x_0 \in X$  and  $x_0 \neq 0$ . Then there is  $f \in X^*$  such that  $f(x_0) = \|x_0\|$  and  $\|f\| = 1$ .

**Proof.** It suffices to show that there is an  $f \in X^*$  such that  $|f(x_0)| = \|x_0\|$ , since we can multiply  $f$  by a constant sign ( $f(x_0)$ ). Let  $X_0 := \operatorname{span}\{x_0\}$ , a subspace of  $X$  and

$$f_0: X_0 \rightarrow \mathbb{K}, k \cdot x_0 \mapsto k \cdot \|x_0\|.$$

Clearly  $\forall k \in \mathbb{K}$ ,  $|f_0(kx_0)| \leq |k|\|x_0\|$  and  $\|kx_0\| = |k|\|x_0\|$ , thus  $\|f_0\|_{X_0^*} \leq 1$ . And the inverse inequality holds, since

$$f(x_0/\|x_0\|) = 1, \|x_0/\|x_0\|\| = 1.$$

Then apply Theorem 9.4, there is a function  $f: X \rightarrow \mathbb{C}$  such that  $f|_{X_0} = f_0 \wedge \|f\|_{X^*} = \|f_0\|_{X_0^*}$ . And

$$f(x_0) = f|_{X_0}(x_0) = f_0(x_0) = \|x_0\|. \quad \square$$

For both Theorem 9.3 and Theorem 9.4, the extension doesn't need to be unique. But there is a result for some unique extension.

**Exercise 9.3.** Let  $X$  be a linear normed space and  $X_0 \hookrightarrow X$  is a dense subspace. Suppose  $f_0 \in X_0^*$ , then there is a unique  $f \in X^*$  such that  $f|_{X_0} = f_0$ . In other words, the following diagram commutes.

$$\begin{array}{ccc} X_0 & \xrightarrow{f_0} & \mathbb{K} \\ & \searrow \iota & \nearrow \exists f \\ & X & \end{array}$$



**Proof.** Existence is ensured by Theorem 9.4.

Uniqueness: if there is another  $f'$  satisfies the commutative diagram, then

$$f|_{X_0} = f_0 = f'|_{X_0},$$

and two continuous functions that coincide on a dense subset must coincide.  $\square$

**Corollary 9.6.** Let  $X$  be a linear normed space and  $x_1, x_2 \in X$  satisfy  $x_1 \neq x_2$ . Then there is  $f \in X^*$  such that  $f(x_1) \neq f(x_2)$ . In other words,  $X^*$  separates points of  $X$ .

**Proof.** Let  $x_0 := x_1 - x_2$  and apply Corollary 9.5.  $\square$

**Corollary 9.7.** Let  $X$  be a linear normed space and  $x_0 \in X$ . If  $\forall f \in X^* : f(x_0) = 0$  then  $x_0 = 0$ .

**Proof.** If  $x_0 \neq 0$ , apply Corollary 9.6 and get a contradiction.  $\square$

Here is a related exercise.

**Exercise 9.4.** Let  $X$  be a linear normed space and  $\{0\} \neq X_0 \hookrightarrow X$ . Suppose  $x_0 \in X$  satisfies  $d(x_0, X_0) = \rho > 0$ . Then there is some  $f \in X^*$  such that  $f(x_0) = \rho$  and  $f|_{X_0} = 0$  (Hint: consider  $\text{span}(\{x_0\} \cup X_0)$ ).

# 10 Week 10

## 10.1 Week 10, Lecture 1

### Recall

[Hahn-Banach Theorem](#) is important, and here is an example of its applications.

**Theorem.** Let  $X$  be a linear normed space,  $X_0 \hookrightarrow X$  and  $f_0 \in X_0^*$ . Then there is some function  $f \in X^*$  such that  $f|_{X_0} = f_0$  and  $\|f\|_{X^*} = \|f_0\|_{X_0^*}$ .

**Proof.** See Theorem 9.4. □

And it has a corollary:

**Corollary 10.1.** Let  $X$  be a linear normed space and  $0 \neq x_0 \in X$ , then there is some  $f \in X^*$  such that  $\|f\|_{X^*} = 1$  and  $f(x_0) = \|x_0\|$ .

**Proof.** See Corollary 9.5. □

Furthermore, recall the [exercise](#) that ensures the completeness of  $X^* = \mathcal{B}(X, \mathbb{K})$ . Here is a corollary shows the duality. Notice that  $\forall f \in X^*, \|f\|_{X^*} = \sup_{\|x\|_X \leq 1} |f(x)|$ , where the sup is taken over all  $x \in X$  such that  $\|x\|_X \leq 1$ .

**Corollary 10.2.** There is another representation of  $\| \cdot \|_X$ , that's  $\forall x \in X$ ,

$$\|x\|_X = \max_{\|f\|_{X^*} \leq 1} |f(x)|,$$

where the max is taken over all  $x \in X$  such that  $\|x\|_X \leq 1$ .

**Remark.** Notice that here we use max, rather than sup, i.e. the sup is accessible.

**Proof.** WLOG, suppose  $x \neq 0$ . Let  $A := \sup_{\|f\|_{X^*} \leq 1} |f(x)|$ . On the one hand,  $\|x\|_X \geq A$ . Because  $\forall f \in X^*$  such that  $\|f\|_{X^*} \leq 1$ , we have

$$|f(x)| \leq \|f\|_{X^*} \|x\|_X \leq \|x\|_X.$$

On the other hand,  $\|x\| \leq A$  and  $A$  is accessible. This follows from Corollary 9.5, which also ensures that  $\exists g \in X^*, \|g\|_{X^*} \leq 1$  with  $|g(x)| = \sup_{\|f\|_{X^*} \leq 1} |f(x)|$ . □

Here is an exercise mentioned last time.

**Exercise 10.1.** Let  $X$  be a linear normed space,  $X_0 \hookrightarrow X$  be a dense subspace of  $X$  and  $f_0 \in X_0^*$ . Then  $\exists! f \in X^*$  such that  $f|_{X_0} = f_0$ .

To do this, apply Theorem 9.2 and notice that two bounded functionals coincide on a dense subset must coincide. Furthermore, the function  $f$  above satisfies  $\|f\|_{X^*} = \|f_0\|_{X_0^*}$ .

Here is an example that shows: the extension from Theorem 9.2 can be not unique.

**Example 23.** Consider the normed space  $(\mathbb{R}^2, \|\cdot\|_1)$ , where

$$\|\cdot\|_1 : (x, y) \mapsto |x| + |y|$$

is equivalent to  $\|\cdot\|_2$  since  $\dim(\mathbb{R}^2) = 2 < \infty$ . In fact,  $(\mathbb{R}^2, \|\cdot\|_1)$  is just  $(\mathbb{R}, \|\cdot\|_1) \oplus (\mathbb{R}, \|\cdot\|_1)$ . Let  $p_2 : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto y$  be the projection to the second coordinate. Then

$$G := p_2^{-1}(0) = \{(x, 0) \in \mathbb{R}^2 : x \in \mathbb{R}\}$$

is a closed subspace of  $\mathbb{R}^2$  that's isometric to  $\mathbb{R}$ . Consider a functional on  $G$  as follows:

$$f : G \rightarrow \mathbb{R}, (x, 0) \mapsto x.$$

It's easy to see that  $\|f\|_{G^*} = 1$ . We can directly construct lots of extensions of  $f$ . To see this, define

$$f_\beta : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x + \beta y$$

for all  $\beta \in \mathbb{R}$ . Then  $\|f_\beta\|_{(\mathbb{R}^2)^*} = 1 \vee |\beta|$ . This can be checked by taking  $(x, y) = (1, 0)$   $(x, y) = (0, 1)$  and using the triangle inequality  $|x + \beta y| \leq |x| + |\beta||y|$ . Thus, for all  $\beta \in [-1, 1]$ ,  $f_\beta \in (\mathbb{R}^2)^*$  is an extension of  $f$  with  $\|f_\beta\|_{(\mathbb{R}^2)^*} = \|f\|_{G^*} = 1$ . Furthermore,  $[-1, 1]$  contains lots of elements, since

$$\text{card}[-1, 1] = \text{card } \mathbb{R} = \text{card}({}^{\mathbb{N}}2) > \text{card } \mathbb{N}. \quad (22)$$

The inequality  $\text{card}({}^{\mathbb{N}}2) > \text{card } \mathbb{N}$  is known as a special case of Cantor's Theorem.

Here is a proof of “ $\mathbb{R}$  is uncountable” using (22).

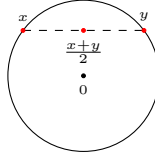
**Proof.** It suffices to show that  ${}^{\mathbb{N}}2$  is uncountable, where 2 means a set of exactly 2 elements,  $\{\emptyset, \{\emptyset\}\}$  or  $\{-1, 1\}$  for example. WLOG, let  $2 = \{-1, 1\}$  then an element of  ${}^{\mathbb{N}}2$  is just a sequence with values being 1 or  $-1$ . Suppose  ${}^{\mathbb{N}}2$  is countable, i.e.  ${}^{\mathbb{N}}2 = \{\alpha(k) : k \in \mathbb{N}\}$  where  $\alpha_k$  is a sequence whose values lie in  $\{-1, 1\}$ . Then consider the sequence

$$\beta : \mathbb{N} \rightarrow \{-1, 1\}, n \mapsto -\alpha(n)_n.$$

In other words,  $\beta$  is a sequence whose values lie in  $\{-1, 1\}$  and satisfies  $\forall n \in \mathbb{N}: \beta_n \neq \alpha(n)_n$ . Hence  $\forall k \in \mathbb{N}: \beta \neq \alpha(k)$ . While  $\beta \in {}^{\mathbb{N}}2$ , which contradicts with  ${}^{\mathbb{N}}2 = \{\alpha(k): k \in \mathbb{N}\}$ .  $\square$

It's possible that Hahn-Banach Theorem gives a unique extension. Here is a proposition about this.

**Definition** (Strictly convex). Let  $X$  be a linear normed space. Then  $X$  is said to be strictly convex, if for all distinct  $x, y \in X$  with  $\|x\| = \|y\| = 1$ , we have  $\|(x+y)/2\| < 1$ .



**Proposition 10.3.** Let  $X$  be a linear normed space. If  $X^*$  is strictly convex, then the extension given by Hahn-Banach Theorem is unique.

**Proof.** Omitted. This is not our main goal.  $\square$

### 10.1.1 Something about Dual Space

In this part,  $p \in [1, \infty]$  and  $q = p' = p/(p-1)$  is the conjugate index of  $p$  unless otherwise specified.

We have studied  $\ell_p$  space and hence it's natural to ask:

**Question 10.1.** What's  $\ell_p^*$  ?

The following example answers a part of this question.

**Example 24.** We have  $\ell_1^* \cong \ell_\infty$ .

**Proof.** I will construct a contraction  $\varphi: \ell_\infty \rightarrow \ell_1^*$  and its inverse that is also a contraction. Then  $\ell_\infty \cong \ell_1^*$ .

Let  $\varphi: \ell_\infty \rightarrow \ell_1^*, \alpha \mapsto \varphi_\alpha$ , where

$$\varphi_\alpha: \ell_1 \rightarrow \mathbb{K}, x \mapsto \sum_{n \geq 1} x_n \overline{\alpha_n}.$$

Hölder's inequality implies that  $\varphi_\alpha$  is well-defined and bounded, with  $\|\varphi_\alpha\|_{\ell_1^*} \leq \|\alpha\|_\infty$ . In other words,  $\varphi_\alpha$  is a bounded linear functional on  $\ell_1$  for each  $\alpha \in \ell_\infty$ . From  $\|\varphi_\alpha\|_{\ell_1^*} \leq \|\alpha\|_\infty$ , we get  $\|\varphi\|_{\mathcal{B}(\ell_\infty, \ell_1^*)} \leq 1$ .

Then we define the inverse of  $\varphi$ . Let

$$e_n := (\underbrace{0, \dots, 0}_{n-1 \text{ terms}}, 1, 0, \dots) \text{ for all } n \in \mathbb{N}$$

and  $\psi: \ell_1^* \rightarrow \ell_\infty, f \mapsto \psi^f$ , where  $\psi_n^f = f(e_n), \forall n \in \mathbb{N}$ . Then  $\psi$  is well-defined. What should be checked is just:  $\forall f \in \ell_1^*, \|\psi^f\|_\infty < \infty$ , which is right since

$$\forall n \in \mathbb{N}: |\psi_n^f| = |f(e_n)| \leq \|f\|_{\ell_1^*} \|e_n\|_{\ell_1} = \|f\|_{\ell_1^*}$$

implies that  $\|\psi^f\|_\infty \leq \|f\|_{\ell_1^*}$  and hence  $\|\psi\|_{\mathcal{B}(\ell_1^*, \ell_\infty)} \leq 1$ . Now we prove that  $\varphi: \ell_\infty \rightarrow \ell_1^*$  is an isometry, i.e.  $(\varphi \circ \psi = \text{id}_{\ell_1^*}) \wedge (\psi \circ \varphi = \text{id}_{\ell_\infty})$ .

- For  $\varphi \circ \psi = \text{id}_{\ell_1^*}$ : let  $f \in \ell_1^*$  be an arbitrary element. Then  $\psi^f$  is a sequence in  $\ell_\infty$ , such that  $\forall n \in \mathbb{N}, \psi_n^f = f(e_n)$ ,

$$\varphi \circ \psi(f) = \varphi(\psi^f)$$

and hence for all  $x \in \ell_1$  :

$$\begin{aligned} [\varphi \circ \psi(f)](x) &= [\varphi(\psi^f)](x) \\ &= \sum_{n \geq 1} x_n \overline{\psi_n^f} \\ &= \sum_{n \geq 1} x_n \overline{f(e_n)} \\ &= \sum_{n \geq 1} f(x_n e_n) \\ &= \lim_{N \rightarrow \infty} \sum_{n=1}^N f(x_n e_n) \\ &= \lim_{N \rightarrow \infty} f\left(\sum_{n=1}^N x_n e_n\right) \end{aligned}$$

Here the limit is the usual limit for sequence in  $\mathbb{K}$  and we will use the continuity of  $f$  to pass the limit to the limit in  $\ell_1$ , in other words, we want to find the limit (with respect to convergence in norm)  $\lim_N \sum_{n=1}^N x_n e_n$  in  $\ell_1$ . It's natural to guess the answer:

just  $x$ . Since

$$\begin{aligned}\left\|\sum_{n=1}^N x_n e_n - x\right\|_1 &= \|(0, \dots, 0, x_{N+1}, x_{N+2}, \dots)\|_1 \\ &= \sum_{n \geq N+1} |x_n| \\ &= \|x\|_1 - \sum_{n=1}^N |x_n|\end{aligned}$$

converges to 0 as  $N \rightarrow \infty$ , and by continuity of  $f$ , we get

$$\lim_{N \rightarrow \infty} f\left(\sum_{n=1}^N x_n e_n\right) = f\left(\lim_{N \rightarrow \infty} \sum_{n=1}^N x_n e_n\right) = f(x).$$

Above all,  $[\varphi \circ \psi(f)](x) = f(x)$  holds for all  $x \in \ell_1$ . Thus,  $\varphi \circ \psi(f) = f$  holds for all  $f \in \ell_1^*$  and hence  $\varphi \circ \psi = \text{id}_{\ell_1^*}$ .

- For  $\psi \circ \varphi = \text{id}_{\ell_\infty}$ : let  $\alpha \in \ell_\infty$  be an arbitrary element. Then for all  $\alpha \in \ell_\infty$

$$\psi \circ \varphi(\alpha) = \psi(\varphi_\alpha)$$

and  $\forall n \in \mathbb{N}$

$$[\psi \circ \varphi(\alpha)]_n = [\psi(\varphi_\alpha)]_n = \varphi_\alpha(e_n) = \alpha_n.$$

The last equality can be checked by definition of  $\varphi_\alpha$  and  $e_n$ . Therefore,  $[\psi \circ \varphi](\alpha) = \alpha$  holds for all  $\alpha \in \ell_\infty$ , i.e.  $\psi \circ \varphi = \text{id}_{\ell_\infty}$ .

Above all,  $\varphi$  is an isometry from  $\ell_\infty$  to  $\ell_1^*$  and  $\ell_\infty \cong \ell_1^*$ .  $\square$

**Remark.** Verification of “ $\varphi$  is a bijection” is a little complex, since the value of  $\varphi \circ \psi$  at a point  $f \in \ell_1^*$  is a map from  $\ell_1$  to  $\mathbb{K}$ . Just keep in mind the domain and codomain of each map, and check equality at each point.

**Example 25.** There is a norm on  $\mathbb{K}^n$  such that

$$(\mathbb{K}^n)^* \cong \mathbb{K}^n,$$

where  $\mathbb{K}$  is equipped with  $\|\cdot\|_2$ .

**Proof.** Consider the orthonormal basis of  $\mathbb{K}^n$ , i.e.

$$e_k := (\underbrace{0, \dots, 0}_{k-1 \text{ terms}}, 1, 0, \dots), \forall k \in [n].$$

Then we set the basis of  $(\mathbb{K}^n)^*$  to be the dual basis of  $\{e_k: k \in [n]\}$ . In other words, we consider the base  $\{\varphi_j: j \in [n]\}$  defined by

$$\forall j, k \in [n]: \varphi_j(e_k) = \delta_k^j.$$

Then for all  $f \in (\mathbb{K}^n)^*$ , we can write

$$f = \sum_{j=1}^n f(e_j) \varphi_j.$$

And we define the norm on  $(\mathbb{K}^n)^*$

$$\| \cdot \|: (\mathbb{K}^n)^* \rightarrow \mathbb{R}, f \mapsto \left( \sum_{j=1}^n |f(e_j)|^2 \right)^{1/2}.$$

It's easy to verify that  $\| \cdot \|$  is a norm. Let

$$F: (\mathbb{K}^n)^* \rightarrow \mathbb{K}^n, f \mapsto (f(e_j))_{j \in [n]} = (f(e_1), \dots, f(e_n)),$$

which will be proved to be an isometry. Clearly  $F$  is surjective and  $\ker F = 0 \in (\mathbb{K}^n)^*$ , i.e.  $F$  is injective. What is left to prove is just

$$\forall f \in (\mathbb{K}^n)^*: \|F(f)\|_2 = \|f\|.$$

Let  $f \in (\mathbb{K}^n)^*$  be given. Then

$$\begin{aligned} \|F(f)\|_2 &= \left\| (f(e_j))_{j \in [n]} \right\|_2 \\ &= \left( \sum_{j=1}^n |f(e_j)|^2 \right)^{1/2} \\ &= \|f\|, \end{aligned}$$

as we wanted. □

Finally, try to finish this exercise by imitating what we have done in Example 24.

**Exercise 10.2.** Prove that  $\ell_p^* \cong \ell_q$  for  $p \in (1, \infty)$ . (Hint: see Section 8.1.1).

**Remark.** We have proved that  $\ell_p^* \cong \ell_q$  for all  $p \in [1, \infty)$ . However, it's not very easy to study  $\ell_\infty^*$ .

## 10.2 Week 10, Lecture 2

### Recall

We have learnt

$$c_0 \hookrightarrow c \hookrightarrow \ell_\infty$$

in Example 4.

### 10.2.1 More Dual Spaces

Today, our first example is the dual space of  $c_0$ .

**Example 26.** For  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ , we have

$$c_0^* \cong \ell_1.$$

**Proof.** To get an isometry, we need two contractions  $\varphi: \ell_1 \rightarrow c_0^*$  and  $\psi: c_0^* \rightarrow \ell_1$  such that  $(\varphi \circ \psi = \text{id}_{c_0^*}) \wedge (\psi \circ \varphi = \text{id}_{\ell_1})$ .

- Construction of  $\varphi$ : consider

$$\varphi: \ell_1 \rightarrow c_0^*, \alpha \mapsto \varphi_\alpha,$$

where  $(c_0^* \ni) \varphi_\alpha: c_0 \rightarrow \mathbb{K}, x \mapsto \sum_{n=1}^{\infty} x_n \alpha_n$ . For all  $\alpha \in \ell_1$ , the functional  $\varphi_\alpha$  is clearly well-defined (by Hölder's inequality) and linear. Moreover, Hölder's inequality implies  $\|\varphi_\alpha\|_{c_0^*} \leq \|\alpha\|_1$  for arbitrary  $\alpha \in \ell_1$ . And linearity of  $\varphi$  is trivial. Thus  $\|\varphi\|_{\mathcal{B}(\ell_1, c_0^*)} \leq 1$ . Above all,  $\varphi$  is a contraction.

- Construction of  $\psi$ : consider

$$\psi: c_0^* \rightarrow \ell_1, f \mapsto \psi^f,$$

where  $\psi^f$  is the sequence  $(f(e_n))_{n \in \mathbb{N}}$ . Here  $\{e_n \in c_0: n \in \mathbb{N}\}$  is still the sequence

$$\forall n \in \mathbb{N}: e_n = (\underbrace{0, \dots, 0}_{n-1 \text{ terms}}, 1, 0, \dots).$$

Now we check that  $\psi^f \in \ell_1$  and  $\|\psi^f\|_1 \leq \|f\|_{c_0^*}$ . For all  $n \in \mathbb{N}$ , let

$$A^{(n)} = \sum_{j=1}^n \text{sign}(f(e_n)) e_n.$$

Then  $A^{(n)} \in c_0$  satisfies  $\|A^{(n)}\|_\infty = 1$  for all  $n \in \mathbb{N}$ . And

$$\forall n \in \mathbb{N}: f(A^{(n)}) = \sum_{j=1}^n \text{sign}(f(e_n)) f(e_n) = \sum_{j=1}^n |f(e_n)|. \quad (23)$$



Then (23) implies  $\|f\|_{c_0^*} \geq \sum_{j=1}^n |f(e_j)|$  holds for all  $n \in \mathbb{N}$ . Let  $n \rightarrow \infty$  and we get

$$\|f\|_{c_0^*} \geq \sum_{j=1}^{\infty} |f(e_j)| = \sum_{j=1}^{\infty} |\psi_n^f| = \|\psi^f\|_1.$$

Therefore,  $\|\psi\|_{\mathcal{B}(c_0^*, \ell_1)} \leq 1$ , i.e.  $\psi$  is a contraction.

Now we check that both  $\varphi, \psi$  are the inverse of each another. For all  $\alpha \in \ell_1$ , we have  $\varphi_\alpha \in c_0^*$  and  $(\psi \circ \varphi)(\alpha) \in \ell_1$ . And  $\forall n \in \mathbb{N}$ ,

$$[(\psi \circ \varphi)(\alpha)]_n = \psi_n^{\varphi_\alpha} = \varphi_\alpha(e_n) = \alpha_n,$$

which is easy to check. And hence  $(\psi \circ \varphi)(\alpha) = \alpha$ . Since  $\alpha \in \ell_1$  is arbitrary, we get  $\psi \circ \varphi = \text{id}_{\ell_1}$  and one direction is done. For another direction, let  $f \in c_0^*$  be an arbitrary functional. Then  $\psi^f \in \ell_1$ ,  $(\varphi \circ \psi)(f) \in c_0^*$ . For an arbitrary element  $x \in c_0$ , we have

$$[(\varphi \circ \psi)(f)](x) = \varphi_{\psi^f}(x) = \sum_{n=1}^{\infty} x_n \psi_n^f.$$

What is left to prove is just  $\sum_{n=1}^{\infty} x_n \psi_n^f = f(x)$ . That's true, since continuity of  $f$  implies

$$\begin{aligned} \sum_{n=1}^{\infty} x_n \psi_n^f &= \sum_{n=1}^{\infty} x_n f(e_n) \\ &= \lim_N \sum_{n=1}^N x_n f(e_n) \\ &= \lim_N f\left(\sum_{n=1}^N x_n e_n\right) \\ &= f\left(\lim_N \sum_{n=1}^N x_n e_n\right). \end{aligned}$$

And

$$\left\| \sum_{n=1}^N x_n e_n - x \right\|_{\infty} = \|(0, 0, \dots, x_{N+1}, x_{N+2}, \dots)\|_{\infty} \rightarrow \limsup_n |x_n| = 0$$

as  $N \rightarrow \infty$ , since  $x \in c_0 \implies \limsup_n |x_n| = \lim_n |x_n| = 0$ . Thus

$$f\left(\lim_N \sum_{n=1}^N x_n e_n\right) = f(x),$$

and hence  $\varphi_{\psi f}(x) = \sum_{n=1}^{\infty} x_n \psi_n^f = f(x)$ . Since  $x$  is arbitrary, we get  $(\varphi \circ \psi)(f) = f$  for all  $f \in c_0^*$ . And  $f$  is also arbitrary, so  $\varphi \circ \psi = \text{id}_{c_0^*}$ .

Above all, both  $\varphi, \psi$  are isometries and  $c_0^* \cong \ell_1$ .  $\square$

**Example 27.** The dual space of  $C[a, b]$  is isometric to  $BV[a, b]$ , the space of functions of bounded variation on  $[a, b]$ . In other words

$$(C[a, b])^* \cong BV[a, b].$$

**Example 28.** For all  $\sigma$ -finite measure space  $(\Omega, \mathcal{F}, \mu)$  we have

$$\forall p \in [1, \infty): (L_p(\Omega, \mathcal{F}, \mu))^* \cong L_q(\Omega, \mathcal{F}, \mu).$$

### 10.2.2 Weak Convergence

Here comes important topics of Functional Analysis. Recall that for a linear normed space  $X$  and  $(x_n)_{n \in \mathbb{N}} \subseteq X, x \in X$ . We say that the sequence  $(x_n)_{n \in \mathbb{N}}$  converges to  $x$  **in norm**, if  $\lim_n \|x_n - x\| = 0$ , denoted by  $(x_n)_{n \in \mathbb{N}} \xrightarrow{\|\cdot\|} x$ . This is the classical convergence mode (hence we usually omit “in norm” and “ $\|\cdot\|$ ”, writing “ $(x_n)_{n \in \mathbb{N}}$  converges to  $x$ ” and “ $(x_n)_{n \in \mathbb{N}} \rightarrow x$ ”) and we will encounter more convergence modes now.

From now on, we don’t distinguish  $=$  and  $\cong$ , i.e. we see two isometric spaces to be same.

**Definition** (Weak Convergence). Let  $X$  be a linear normed space and  $(x_n)_{n \in \mathbb{N}} \subseteq X, x \in X$ . If for all  $f \in X^*$  we have  $\lim_n f(x_n) = f(x)$ , then we say  $(x_n)_{n \in \mathbb{N}}$  converges to  $x$  weakly, written  $(x_n)_{n \in \mathbb{N}} \xrightarrow{w} x$  or  $(x_n)_{n \in \mathbb{N}} \xrightarrow{\omega} x$  or  $(x_n)_{n \in \mathbb{N}} \rightharpoonup x$ .

And here is a similar but different notion

**Definition** (Weak\* convergence). Let  $X$  be a linear normed space and  $(f_n)_{n \in \mathbb{N}} \subseteq X^*, f \in X^*$ . If for all  $x \in X$  we have  $\lim_n f_n(x) = f(x)$ , then we say  $(f_n)_{n \in \mathbb{N}}$  converges in the weak star topology to  $f$ , written  $(f_n)_{n \in \mathbb{N}} \xrightarrow{w^*} f$  or  $(f_n)_{n \in \mathbb{N}} \xrightarrow{\omega^*} f$  or  $(f_n)_{n \in \mathbb{N}} \rightharpoonup^* f$ .

**Remark.** Weak convergence is a property of sequence in  $X$  while weak\* is a property of sequence in  $X^*$ .

Here are some examples.

**Example 29.** Consider the linear normed space  $\ell_p$ , where  $p \in [1, \infty)$ . The sequence  $(e_n)_{n \in \mathbb{N}} \subseteq \ell_p$  converges weakly to  $0(\in \ell_p)$  but doesn’t converge to 0 in norm.

**Proof.** We prove as follows

- Weak convergence: let  $f \in \ell_p^*$ , then it suffices to prove

$$\lim_n f(e_n) = f(0) = 0.$$

While  $f$  has the representation

$$\ell_q \ni \psi^f = (f(e_n))_{n \in \mathbb{N}}$$

and hence  $\lim_n f(e_n) = 0$  as we wanted.

- Convergence in norm: since  $\|e_n\|_p = 1$  for all  $n \in \mathbb{N}$ . □

It's easy to prove that

**Theorem 10.4.** Let  $X$  be a linear normed space and  $(x_n)_{n \in \mathbb{N}} \subseteq X$ , then

$$(x_n)_{n \in \mathbb{N}} \xrightarrow{\|\cdot\|} x \implies (x_n)_{n \in \mathbb{N}} \xrightarrow{*} x.$$

Now we're going to the bi-dual space, for which we need some notions relevant to natural embedding.

**Definition** (Induced Functional). Let  $X$  be a linear normed space. For all  $x \in X$ , we define a functional on  $X^*$ , denoted by  $\hat{x}$  with rules

$$\hat{x}: X^* \rightarrow \mathbb{K}, f \mapsto f(x).$$

From this definition, we have

**Claim.** For all  $x \in X$ , the functional  $\hat{x}$  is bounded, i.e.  $\hat{x} \in (X^*)^*$ .

**Proof.** For all  $f \in X^*$ , we have

$$|\hat{x}(f)| = |f(x)| \leq \|f\|_{X^*} \|x\|,$$

which implies  $\|\hat{x}\|_{X^{**}} \leq \|x\|$ . □

For convenience, we define

**Definition** (Bidual Space). For a linear normed space, the space  $(X^*)^*$  is called the bidual space of  $X$ , usually denoted by  $X^{**}$ .

The induced functional connects the space  $X$  and  $X^{**}$  in the following meaning.

**Definition** (Natural Embedding). Let  $\widehat{X} := \{\hat{x} : x \in X\}$ . The mapping  $\iota$  defined as follows is called the natural embedding of  $X$

$$\iota : X \rightarrow \widehat{X}, x \mapsto \hat{x}.$$

For  $x \in X$ ,  $\hat{x}$  is also denoted by  $\iota_x$ .

From definition, we see that  $\iota$  is surjective. Moreover, we have proved that  $\iota$  is contractive. In fact,  $\iota$  is an isometry. It suffices to show that  $\|\iota(x)\| \geq \|x\|$  for all  $x \in X$ . By definition

$$\|\iota_x\|_{X^{**}} = \sup_{\substack{f \in X^* \\ \|f\|_{X^*} \leq 1}} |\iota_x(f)| = \sup_{\substack{f \in X^* \\ \|f\|_{X^*} \leq 1}} |f(x)| \geq \|x\|,$$

the last inequality of which follows from a corollary of Hahn-Banach Theorem, see Corollary 9.5. Thus, for all  $x \in X$  we have  $\|\iota_x\|_{X^{**}} = \|x\|$ . Above all, we have  $X \cong \widehat{X}$  where the isometry is  $\iota$ .

**Theorem 10.5.** Let  $X$  be a linear normed space and  $(f_n)_{n \in \mathbb{N}} \subseteq X^*, f \in X^*$ . If  $(f_n)_{n \in \mathbb{N}} \rightharpoonup f$  in  $X^*$ , then  $(f_n)_{n \in \mathbb{N}} \xrightarrow{*} f$ .

**Proof.** Suppose  $(f_n)_{n \in \mathbb{N}} \rightharpoonup f$ , i.e.  $\forall F \in X^{**} : \lim_n F(f_n) = F(f)$ . Then for any  $x \in X$ ,  $\iota_x \in X^{**}$  and hence  $\lim_n \iota_x(f_n) = \iota_x(f)$ , i.e.  $\lim_n f_n(x) = \lim_n f(x)$ . Then by arbitrariness of  $x$ , we have proved  $(f_n)_{n \in \mathbb{N}} \xrightarrow{*} f$ .  $\square$

**Remark.** This theorem is another natural description of the inclusion  $X \hookrightarrow X^{**}$

The reversed proposition of Theorem 10.5 is wrong, as the following counter example.

**Example 30.** Here we give an example, with a linear normed space  $X$  and  $(f_n)_{n \in \mathbb{N}} \subseteq X^*, f \in X^*$  such that

$$((f_n)_{n \in \mathbb{N}} \xrightarrow{*} f) \wedge ((f_n)_{n \in \mathbb{N}} \not\rightarrow f).$$

WLOG, set  $f = 0$ , else we can replace  $(f_n)_{n \in \mathbb{N}}$  by  $(f_n - f)_{n \in \mathbb{N}}$ . We want something like the following graph

$$\begin{array}{ccc} 0 & & X = c_0 \\ \textcolor{red}{\uparrow} & & \\ * & & \\ f_n & \in & X^* = \ell_1 \\ \textcolor{red}{\downarrow} & & \\ 0 & & X^{**} = \ell_\infty \end{array}$$

It's natural to consider the case that  $X \neq X^{**}$  for which we have learnt a example:  $X = c_0$  with  $X^* = \ell_1, X^{**} = \ell_\infty$ . Consider the sequence

$$(e_n)_{n \in \mathbb{N}} \subseteq c_0,$$

where the  $m$ -th term of  $e_n$  is  $\delta_n^m$  for all  $m, n \in \mathbb{N}$ .

**Verification.** For convenience, we denote  $\varphi_a$  by  $a$ , i.e.  $a(x) = \varphi_a(x) = \sum_{n \geq 1} a_n x_n$  for all  $x$ . Then

- $(e_n)_{n \in \mathbb{N}} \xrightarrow{*} 0$ : for all  $\varphi_a \in c_0^*$  where  $a \in c_0$ , we have

$$\lim_n e_n(a) = \lim_n \sum_{j=1}^{\infty} \delta_n^j a_j = \lim_n a_n = 0$$

since  $a \in c_0$ . Thus  $(e_n)_{n \in \mathbb{N}} \xrightarrow{*} 0$ .

- $(e_n)_{n \in \mathbb{N}} \not\xrightarrow{*} 0$ : consider the sequence  $\mathbf{1} = (1)_{n \in \mathbb{N}} \in \ell_\infty$ , i.e. the sequence whose all elements are 1. Then

$$\lim_n \mathbf{1}(e_n) = \lim_n \sum_{j=1}^{\infty} \delta_n^j = \lim_n 1 = 1 \neq 0.$$

Thus  $(e_n)_{n \in \mathbb{N}} \not\xrightarrow{*} 0$ . □

Weak convergence defines weak limit, which is unique.

**Proposition 10.6.** Let  $X$  be a linear normed space. Suppose

$$((x_n)_{n \in \mathbb{N}} \rightharpoonup x) \wedge ((x_n)_{n \in \mathbb{N}} \rightharpoonup y)$$

where  $x, y \in X, (x_n)_{n \in \mathbb{N}} \subseteq X$ . Then  $x = y$ .

**Proof.** For all  $f \in X^*$ , we have

$$\lim_n f(x_n) = f(x), \lim_n f(x_n) = f(y)$$

from the definition of weak convergence. Since  $(f(x_n))_{n \in \mathbb{N}}$  is a sequence in  $\mathbb{K}$ , whose limit is unique, we have  $f(x) = f(y)$ . While  $f$  is arbitrary, if  $x \neq y$ , Hahn-Banach Theorem implies that there is some  $f_0 \in X^*$  that distinguishes  $x$  and  $y$ , contradiction. Above all,  $x = y$ . □

Finally, here is a theorem that describes weak convergence.

**Theorem 10.7.** Let  $X$  be a linear normed space. Suppose  $(x_n)_{n \in \mathbb{N}} \subseteq X$  and  $x \in X$ . Then  $(x_n)_{n \in \mathbb{N}} \rightharpoonup x$  if and only if the following 2 conditions hold:

1.  $\sup_n \|x_n\| < \infty$ ;
2.  $\exists G \subseteq X^*$  with  $\overline{\text{span } G} = X^*$  such that

$$\forall g \in G: (g(x_n))_{n \in \mathbb{N}} \rightarrow g(x).$$

**Proof.** Sufficiency: for all  $f \in \text{span } G$ , we have  $f = \sum_{j=1}^m k_j g_j$ , where  $g_j \in G, k_j \in \mathbb{K} (\forall j \in [m])$ . Thus

$$\lim_n f(x_n) = \lim_n \sum_{j=1}^m k_j g_j(x_n) = \sum_{j=1}^m k_j \lim_n g_j(x_n) = \sum_{j=1}^m k_j g_j(x) = f(x).$$

Therefore we can suppose  $G \hookrightarrow X$  with  $\overline{G} = X^*$ . By the density, for all  $\varepsilon > 0$  and all  $f \in X^*$ , there is  $g \in G$  such that  $\|g - f\|_{X^*} < \varepsilon$ . Let  $M := \sup_n \|x_n\|$  and apply the triangle inequality:

$$\begin{aligned} & |f(x_n) - f(x)| \\ & \leq |f(x_n) - g(x_n)| + |g(x_n) - g(x)| + |g(x) - f(x)| \\ & \leq \|f - g\|_{X^*} \|x_n\|_X + |g(x_n) - g(x)| + \|g - f\|_{X^*} \|x\|_X \\ & \leq \varepsilon M + |g(x_n) - g(x)| + \varepsilon \|x\|_X. \end{aligned}$$

Since  $(g(x_n))_{n \in \mathbb{N}} \rightarrow g(x)$ , there is  $N \in \mathbb{N}$  such that

$$\forall n > N: |g(x_n) - g(x)| < \varepsilon.$$

Then  $\forall n > N$

$$|f(x_n) - f(x)| \leq (M + \|x\|_X + 1)\varepsilon,$$

which implies  $\lim_n f(x_n) = f(x)$  since  $\varepsilon > 0$  is arbitrary. Therefore,  $(x_n)_{n \in \mathbb{N}} \rightarrow x$  since  $f \in X^*$  is arbitrary.

Necessity: For the first condition, consider the natural embedding  $\iota: X \rightarrow X^{**}$ . Then for all  $f \in X^*$ , the sequence  $(f(x_n))_{n \in \mathbb{N}}$  in  $\mathbb{K}$  is convergent and hence bounded. Thus

$$\forall f \in X^*: \sup_{n \in \mathbb{N}} |f(x_n)| = \sup_{n \in \mathbb{N}} |\widehat{x_n}(f)| < \infty.$$

Then, Theorem 7.6 implies that  $\sup_n \|\widehat{x_n}\|_{X^{**}}$  is finite, i.e.  $\sup_n \|x_n\|$  is finite. For the second condition, consider  $G = X$ .  $\square$

Here are some examples of application of Theorem 10.7.

**Example 31.** For  $p \in (1, \infty)$ , consider the space  $\ell_p$ . Then Theorem 10.7 means that: suppose  $(X^{(n)})_{n \in \mathbb{N}} \subseteq \ell_p$  and  $X \in \ell_p$ , then  $(X^{(n)})_{n \in \mathbb{N}} \rightharpoonup X$  if and only if the following 2 conditions hold.

1.  $\sup_n \|X^{(n)}\|_p < \infty$ ;
2.  $\forall i \in \mathbb{N}: \lim_n X_i^{(n)} = X_i$ .

**Proof.** Sufficiency: take  $G = \{e_n: n \in \mathbb{N}\}$  and the second point is done, if  $\text{span } G = \ell_q$  and that's true. To see this, let  $x \in \ell_q$  be given. Then for all  $\varepsilon > 0$ , there is some  $N \in \mathbb{N}$  such that

$$\sum_{j>N} |x_j|^q < \varepsilon^q,$$

i.e.

$$\left\| x - \sum_{j=1}^N x_j e_j \right\|_q = \left( \sum_{j>N} |x_j|^q \right)^{1/q} < \varepsilon.$$

Then apply Theorem 10.7.

Necessity: it suffices to show the second condition, since the first condition is guaranteed by Theorem 10.7. For all  $i \in \mathbb{N}$ , consider  $e_i \in \ell_q$  then we have

$$X_i = e_i(X) = \lim_n e_i(X^{(n)}) = \lim_n \sum_{j=1}^{\infty} \delta_j^i X_j^{(n)} = \lim_n X_i^{(n)}. \quad \square$$

**Remark.** Here we restrict  $p \notin \{1, \infty\}$ . Since

- For  $p = 1$ , then  $q = \infty$  and  $\text{span } G$  is not dense in  $\ell_q$ ;
- For  $p = \infty$ , the dual space  $\ell_p^*$  is almost unknown.

**Example 32.** Let a  $\sigma$ -finite measure space  $(\Omega, \mathcal{F}, \mu)$  and  $p \in (1, \infty)$  be given. Suppose  $(f_n)_{n \in \mathbb{N}} \subseteq L_p$ ,  $f \in L_p$ , then  $(f_n)_{n \in \mathbb{N}} \rightharpoonup f$  if and only if the following 2 conditions hold.

1.  $\sup_n \|f_n\|_p < \infty$ ;
2.  $\forall E \in \mathcal{F}$  with  $\mu(E) < \infty$ , we have

$$\lim_n \int_{\Omega} f_n \chi_E \, d\mu = \int_{\Omega} f \chi_E \, d\mu.$$

**Proof.** Sufficiency: take  $G = \{\chi_E: (E \in \mathcal{F}) \wedge (\mu(E) < \infty)\}$ , then  $\text{span } G$  is the set of all integrable simple functions on  $\Omega$ , which is dense in  $L_p$ . Then apply Theorem 10.7.

Necessity: The first condition is guaranteed by Theorem 10.7. And the second condition is also satisfied. Let  $E \in \Omega$  such that  $\mu(E) < \infty$  be given, then  $\chi_E \in L_q$  and hence

$$\lim_n \int_{\Omega} f \chi_E \, d\mu = \lim_n \chi_E(f_n) = \chi_E(f) = \int_{\Omega} f \chi_E \, d\mu,$$

where  $\chi_E$  means the functional

$$\varphi_{\chi_E} : L_p \rightarrow \mathbb{K}, f \mapsto \int_{\Omega} f \chi_E \, d\mu. \quad \square$$

**Remark.** Here we restrict  $p \notin \{1, \infty\}$  for the same reason as the previous example. And we used the fact that  $L_p^* \cong L_q$ , which needs the assumption that  $(\Omega, \mathcal{F}, \mu)$  is  $\sigma$ -finite.



# 11 Week 11

## 11.1 Week 11, Lecture 1

Something about biduality: let  $X$  be a linear normed space, then the space  $X^{**}$  is complete and the natural embedding  $\iota: X \rightarrow X^{**}$  is an injection keeping norms. Thus, the pair  $(X^{**}, \iota)$  is a completion of  $X$ , which ensures the existence of completion.

### 11.1.1 More about Weak Convergence and Weak-star Convergence

Let  $X$  be a linear normed space and  $(x_n)_{n \in \mathbb{N}} \subseteq X, x \in X$ . There are 3 modes of convergence on  $X$ :

$$\begin{cases} (x_n)_{n \in \mathbb{N}} \rightarrow x \iff \lim_n \|x_n - x\| = 0, \\ (x_n)_{n \in \mathbb{N}} \rightharpoonup x \iff \forall f \in X^*: \lim_n f(x_n) = f(x), \\ (x_n)_{n \in \mathbb{N}} \xrightarrow{*} x \iff \forall y \in X_*: \lim_n x_n(y) = \lim_n x(y), \end{cases}$$

where  $X_*$  is a linear normed space such that  $(X_*)^* = X$ , called the pre-dual space of  $X$ , if it exists (so we don't talk the weak-\* convergence on  $X$  is not the dual space of any linear normed space). The definition of  $(x_n)_{n \in \mathbb{N}} \xrightarrow{*} x$  is just viewing  $X$  as the dual space of  $X_*$ , as we learnt late week.

**Example 33.** Consider  $c_0^* = \ell_1, \ell_1^* = \ell_\infty$ . Then  $c_0$  is the pre-dual space of  $\ell_1$  and we can discuss the 3 modes of convergence on  $\ell_1$ :  $\rightarrow, \rightharpoonup, \xrightarrow{*}$ .

Weak convergence and weak-star convergence generalize the notion of convergence, and in fact, they generalize the notion of boundedness.

**Definition (Boundedness).** Let  $X$  be a linear normed space and  $A \subseteq X$ , then  $A$  is said to be bounded if

$$\exists M > 0: \forall x \in A, \|x\| < M.$$

**Definition (Weak Boundedness).** Let  $X$  be a linear normed space and  $A \subseteq X$ , then  $A$  is said to be weakly bounded, if

$$\forall f \in X^* \exists M_f > 0 \text{ such that } \forall x \in A, |f(x)| \leq M_f.$$

**Definition (Weak \*- Boundedness).** Let  $X$  be a linear normed space and  $A \subseteq X$ , then  $A$  is said to be weak-star bounded, if

$$\forall y \in X_* \exists M_y > 0 \text{ such that } \forall x \in A, |x(y)| \leq M_y.$$

Here is a simple exercise about them.

**Exercise 11.1.** For a linear normed space  $X$  and  $A \subseteq X$ , show that  $A$  is bounded if and only if  $A$  is weak-bounded.

Our classmate, C Li gave a convergence mode:

**Definition.** Let  $X$  be a linear normed space and  $(x_n)_{n \in \mathbb{N}} \subseteq X, x \in X$ . If  $\forall f \in X^*, (f_n)_{n \in \mathbb{N}} \subseteq X^*$  such that  $(f_n)_{n \in \mathbb{N}} \xrightarrow{*} f$ , we have  $\lim_n f_n(x_n) = f(x)$ , then  $(x_n)_{n \in \mathbb{N}}$  converges to  $x$ . We write  $(x_n)_{n \in \mathbb{N}} \xrightarrow{l} x$  to mean  $(x_n)_{n \in \mathbb{N}}$  is convergent to  $x$  in this mode.

**Claim.** This convergence mode is equivalent to weak convergence.

**Proof.** Suppose  $(x_n)_{n \in \mathbb{N}} \xrightarrow{l} x$ . Let an arbitrary functional  $f \in X^*$  be fixed. Define a sequence  $(f_n)_{n \in \mathbb{N}} \subseteq X^*$ :  $\forall n \in \mathbb{N}, f_n = f$ . Then  $(f_n)_{n \in \mathbb{N}} \xrightarrow{*} f$  and hence

$$f(x) = \lim_n f_n(x_n) = \lim_n f(x_n).$$

Since  $f$  is arbitrary, we have  $(x_n)_{n \in \mathbb{N}} \rightharpoonup x$ .

Suppose  $(x_n)_{n \in \mathbb{N}} \rightharpoonup x$ . Let arbitrary  $f \in X^*, (f_n)_{n \in \mathbb{N}} \subseteq X^*$  such that  $(f_n)_{n \in \mathbb{N}} \xrightarrow{*} f$  be given. Let  $\varepsilon > 0$  be an arbitrary number. For all  $m, n \in \mathbb{N}$ :

$$|f_n(x_n) - f(x)| \leq |f_n(x_n) - f(x_n)| + |f(x_n) - f_m(x_n)|$$

+

□

Weak topology is a topological notion:

**Definition** (Weak Topology). Let  $X$  be a topological space and  $\mathcal{F}$  be a family of functions  $f: X \rightarrow \text{cod}(f)$ , where  $\text{cod}(f)$  is a topological space for each  $f \in \mathcal{F}$ . The weak topology of  $X$  with respect to  $\mathcal{F}$  is the smallest topology such that for all  $f \in \mathcal{F}$  is continuous.

**Example 34.** For a family of topological space  $\{X_i: i \in I\}$ , the product topology on  $X := \prod_{i \in I} X_i$  is the weak topology of  $X$  with respect to the family of projection  $\{\pi_i: X \rightarrow X_i \mid i \in I\}$ .

For a linear normed space, we usually mean the weak topology with respect to  $X^*$  to be the weak topology of  $X$ . Similarly we define the weak\*-topology on  $X^*$  to be the weak topology of  $X^*$  with respect to  $\widehat{X}$ . In other words, the weak\*-topology on  $X^*$  is the smallest topology such that for all  $x \in X$ , the functional  $\hat{x}$  is continuous on  $X^*$ .

Now, we can define closure for  $A \subseteq X$  with respect to each topology, which is important to check whether a set is closed or not. To be clear, let  $\| \cdot \|, \cdot^\omega, \cdot^{\omega*}$  denote the closure operator of norm topology, weak topology and weak\* topology (if pre-dual space exists) respectively. There is a natural question:

**Question 11.1.** Let a linear normed space  $X$  and a subset  $X_0$  be given. What's the order relation for this 3 closure:  $\overline{X_0}^{\| \cdot \|}, \overline{X_0}^\omega, \overline{X_0}^{\omega*}$ ?

**Answer and Proof.** When  $X = (X_*)^*$  exists, the relation is

$$\overline{X_0}^{\| \cdot \|} \subseteq \overline{X_0}^\omega \subseteq \overline{X_0}^{\omega*}.$$

It suffices to see that  $\mathcal{T}_{\| \cdot \|} \subseteq \mathcal{T}_\omega \subseteq \mathcal{T}_{\omega*}$ , where the  $\mathcal{T}$  means the topology induced by its subscript. We know  $\mathcal{T}_{\| \cdot \|} \supseteq \mathcal{T}_\omega$  since  $\forall f \in X^*$ , we know  $f$  is continuous when  $X$  is equipped with  $\mathcal{T}_{\| \cdot \|}$ . Moreover, we know  $\mathcal{T}_\omega \supseteq \mathcal{T}_{\omega*}$  since  $\widehat{X_*} \subseteq (X_*)^{**} = X^*$ , as we wanted.

When  $X = (X_*)^*$  doesn't exist, we have just  $\overline{X_0}^{\| \cdot \|} \subseteq \overline{X_0}^\omega$ .  $\square$

### 11.1.2 Conjugate Operators

For 2 linear normed spaces  $X, Y$  and a map  $T: \mathcal{L}(X, Y)$ , we have already seen the dual spaces of  $X, Y$ . We want a linear operator  $T^*$  such that  $\forall x \in X \forall f \in Y^*$ , we have  $(T^*f)(x) = f(Tx)$ . Also we denote  $f(x) =: \langle f, x \rangle$  for  $x \in X, f \in Y^*$  as the notation of inner product. Thus, we want  $T^*$  such that

$$\begin{array}{ccc} x \in X & \xrightarrow{T} & Y \ni Tx \\ T^*f \in X^* & \xleftarrow{T^*} & Y^* \ni f \end{array} \quad \langle f, Tx \rangle = \langle T^*f, x \rangle.$$

Such  $T^*$  exists if  $T \in \mathcal{B}(X, Y)$ .

**Theorem 11.1.** Let 2 linear normed spaces  $X, Y$  and a map  $T \in \mathcal{B}(X, Y)$  be given, then:

1. such  $T^*$  exists and is unique;
2. we have  $\|T\| = \|T^*\|$ .

**Proof.** Existence and Uniqueness: we define  $T^*$  as the following commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{T} & Y \\ & \searrow T^*f & \downarrow f \\ & & \mathbb{K} \end{array}$$

i.e.  $T^*: Y^* \rightarrow X^*, f \mapsto f \circ T$ . For all  $f \in Y^*$  we have  $f \circ T \in X^*$  since  $\|f \circ T\| \leq \|f\| \cdot \|T\|$ . This definition satisfies:

$$\forall x \in X, \forall f \in Y^*: \langle f, Tx \rangle = \langle T^*f, x \rangle,$$

which can be easily verified. Existence is ensured. Furthermore, if another map  $F: Y^* \rightarrow X^*$  satisfies:

$$\forall x \in X, \forall f \in Y^*: \langle f, Tx \rangle = \langle Ff, x \rangle,$$

then we have  $F = T^*$ . To see this, let an arbitrary  $f \in Y^*$  be given and we prove that  $T^*f = Ff$ . For all  $x \in X$ :

$$\langle Ff, x \rangle = \langle f, Tx \rangle = \langle T^*f, x \rangle,$$

thus  $Ff = T^*f$ . Above all,  $F = T^*$  and hence  $T^*$  is unique.

Norms coincide: one direction was done since  $\|f \circ T\| \leq \|f\| \cdot \|T\|$ . For another direction: apply Corollary 10.2

$$\begin{aligned} \|T\| &= \sup_{\substack{x \in X \\ \|x\| \leq 1}} \|Tx\|_Y \\ &= \sup_{\substack{x \in X \\ \|x\| \leq 1}} \sup_{\substack{f \in Y^* \\ \|f\| \leq 1}} |f(Tx)| \\ &= \sup_{\substack{x \in X \\ \|x\| \leq 1}} \sup_{\substack{f \in Y^* \\ \|f\| \leq 1}} |(T^*f)(x)| \\ &\leq \sup_{\substack{x \in X \\ \|x\| \leq 1}} \sup_{\substack{f \in Y^* \\ \|f\| \leq 1}} \|T^*\| \cdot \|f\| \cdot \|x\| \\ &\leq \|T^*\|, \end{aligned}$$

as we wanted. Above all,  $\|T^*\| = \|T\|$ . □

**Remark.** You should check that: for all linear normed spaces  $X, Y, Z$  and bounded linear operators  $X \xrightarrow{f} Y \xrightarrow{g} Z$ , we have  $g \circ f$  is a bounded linear map with  $\|g \circ f\| \leq \|g\| \cdot \|f\|$ .

Here is an example from Linear Algebra.

**Example 35.** Let a linear map  $T: \mathbb{K}^n \rightarrow \mathbb{K}^m$  be given, then it must be bounded. Thus there is  $T^*: (\mathbb{K}^m)^* \rightarrow (\mathbb{K}^n)^*$ . Though we have proved that  $(\mathbb{K}^n)^* \cong \mathbb{K}^n$ , we won't apply the result here. Consider the standard basis of  $\mathbb{K}^n$ :  $\{e_j: j \in [n]\}$ , where

$$\forall j \in [n]: e_j = (\underbrace{0, \dots, 0}_{j-1 \text{ terms}}, 1, \dots).$$

Similarly, consider the basis of  $\mathbb{K}^m$ :  $\{\mu_k : k \in [m]\}$ . Then the operator  $T$  is corresponded to a matrix  $M_T$ , since

$$\forall j \in [n]: Te_j = \sum_{k=1}^m a_{j,k} \mu_k,$$

whose corresponding matrix is  $M_T = (a_{j,k})_{\substack{j \in [n] \\ k \in [m]}}$ . And now we consider the basis of  $(\mathbb{K}^n)^*$  and  $(\mathbb{K}^m)^*$  to get the matrix of  $T^*$ . There is a related exercise, see Exercise 9.4. Let

$$Y_k := \text{span}\{e_j : j \in [n] \setminus \{k\}\}, \forall k \in [n]$$

then Exercise 9.4 implies  $\exists e_k^* \in (\mathbb{K}^n)^*$  such that  $e_k^*(e_k) = d(e_k, Y_k) > 0$  and  $e_k^*|_{Y_k} = 0$ . In fact,  $d(e_k, Y_k) = 1$ . Thus  $e_k^*(e_j) = \delta_j^k$ .

**Definition** (Dual Basis). For linear normed space  $\mathbb{K}^n$ , the basis of  $(\mathbb{K}^n)^*$  defined as above, i.e.  $\{e_k^* \in (\mathbb{K}^n)^* : k \in [n]\}$  is called the dual basis of  $\{e_k \in \mathbb{K}^n : k \in [n]\}$ .

Now we are going to find  $M_{T^*}$ . Suppose  $M_{T^*} = (b_{j,k})_{\substack{j \in [m] \\ k \in [n]}}$ . From

$$\begin{array}{ccc} T^* & : & (\mathbb{K}^m)^* \longrightarrow (\mathbb{K}^n)^* \\ & & \cup \qquad \qquad \qquad \cup \\ & & \{\mu_j^* : j \in [m]\} \qquad \qquad \{e_k^* : k \in [n]\} \end{array}$$

we get  $T^* \mu_j^* = \sum_{l=1}^n b_{j,l} e_l^*$  for all  $j \in [m]$ . From the definition of  $T^*$ : we have

$$\langle T^* \mu_j^*, e_l \rangle = \langle \mu_j^*, Te_l \rangle,$$

and apply  $T^* \mu_j^* = \sum_{k=1}^n b_{j,k} e_k^*$ ,  $Te_l = \sum_{k=1}^m a_{l,k} \mu_k$  to get

$$b_{j,l} = a_{l,j}$$

for all  $l \in [n], j \in [m]$ . In other words,  $M_{T^*} = {}^t M_T$ .

Example 35 means: transpose matrices are the special case of dual operators. Recall that, for all matrix  $M$ , we have  ${}^{tt}M = M$  and hence

$$\forall T \in \mathcal{B}(X, Y): T^{**} = T,$$

in some sense for 2 arbitrary linear normed spaces  $X, Y$ . Above all

**Corollary 11.2.** For arbitrary linear normed spaces  $X, Y, Z$  and  $T \in \mathcal{B}(X, Y), S \in \mathcal{B}(Y, Z)$ :

1.  $T^{**} \circ \iota_X = \iota_Y \circ T$ , where  $\iota_X, \iota_Y$  is the natural embedding of  $X, Y$  respectively;
2.  $(\text{id}_X)^* = \text{id}_{X^*}$ ;
3.  $(T \circ S)^* = S^* \circ T^*$ .

**Proof.** What we want is: the following diagram commutes

$$\begin{array}{ccc} X & \xrightarrow{\iota_X} & X^{**} \\ T \downarrow & & \downarrow T^{**} \\ Y & \xrightarrow{\iota_Y} & Y^{**} \end{array}$$

Remember that  $Y^{**} = (Y^*)^*$ . For all  $x \in X, f \in Y^*$ , we have

$$\begin{aligned} [(\iota_Y \circ T)(x)](f) &= [\iota_Y(Tx)](f) \\ &= [\widehat{Tx}](f) \\ &= f(Tx) \\ &= (f \circ T)(x), \end{aligned}$$

and

$$\begin{aligned} [(T^{**} \circ \iota_X)(x)](f) &= [T^{**}(\hat{x})](f) \\ &= \langle T^{**}(\hat{x}), f \rangle \\ &= \langle \hat{x}, T^* f \rangle \\ &= \langle T^* f, x \rangle \\ &= \langle f \circ T, x \rangle \\ &= (f \circ T)(x). \end{aligned}$$

Therefore,

$$[(\iota_Y \circ T)(x)](f) = [(T^{**} \circ \iota_X)(x)](f)$$

holds for all  $f \in Y^*, x \in X$ , i.e.

$$[(\iota_Y \circ T)(x)] = (T^{**} \circ \iota_X)(x)$$

holds for all  $x \in X$ . Hence

$$\iota_Y \circ T = T^{**} \circ \iota_X.$$

The rest is easy. □

## 11.2 Week 11, Lecture 2

### Recall

We studied conjugate operators last time: for  $f \in Y^*$  and  $T \in \mathcal{B}(X, Y)$ , the functional  $T^*f$  is defined by

$$\langle T^*f, x \rangle = \langle f, Tx \rangle, \forall x \in X.$$

### 11.2.1 Compact Operators and Finite-rank Operators

This lecture was given by our classmate Kangwen Zhang.

**Definition** (Compact Operator). Let  $X, Y$  be 2 linear normed spaces and  $T \in \mathcal{L}(X, Y)$ . If for all  $A \subseteq X$  that is bounded, we have  $T(A)$  is relatively compact in  $Y$ , then  $T$  is said to be a **compact operator**. The set of all compact operators from  $X$  to  $Y$  is denoted by  $\mathcal{C}(X, Y)$ . In the case that  $Y = X$ , the set  $\mathcal{C}(X, Y)$  is written  $\mathcal{C}(X)$ .

**Remark.** The following statements are equivalent:

1.  $T$  is compact;
2.  $T(B_X)$  is relatively compact, where  $B_X := \{x \in X : \|x\| < 1\}$ ;
3. For all bounded sequence  $(x_n)_{n \in \mathbb{N}} \subseteq X$ , there is a subsequence  $(x_{n_k})_{k \in \mathbb{N}}$  such that  $(Tx_{n_k})_{k \in \mathbb{N}}$  converges in  $Y$ .

**Remark.** The  $\mathcal{C}$  of  $\mathcal{C}(X, Y)$  means “compactness”, not “continuity”.

**Definition** (Finite-Rank Operator). Let 2 linear normed spaces  $X, Y$  be given. If  $T \in \mathcal{B}(X, Y)$  satisfies  $\dim(\text{Im } T) < \infty$ , then  $T$  is called an operator of **finite rank**. The set of all finite-rank operators from  $X$  to  $Y$  is denoted by  $\mathcal{F}_r(X, Y)$ . In the case that  $Y = X$ , the set  $\mathcal{F}_r(X, Y)$  is written  $\mathcal{F}_r(X)$ .

**Remark.** There are some books that don’t require finite-rank operators to be bounded. We require this, since we don’t care much about operators that are not continuous.

We don’t need a compact operator to be continuous in the definition, because

**Proposition 11.3.** For 2 linear normed spaces  $X, Y$ , we have

$$\mathcal{C}(X, Y) \subseteq \mathcal{B}(X, Y).$$

**Proof.** suppose  $T \in \mathcal{C}(X, Y)$ , then  $T(B_X)$  is relatively compact, i.e.  $\overline{T(B_X)}$  is compact and hence bounded. Therefore,  $T(B_X)$  is bounded and so is  $T$ .  $\square$

In fact, we have

**Proposition 11.4.** For 2 linear normed spaces  $X, Y$ , we have

$$\mathcal{C}(X, Y) \hookrightarrow B(X, Y).$$

**Proof.** For addition: let  $S, T \in \mathcal{C}(X, Y)$  be 2 arbitrary compact operators. Let  $(x_n)_{n \in \mathbb{N}} \subseteq X$  be an arbitrary bounded sequence. Then there is a subsequence (still denoted by  $(x_n)_{n \in \mathbb{N}}$ ) such that  $(Tx_n)_{n \in \mathbb{N}}$  converges in  $Y$ . Furthermore, since  $S$  is compact: there is a subsequence  $(x_{n_k})_{k \in \mathbb{N}}$  such that  $(Sx_{n_k})_{k \in \mathbb{N}}$  converges in  $Y$ . Therefore,  $(x_{n_k})_{k \in \mathbb{N}}$  is a subsequence of the original  $(x_n)_{n \in \mathbb{N}}$  such that  $((T + S)x_{n_k})_{k \in \mathbb{N}}$  converges. Above all,  $T + S$  is compact.

For multiplication with scalars: let  $T \in \mathcal{C}(X, Y)$  and  $\lambda \in \mathbb{K}$ . Let  $(x_n)_{n \in \mathbb{N}} \subseteq X$  be an arbitrary bounded sequence. Then there is a subsequence such that  $(Tx_{n_k})_{k \in \mathbb{N}}$  converges in  $Y$  and then  $(\lambda T(x_{n_k}))_{k \in \mathbb{N}}$  also converges. Therefore,  $\lambda T \in \mathcal{C}(X, Y)$ .  $\square$

**Remark.** Proof of addition part is similar to the proof this proposition: a bounded sequence  $(x_n, y_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^2$  has a convergent subsequence. Let  $(x_{n_k})_{k \in \mathbb{N}} \subseteq (x_n)_{n \in \mathbb{N}}$  be a convergent subsequence, then  $(y_{n_k})_{k \in \mathbb{N}}$  is a bounded sequence in  $\mathbb{R}$ . Take a convergent subsequence of  $(y_{n_k})_{k \in \mathbb{N}}$  and we're done.

Similar to Exercise 8.1, we have

**Proposition 11.5.** Let a linear normed space  $X$  and a Banach space  $Y$  be given. The space  $\mathcal{C}(X, Y)$  is a closed subspace of  $\mathcal{B}(X, Y)$ .

**Remark.** This implies that  $\mathcal{C}(X, Y)$  is a Banach space whenever  $Y$  is a Banach space, since  $\mathcal{B}(X, Y)$  is a Banach space.

**Proof.** Let  $(T_n)_{n \in \mathbb{N}} \subseteq \mathcal{C}(X, Y)$  be a convergent sequence of compact operators with limit  $T \in \mathcal{B}(X, Y)$ . We want  $T \in \mathcal{C}(X, Y)$ . Since  $Y$  is a Banach space,  $T(B_X) \subseteq Y$  is relatively compact if and only if  $T(B_X)$  is totally bounded (see Corollary 6.4). Let  $\varepsilon > 0$  be an arbitrary number. Suppose  $N \in \mathbb{N}$  satisfies:  $\forall n > N, \|T_n - T\| < \varepsilon$ . Fix some  $n > N$  and we have  $T_n(B_X)$  is relatively compact and hence totally bounded in  $Y$ , i.e.  $\exists \{x_j \in B_X \mid j \in [m]\}$  such that

$$T_n(B_X) \subseteq \bigcup_{j \in [m]} B(T_n x_j, \varepsilon).$$



Now we have a claim:  $\{Tx_j \mid j \in [m]\}$  is a  $4\varepsilon$ -net of  $T(B_X)$ . To see this, let an arbitrary element  $x \in B_X$  be given, then there is some  $j \in [m]$  such that  $\|T_n x_j - T_n x\| < \varepsilon$ . Thus

$$\begin{aligned}\|Tx_j - Tx\| &\leq \|Tx_j - T_n x_j\| + \|T_n x_j - T_n x\| + \|T_n x - Tx\| \\ &\leq \|T - T_n\| \|x_j\| + \varepsilon + \|T_n - T\| \|x\| \\ &\leq \varepsilon + \varepsilon + \varepsilon \\ &< 4\varepsilon.\end{aligned}$$

where we used triangle inequality, definition of operator norm, definition of unit ball and so on. Therefore,  $\{Tx_j \in T(B_X) \mid j \in [m]\}$  is a  $4\varepsilon$ -ball of  $T(B_X)$  and hence  $T(B_X)$  is totally bounded.  $\square$

Now we prove that: an operator of finite rank must be compact.

**Proposition 11.6.** Let 2 linear normed spaces  $X, Y$  be given. We have

$$\mathcal{F}_r(X, Y) \subseteq \mathcal{C}(X, Y).$$

**Proof.** Since  $\dim(\text{Im } T) < \infty$ , we know that  $\text{Im } T$  is complete. Then for all bounded set  $A \subseteq X$ , the set  $T(A) \subseteq \text{Im } T$  is bounded and hence  $\overline{T(A)}$  is bounded. While  $\dim(\text{Im } T) < \infty$ , we know  $\overline{T(A)}$  is compact from Theorem 6.8. Thus  $T(E)$  is relatively compact and hence  $T$  is compact.  $\square$

In fact, the set of all operators of finite rank is a subspace of the space  $\mathcal{C}(X, Y)$ .

**Proposition 11.7.** Let 2 linear normed spaces  $X, Y$  be given. Then

$$\mathcal{F}_r(X, Y) \hookrightarrow \mathcal{C}(X, Y).$$

**Proof.** Clearly  $0 \in \mathcal{F}_r(X, Y)$ . We prove addition only.  $\square$

It's natural to consider the case that the codomain of an operator is a finite dimensional space:

**Proposition 11.8.** Let 2 finite-dimensional linear normed spaces  $X, Y$  be given. Then we have

$$\mathcal{L}(X, Y) = \mathcal{F}_r(X, Y) = \mathcal{C}(X, Y).$$

**Remark.** The space  $X$  is also needed to be finite-dimensional. To see this, set  $Y = \mathbb{K}$  and find some space  $X$  that  $X^\# \neq X^*$ .

**Proof.** It suffices to prove that  $\forall T \in \mathcal{L}(X, Y)$ , we have  $T \in \mathcal{F}_r(X, Y)$ . Since  $\dim(\text{Im } T) \leq \dim(Y) < \infty$ , what needs to be proved is just  $T \in \mathcal{B}(X, Y)$ . It's clear that  $\mathcal{L}(\mathbb{K}^m, \mathbb{K}^n) = \mathcal{B}(\mathbb{K}^m, \mathbb{K}^n)$  and  $\mathcal{L}(X, Y) = \mathcal{B}(X, Y)$  follows from the following commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{T} & Y \\ \varphi \uparrow & & \uparrow \psi \\ \mathbb{K}^m & \xrightarrow[\exists! L_T]{} & \mathbb{K}^n \end{array}$$

where  $\varphi, \psi$  are isomorphisms. In other words,  $T = \psi^{-1} \circ L_T \circ \varphi$  is composition of bounded linear maps.  $\square$

The following proposition tells some structure about the 3 spaces  $\mathcal{C}(X, Y)$ ,  $\mathcal{F}_r(X, Y)$  and  $\mathcal{B}(X, Y)$ .

**Proposition 11.9.** Let 3 linear normed spaces  $X, Y, Z$  and bounded linear operators  $T: X \rightarrow Y, S: Y \rightarrow Z$  be given. Then  $S \circ T$  is compact whenever one of  $S, T$  is compact.

**Proof.** We consider the following 2 cases respectively.

**Case 1:**  $S$  is compact. Since  $T(B_X)$  is a bounded subset of  $Y$  and  $T$  is compact, we know that  $S(T(B_X))$  is relatively compact, i.e.  $S \circ T(B_X)$  is relatively compact. Thus  $T \circ S$  is compact.

**Case 2:**  $T$  is compact. For all bounded sequence  $(x_n)_{n \in \mathbb{N}} \subseteq X$ , the sequence  $(Tx_n)_{n \in \mathbb{N}}$  is bounded in  $Y$ , then

$$(S(Tx_{n_k}))_{k \in \mathbb{N}} = ((S \circ T)x_{n_k})_{k \in \mathbb{N}}$$

converges in  $Z$  for some subsequence  $(x_{n_k})_{k \in \mathbb{N}}$ . Thus  $S \circ T$  is compact.  $\square$

**Corollary 11.10.** Therefore,  $\mathcal{C}(X)$  is a two-sided ideal of the ring  $\mathcal{B}(X)$ .

Here is an example of compact operator.

**Example 36.** Consider the infinite dimensional matrix  $T = (a_{i,j})_{i,j \in \mathbb{N}}$  with elements such that  $\sum_{i,j \geq 1} |a_{i,j}|^2 < \infty$ . The operator

$$T: \ell_2 \rightarrow \ell_2, x \mapsto Tx := \left( \sum_{j=1}^{\infty} a_{n,j} x_j \right)_{n \in \mathbb{N}}$$

is compact.

**Proof.** We prove that  $T$  is a limit of a sequence of compact operators sequence, and  $T$  is bounded then Proposition 11.5 can be applied.

First,  $T$  is bounded: from Hölder's inequality, we get

$$\forall n \in \mathbb{N}: |Tx_n| = \left| \sum_{j=1}^{\infty} a_{n,j} x_j \right| \leq \|x\|_2 \left( \sum_{j=1}^{\infty} |a_{n,j}|^2 \right)^{1/2},$$

and hence

$$\|Tx\|_2^2 \leq \|x\|_2^2 \left( \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{i,j}|^2 \right).$$

Above all

$$\|T\| \leq \left( \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} |a_{i,j}|^2 \right)^{1/2}. \quad (24)$$

Secondly, let  $(T_n)_{n \in \mathbb{N}}$  be defined as follows

$$\forall n \in \mathbb{N}, T_n: \ell_2 \rightarrow \ell_2, x \mapsto \left( \sum_{j=1}^n a_{1,j} x_j, \dots, \sum_{j=1}^n a_{n,j} x_j, 0, \dots \right).$$

Then  $\dim(\text{Im } T_n) \leq n$  and  $T_n$  is bounded (since  $\|T_n\| \leq \|T\|$ ). Hence

$$(T_n)_{n \in \mathbb{N}} \subseteq \mathcal{F}_r(\ell_2) \hookrightarrow \mathcal{C}(\ell_2).$$

Let suitable elements in (24) be 0 and we get

$$\begin{aligned} \|T - T_n\| &\leq \left( \sum_{i=n+1}^{\infty} \sum_{j=n+1}^{\infty} |a_{i,j}|^2 \right)^{1/2} \\ &= \left( \sum_{i,j \geq 1} |a_{i,j}|^2 - \sum_{1 \leq i,j \leq n} |a_{i,j}|^2 \right)^{1/2} \\ &\rightarrow 0 (n \rightarrow \infty). \end{aligned}$$

Thus  $T = \lim_n T_n$  and hence  $T \in \mathcal{C}(\ell_2)$ . □

For compact operators, relation between convergence modes is different.

**Proposition 11.11.** Suppose  $X, Y$  are 2 linear normed spaces and  $T \in \mathcal{C}(X, Y)$ . Then for all  $X \supseteq (x_n)_{n \in \mathbb{N}} \rightharpoonup x$ , we have  $(Tx_n)_{n \in \mathbb{N}} \rightarrow Tx$ .

**Remark.** In general: if  $T \in \mathcal{B}(X, Y)$ , we have  $(Tx_n)_{n \in \mathbb{N}} \rightharpoonup Tx$ .

**Proof.** For an arbitrary  $f \in Y^*$ , we have  $T^*f = f \circ T \in X^*$ . By definition of weak convergence:

$$\lim_n T^*f(x_n) = \lim_n T^*f(x),$$

i.e.

$$\lim_n f(Tx_n) = f(Tx).$$

Since  $f \in Y^*$  is arbitrary, we proved  $(Tx_n)_{n \in \mathbb{N}} \rightharpoonup Tx$ . □

And there is an example such that  $(x_n)_{n \in \mathbb{N}} \rightharpoonup x$  and  $(Tx_n)_{n \in \mathbb{N}} \not\rightharpoonup Tx$ , where  $T$  is a bounded operator.

**Example 37** (Counter-example). Consider  $X = c_0$  and  $(e_n)_{n \in \mathbb{N}} \subseteq c_0$ , where  $e_n$  is the sequence whose all elements are 0 except for the  $n$ -th element being 1. Then consider  $\text{id}_{c_0}: c_0 \rightarrow c_0$ , we have  $(e_n)_{n \in \mathbb{N}} \rightharpoonup 0$  but  $(e_n)_{n \in \mathbb{N}} \not\rightharpoonup 0$ . See Example 29.

Thus, compact operators are special.

**Proof of Proposition 11.11.**

□

There is a important theorem about compact operators and conjugate operators.

**Theorem 11.12** (Schauder). Let  $X, Y$  be 2 given linear normed spaces and  $T \in \mathcal{C}(X, Y)$ . Then  $T^* \in \mathcal{C}(Y^*, X^*)$ .

For another direction: if  $Y$  is complete, then  $T^* \in \mathcal{C}(Y^*, X^*)$  implies  $T \in \mathcal{C}(X, Y)$ .

**Proof.**

□

Here is a result for finite-rank operators, similar to Proposition 11.9.

**Proposition 11.13.** Let 3 linear normed spaces  $X, Y, Z$  and bounded linear operators  $T: X \rightarrow Y, S: Y \rightarrow Z$  be given. Then  $S \circ T$  is finite-rank whenever one of  $S, T$  is of finite rank.

**Proof.**

□

**Corollary 11.14.**  $\mathcal{F}_r(X)$  is a two-sided ideal of the ring  $\mathcal{B}(X)$ .

## 12 Week 12

### 12.1 Week 12, Lecture 1

This lecture was given by Lingxuan Wu.

#### 12.1.1 Inner Product and Inner Product Space

**Definition** (Semi-inner Product). Let  $X$  be a vector space over the field  $\mathbb{K}$ . A map  $u: X \times X \rightarrow \mathbb{K}$  is said to be a **semi-inner product**, if  $\forall x, y, z \in X$  and  $\forall \alpha, \beta \in \mathbb{K}$ , the following properties are satisfied:

1. linearity:  $u(\alpha x + \beta y) = \alpha u(x, y) + \beta u(x, z)$ ;
2. semi-positive definite:  $u(x, x) \geq 0$ ;
3. conjugate-symmetry:  $u(x, y) = \overline{u(y, x)}$ .

**Definition** (Inner Product). An inner product is a semi-inner product that is positive definite. In other words:  $u$  is an inner product if  $u$  is a semi-inner product and  $u(x, x) = 0 \implies x = 0$ .

**Remark.** Let  $u$  be a semi-inner product on  $X$ .

1. The operator  $u$  is conjugate-linear (or anti-linear) for the second entry. This follows from linearity and conjugate-symmetry.
2. If one of  $x, y = 0$ , then  $u(x, y) = 0$ .

**Definition.** A vector space with an inner product is called an inner product space.

**Notation.** Suppose there is an inner product  $u: X \times X \rightarrow \mathbb{K}$ , then  $u(x, y)$  is also denoted by  $\langle x, y \rangle$ ,  $(x, y)$ ,  $(x | y)$  or  $\langle x | y \rangle$ . The last 2 symbols are usually used in **Quantum Mechanics**, also known as Dirac notation. And I would like to use  $\langle x, y \rangle$ .

Here are some examples of inner product spaces.

**Example 38.** The space of square-summable sequence  $\ell_2$  is an inner product space, whose inner product is

$$\langle \cdot, \cdot \rangle : \ell_2 \times \ell_2 \rightarrow \mathbb{K}, (x, y) \mapsto \sum_{n \geq 1} x_n \bar{y}_n. \quad (25)$$

Hölder's inequality implies that (25) is well-defined.

**Example 39.** Let  $(\Omega, \Sigma, \mu)$  be a measure space. Then the space of square-integrable functions  $L_2(\Omega, \Sigma, \mu)$  (denoted by  $L_2$  for short) is an inner product space, whose inner product is

$$\langle \cdot, \cdot \rangle : L_2 \times L_2 \rightarrow \mathbb{K}, (x, y) \mapsto \int_{\Omega} f \bar{g} \, d\mu.$$

Hölder's inequality implies that  $\langle \cdot, \cdot \rangle$  is well-defined. Example 38 is a special case of this example.

**Remark.** This is a classical Hilbert space in Quantum Mechanics, whose elements are wave functions.

**Theorem 12.1** (Cauchy-Schwarz Inequality). Let  $X$  be an inner product space. For all  $x, y \in X$ , we have

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle.$$

**Proof.** Let  $\alpha \in \mathbb{K}$  be an arbitrary element. Then the □

**Remark.** The Cauchy-Schwarz inequality comes to be a equality if and only if  $x, y$  are linearly dependent.

**Proof.** □

Cauchy-Schwarz inequality implies that an inner product space is a normed space:

**Corollary 12.2.** An inner product  $\langle \cdot, \cdot \rangle$  on a vectors space  $X$  induces a norm on  $X$  as follows:

$$\| \cdot \| : X \rightarrow \mathbb{R}, x \mapsto \sqrt{\langle x, x \rangle}.$$

**Proof.** Just verify the axioms:

- 
- □

**Corollary 12.3.** The inner product of an inner product space is a continuous function on  $(X \times X, \| \cdot \|_1)$ . Here the space  $X \times X$  is equipped with only product topology.

**Proof.** Since  $X$  is first countable, we know  $X \times X$  is first countable. Then we prove that  $\langle \cdot, \cdot \rangle$  keeps limit of sequences. Let  $x, y \in X$  be

given and  $(x_n)_{n \in \mathbb{N}} \rightarrow x, (y_n)_{n \in \mathbb{N}} \rightarrow y$ . Then Cauchy-Schwarz inequality implies

$$\begin{aligned}
 |\langle x_n, y_n \rangle - \langle x, y \rangle| &\leq |\langle x_n, y_n \rangle - \langle x_n, y \rangle| + |\langle x_n, y \rangle - \langle x, y \rangle| \\
 &= |\langle x_n, y_n - y \rangle| + |\langle x_n - x, y \rangle| \\
 &\leq \|x_n\| \|y_n - y\| + \|x_n - x\| \|y\| \\
 &\leq \sup_{j \in \mathbb{N}} \|x_j\| \|y_n - y\| + \|x_n - x\| \|y\| \\
 &\rightarrow 0 (n \rightarrow \infty).
 \end{aligned}$$

Since  $(x_n)_{n \in \mathbb{N}}$  converges in  $X$  implies that  $(x_n)_{n \in \mathbb{N}}$  is bounded in  $X$ .  $\square$

Inner product spaces have some interesting properties:

**Theorem 12.4** (Polar Identity). Let  $X$  be an inner product space over  $\mathbb{K}$  and  $x, y \in X$  are arbitrary elements.

- If  $\mathbb{K} = \mathbb{R}$ :

$$\langle x, y \rangle = \frac{\langle x + y, x + y \rangle - \langle x - y, x - y \rangle - \langle y, y \rangle}{4}.$$

- If  $\mathbb{K} = \mathbb{C}$ :

$$\langle x, y \rangle = \sum_{k=0}^3 \frac{i^k}{4} \langle x + i^k y, x + i^k y \rangle.$$

**Remark.** Polar identity is important, since it rewrites the inner product of 2 elements as sum of inner product of same elements.

**Exercise 12.1.** Prove Theorem 12.4.

**Theorem 12.5** (Parallelogram Law). Let  $X$  be an inner product space over  $\mathbb{K}$ . For all  $x, y \in X$ , we have

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2). \quad (26)$$

**Proof.** Write the norm in the form of inner product and apply linearity.  $\square$

**Remark.** This theorem has geometric meaning as follows:

In fact, a norm of a linear normed space is induced by an inner product if and only if it satisfies (26).

**Theorem 12.6.** Let  $X$  be a linear normed space over  $\mathbb{K}$  whose norm satisfies (26), then there is an inner product  $\langle \cdot, \cdot \rangle$  on  $X$  such that  $\|\cdot\|$  is induced by  $\langle \cdot, \cdot \rangle$ .

**Proof.**

$\square$

### 12.1.2 Hilbert Space

**Definition** (Hilbert Space). An inner product space  $H$  is said to be a Hilbert space, if  $H$  is complete with respect to the norm induced by the inner product.

## 12.2 Week 12, Lecture 2

### 12.2.1 Orthogonality

In this lecture, we assume that  $\mathcal{H}$  is a Hilbert space and  $H$  is an inner product space.

**Definition.** Let  $x, y \in H$ . We say  $x$  is **orthogonal to**  $y$ , denoted by  $x \perp y$ , if  $\langle x, y \rangle = 0$ .

**Definition.** Let  $A, B \subseteq H$ . We say  $A$  is **orthogonal to**  $B$ , denoted by  $A \perp B$ , if  $\langle x, y \rangle = 0$  holds for all  $x \in A, y \in B$ . For the case that  $A = \{x\}$  is a singleton, we write  $x \perp B$  instead of  $A \perp B$ .

**Remark.** We have  $A \perp \emptyset$  for all  $A \subseteq H$  as a “vacuous truth”.

**Definition.** A subset  $\mathcal{E} \subseteq H$  is said to be an **orthogonal set**, if for all  $x \in \mathcal{E}$ , we have  $x \perp \mathcal{E} \setminus \{x\}$ . A subset  $\mathcal{E}$  is said to be an **orthonormal set**, if it's an orthogonal set and for all  $x \in \mathcal{E}$ , we have  $\|x\| = 1$ , i.e.  $\langle x, x \rangle = 1$ .

**Example 40.** Consider  $H = \mathbb{K}^2$ . The set  $\{e_1 = (1, 0), e_2 = (0, 1)\}$  is an orthonormal set.

**Example 41.** Consider  $H = \ell_2$  over the field  $\mathbb{K}$ . The set  $\{e_n : n \in \mathbb{N}\}$  is an orthonormal set.

As a generalization of the classical Pythagoras Theorem on  $\mathbb{R}^2$ , we have

**Theorem 12.7** (Pythagoras). Let  $\mathcal{E} \subseteq H$  be an orthogonal set. Then for all  $x_j \in \mathcal{E}$  and  $k_j \in \mathbb{K}$ , where  $j \in [n]$ , we have

$$\left\| \sum_{j=1}^n k_j x_j \right\|^2 = \sum_{j=1}^n |k_j|^2 \|x_j\|^2. \quad (27)$$

**Proof.** Prove by mathematical induction. For  $n = 1$ , it's trivial. For  $n = 2$ , we have . Then suppose (27) holds for  $n$ , now we prove (27) is true for  $n + 1$ :

$$\left\| \sum_{j=1}^{n+1} k_j x_j \right\|^2 = \left\| \sum_{j=1}^n k_j x_j + k_{n+1} x_{n+1} \right\|^2,$$



and  $x_{n+1}$  is orthogonal to  $\sum_{j=1}^n k_j x_j$ , thus by the case  $n = 2$  and assumption about the case for  $n$ :

$$\left\| \sum_{j=1}^{n+1} k_j x_j \right\|^2 = \left\| \sum_{j=1}^n k_j x_j \right\|^2 + \|k_{n+1} x_{n+1}\|^2 = \sum_{j=1}^{n+1} \|k_j x_j\|^2. \quad \square$$

**Theorem 12.8** (Bessel's Inequality). Let  $\{e_n : n \in \mathbb{N}\}$  be an orthonormal set. Then for all  $h \in H$ :

$$\sum_{n \geq 1} |\langle h, e_n \rangle|^2 \leq \|h\|^2.$$

**Proof.** For all  $n \in \mathbb{N}$ , consider the element

$$h_n := h - \sum_{j=1}^n \langle h, e_j \rangle e_j.$$

Then for all  $k \in [n]$ :  $h_n \perp e_k$ , since

$$\langle h_n, e_k \rangle = \langle h, e_k \rangle - \sum_{j=1}^n \langle h, e_j \rangle \langle e_j, e_k \rangle = 0.$$

Apply Theorem 27 and we get that:

$$\|h\|^2 = \left\| h_n + \sum_{j=1}^n \langle h, e_j \rangle e_j \right\|^2 = \|h_n\|^2 + \sum_{j=1}^n |\langle h, e_j \rangle|^2 \geq \sum_{j=1}^n |\langle h, e_j \rangle|^2,$$

holds for all  $n \in \mathbb{N}$ . Pass  $n \rightarrow \infty$  and we're done.  $\square$

We can generalize this theorem to the case where  $\{e_n : n \in \mathbb{N}\}$  is replaced by an orthonormal set  $\{e_\alpha : \alpha \in I\}$  that doesn't need to be countable. For this case, the sum  $\sum_{n \geq 1}$  should be replaced by  $\sum_{\alpha \in I}$  but it's still meaningful/convergent (and in fact, the sum is still countable sum, see the following Corollary), whose limit can be defined by net, filter or integration with respect to counting measure.

**Corollary 12.9.** Let  $\{e_\alpha : \alpha \in I\}$  be an orthonormal set in  $H$ . Then for all  $h \in H$ , we have

$$\sum_{\alpha \in I} |\langle h, e_\alpha \rangle|^2 \leq \|h\|^2$$

and  $\text{card } \mathcal{E} \leq \text{card } \mathbb{N}$ , where  $\mathcal{E} := \{\alpha \in I : \langle h, e_\alpha \rangle \neq 0\}$ .

**Proof.** Consider  $\mathcal{E}_n := \{\alpha \in I : |\langle h, e_\alpha \rangle| > 1/n\}$ . Then  $\mathcal{E} = \bigcup_{n \geq 1} \mathcal{E}_n$ . It suffices to prove that for all  $n \in \mathbb{N}$ , the set  $\mathcal{E}_n$  is at most countable. In fact,  $\mathcal{E}_n$  is a finite set for all  $n \in \mathbb{N}$ . To see this, let an arbitrary  $n \in \mathbb{N}$  be given. Then apply the same argument in the proof of Theorem 27: for all  $\alpha_1, \dots, \alpha_p \in \mathcal{E}_n$ :

$$\frac{p}{n^2} < \sum_{j=1}^p \langle h, e_{\alpha_j} \rangle \leq \|h\|^2,$$

thus  $p \leq n^2 \|h\|^2$ . Therefore, we can pick at most  $\left\lfloor n^2 \|h\|^2 \right\rfloor$  elements in  $\mathcal{E}_n$ . Thus  $\mathcal{E}_n$  is finite.  $\square$

**Theorem 12.10** (Gram-Schmidt). Let  $\{h_n : n \in \mathbb{N}\}$  be a linearly independent subset of  $H$ . Then there is an orthonormal set  $\{e_n : n \in \mathbb{N}\}$  such that for all  $n \in \mathbb{N}$ :

$$\text{span}\{h_j : j \in [n]\} = \text{span}\{e_j : j \in [n]\}.$$

**Proof.** Prove by mathematical induction.

For  $n = 1$ , the linear independence implies that  $h_1 \neq 0$ , thus define  $e_1 := h_1 / \|h_1\|$  and we're done. Suppose the proposition is true for  $n$ , i.e. there is already an orthonormal set  $\{e_j : j \in [n]\}$  such that

$$\text{span}\{h_j : j \in [n]\} = \text{span}\{e_j : j \in [n]\}.$$

Then we define a vector

$$\widehat{e_{n+1}} := h_{n+1} - \sum_{j=1}^n \langle h_{n+1}, e_j \rangle e_j,$$

which is orthogonal to  $\{e_j : j \in [n]\}$ . Moreover, we will prove  $\widehat{e_{n+1}} \neq 0$  and hence we can define  $e_{n+1} := \widehat{e_{n+1}} / \|\widehat{e_{n+1}}\|$ . If  $\widehat{e_{n+1}} = 0$ , then

$$h_{n+1} = \sum_{j=1}^n \langle h_{n+1}, e_j \rangle e_j,$$

which implies  $h_{n+1} \in \text{span}\{e_j : j \in [n]\}$ , i.e.  $h_{n+1} \in \text{span}\{h_j : j \in [n]\}$  by assumption. Contradiction with  $\{h_n : n \in \mathbb{N}\}$  being a linearly independent. Thus we can define  $e_{n+1} := \widehat{e_{n+1}} / \|\widehat{e_{n+1}}\|$  and  $\{e_j : j \in [n+1]\}$  is an orthonormal set in  $H$ .

Finally, we prove that

$$\text{span}\{h_j : j \in [n+1]\} = \text{span}\{e_j : j \in [n+1]\}. \quad (28)$$

Since

$$\forall k \in [n]: e_k \in \text{span}\{h_j: j \in [n]\} \subseteq \text{span}\{h_j: j \in [n+1]\},$$

and hence

$$e_{n+1} = \frac{h_{n+1} - \sum_{j=1}^n \langle h_{n+1}, e_j \rangle e_j}{\|\widehat{e_{n+1}}\|} \in \text{span}\{h_j: j \in [n+1]\}.$$

Therefore, we proved  $\supseteq$  part of (28). The  $\subseteq$  part follows from

$$h_{n+1} = \|\widehat{e_{n+1}}\| e_{n+1} + \sum_{j=1}^n \langle h_{n+1}, e_j \rangle e_j$$

and the assumption.  $\square$

**Theorem 12.11.** Let  $\{e_n: n \in \mathbb{N}\}$  be an orthonormal set in  $\mathcal{H}$  and  $h \in \mathcal{H}$  be an arbitrary fixed element. Then the following statements are equivalent.

1. the space  $\overline{\text{span}\{e_n: n \in \mathbb{N}\}}$  contains  $h$ ;
2. we have the representation:  $h = \sum_{n \geq 1} \langle h, e_n \rangle e_n$ ;
3. Parseval's Identity holds:  $\|h\|^2 = \sum_{n \geq 1} |\langle h, e_n \rangle|^2$ .

**Proof.** 1  $\implies$  2:

2  $\implies$  3:

3  $\implies$  1:  $\square$

### Further Topics: Von Neumann Algebra

Let  $\mathcal{M} \hookrightarrow \mathcal{B}(\mathcal{H})$  be a closed subspace. We can analysis the subspace as we studied the subspace  $L_\infty \hookrightarrow L_0$ , where  $L_0$  is the space of all measurable functions. They have clear different properties: commutativity.

# 13 Week 13

## 13.1 Week 13, Lecture 1

This lecture was given by Ruirui Chen.

We assume that  $H$  is an inner product space and  $\mathcal{H}$  is a Hilbert space (over  $\mathbb{K}$ ) in this lecture.

### Recall

We have studied

1. Two elements  $x, y \in H$  is said to be orthogonal, if and only if  $\langle x, y \rangle = 0$ . In other words,  $x \perp y \iff \langle x, y \rangle = 0$ .
2. An element and a set  $M \subseteq H$  is said to be orthogonal, if and only if  $\forall M \in M \langle x, y \rangle = 0$ . In other words,  $x \perp y \iff (\forall y \in M \langle x, y \rangle = 0)$ .

And we define

**Definition.** The orthogonal component of  $M \subseteq H$  is  $M^\perp := \{x \in H : x \perp M\}$ .

The following proposition follows from definition.

**Proposition 13.1.** We have

1. If  $\overline{M} = H$ , then  $M^\perp = \{0\}$ ;
2. For all  $M \subseteq H$ :  $M \cap M^\perp = \emptyset$ ;
3. The orthogonal component of  $M \subseteq H$  is naturally a closed subspace of  $H$ .

### Proof.

1. If  $\overline{M} = H$ , take an arbitrary element  $y \in M^\perp$ , then for all  $x \in H$ , there is a sequence  $M \ni (x_n)_{n \in \mathbb{N}} \rightarrow x$  by the density. Since  $\langle \cdot, \cdot \rangle$  is continuous:

$$\langle y, x \rangle = \lim_{n \rightarrow \infty} \langle y, x_n \rangle = \lim_n 0 = 0.$$

Thus,  $y \perp x$  for all  $x \in H$ , i.e.  $y \perp H$ . Thus  $y \perp y \implies y = 0$ .

2. For all  $M \subseteq H$ : let  $x \in M \cap M^\perp$ , then

$$(M \ni) x \perp x (\in M^\perp),$$

i.e.  $\langle x, x \rangle = 0 \implies x = 0$ .

3. We have

$$M^\perp = \bigcap_{y \in M} \{x \in H: \langle x, y \rangle = 0\}.$$

It suffices to show that  $\forall y \in M$ , the set

$$\{y\}^\perp = \{x \in H: \langle x, y \rangle = 0\}$$

is a closed subspace of  $H$ . That's trivial since it's just the kernel of  $f_y: H \rightarrow \mathbb{K}, x \mapsto \langle x, y \rangle$ .  $\square$

### 13.1.1 Orthonormal Basis

**Theorem 13.2** (Riesz–Fréchet). Let  $\mathcal{H}$  and an orthonormal subset  $E \subseteq \mathcal{H}$  be given. Then for all  $\alpha \in \ell_2$ , there is a unique element  $x \in \overline{\text{span } E}$  such that

$$x = \sum_{n \geq 1} \alpha_n e_n.$$

**Proof.**

$\square$

**Definition.** Let  $\mathcal{H}$  and an orthonormal subset  $E \subseteq \mathcal{H}$  be given. For all  $x \in \mathcal{H}$ ,

**Question 13.1.** Does orthonormal basis exist? If so, is it unique?

**Theorem 13.3.** Let  $\mathcal{H}$  and an orthonormal subset  $E \subseteq \mathcal{H}$  be given. The following statements are equivalent.

1.

**Proof.**

$\square$

And we answer the existence part of Question 13.1.

**Theorem 13.4.** Let  $\mathcal{H}$  be given. Then it must have an orthonormal basis.

**Proof.**

$\square$

**Example 42.** The space  $\ell_2$  has the familiar completely orthonormal basis  $(e_n)_{n \in \mathbb{N}}$ .

**Example 43.** The space  $L_2[-\pi, \pi]$  (denoted by  $L_2$  for short), with the inner product

$$\langle \cdot, \cdot \rangle : L_2 \times L_2 \rightarrow \mathbb{K}, (f, g) \mapsto \frac{1}{2\pi} \int_{[-\pi, \pi]} f \bar{g} \, dm.$$

has the basis

$$E = (e_n)_{n \in \mathbb{Z}}, \text{ where we define } e_n : [-\pi, \pi] \rightarrow \mathbb{C}, x \mapsto e^{inx}.$$

**Proof.** We need some lemmas to finish this proof. Let  $C[-\pi, \pi]$  denote the space of all continuous functions on  $[-\pi, \pi]$ , and  $C_{2\pi}$  be the space as defined in [Week 8 Lecture 1](#) (but we restrict them on  $[-\pi, \pi]$  in this example). Furthermore, let  $T[-\pi, \pi] := \text{span } E$  be the space of all trigonometric functions.

**Lemma 13.5.** The space  $C[-\pi, \pi]$  is dense in  $L_2$ .

**Proof of Lemma 13.5.** It suffices to prove the case that  $f \in L_2$  is non-negative. And  $\square$

**Lemma 13.6.** The space  $C_{2\pi}$  is dense in  $L_2$ .

**Proof of Lemma 13.6.** It suffices to show that  $C_{2\pi}$  is dense in  $C[-\pi, \pi]$ . Let an arbitrary element  $f \in C[-\pi, \pi]$  be given. For all  $\delta > 0$ , consider

$$f_\delta : [-\pi, \pi] \rightarrow \mathbb{C},$$

$\square$

And we need

**Theorem (Weierstrass).** The space  $T[-\pi, \pi]$  is dense in  $C[-\pi, \pi]$ , with respect to the infinity norm  $\|\cdot\|_\infty$ .

$\square$

**Remark.**

**Theorem 13.7.** Let  $\mathcal{H}$  and an orthonormal basis  $E \subseteq \mathcal{H}$  be given.

1. The space  $\mathcal{H}$  is separable if and only if  $\mathcal{H}$  has an orthonormal basis;
2. If  $\text{card } E = \text{card } \mathbb{N}$ , then  $\mathcal{H} \cong \ell_2$ ;
3. If  $\text{card } E = n \in \mathbb{N}$ , then  $\mathcal{H} \cong \mathbb{K}^n$ .

**Proof.**

$\square$

# A Hamel Basis

**Definition** (Partial order, Partially ordered set). A binary relation on  $X$  is called an **partial order** on  $X$  if it satisfies

1.  $x \prec y \wedge y \prec z \implies x \prec z$ ;
2.  $\forall x \in X \ x \prec x$ ;
3.  $x \prec y \wedge y \prec x \implies x = y$ .

A set with an order is called an **partially ordered set**.

**Remark.** In fact, an order on  $X$  can be defined as a binary relation, i.e. a subset of  $X \times X$ . But we don't care this now.

**Definition** (Total order, Totally ordered set). An order is said to be **linear**, if  $\forall x, y \in X (x \prec y \vee y \prec x)$ . A set with a linear order is called a **totally ordered set**.

**Definition** (Bound, Bounded set, Maximal element). Let  $X$  be an ordered set and  $Y \subseteq X$ . An element  $x \in X$  is called a **bound** for  $Y$  if  $y \prec x (\forall y \in Y)$  and at the same time  $Y$  is called a **bounded set**. An element  $m \in X$  is called a **maximal element** if  $\forall y \in X \neg (m \prec y)$ .

**Axiom** (Zorn's lemma). Let  $X$  be an ordered set with the following property: every totally ordered subset of  $X$  (in the sense of the order induced by the initial order of  $X$ ) is bounded. Then there is at least one maximal element in  $X$ .

This is equivalent to the **Axiom of Choice**, which cannot be proved from the other axioms of set theory. To define base, we need the notion of linear independence.

**Definition** (Linearly Independent, Hamel Base). Let  $V$  be a linear space over  $\mathbb{K}$ . A system of vectors of  $V$  is called **linearly independent** if every finite subsystem of this system is linearly independent (i.e. every finite combination gives 0 if and only if all coefficients are 0).

A family of vectors  $\{e_i \in V : i \in I\}$  is called a **Hamel basis** of  $V$ , if  $\forall x \in V, x \neq 0$  can be uniquely represented as a (finite) linear combination of vectors in  $\{e_i : i \in I\}$ .

**Theorem A.1** (Existence of Hamel base). Each linear space  $V$  (over an arbitrary field) has a Hamel Base.

**Proof.** To use **Zorn's lemma**, we need to construct an ordered set whose maximal element can be a Hamel basis of  $V$ . Thus, consider

$$\mathcal{D} := \{X \subseteq V : X \text{ is linearly independent}\}$$

with the order:  $\forall A, B \in \mathcal{D} : A \prec B \iff A \subseteq B$ . Given an arbitrary totally ordered set  $\mathcal{A} \subseteq \mathcal{D}$ , we have a bound for  $\mathcal{A}$ , just  $\bigcup \mathcal{A}$ . To show that  $\bigcup \mathcal{A} \in \mathcal{D}$ , taking arbitrary  $e_1, e_2, \dots, e_n \in \bigcup \mathcal{A}$  such that  $e_j \in X_j \in \mathcal{A}$  for all  $j = 1, 2, \dots, n$ . Since  $\mathcal{A}$  is totally ordered, we can suppose  $X_1 \subseteq X_2 \subseteq \dots \subseteq X_n$  and hence  $e_j \in X_n (\forall j = 1, 2, \dots, n)$ . Since  $X_n \in \mathcal{A} \subseteq \mathcal{D}$ ,  $X_n$  is linear independent and hence  $e_1, e_2, \dots, e_n$  is linearly independent. Therefore,  $\bigcup \mathcal{A} \in \mathcal{D}$ . Now apply [Zorn's lemma](#) and we know there is a maximal element  $B$  in  $\mathcal{D}$ . And  $B$  is a Hamel basis. To show this, it suffices to prove that every element in  $V$  lies in  $\text{span}(B)$ . If there is an element  $v \in V$  such that  $v \notin \text{span}(B)$ , i.e.  $B \cup \{v\}$  is linearly independent. This is impossible by the definition of the maximal element.  $\square$

In fact, we can define Hamel bases for an arbitrary linear space over an arbitrary field such as  $\mathbb{Q}$  and  $\mathbb{F}_p$  for some prime  $p$ .

Here I explain why we can view a vector space over  $\mathbb{C}$  as a vector space over  $\mathbb{R}$ .

**Proposition A.2.** Let  $V$  be a vector space over  $\mathbb{C}$ , then there is a real vector space  $W$  and a  $\mathbb{R}$ -linear bijection  $\varphi: V \rightarrow W$ .

**Proof.** Let  $\{v_\alpha\}_{\alpha \in I}$  be a base for  $V$ . Consider the set  $W := V \times iV$ , where  $iV = \{iv : v \in V\}$  is equipped with the natural real linear structure. Now I define a linear structure on  $W$  such that  $W$  is a  $\mathbb{R}$ -linear space. Then, define the mapping  $\varphi: V \rightarrow W$  by

$$\forall \alpha \in I, \forall z \in \mathbb{C} : \varphi(zv_\alpha) := (\text{Re}(z)v_\alpha, \text{Im}(z)v_\alpha).$$

Extend  $\varphi$  to  $V$  keeping  $\mathbb{R}$ -linear. Then  $\varphi|_{\text{Im}(\varphi)}$  is what we wanted.  $\square$



## B Banach Functor

This appendix comes from [3].

Recall [the exercise](#):

**Exercise B.1.** If  $Y$  is a Banach space and  $X$  is a linear normed space, then  $\mathcal{B}(X, Y)$  is a Banach space. Especially,  $X^*$  is a Banach space.

**Proof.** Let  $(u_n)_{n \in \mathbb{N}} \subseteq \mathcal{B}(X, Y)$  be a Cauchy sequence. Thus

$$\lim_{m, n} \|u_n - u_m\|_{\mathcal{B}(X, Y)} = 0.$$

Taking an arbitrary  $x \in X$ , we have

$$\begin{aligned} \|u_n x - u_m x\|_Y &\leq \|(u_n - u_m)x\|_Y \\ &\leq \|u_n - u_m\|_{\mathcal{B}(X, Y)} \|x\|_X \\ &\rightarrow 0 (m, n \rightarrow \infty). \end{aligned} \tag{29}$$

Therefore,  $(u_n x)_{n \in \mathbb{N}}$  is a Cauchy sequence in  $Y$ . Since  $Y$  is a Banach space, we know  $(u_n x)_{n \in \mathbb{N}}$  converges to some point in  $Y$ . Thus we can define a map

$$u: X \rightarrow Y, x \mapsto \lim_n u_n(x).$$

And now we prove that  $(u_n)_{n \in \mathbb{N}} \rightarrow u$  in  $\mathcal{B}(X, Y)$ . This proof is similar to the proof of “uniform limit of a continuous function sequence is continuous”, see [this proof](#).

By definition,  $\forall \varepsilon > 0 \exists N \in \mathbb{N}$  such that  $\forall m, n \geq N$  we have

$$\|u_n - u_m\|_{\mathcal{B}(X, Y)} < \varepsilon,$$

which implies

$$\|u_n x - u_m x\|_Y = \|(u_n - u_m)x\|_Y \leq \varepsilon \|x\|, \forall x \in X.$$

Let  $m \rightarrow \infty$ , by the continuity of  $\|\cdot\|_Y$ , we have

$$\|(u_n - u)x\| = \|u_n x - u x\|_Y \leq \varepsilon \|x\|, \forall x \in X.$$

Therefore,  $\|u_n - u\| \leq \varepsilon$  holds for all  $n > N$ . That is  $(u_n)_{n \in \mathbb{N}} \rightarrow u$ .  $\square$

Now we can define

**Definition.** The contravariant functor

$$\begin{aligned} * : \text{Nor} &\longrightarrow \text{Ban}, \\ \text{Ob}(\text{Nor}) \ni X &\longmapsto X^*, \\ \text{Mor}(\text{Nor}) \ni \varphi : X_1 \rightarrow X_2 &\longmapsto \varphi^* : X_2^* \rightarrow X_1^*. \end{aligned} \tag{30}$$

Where  $\varphi^* : X_2^* \rightarrow X_1^*, f \mapsto f \circ \varphi$ .

**Remark.** Banach Functor is a special case of the functor  $\mathcal{B}(\cdot, Y)$  where  $Y$  is a Banach space, defined as

$$\begin{aligned}\mathcal{B}(\cdot, Y): \text{Nor} &\longrightarrow \text{Ban}, \\ \text{Ob}(\text{Nor}) \ni X &\longmapsto \mathcal{B}(X, Y), \\ \text{Mor}(\text{Nor}) \ni \varphi: X_1 \rightarrow X_2 &\longmapsto \mathcal{B}(\varphi, Y): \mathcal{B}(X_1, Y) \rightarrow \mathcal{B}(X_2, Y).\end{aligned}$$

Here  $\mathcal{B}(\varphi, Y): \mathcal{B}(X_1, Y) \rightarrow \mathcal{B}(X_2, Y), f \mapsto f \circ \varphi$ .

Banach functor is surely a functor.

**Proof.** It suffices to prove that  $(\text{id}_X)^* = \text{id}_{X^*}$  and  $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$ . To prove two maps are the same, we should prove that they coincide at every point.

- Given  $X \in \text{Ob}(\text{Nor})$ , we have

$$\forall f \in X^* : (\text{id}_X)^*(f) = f \circ \text{id}_X = f = \text{id}_{X^*}(f).$$

Thus  $(\text{id}_X)^* = \text{id}_{X^*}$ , since  $\text{id}_{X^*}$  is uniquely determined by this property.

- Given  $X_1 \xleftarrow{\varphi} X_2 \xleftarrow{\psi} X_3$ . Notice that  $\text{dom}((\varphi \circ \psi)^*) = X_1^*$  and for any  $f \in X_1^*$ , we have

$$\begin{aligned}(\varphi \circ \psi)^*(f) &= f \circ (\varphi \circ \psi) \\ &= (f \circ \varphi) \circ \psi \\ &= (\varphi^*(f)) \circ \psi \\ &= \psi^*(\varphi^*(f)) \\ &= (\psi^* \circ \varphi^*)(f).\end{aligned}$$

This means  $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$ .

Now, What is needed to check is just  $\varphi^* \in \mathcal{B}(X_2^*, X_1^*)$ , and this is true since  $\forall f \in X_2^*$

$$\|\varphi^*(f)\|_{X_1^*} = \|f \circ \varphi\|_{\mathcal{B}(X_1, \mathbb{K})} \leq \|f\|_{X_2^*} \|\varphi\|_{\mathcal{B}(X_1, X_2)},$$

hence  $\|\varphi^*\|_{\mathcal{B}(X_2^*, X_1^*)} \leq \|\varphi\|_{\mathcal{B}(X_1, X_2)}$ . □

If we restrict  $*$ :  $\text{Nor} \rightarrow \text{Ban}$  to the full subcategory  $\text{Ban}$ , and call it **Banach adjointness functor**, then we can consider the composition of  $*$ :  $\text{Ban} \rightarrow \text{Ban}$ , i.e.  $**$ :  $\text{Ban} \rightarrow \text{Ban}$ , which is covariant. On object,  $X^{**}$  is the usual second dual space; on morphism,  $T^{**}$  is the usual

second dual operator. This is the main topic about “the dual theory of Banach space”.

Similarly we can define a covariant functor

$$\begin{aligned}\mathcal{B}(\mathbb{K}, \_): \mathbf{Ban} &\longrightarrow \mathbf{Ban} \\ X &\longmapsto \mathcal{B}(\mathbb{K}, X), \\ (f: X \rightarrow Y) &\longmapsto (\mathcal{B}(\mathbb{K}, f): \mathcal{B}(\mathbb{K}, X) \rightarrow \mathcal{B}(\mathbb{K}, Y)),\end{aligned}$$

where

$$\mathcal{B}(\mathbb{K}, f): \mathcal{B}(\mathbb{K}, X) \rightarrow \mathcal{B}(\mathbb{K}, Y), \psi \mapsto f \circ \psi.$$

**Proposition B.1.** Show that  $\mathcal{B}(\mathbb{K}, \_)$  is naturally equivalent to  $\text{id}_{\mathbf{Ban}}$ , the identity functor of  $\mathbf{Ban}$ .

**Proof.** Define the natural transformation  $\theta = \{\theta_X: X \in \text{Ob}(\mathbf{Ban})\}$  as follows

$$\theta_X: \mathcal{B}(\mathbb{K}, X) \rightarrow X, \varphi \mapsto \varphi(1),$$

where 1 is just the multiplicative identity of  $\mathbb{K}$ . We check that  $\theta_X$  is a

- injection:  $\ker \theta_X = \{0\}$ . From definition:  $\theta_X(\varphi) = 0$  if and only if  $\varphi(1) = 0$ , while  $\varphi \in \mathcal{B}(\mathbb{K}, X)$  so  $\varphi(1) = 0$  if and only if  $\varphi = 0$ ;
- surjection:  $\text{Im } \theta_X = X$ . For arbitrary  $x \in X$ , define  $\varphi_x: \mathbb{K} \rightarrow X, z \mapsto zx$ . Then  $\|\varphi_x\| = \|x\|$  and  $\theta_X(\varphi_x) = x$  as we want. Furthermore, we know that  $\theta_X$  keeps norms.

Thus  $\theta_X$  is an isomorphism for each  $X \in \text{Ob}(\mathbf{Ban})$ .

Then we check that the following diagram commutes

$$\begin{array}{ccccc}\mathcal{B}(\mathbb{K}, X) & \longrightarrow & \text{id}_{\mathbf{Ban}}(X) & \equiv & X \\ \mathcal{B}(\mathbb{K}, f) \downarrow & & \text{id}_{\mathbf{Ban}}(f) \downarrow & \equiv & \downarrow f \\ \mathcal{B}(\mathbb{K}, Y) & \longrightarrow & \text{id}_{\mathbf{Ban}}(Y) & \equiv & Y\end{array}$$

Given  $\varphi \in \mathcal{B}(X, C)$ , on the one hand

$$(f \circ \theta_X)(\varphi) = f(\theta_X(\varphi)) = f(\varphi(1)).$$

On the other hand

$$(\theta_Y \circ \mathcal{B}(f, \mathbb{K}))(\varphi) = \theta_Y\left((\mathcal{B}(f, \mathbb{K}))(\varphi)\right) = \theta_Y(f \circ \varphi) = (f \circ \varphi)(1).$$

We’re done since  $f(\varphi(1)) = (f \circ \varphi)(1)$ . □

**Remark.** The definition of  $\theta_X$  is natural since this is the simplest element of  $\mathcal{B}(\mathcal{B}(\mathbb{K}, X), X)$ .

Now we explain the reason of name “Natural Embedding”, whose definition can be found [here](#). To view this better, we should view the “Biduality” as a functor, which can be defined as a composition of a functor.

**Definition.** Let  $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$  be three categories and  $F: \mathcal{C} \rightarrow \mathcal{D}, G: \mathcal{D} \rightarrow \mathcal{E}$  are two functors. The composition functor of  $F$  and  $G$ , denoted by  $G \circ F$  is defined as follows.

$$\begin{aligned} G \circ F: \mathcal{C}_1 &\longrightarrow \mathcal{C}_3, \\ \text{Ob}(\mathcal{C}_1) \ni A &\mapsto G(F(A)), \\ \text{Mor}(\mathcal{C}_1) \ni \varphi: X \rightarrow Y &\mapsto G(F(\varphi)): G(F(X)) \rightarrow G(F(Y)). \end{aligned}$$

Thus the bidual functor  $**$  is just the composition of  $*$ :  $\text{Nor} \rightarrow \text{Ban}$  and itself (more precisely,  $** := * \circ \iota \circ *$ , where  $\iota: \text{Ban} \rightarrow \text{Nor}$  is the identity functor that keeps everything the same). Here is a easy exercise that ensures something like  $(X^{**})^* = (X^*)^{**}$ .

**Exercise B.2.** Show that the composition of functor is associative.

Now we explain the name of “Natural Embedding”. For convenience, here we view  $**$  as a functor from  $\text{Nor}$  to itself.

**Proposition B.2.** Consider the category  $\text{Nor}$  and two functors

$$\begin{aligned} ** &: \text{Nor} \rightarrow \text{Nor}, \\ \text{id} &: \text{Nor} \rightarrow \text{Nor}, \end{aligned}$$

where  $\text{id}$  is  $\text{id}_{\text{Nor}}$  for short. Then the family of natural embedding

$$\{\iota_X: X \rightarrow X^{**}, x \mapsto \hat{x} \mid X \in \text{Ob}(\text{Nor})\}$$

is a natural transformation from  $\text{id}$  to  $**$ .

**Proof.** See Corollary [11.2](#). □

## C Completion is unique up to a unique isomorphism

To ensure the uniqueness of completion, we need to change the definition of completion.

**Definition.** A completion of a linear normed space is a pair  $(\widehat{X}, \iota)$  such that  $\widehat{X}$  is a Banach space and  $\iota: X \hookrightarrow \widehat{X}$  is an isometry with dense image.

**Theorem C.1** (Uniqueness of completion). The completion of a linear normed space  $X$  is unique up to a unique isometry (that coincides with the two inclusions). That is, if  $\widehat{X}, Y$  with isometric inclusion map  $\iota, \iota'$  respectively are completions of  $X$ , then the following diagram commutes

$$\begin{array}{ccc} & X & \\ \iota \swarrow & & \searrow \iota' \\ \widehat{X} & \xrightarrow{\exists \varphi} & Y \end{array}$$

**Proof.** Consider the corestriction of  $\iota$ , that is  $\iota_0 := \iota|_{\iota(X)}$ . Clearly  $\iota_0$  is an isometry from  $X$  to  $\iota(X)$  (which is dense in  $\widehat{X}$ ). Now we define a map  $\varphi_0$  by the following diagram (i.e.  $\varphi_0 := \iota' \circ (\iota_0)^{-1}$ )

$$\begin{array}{ccc} & X & \\ (\iota_0)^{-1} \swarrow & & \searrow \iota' \\ \iota(X) & \xrightarrow{\varphi_0} & Y \end{array}$$

Now  $\varphi_0$  is linear and keeps norm. Since  $\iota(X)$  is dense in  $\widehat{X}$ ,  $Y$  is complete and  $\varphi_0$  is continuous ( $\varphi_0$  keeps norm and hence is continuous), we can extend  $\varphi_0$  to a continuous map  $\varphi: \widehat{X} \rightarrow Y$  (see [this exercise](#), to prove which, it suffices to displace  $X_0, \mathbb{K}$  with  $\iota(X), Y$  respectively).

To show that  $\varphi$  is an isometry, we should show that:

1.  $\varphi$  is linear;
2.  $\varphi$  keeps norm.
3.  $\varphi$  is surjective;

First,  $\varphi$  is linear. We have proved this in [the exercise](#).

Second,  $\varphi$  keeps norm.  $\iota(X)$  is dense in  $\widehat{X}$ , so  $\forall x \in \widehat{X}$ ,  $\exists (x_n)_{n \in \mathbb{N}} \subseteq \iota(X)$  such that  $(x_n)_{n \in \mathbb{N}} \rightarrow x$ . Then

$$\begin{aligned}
\|x\| &= \left\| \lim_n x_n \right\| \\
&= \lim_n \|x_n\| && \text{(continuity of } \|\cdot\| \text{)} \\
&= \lim_n \|\varphi_0(x_n)\| && (\varphi_0 \text{ is an isometry)} \\
&= \lim_n \|\varphi(x_n)\| && (\varphi|_{\iota(X)} = \varphi_0) \\
&= \left\| \lim_n \varphi(x_n) \right\| && \text{(continuity of } \|\cdot\| \text{)} \\
&= \|\varphi(x)\| && \text{(continuity of } \varphi \text{).}
\end{aligned}$$

Thirdly,  $\varphi$  is surjective.  $\forall y \in Y$ , by the density of  $\iota'(X)$ ,  $\exists (y_n)_{n \in \mathbb{N}} \subseteq \iota'(X)$  such that  $(y_n)_{n \in \mathbb{N}} \rightarrow y$ . And  $\forall n \in \mathbb{N}$ , let  $x_n := \varphi_0^{-1}(y_n)$  then  $(x_n)_{n \in \mathbb{N}} \subseteq \iota(X) \subseteq \widehat{X}$  is well-defined and Cauchy (since  $(y_n)_{n \in \mathbb{N}}$  is Cauchy and  $\varphi$  keeps norm). Now

$$y = \lim_n y_n = \lim_n \varphi(x_n) = \varphi(\lim_n x_n) = \varphi(x).$$

The last equality used the completeness of  $\widehat{X}$ . Therefore,  $\varphi$  is surjective.

Above all,  $\varphi$  is an isometry. If there is another isometry  $\phi: \widehat{X} \rightarrow Y$  such that the diagram commutes, then  $\varphi|_{\iota(X)} = \phi|_{\iota(X)} = \varphi_0$ . Then  $\varphi$  and  $\phi$  coincide on a dense subset of  $\widehat{X}$  and hence  $\varphi = \phi$ .  $\square$

## D Riesz Representation Theorem and so on

This Lecture was given by me.

### D.1 Riesz Representation Theorem

We have studied duality of linear normed spaces: dual spaces and conjugate operators. Hilbert spaces are special linear normed spaces; let  $H$  be a Hilbert space (an inner product is enough, in fact), then for all  $y \in H$ , there is a bounded functional on  $H$  induced by  $y$ :

$$H^* \ni \varphi_y: H \rightarrow \mathbb{K}, x \mapsto \langle x, y \rangle.$$

Then  $\varphi: H \rightarrow H^*, y \mapsto \varphi_y$  is conjugate-linear and keeps norms. The map  $\varphi$  depends on  $H$ , also denoted by  $\varphi^H$  for clarity, is a conjugate-linear injection satisfying  $\|\cdot\|_{H^*} \circ \varphi = \|\cdot\|_H$ .

**Theorem D.1** (Riesz Representation Theorem). Let  $H$  be a Hilbert space, then  $\forall f \in H^* \exists! y_f \in H$  such that

$$\forall x \in H: f(x) = \langle x, y_f \rangle.$$

Furthermore, we have  $\|f\|_{H^*} = \|y_f\|_H$ .

**Proof.** □

**Remark.** Here we can't change the "a Hilbert space" to "an inner product space" in Theorem D.1. In other words, completeness is necessary in some sense.

**Example 44** (Not Complete Space). Consider the inner product space  $c_c \hookrightarrow \ell_2$ , where

$$c_c := \{x = (x_n)_{n \in \mathbb{N}} \in {}^{\mathbb{N}}\mathbb{K}: x_n = 0 \text{ for all but finitely many } n\}$$

and  $\ell_2$  is equipped with the usual inner product. Then the functional

$$f: c_c \rightarrow \mathbb{K}, x \mapsto \sum_{n=1}^{\infty} x_n/n$$

is bounded, with norm  $\pi/\sqrt{6}$  but there is no element  $y \in c_c$  such that

$$\forall x \in c_c: f(x) = \langle x, y \rangle.$$

**Definition** (Conjugate-Isometry). Let  $H, K$  be 2 linear normed spaces. If  $f: X \rightarrow Y$  be a conjugate-linear bijection such that

$$\forall x \in X: \|f(x)\|_Y = \|x\|_X,$$

then  $f$  is said to be a **conjugate-isometry**.

**Corollary D.2.** For a Hilbert space  $H$ , there is a conjugate-isometry:

$$\varphi: H \rightarrow H^*, y \mapsto \varphi(y) = \varphi_y,$$

where the map  $\varphi_y$  is defined as follows

$$\varphi_y: H \rightarrow \mathbb{K}, x \mapsto \langle x, y \rangle,$$

for all  $y \in H$ .

**Proof.** □

**Question D.1.** For an inner product space  $X$ , if it's complete, i.e. a Hilbert space then there is a conjugate-isometry from  $X$  to  $X^*$ . Does the reversed proposition holds? In other words, if an inner product space  $X$  satisfies that there is a conjugate-isometry from  $X$  to  $X^*$ , then  $X$  is a Hilbert space.

We will answer this later.

**Proposition D.3.** Let  $H$  be a Hilbert space and  $\varphi$  be the bijection of above corollary. Then the norm (as a operator norm) of  $H^*$  is induced by the following inner product:

$$\langle \cdot, \cdot \rangle_{H^*} : H^* \times H^* \rightarrow \mathbb{K}, (\varphi(y), \varphi(x)) \mapsto \langle x, y \rangle_H, \quad (31)$$

and  $H^*$  is a Hilbert space.

**Proof.** It suffices to prove that  $\| \cdot \|_{H^*}$  satisfies the Parallelogram Law and apply Theorem 12.6. Consider the Theorem 12.4: for all  $x, y \in H$

$$\begin{aligned} & \|\varphi(y) + \varphi(x)\|^2 + \|\varphi(y) - \varphi(x)\|^2 \\ &= \|\varphi(y+x)\|^2 + \|\varphi(y-x)\|^2 && \text{(linearity)} \\ &= \|y+x\|^2 + \|y-x\|^2 && \text{(isometry)} \\ &= 2(\|y\|^2 + \|x\|^2) \\ &= 2(\|\varphi(y)\|^2 + \|\varphi(x)\|^2). && \text{(isometry)} \end{aligned}$$

Then Theorem 12.6 implies that  $\| \cdot \|_{H^*}$  is induced by an inner product  $\langle \cdot, \cdot \rangle_{H^*}$ . Now we prove (31). WLOG, let  $\mathbb{K} = \mathbb{R}$ . Then apply the polar identity: for all  $x, y \in H$

$$\begin{aligned} \langle \varphi(y), \varphi(x) \rangle_{H^*} &= \frac{\|\varphi(y) + \varphi(x)\|^2 - \|\varphi(y) - \varphi(x)\|^2}{4} \\ &= \frac{\|\varphi(y+x)\|^2 - \|\varphi(y-x)\|^2}{4} \\ &= \frac{\|y+x\|^2 - \|y-x\|^2}{4} \\ &= \langle y, x \rangle = \langle x, y \rangle. \end{aligned}$$



For the case  $\mathbb{K} = \mathbb{C}$ , proof is similar:

$$\begin{aligned}
\langle \varphi(y), \varphi(x) \rangle_{H^*} &= \frac{1}{4} \sum_{\alpha^4=1} \alpha \|\varphi(y) + \alpha \varphi(x)\|^2 \\
&= \frac{1}{4} \sum_{\alpha^4=1} \alpha \|\varphi(y + \bar{\alpha}x)\|^2 \\
&= \frac{1}{4} \sum_{\alpha^4=1} \alpha \|y + \bar{\alpha}x\|^2 \\
&= \frac{1}{4} \sum_{\beta^4=1} \beta \|y + \beta x\|^2 \\
&= \overline{\langle y, x \rangle} \\
&= \langle x, y \rangle.
\end{aligned}$$

The space  $H^*$  is the dual space of  $H$  and must be complete.  $\square$

**Theorem D.4.** Every Hilbert space is reflexive. In other words, for all Hilbert space  $H$ , the canonical embedding  $\iota_H: H \rightarrow H^{**}$  is an isometry. And  $\iota_H = \varphi^{H^*} \circ \varphi^H$ .

**Proof.** Theorem D.1 gives an conjugate-isometry  $\varphi^H: H \rightarrow H^*$ . We have proved that  $H^*$  is a Hilbert space, thus there is an conjugate-isometry  $\varphi^{H^*}$  from  $H^*$  to  $H^{**} = (H^*)^*$ .

**Claim.** The natural embedding is just the composition of 2 isometries:  $\iota_H = \varphi^{H^*} \circ \varphi^H$ . In other words, the following diagram commutes

$$\begin{array}{ccc}
H & \xrightarrow{\iota_H} & H^{**} \\
\searrow \varphi^H & & \nearrow \varphi^{H^*} \\
& H^* &
\end{array}$$

The theorem is proved as long as the claim is proved.

Taking arbitrary  $x \in H$  and  $H^* \ni f = \varphi(y)$ , where  $y \in H$ , we have

$$[\iota_H(x)](f) = \hat{x}(f) = f(x) = \langle x, y \rangle_H,$$

and

$$\begin{aligned}
[(\varphi^{H^*} \circ \varphi^H)(x)](f) &= [\varphi^{H^*}(\varphi^H(x))](f) \\
&= \langle f, \varphi^H(x) \rangle_{H^*} \\
&= \langle \varphi^H(y), \varphi^H(x) \rangle_{H^*} \\
&= \langle x, y \rangle_H.
\end{aligned}$$

Thus

$$\iota_H(x) = (\varphi^{H^*} \circ \varphi^H)(x)$$

holds for all  $x \in H$ . And hence  $\iota_H = \varphi^{H^*} \circ \varphi^H$ .  $\square$

Now we answer Question D.1.

**Answer of Question D.1.** That's true (and this proof is almost the same as the proof of Theorem D.4). In other words, Riesz's Theorem states a property that only Hilbert spaces enjoy.

We prove that there is an conjugate-isometry from  $X^*$  to another Hilbert space. Let  $\varphi: X \rightarrow X^*$  be the conjugate-isometry, Let  $Y := X^*$ , and  $Y^*$  be the dual space of  $Y$ , then both of which are Hilbert spaces. There is an conjugate-isometry

$$\varphi^Y: Y \rightarrow Y^*, f \mapsto \varphi_f^Y = \langle \cdot, f \rangle_{H^*}$$

given by Theorem D.1. Just like what we do in the proof of Theorem D.1, we can define an inner product on  $Y^*$  such by

$$\forall g, f \in Y: \langle \varphi^Y(g), \varphi^Y(f) \rangle_{Y^*} := \langle f, g \rangle_Y,$$

which induces the norm on  $Y^*$ . Then consider  $\mu := \varphi^Y \circ \varphi: X \rightarrow Y^*$ , which is an isometry. Thus  $X \cong Y^*$  and hence  $X$  is a Hilbert space.  $\square$

## D.2 Application

We introduce some notions similar to inner product (just like the relation between sub-linear functional and functional).

**Definition** (Bilinearity, Sesquilinearity and so on). Let  $H, K$  be 2 linear normed spaces. A map  $\psi: H \times K \rightarrow \mathbb{K}$  is said to be

1. a **bilinear functional**, if for all  $x, y \in X, z, w \in Y$  and  $k, l \in \mathbb{K}$ , we have

$$\begin{aligned}\psi(kx + ly, z) &= k\psi(x, z) + l\psi(y, z), \\ \psi(x, kz + lw) &= k\psi(x, z) + l\psi(x, w).\end{aligned}$$

2. a **sesquilinear functional**, if for all  $x, y \in X, z, w \in Y$  and  $k, l \in \mathbb{K}$ , we have

$$\begin{aligned}\psi(kx + ly, z) &= k\psi(x, z) + l\psi(y, z), \\ \psi(x, kz + lw) &= \bar{k}\psi(x, z) + \bar{l}\psi(x, w).\end{aligned}$$

3. a **bounded functional** on  $H \times K$ , if  $\psi$  is bilinear or sesquilinear and the supremum  $\sup_{\substack{\|y\| \leq 1 \\ \|x\| \leq 1}} |\psi(x, y)|$  is finite. Equivalently, there is  $M > 0$  such that

$$\forall x \in X, y \in Y: |\psi(x, y)| \leq M \|x\| \cdot \|y\|.$$

And denote  $\sup_{\substack{\|y\| \leq 1 \\ \|x\| \leq 1}} |\psi(x, y)|$  by  $\|\psi\|$  ( $\|\cdot\|$  is a norm).

4. a Hermitian functional, if  $X = Y$  and

$$\forall x, y \in X: \psi(x, y) = \overline{\psi(y, x)}.$$

**Remark.** Inner product is a special sesquilinear functional.

**Remark.** For saving time, this will be given the latest. For the boundedness of  $\psi$ , we don't need  $H \times K$  to have a norm. The boundedness is an abstract general version of Cauchy Schwarz inequality. Here is another possible interpretation for the case that  $\psi$  is bilinear: consider the tensor product of  $H$  and  $K$ , i.e. a pair  $(X \otimes Y, \theta)$  satisfying the universal property of tensor product. Then the following diagram commutes:

$$\begin{array}{ccc} H \times K & & (x, y) \xrightarrow{\psi} \psi(x, y) \\ \theta \downarrow & \searrow \psi & \downarrow \theta \quad \parallel \\ H \otimes K & \xrightarrow{\Psi} & K \quad x \otimes y \xrightarrow{\Psi} \Psi(x \otimes y) \end{array}$$

Here  $H \otimes K$  is a linear normed space whose norm satisfies

$$\forall x \in H, y \in K: \|x \otimes y\| = \|x\| \cdot \|y\|.$$

Let

$$B_E \otimes F \supseteq S := \left\{ \theta(x, y) = x \otimes y \in X \otimes Y: x \in B_E, y \in B_F \right\},$$

then

$$\sup_{\substack{\|x\| \leq 1 \\ \|y\| \leq 1}} |\psi(x, y)| = \sup_{\substack{\|x\| \leq 1 \\ \|y\| \leq 1}} |(\Psi \circ \theta)(x, y)| = \sup_{x \otimes y \in S} |\Psi(x \otimes y)|.$$

And in fact, the convex hull of  $S$  is just  $B_E \otimes F$ . Thus

$$\sup_{u \in S} |\Psi(u)| = \sup_{u \in B_E \otimes F} |\Psi(u)| = \|\Psi\|_{(E \otimes F)^*}.$$

In other words, the norm  $\psi$  is just the norm of the bounded linear functional  $\Psi$ .

**Theorem D.5** (Representation of sesquilinear functional). Let  $H, K$  be 2 Hilbert spaces. Then  $\psi: H \times K \rightarrow \mathbb{K}$  is a sesquilinear functional if and only if  $\exists T \in \mathcal{B}(H, K)$  such that for all  $x \in H, y \in K$ :

$$\psi(x, y) = \langle Tx, y \rangle_K.$$

Furthermore, if  $T$  exists then it's unique, and satisfies:  $\|T\| = \|\psi\|$ .

**Proof.** For all  $x \in H$ , consider the map  $f_x: K \rightarrow \mathbb{K}, y \mapsto \overline{\psi(x, y)}$ . Then  $f_x \in K^*$ . Riesz Representation Theorem implies that there is a unique element, which will be denoted by  $Tx$ , such that for all  $x \in H, y \in K$ :

$$f_x(y) = \langle y, Tx \rangle_K,$$

i.e. for all  $x \in H, y \in K$ :

$$\psi(x, y) = \langle Tx, y \rangle_K.$$

Applying the uniqueness of Riesz representation Theorem, we see that  $T$  is linear. For boundedness: ...  $\square$

**Remark.** To get an element in  $K$  representing a bounded linear functional, we should give a bounded linear functional on  $K$ . Thus we fix  $x \in X$  and consider  $\psi(x, \cdot)$ .

**Theorem D.6.** Let  $H, K$  be 2 Hilbert spaces and  $T \in \mathcal{B}(H, K)$ . There is a unique bounded operator  $S \in \mathcal{B}(K, H)$  such that

$$\forall x \in H, y \in K: \langle Tx, y \rangle_K = \langle x, Sy \rangle_H.$$

**Proof.** Let

$$\psi: K \times H \rightarrow \mathbb{K}, (y, x) \mapsto \overline{\langle Tx, y \rangle_K}.$$

Then  $\psi$  is a sesquilinear bounded functional. Apply Theorem D.5.  $\square$

**Definition** (Adjoint Operator). Let  $H, K$  be 2 Hilbert spaces and  $T \in \mathcal{B}(H, K)$ . The operator  $S$  uniquely determined by

$$\forall x \in H, y \in K: \langle Tx, y \rangle_K = \langle x, Sy \rangle_H.$$

is called the adjoint operator of  $T$ , which will be denoted by  $T^{h*}$ .

**Remark.** For convenience, we will denote the conjugate operator of  $T$  by  $T^{b*}$  and the adjoint operator of  $T$  by  $T^*$ .

**Example 45.** Consider the Hilbert space  $H = \mathbb{K}^m, K = \mathbb{K}^n$ . Let an operator  $T \in \mathcal{B}(H, K)$  be given, whose corresponding matrix under the standard bases

$$\mathbb{K}^m \supseteq \{\alpha_j : j \in [m]\}, \mathbb{K}^n \supseteq \{\beta_k : k \in [n]\}$$

is  $M_T = (a_{j,k})_{\substack{j \in [m] \\ k \in [n]}}$ , then the matrix corresponds to  $T^*$  is just  $M_{T^*} = M'_T$ , where  $'$  means conjugate transpose.

**Proof.** From definitions: □

**Example 46.** Consider a Hilbert space  $H$  and a closed subspace  $\emptyset \neq E \hookrightarrow H$  with an inclusion map  $i_E : E \rightarrow H, x \mapsto x$ . We have studied the projection operator  $P_E : H \rightarrow E$ . Try to prove:  $i_E^* = P_E$ .

**Proof.** From definitions: for all  $x \in E, y \in H$ :

$$\langle i_E x, y \rangle_H = \langle x, i_E^* y \rangle_E.$$

Equivalently: for all  $x \in E, y \in H$ :

$$\langle x, y - i_E^* y \rangle_H = 0.$$

Thus  $y - i_E^* y \perp x (\forall x \in E)$ , i.e.  $y - i_E^* y \perp E$ . Thus  $i_E^* y = P_E(y)$ . We're done since  $y \in H$  is arbitrary. □

There is an interesting relation between adjoint and conjugate operators:

**Theorem D.7.** Let  $H, K$  be 2 Hilbert spaces and  $T \in \mathcal{B}(H, K)$ . Then  $T^* \in \mathcal{B}(K, H)$  is the unique bounded operator such that the following diagram commutes:

$$\begin{array}{ccc} H & \xleftarrow{T^*} & K \\ \varphi^H \downarrow & & \downarrow \varphi^K \\ H^* & \xleftarrow{T^{b*}} & K^* \end{array}$$

**Proof.** We check that  $T^{b*} \circ \varphi^K = \varphi^H \circ T^*$ . Let an arbitrary element  $x \in K$  be given. On the one hand:

$$\varphi^H \circ T^*(x) = \varphi_{T^*x}^H$$

and hence for all  $y \in H$

$$[\varphi^H \circ T^*(x)](y) = \varphi_{T^*x}^H(y) = \langle y, T^*x \rangle_H = \langle Ty, x \rangle_K.$$

On the other hand:

$$(T^{b*} \circ \varphi^K)(x) = T^{b*}(\varphi_x^K) = \varphi_x^K \circ T$$

and hence

$$[(T^{b*} \circ \varphi^K)(x)](y) = [\varphi_x^K \circ T](y) = \varphi_x^K(Ty) = \langle Ty, x \rangle_K.$$

Since  $y$  is arbitrary, we have  $(T^{b*} \circ \varphi^K)(x) = (\varphi^K \circ T^*)(x)$ . And  $x$  is arbitrary, therefore  $T^{b*} \circ \varphi^K = \varphi^K \circ T^*$ .  $\square$

**Remark.** This theorem means that  $\{\varphi^H : H \in \text{Ob}(\text{Hil})\}$  is a natural transformation from the Banach functor to the Hilbert functor.

**Remark.** We can also define the adjoint operator by this. In other words: both  $\varphi^H, \varphi^K$  are conjugate isometries, we define  $T^*$  just by

$$T^* := (\varphi^H)^{-1} \circ T^{b*} \circ \varphi^K.$$

Linearity and boundedness are ensured by conjugate linearity and linearity of the maps and the fact that they are all bounded.

Just like we discuss Hermitian matrices in Linear Algebra, we discuss self-adjoint operator.

**Definition (Self-Adjoint).** Let  $H$  be a Hilbert space and  $T \in \mathcal{B}(H)$ . Then  $T$  is said to be a **self-adjoint operator**, if  $T^* = T$ .

**Theorem D.8.** Let  $H$  be a Hilbert space and  $T \in \mathcal{B}(H)$ . The following statements are equivalent.

1. the operator  $T$  is self-adjoint;
2. the functional induced by  $T$  is Hermitian:

$$\psi_T : H \times H \rightarrow \mathbb{K}, (x, y) \mapsto \langle Tx, y \rangle.$$

If  $\mathbb{K} = \mathbb{C}$ , they are both equivalent to

3. for all  $x \in H$ :  $\psi_T(x, x) = \langle Tx, x \rangle \in \mathbb{R}$ .

We need a general version of Polar Identity for this.

**Theorem D.9 (Polar identity).** Let  $H$  be a linear (normed) space. If  $\mathcal{S}$  is a sesquilinear functional  $\mathcal{S} : H \times H \rightarrow \mathbb{C}$  and  $\mathcal{Q} : H \rightarrow \mathbb{K}, x \mapsto \mathcal{S}(x, x)$ , then for all  $x, y \in H$  we have

$$\mathcal{S}(x, y) = \sum_{k=0}^3 \frac{i^k}{4} \mathcal{Q}(x + i^k y).$$

**Proof.** Just like Theorem 12.4. □

**Proof of Theorem D.8.** Denote  $\psi_T$  by  $\psi$  for convenience.

1  $\iff$  2: suppose  $T$  is self-adjoint, i.e. for all  $x, y \in H$ :

$$\langle x, Ty \rangle = \langle Tx, y \rangle.$$

Thus,  $\psi(x, y) = \overline{\psi(y, x)}$  holds for all  $x, y \in H$ . Therefore  $\psi$  is Hermitian. And vice versa.

2  $\implies$  3: suppose  $\psi$  is Hermitian. Then for all  $x \in H$ , we have

$$\langle Tx, x \rangle = \psi(x, x) = \overline{\psi(x, x)} = \overline{\langle Tx, x \rangle}$$

thus  $\langle Tx, x \rangle \in \mathbb{R}$ .

From now on, suppose  $\mathbb{K} = \mathbb{C}$ . 3  $\implies$  2: consider the Polar Identity (general version) and we have: for all  $x, y \in H$ :

$$\begin{aligned} 4\psi(x, y) &= \sum_{k=0}^3 i^k \psi(x + i^k y, x + i^k y). \\ 4\psi(y, x) &= \sum_{k=0}^3 i^k \psi(y + i^k x, y + i^k x). \end{aligned}$$

Then apply  $\psi(z, z) \in \mathbb{R}$  for all  $z \in H$ , check

$$\operatorname{Re}(4\psi(x, y)) = \operatorname{Re}(4\psi(y, x)), \operatorname{Im}(4\psi(x, y)) = -\operatorname{Im}(4\psi(y, x))$$

and we're done. □

## References

- [1] S. Axler. *Measure, Integration & Real Analysis*. Graduate Texts in Mathematics. Springer International Publishing, 2019. ISBN: 9783030331429.
- [2] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.
- [3] A. Ya. Helemskii. *Lectures and exercises on functional analysis*. Translations of mathematical monographs 233. AMS Bookstore, 2006. ISBN: 9780821840986.
- [4] J.L. Kelley. *General Topology*. Graduate Texts in Mathematics. Springer New York, 1975.
- [5] J. Michael Steele. *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities (Maa Problem Books Series.)* illustrated edition. Maa Problem Books Series. Cambridge University Press, 2004.