Functional Analysis Class Notes

Typer: Hao Fan HaoFan271@gmail.com

2022 Autumn Semester, 1-16 Weeks
Version: 2022-12-21 06:44:03+08:00

中南大學

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 $\mathrm{Re}, \mathrm{Im} \colon \mathbb{C} \to \mathbb{R}$ 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 $\{k \in \mathbb{N} \colon 1 \leq k \leq n\}$ 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,

$$\mathbb{C} := {}^{\mathbb{R}[t]}\!\big/\!_{(1+t^2)}$$

  means $\mathbb{C}$ is defined as a quotient ring.
- For each set A, the identity map is $\mathrm{id}_A \colon A \to A, a \mapsto a$. It's also used in the case that $A \subseteq X$, then $\mathrm{id}_A \colon A \to X, a \mapsto a$.
- For a mapping $f \colon A \to B$, we write $A = \mathrm{dom}(f), B = \mathrm{cod}(f)$. 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

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

  and maximum function

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

- Subtraction of sets A, B is $A \setminus B := \{x \in A : x \notin B\}$. We write the union of A and B by $A \bigsqcup B$ if $A \cap B = \varnothing$.
- For a set A, $\mathcal{P}(A)$ means the power set of A.

  For a set A, and its subset B, the mapping χ_B is defined by

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

  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 $x\colon \mathbb{N} \to X, n \mapsto x_n$ (in other words, $x \in {}^{\mathbb{N}}X$), and $x\colon \mathbb{N} \to X$ 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 $f, g\colon X \to V$ be given, we define the function

$$f + g\colon X \to V, x \mapsto f(x) + g(x).$$

- For $f\colon X \to \mathbb{K}$ 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 $\lim_{n\to\infty}$ for short.

- Given a linear map $f\colon X \to Y$ 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

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

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

- The sign function is defined as follows

$$\operatorname{sign}\colon \mathbb{K} \to \mathbb{K}, z \mapsto \begin{cases} 0, & z = 0; \\ \overline{z}/|z|, & z \neq 0. \end{cases}$$

- The rounding down function is

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

- Let $(\Omega, \mathcal{F})$ be a measurable space. We say a function $f\colon \Omega \to \mathbb{R}$ 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 σ-algebra.

- The symbol $\square$ means that a proof or solution ends. The symbol $\blacksquare$ means that an example or a remark ends (I wouldn't use this everytime).

- Somewhere you can see color different, that is reminding you to think about what here should be. (Just like $1 + 1 = 2$.)

Contents

0 Introduction

Here is something about this note.

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:
    - Projection Theorem;
    - Projection Operator.
  - Riesz Representation Theorem (on Hilbert space):

- Applications:
  * Sesquilinear functional and Representation Theorem;
  * Hilbert Adjoint Operator:
  * Self-Adjoint Operator;

1 Week 1

1.1 Lecture 1-1

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 $+ \colon X \times X \to X$ is called the addition on X, and $\cdot : \mathbb{K} \times X \to X$ 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 $\varphi \colon X \to Y$ 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	$\mathsf{Lin}_{\mathbb{K}}$	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 $\| \ \| \colon X \to \mathbb{R}_{\geq 0}$ 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 $\| \ \|$ is called a **norm** over X, and $(X, \| \ \|$ is called a **linear normed space**.

Remark 1.1. There is some similar weaker definitions:

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

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

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

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

$$\| \|: \mathbb{R}^n \to \mathbb{R}_{\geq 0}, x = (x_1, \ldots, x_n) \mapsto \Big(\sum_{j=1}^{n} x_j^2\Big)^{1/2} (= d(x, 0)).$$

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)**.** $\big(C([a,b],\mathbb{K}), \max_{[a,b]}| |\big)$ 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:

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

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 1.2. 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 there is a sequence $(x_n)_{n\in\mathbb{N}} \subseteq [a,b]$ converges to x, then:
$$\begin{aligned}\lim_{n\to\infty}(f+g)(x_n) =& \lim_{n\to\infty}\big(f(x_n) + g(x_n)\big)\\ =& \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n)\\ =&f(x) + g(x)\\ =&(f+g)(x).\end{aligned}$$

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

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

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

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

$$\begin{tikzcd}
{[a,b]} \arrow[r, "\langle f,g\rangle"] \arrow[dr, "f+g"'] & \mathbb{R}\times\mathbb{R} \arrow[d, "+"] \\
& \mathbb{R}
\end{tikzcd}
\qquad
\begin{tikzcd}
x \arrow[r, mapsto] \arrow[dr, mapsto] & (f(x),g(x)) \arrow[d, mapsto] \\
& f(x)+g(x)
\end{tikzcd}$$

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 $+\colon\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ 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 $f\colon X\to\mathbb{C}$ is continuous if and only if both $\mathrm{Re}(f),\mathrm{Im}(f)$ are continuous.

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

$$\begin{aligned}
\ell_p &:=\Big\{(a_n)_{n\in\mathbb{N}}:\sum_{n\geq1}|a_n|^p<\infty\Big\}, \text{(for } p<\infty)\\
\ell_\infty &:=\Big\{(a_n)_{n\in\mathbb{N}}:\sup_{n\in\mathbb{N}}|a_n|<\infty\Big\}.
\end{aligned}$$

And norms are

$$\begin{aligned}
\|a\|_p &:=\Big(\sum_{n\geq1}|a_n|^p\Big)^{1/p}, && \text{(for } p<\infty)\\
\|a\|_\infty &:=\sup_{n\in\mathbb{N}}|a_n|. && \text{(Here } a \text{ means } (a_n)_{n\in\mathbb{N}})
\end{aligned}$$

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

Proof. Clearly ℓ_∞ is a vector space. Now we prove $\|\ \|_\infty$ 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 1.3. In general, the inequality $\|f + g\|_p \leq \|f\|_p + \|g\|_p (p \geq 1)$ is called the Minkowski's inequality.

Example 4. ℓ_∞ 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 Lecture 1-2

1.2.1 Lebesgue integrable function spaces

Recall the left problem: Minkowski's inequality, which makes $(\ell_p, \| \|_p)$ 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:

$$\|ab\|_1 \leq \|a\|_p \|b\|_q, \tag{1}$$

Remark 1.4. $q = p/(p-1)$ is also called the dual index of p, usually denoted by p'.

Remark 1.5. 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? [6] Lagerange's Idendity, 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,"

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

(2) proves Cauchy's inequality as follows.

Proof of Cauchy's inequality from (2). Without lost 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 $\sum_{n \geq 1}(a_n')^2 = \sum_{n \geq 1}(b_n\prime)^2 = 1$. Now (2) implies

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

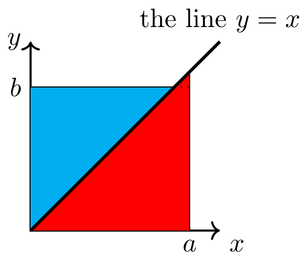


Figure 1: Area meaning of (2)

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

$$\sum_{n\geq 1} a_n b_n \leq \Big(\sum_{n\geq 1} a_n^2\Big)^{1/2}\Big(\sum_{n\geq 1} b_n^2\Big)^{1/2}. \qquad \square$$

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

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

(3) is the perfect analog of the "humble boun"(2).

Proof of (2). There is many ways to to this, see[6]. 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

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

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. $\square$

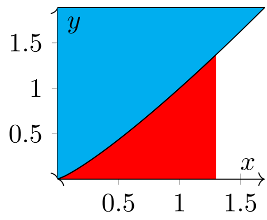


Figure 2: Fucking Area meaning of (3)

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

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

Proof of Minkowski's inequality.

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

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 $p - p/q = 1$. $\square$

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

Definition (σ-algebra)**.** A **σ-algrbra** on a set Ω is a subset Ω, satisfying:

1. $\Omega, \varnothing \in \mathcal{F}$;
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 Ω is an aritrary set and $\mathcal{F}$ is a σ-algebra over Ω. Elements of $\mathcal{F}$ is called **measurable sets** of $(\Omega, \mathcal{F})$.

Definition (Measure, Measure space)**.** A **measure** is a σ-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 μ 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 [3].

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:

$$f \longmapsto \int f^+ \,\mathrm{d}\mu - \int f^- \,\mathrm{d}\mu,$$

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

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

Fact. The proof of Minkowski's inequalityover ℓ_p actually proved the Minkowski's inequalityof **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 $\mu\colon \mathcal{F} \to [0, \infty], A \mapsto \sharp A$. Where $\sharp A = \infty$ if A is an infinite set, and $\sharp A = n$ if A has exactly n elements. μ is called the **counting measure** over $(\Omega, \mathcal{F})$.

Remark 1.6. It can be shown that,[1] for real sequence $(a_n)_{n\in\mathbb{N}}$ (equivalent to a function $a\colon \mathbb{N} \to \mathbb{R}$), we have

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

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

Remark 1.7. We can also prove Minkowski's inequalityof L_p by using the $L_{p'}$. Since

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

2 Week 2

2.1 Lecture 2-1

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:

$$\oplus : {}^{X}\!/_{X_0} \times {}^{X}\!/_{X_0} \to {}^{X}\!/_{X_0}, ([x], [y]) \mapsto [x + y];$$
$$\odot : \mathbb{K} \times {}^{X}\!/_{X_0} \to {}^{X}\!/_{X_0}, ([x], k) \mapsto [kx].$$

Where $[x + y]$ means the addtion (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. $\varnothing$. Since (3) implies $[x] \cap [y] \neq \varnothing$ 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 $a = y + \big(w + (x - y)\big)$ and $\big(w + (x - y)\big) \in X_0$ so $a \in [y]$. Above all, $[x] \subseteq [y]$. It is the same to know $[y] \subseteq [x]$.

3. Since

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

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 $\operatorname{codim}_X X_0 := \dim {}^{X}\!/_{X_0}$. Also denoted by just $\operatorname{codim}(X_0)$ if there is no confusion.

Claim. $\operatorname{codim}_c c_0 = 1$.

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 2.1. There is an isomorphism from ${}^{c}\!/_{c_0}$ to $\mathbb{K}$: $[(x_n)_{n\in\mathbb{N}}] \mapsto \lim_n x_n$.

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

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

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

$$ \|\ \|\colon {}^{X}\!/_{X_0} \to \mathbb{R}_{\geq 0}, [x] \mapsto \inf_{y \in X_0} \|x - y\|. $$

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

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

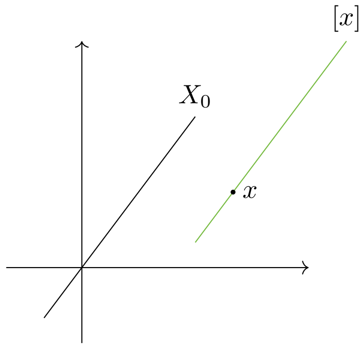


Figure 3: X, X_0 and points of ${}^{X}\!/_{X_0}$

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 Theorem 2.1 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

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

Therefore we have

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

Since ε 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_ε may not coincide with w_ε. To fix this, write

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

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, $\| \ \|$ is actually a norm. □

Remark 2.2. We define the topology of linear normed space as follows:

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

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 2.3. A quotient semi-norm in ${}^{X}\!/_{X_0}$ is a norm if and only if X_0 is closed.

2.2 Lecture 2-2

2.2.1 Metric Spaces

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

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

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

Remark 2.4. Every metric space has a topology, we will discuss this later.

Remark 2.5. Let's compare normed spaces and metric spaces: normed space need linear structures but metric spaces don't need. A normed space $(X, \|\ \|)$ is naturally a metric space by the metric induced by norm $d\colon X \times X \to \mathbb{R}, (x,y) \mapsto \|x - y\|$.

Remark 2.6. Let X be an arbitrary set, we can define a metric on X by the Kronecker symbol δ.

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

$$d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \big((x_1, \ldots, x_n), (y_1, \ldots, y_n)\big) \mapsto \Big(\sum_{j=1}^{n} (x_j - y_j)^2\Big)^{1/2}.$$

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

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

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 $f\colon [0,\infty) \to [0,1), x \mapsto x/(1+x)$. 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 Ω 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 μ-almost everywhere. Now consider the quotient space ${}^{\mathcal{L}_0}\!/_{\mathcal{Z}}$. We define

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

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

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

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 $f - g = f' - g'$ μ-almost everywhere, and hence $\frac{|f-g|}{1+|f-g|} = \frac{|f'-g'|}{1+|f'-g'|}$ μ-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 $(x_n)_{n\in\mathbb{N}} \to x$ if there is no ambiguity.

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

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

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

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

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

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

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

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:

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

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

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

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

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

Now, from the dominated convergence theorem (1_Ω being the dominated function, here we need $\mu(\Omega) < \infty$), we have:

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

□

Topology of metric spaces

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

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

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

Remark 2.7. 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. □

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

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

Definition (Continuity for maps between Metric Spaces)**.** Let (X, d_X), (Y, d_Y) be two metric spaces. We say $f \colon X \to Y$ is continuous at $x \in X$, if $\forall \varepsilon > 0 \ \exists r > 0$ such that $f\big(B(x, r)\big) \subseteq B\big(f(x), \varepsilon\big)$. 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 $f \colon X \to Y$ 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 $(x_n)_{n\in\mathbb{N}} \to x$. $\forall\varepsilon > 0$, by continuity of f at x, $\exists r > 0$ such that $f\big(B(x,r)\big) \subseteq B\big(f(x),\varepsilon\big)$. 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 $\forall n > N$ $f(x_n) \in B\big(f(x),\varepsilon\big)$. 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,

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

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

Definition (Continuity for maps between Topological Spaces)**.** Let $(X,\mathcal{T})$, $(Y,\mathcal{U})$ be two topological spaces. We say $f\colon X \to Y$ is continuous if $\forall O \in \mathcal{U}$ $f^{-1}(O) \in \mathcal{T}$.

Theorem (Equivalence of Definitions of Continuity)**.** $f\colon (X,d) \to (Y,d)$ is continuous if and only if $f\colon (X,\mathcal{T}_{d_X}) \to (Y,\mathcal{T}_{d_Y})$ is continuous.

Remark 2.8. Here we mean $f\colon (X,d) \to (Y,d)$ is continuous, if it satisfies the definition of continuous maps between metric spaces. And "$f\colon (X,\mathcal{T}_{d_X}) \to (Y,\mathcal{T}_{d_Y})$ is continuous" means it satisfies the definition of continuous maps between topological spaces.

Proof. Suppose $f\colon (X,d) \to (Y,d)$ 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 $f^{-1}\big(B(y,r)\big) \in \mathcal{T}_{d_X}$. Suppose $f^{-1}\big(B(y,r)\big) \neq \varnothing$, 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 $f\big(B(x_1,\delta)\big) \subseteq B\big(f(x_1),r_1\big) \subseteq B(y,r)$. Therefore $B(x_1,\delta) \subseteq f^{-1}\big(B(y,r)\big)$. This means $f^{-1}\big(B(y,r)\big)$ contains a neighbourhood for each point of itself, and hence $f^{-1}\big(B(y,r)\big)$ is open.

Suppose $f\colon (X,\mathcal{T}_{d_X}) \to (Y,\mathcal{T}_{d_Y})$ is continuous. Then $\forall x \in X$, $f^{-1}\big(B(f(x),r)\big)$ is open for all $r > 0$. $x \in f^{-1}\Big(B\big(f(x),r\big)\Big)$ and $f^{-1}\big(B(f(x),r)\big)$ 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

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

We're done. $\square$

3 Week 3

3.1 Lecture 3-1

Recall

Every linear normed space $(X, \| \ \|)$ has a metric (induced by its norm) $d\colon X \times X \to \mathbb{R}, (x, y) \mapsto \|x - y\|$. 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 $\| \ \|$ such that d is induced from $\| \ \|$? If there is a norm that we want, it is clear that $\| \ \|\colon X \to \mathbb{R}, x \mapsto \|x\| := d(x, 0)$. We want $\| \ \|$ is a norm, so it should satisfy:

1. $\| \ \| \geq 0$ 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 $\| \ \|$ satisfies the triangle inequality.

Proof.

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

We're done. $\square$

Here comes an important notion of functional analysis.

3.1.1 Banach Space

Definition (Banach Space)**.** A **complete** linear normed space $(X, \| \ \|)$ 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 3.1. 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

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

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

Example 14. Let $d\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}, (x, y) \mapsto |x - y|$ 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 $\big(C[0, 1], \| \ \|_{L_1}\big)$, where $\| \ \|_{L_1}$ means the norm

$$\| \ \|_{L_1} : C[0, 1] \to \mathbb{R}, f \mapsto \int_{[0,1]} |f| \, \mathrm{d}m.$$

This is a norm, since $\|f\|_{L_1} = 0 \iff |f| = 0$ 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 $\| \ \|_{L_1}$) in L_1.

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

Remark 3.2. 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 $P[a, b] :=$ {Polynomial functions defined on$[a, b]$}, then the linear normed space $(P[a, b], \max_{[a,b]}| \ |)$ is incomplete. Since

$\exists f \in C[a,b]$ such that f is not a polynomial, such as $f = \exp|_{[a,b]}$. Suppose $\exp\colon \mathbb{R} \to \mathbb{R}$ 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

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

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

Now we compare two normed spaces sharing the underlying set $C[a,b]$. $C[a,b]$ means the normed space $(C[a,b], \max_{[a,b]}|\ |)$ 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	$\max_{[a,b]}\lvert\ \rvert$	$\lVert\ \rVert_p$
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 $\big(f_n(x)\big)_{n\in\mathbb{N}}$ is a Cauchy sequence. By the completeness of $\mathbb{R}$, $\big(f_n(x)\big)_{n\in\mathbb{N}}$ converge. Then we can define a function

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

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

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

$\forall n > N$, 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 $m \to \infty$, using the continuity of $|\ |$ (to change the order of $\lim_m$ and $|\ |$)

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

which is equivalent to

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

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

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 $f_{n_j} \to f$ μ-a.e.. Therefore, by **Fatou's Lemma**:

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

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

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

and

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

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

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

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

If $f \in L_p$, we are done. f is a μ- 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

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

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

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

Let $m \to \infty$

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

and hence

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

Let $n \to \infty$ and use

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

We're done. $\square$

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

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

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

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

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

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

$$|f_m - f_n|_* \mu[2^{-j}, \infty) < 2^{-j} \ \forall m, n \ge n_j.$$

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

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

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

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

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

Exercise 3.1. Prove that ℓ_p is complete when $1 \le p < \infty$.

Suppose $(X, \| \ \|)$ 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

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

Proposition 3.1. $(X, \| \ \|)$ 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

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

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 $\left(x_{n_j}\right)_{j\in\mathbb{N}}$ such that

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

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

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

Therefore, $\sim_{n\geq 1} \|y_j\| < \infty$. 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 Lecture 3-2

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 μ-a.e..

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

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

2. About quotient space. Given a normed space $(X, \|\ \|)$ 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 $(p, p(x) = 0 \not\Longrightarrow x = 0)$. 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 Remark 2.3

   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. □

Now, Remark 2.3 ensures that $\|\ \| : X/X_0, [x] \mapsto p(x)$ is a norm on X/X_0.

Proof. It should be verified that p is well-defined (though this should have been proved in Remark 2.3). 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) $\|[x]\| = 0 \iff p(x) = 0 \iff x \in X_0 = [0] \iff [x] = [0] (\in {}^{X}\!/_{X_0})$.

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

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

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

3.2.1 Completion

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

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

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

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

Remark 3.3. 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 $(X, \| \; \|)$ be a liner 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

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

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

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

Remark 3.4. 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 $\iota\colon X \to \widehat{X}$ satisfies $\forall x \in X\colon \|\iota(x)\| = \|x\|$ 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 $p\colon \mathcal{E} \to \mathbb{R}, x (= (x_n)_{n\in\mathbb{N}}) \mapsto \lim_n \|x_n\|$. 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 semimorn 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 $\| \ \|\colon \widehat{X} \to \mathbb{R}, x + N \mapsto p(x)$.

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

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

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

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

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 $\forall [x] = (x_1, \ldots, x_n, \ldots) + N \in \widehat{X}$, there is a sequence in X_0 converge to X. Let

$$[x]^{(m)}\colon \mathbb{N} \to [(x_m)_{n\in\mathbb{N}}] = (x_m, \ldots, x_m, \ldots) + N,$$

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

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

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

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

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

We find that

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

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 $([x]^{(n)})_{n\in\mathbb{N}} \to [y] = (y_1, \ldots, y_n, \ldots) + N$. By definition of $\|\ \|_{\widehat{X}}$

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

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

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

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

Remark 3.5. 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 $\| \ \|$ 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 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 Lecture 4-1

Recall

No recall today.

4.1.1 Exercise course

We have only 3 exercises this course.

Question 4.1. Let $(X, \| \|)$ 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. $T\colon X \to X$ 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 4.1. This result doesn't hold when $\lambda = 1$. To see this, consider

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

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

Question 4.3. Let $(X, \| \|)$ be a linear normed space. Then X is a Banch 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 $\lim_n \operatorname{diam}(A_n) = 0$.

There are answers in the next section.

4.2 Lecture 4-2

Recall

For all l.n.s $(X, \| \|)$, 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 4.2. 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 Question 4.1. 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 $(x_n)_{n\in\mathbb{N}} \to x$. Now, X_0 is closed and hence $x \in X_0$. Thus, $(x_n)_{n\in\mathbb{N}} \to x \in X_0$. 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. □

Proof of Question 4.2. Let a be an arbitrary point in X. Define a sequence inductively:

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

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

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

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

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

The triangle inequality implies that

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

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 $(x_n)_{n\in\mathbb{N}} \to x$. 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. Supppose 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. □

Proof of Question 4.3. I think this proof is similar to the proof of [5, 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 $\forall n \in \mathbb{N}\ A_n \neq \varnothing$. 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

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

Therefore, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Then the completeness of X ensures that $\exists a \in X$ such that $(x_n)_{n\in\mathbb{N}} \to a$. $\forall n \in \mathbb{N}$, 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 $\bigcap_{n\geq 1} A_n$ cann't have more than 1 elements. If so, $\exists y \in A_n \forall n \in \mathbb{N}$ and hence $\operatorname{diam}(A_n) \geq d(x, y) \geq 0$. 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 $\operatorname{diam}(\overline{A_n}) = \operatorname{diam}(A) \to 0$ since $(x_n)_{n\in\mathbb{N}}$ is Cauchy. The reason of $\operatorname{diam}(\overline{A_n}) = \operatorname{diam}(A_n)$ is written in remark. Therefore, $\exists! a \in \bigcap_{n\geq 1} A_n$. Now, it suffices to show that $(x_n)_{n\in\mathbb{N}} \to a$. This follows from

$$d(x_n, a) \leq \operatorname{diam}(\overline{A_n}) \to 0(n \to \infty).$$ □

Remark 4.3. $\forall n \in \mathbb{N}$, we want to show that $\operatorname{diam}(\overline{A_n}) = \operatorname{diam}(A_n)$. 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 $\|x_\varepsilon - y_\varepsilon\| \leq \operatorname{diam}(A)$,

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

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

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

And ε is arbitrary, so

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

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 $T\colon X \to X$ is called a contraction if $\exists \lambda \in (0, 1)$ such that $d(T(x), T(y)) \leq \lambda d(x, y)$.

Remark 4.4. Every linear normed space $(X, \| \,\|)$ has the natural metric $d(x, y) = \|x - y\|$ and hence a contraction on $(X, \| \,\|)$ means $T\colon X \to X$, such that $\exists \lambda \in (0, 1), \forall x, y \in X$

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

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. $\square$

Let's have some applications. Suppose X is a Banach space and $U\colon X \to X$. 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 $T\colon X \to X, x \mapsto U(x) + x - y$. And

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

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. $(U - \mathrm{id})(x) - y = 0$. So the discussion above tells us that we should consider $T = U - \mathrm{id} + \mathrm{id} - y = U - y$. Let $x_1 \in X$ be an arbitrary point. Define $T(x) = U(x) - y$ 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 Lecture 5-1

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

Recall

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

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

5.1.1 Bounded Linear Operators/Maps

Here is the definition of Bounded linear operators/maps

Definition (Bounded linear operators/maps)**.** $T\colon X \to Y$ is said to be bounded, if $\exists C > 0$ such that $\| \|_Y \circ T \leq C\| \|_X$, equivalently $\|Tx\|_Y \leq C\|x\|_X$, $\forall x \in X$. 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 5.1. $\exists C > 0 : \|Tx\|_Y \leq C\|x\|_X, \forall x \in X$ is **not** equivalent to $\forall x \in X \exists C > 0 : \|Tx\|_Y \leq C\|x\|_X$.

Remark 5.2. 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 5.3. 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

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

1 $\Longrightarrow$ 2: is done automatically.

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

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

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

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

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

3 $\Longrightarrow$ 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 $\bigcup_{0\le t\le 1} tA$, 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 seuqnce $(y_n)_{n\in\mathbb{N}} \subseteq A$ 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 $\Longrightarrow$ 1: T is bounded, then T is uniformly continuous. □

Remark 5.4. There is another proof for 3 $\Longrightarrow$ 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

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

and

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

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

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

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

Remark 5.5. $(\mathcal{B}(X, Y), \| \ \|)$ is a linear normed space.

Remark 5.6. Equivalent definitions:

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

Proof. Since

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

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 5.7. 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

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

Show that:

1) T is bounded;

2) $\|T\| = \|a\|_1$.

Proof.

1) Recall that $c_0 \hookrightarrow \ell_\infty$ is equipped with the norm $\| \ \|_\infty = \sup_{\mathbb{N}} | \ |$. $\forall x \in c_0$, 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 $\|Tx\|_{\ell_1} \leq C\|c\|_\infty$. 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 $\| \ \|_1$

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

In particular

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

Now consider

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

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 $\|T\| \geq \|a\|_1$.

$\square$

Remark 5.8. 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

$$f \colon X \to \mathbb{K}, x \mapsto \int_0^{1/2} x \, \mathrm{d}m - \int_{1/2}^1 x \, \mathrm{d}m.$$

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

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

2) $\|f\| = 1$.

5.2 Lecture 5-2

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

$$\int_0^{1/2} x \, \mathrm{d}m = 1/2, \int_{1/2}^1 x \, \mathrm{d}m = -1/2.$$

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

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

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

$$T_\alpha \colon L_\infty(\Omega) \to L_1(\Omega), x \mapsto \alpha \cdot x,$$

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,

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

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

Remark 5.9. We have proved this for μ 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 $T \colon \mathbb{K}^n \to \mathbb{K}^n, x \mapsto Tx$ is bounded.

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

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

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

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

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

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

which is a bijection from definition of base. And φ is bounded, since

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

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

Now we prove that $\Phi := \varphi^{-1}$ is bounded. Given $(x_1, \ldots, x_n) \in \mathbb{K}^n$ 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 $\Phi(\sum_{j=1}^n x_j \alpha_j) = (x_1, \ldots, x_n)$ lies in some ball of $\mathbb{K}^n$. $\{\alpha_1, \ldots, \alpha_n\}$ 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

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

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

$$T_\alpha \colon L_p(\Omega) \to L_p(\Omega), x \mapsto \alpha \cdot x.$$

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

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

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

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

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

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

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

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 μ is defined as the unique measure such that

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

where $\operatorname{card}(A)$ is the number of elements of the set A when A is finite, and ∞ 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

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

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

Therefore, the operator

$$T \colon L_\infty(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu) \to \mathcal{B}\big(L_p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \nu)\big)$$

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

6 Week 6

6.1 Lecture 6-1

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 6.1. 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 $(x_{n_k})_{k \in \mathbb{N}} \to x \in A$.

Definition (ε-net)**.** Let (X, d) be a metric space. $E \subseteq X$ is called an ε-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 ε-net of A.

Remark 6.2. 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 $a_1, \ldots, a_m \in A$ such that

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

Let $N := N_{a_1} \vee \cdots N_{a_m}$ then $x_N \notin B(a_k, \varepsilon_{a_k}) \forall 1 \leq k \leq m$. But $x_N \in A = \bigcup_{k=1}^{m} B(x_k, \varepsilon_{a_k})$. 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 λ 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 λ 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 $x_3, \ldots, x_n, \ldots$ 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

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

Recall the selection of λ, 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 6.3. 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 ${x_{n_k}}_{k\in\mathbb{N}} \to x$ 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 ${x_{n_k}}_{k\in\mathbb{N}} \to x$ whenever ${y_{n_k}}_{k\in\mathbb{N}} \to x$. 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 $(y_{n_k})_{k\in\mathbb{N}} \to x \in X$, and hence $(x_{n_k})_{k\in\mathbb{N}} \to x \in X$ 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 $(x_n)_{n\in\mathbb{N}} \subseteq A$. WLOG, suppose $(x_n)_{n\in\mathbb{N}}$ has infinite distinct terms, else we're done. $\forall \varepsilon > 0$ there is a finite ε-net of A.

Thus for each $k \in \mathbb{N}$ there is a finite ε-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 ε_0-net. Thus pick an arbitrary point $x_1 \in \mathbb{N}$ and $X \setminus B(x_1, \varepsilon_0) \neq \varnothing$ (since A has no finite ε_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

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

(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 [5, Chapter 5, Thm 14] implies this. □

Remark 6.4. The inverse proposition of 2 is **incorrect**. Consider $(\mathbb{R}, d_1)$ where d_1 is defined as

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

where

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

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 Lecture 6-2

Recall

Let X, Y be two linear normed spaces.

- $T\colon X \to Y$ 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. $\exists T\colon X \to Y$ 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 6.5. In the category $\mathsf{Vect}_{\mathbb{K}}$, an isomorphism is a linear bijection and vice versa. In the category Nor: $\mathrm{Ob}(\mathsf{Nor})$ are normed spaces and $\mathrm{Mor}(\mathsf{Nor})$ are bounded linear maps. An isomorphism in Nor is a linear homeomorphism. In the category Nor_1: $\mathrm{Ob}(\mathsf{Nor}_1)$ are normed spaces and $\mathrm{Mor}(\mathsf{Nor}_1)$ are contraction operators. An isomorphism in 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 $(X, \| \ \|_1)$, $(X, \| \ \|_2) \in \mathrm{Ob}(\mathsf{Nor})$. We say $\| \ \|_1$ is equivalent to $\| \ \|_2$, if $\exists a, b > 0$ such that

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

Remark 6.6. $\sim$ 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

$$\mathrm{id}\colon (X, \| \ \|_2) \to (X, \| \ \|_1), x \mapsto x$$

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 $X \in \mathrm{Ob}(\mathsf{Nor})$ with $\dim(X) = n < \infty$, then $X \simeq \mathbb{K}^n$.

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

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

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

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

Then let

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

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

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

i.e.

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

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

Remark 6.7. Consider $\min \Phi|_S$ is natural, just like

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

Corollary 6.6. Let $(X, \| \ \|) \in \mathrm{Ob}(\mathsf{Nor})$.

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 ℓ_2 and its base $\{e_n : n \in \mathbb{N}\}$, where

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

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 $(x_n)_{n\in\mathbb{N}} \to e_m \in B$.
- B is bounded: since $\operatorname{diam}(B) = \sqrt{2}$.
- 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. □

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 $y = \bar{x} - \|\bar{x} - x'\|x \in X_0$. □

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 $\operatorname{span}\{x_1\}$ (this is a closed linear subspace, see the third corollary of Theorem 6.5). From Lemma 6.7, there is $x_2 \in X \setminus \operatorname{span}\{x_1\}$ such that $\|x_2\| = 1$ and $d(x_2, \operatorname{span}\{x_1\}) > 1/2$. Then consider the closed linear subspace $\operatorname{span}\{x_1, x_2\}$ (that is closed by the same reason as $\operatorname{span}\{x_1\}$), $\operatorname{span}\{x_1, x_2\} \neq X$ and Lemma 6.7 implies that there is $x_3 \in X \setminus \operatorname{span}\{x_1, x_2\}$ such that $\|x_3\| = 1$ and $d(x_3, \operatorname{span}\{x_1, x_2\}) > 1$. 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 Lecture 7-1

7.1.1 Construct more Linear Normed Spaces

Let $(X_i, \| \ \|_{X_i}) \in \mathrm{Ob}(\mathsf{Nor}), 1 \le i \le n$. Define

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

with operations

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

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

$$\| \ \|_X : X \to \mathbb{R}, (x_1, \ldots, x_n) \mapsto \Big(\sum_{i=1}^{n} \|x_i\|_{X_i}^p \Big)^{1/p}.$$

At the case of $p = \infty$, $\|x\|$ should be interpreted like $\| \ \|_\infty$. To see that $\| \ \|$ 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 ℓ_p, since

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

The first inequality comes from the triangle inequality of $\| \ \|_X$ and the second inequality comes from the Minkowski's Inequality:

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

Now, we have some questions

Question 7.1.

1) Let $p_i \colon X \to X_i, (x_1, \dots, x_n) \mapsto x_i$ be the projection to the i- coordinate. Show that p_i is continuous and $\|p_i\| = 1$.

2) $X = \bigtimes_{i=1}^n X_i$ 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

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

Thus $\|p_i\| \leq 1$. On the other hand: taking $x = (0, \dots, x_i, \dots, 0) \in X$ with $\alpha_i \neq 0 \in X_i$, 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

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

in X. Then

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

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

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

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

$$\iota_i \colon X_i \to E_i, x \mapsto (0, \dots, x, 0, \dots).$$

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 $X \in \mathrm{Ob}(\mathsf{Nor})$ and $f \in \mathcal{L}(X, \mathbb{K})$ is an unbounded linear functional. Then

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

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 ball to a bounded set and hence f is bounded). Taking $x = \frac{\alpha}{f(x')}x'$, we're done since

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

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

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

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) $\Longrightarrow$ 2) $\{0\}$ is closed in $\mathbb{K}$ and **2)** follows from the topological definition of continuity.

2) $\Longrightarrow$ 1) $\ker f$ is closed and hence is not dense in X since $f \neq 0$. Therefore

$$\exists x_0 \in X \exists r > 0 (B(x_0, r) \cap \ker f = \varnothing.) \tag{13}$$

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 $f\big(B(0,r)\big) = \mathbb{K}$. 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 $B(x_0,r) \cap \ker f = \varnothing$.

□

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 $f\colon \ell_2 \to \mathbb{K}$ is not well-defined. There is an element in ℓ_2:

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

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

The set in **1)** can be proved to be not closed by the theory of **Hilbert Space**.

Proof. First, $M^\perp = \{0\}$. Let $x \in M^\perp$, then for all $n \in \mathbb{N}$:

$$x \perp e_1 - \sqrt{n}e_n \in M \implies x_n = 1/\sqrt{n}x_1.$$

Thus,

$$\infty > \sum_{n\geq 1} |x_n|^2 = |x_1|^2 \sum_{n\geq 1} 1/n \implies x_1 = 0,$$

and hence $x - 0$. □

2) Let

$$f\colon \ell_2 \to \mathbb{K}, x \mapsto \sum_{n=1}^{\infty} \frac{1}{n} x_n. \tag{14}$$

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

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

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

Remark 7.1. We have $\|f\| = \pi/\sqrt{6}$ by taking

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

since $\|x\|_{\ell_2} = \pi/\sqrt{6}$ and $|f(x)| = \pi^2/6$.

Remark 7.2. In fact,

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

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

7.2 Lecture 7-2

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 $f: X \to Y$ 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. □

Theorem 7.4 implies

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

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

Theorem 7.6 (Banach-Steinhaus)**.** Let $(X, \| \ \|_X)$ be a Banach space, $(Y, \| \ \|_Y) \in \mathrm{Ob}(\mathsf{Nor})$ 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 $\| \|_I$ be a new norm on X, defined as

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

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

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

And (15) is equivalent to

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

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

Since $(X, \| \|_X)$ is a Banach space, (16) implies that $(x_n)_{n\in\mathbb{N}} \xrightarrow{\| \|_X} x \in X$. Now we prove that $(x_n)_{n\in\mathbb{N}} \xrightarrow{\| \|_I} x \in X$ 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 $m \to \infty$, the continuity of $\| \|_Y$ and T_λ (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,

$$\forall n > N, \sup \|T_\lambda x_n - T_\lambda x\|_Y \leq \varepsilon.$$

Equivalently,

$$\lim_n \sup \|T_\lambda x_n - T_\lambda x\| = 0,$$

which was what we wanted. $\square$

7.2.2 Baire Category Theorem

Definition (G_δ-set, F_σ-set)**.** Let $(X, \mathcal{T}) \in \mathrm{Ob}(\mathsf{Top})$.

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

Remark 7.3. Here "G" is German (Gebiet) and "F" is French (Fermé).

Definition (First Category Set)**.** Let $(X, \mathcal{T}) \in \mathrm{Ob}(\mathsf{Top})$. $A \subseteq X$ is called a **set of the first category**, if $A \subseteq B$ for some F_σ set B with $\mathring{B} = \varnothing$.

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

Definition (Baire Space)**.** Let $(X, \mathcal{T}) \in \mathrm{Ob}(\mathsf{Top})$. $(X, \mathcal{T})$ 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 $(X, \mathcal{T}) \in \mathrm{Ob}(\mathsf{Top})$. $(X, \mathcal{T})$ 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 $\varnothing \neq U \in \mathcal{T}$, now we show that $O \cap U \neq \varnothing$.

Since O_1 is dense in X, we have $O_1 \cap U \neq \varnothing$ 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 $\mathring{F_1} \neq \varnothing$, $F_1 \subseteq O_1 \cap U$ and $\operatorname{diam} F_1 = r =: r_1$.

Since O_2 is dense in X and $\mathring{F_1} \neq \varnothing$, we have $O_2 \cap \mathring{F_1} \neq \varnothing$ and thus $\exists x_2 \in O_1 \cap U$. Moreover, $\exists r_2 > 0 \wedge r_2 < r_1/2$ such that $B(x_2, r_2) \subseteq O_2 \cap \mathring{F_1}$. Let $F_2 := \overline{B(x_2, r_2)}$. Then $\mathring{F_2} \neq \varnothing$, $F_2 \subseteq O_2 \cap U$ and $\operatorname{diam} F_2 < r_1/2$.

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

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

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

Therefore,

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

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

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 $\bigcup_{n \geq 1} \mathring{F}_n$ is dense in X.

Proof.

1) For $A \subseteq X$, let cl_A means the closure operator with respect to the subspace topology of A. Similarly, int means the interior operator.

Suppose $\Omega \subseteq X$ is open. Given $(O_n)_{n\in\mathbb{N}} \subseteq \Omega$ such that $\operatorname{cl}_\Omega(O_n) = \Omega(\forall n \in \mathbb{N})$, i.e. O_n is dense in Ω for all $n \in \mathbb{N}$. Since $\operatorname{cl}_\Omega(O_n) = \Omega \cap \overline{O_n}$, 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 $O_n \cup \left(\overline{\Omega}\right)^c$ is dense in X for all $n \in \mathbb{N}$. Therefore,

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

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

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

and hence

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

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

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

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

Above all, Ω is a Baire space.

2) Let $\Omega \neq \varnothing$ be an arbitrary open set in X. Then 7.8 implies that Ω 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 $\mathrm{int}(\Omega \cap F_n) \neq \varnothing$. Since "the interior of intersection is the intersection of union" and Ω is open, we have $\Omega \cap \mathring{F}_n \neq \varnothing$. Therefore

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

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

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

Proof of Theorem 7.6. Let

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

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

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

and $\|\ \|_Y, T_\lambda (\forall \lambda \in \Gamma)$ 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 $\mathring{M}_k \neq \varnothing$. There is $x_0 \in \mathring{F}_k$ and $r > 0$ such that $B_X(x_0, r) \subseteq \mathring{F}_k$. 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 $T\big(B_X(x_0,r)\big) \subseteq \overline{B_Y(0,k+M(x_0))}$ holds for all $\lambda \in \Gamma$, which implies

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

Above all, $\sup_{\lambda\in\Gamma}\|T_\lambda\| \leq (k+M(x_0))/r < \infty$. $\square$

Remark 7.4. To give $X = \bigcup_{n\geq 1} F_n$, we need the assumption

$$\forall x \in X\colon M(x) < \infty.$$

8 Week 8

8.1 Lecture 8-1

Recall

We have proved Theorem 7.8 by Open mapping theorem. We used the corollary 7.5 and proved that $\mathrm{cod}(\mathrm{id}_X)$ is a Banach space, where

$$\mathrm{id}_X : (X, \| \|_X) \to (X, \| \|_X + \sup_{\lambda\in\Gamma} \| \|_Y \circ T_\lambda)$$

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

$$\forall x \in X, (T_n x)_{n\in\mathbb{N}} \xrightarrow{\| \|_Y} Tx,$$

denoted as $(T_n)_{n\in\mathbb{N}} \xrightarrow{s} T$.

Remark 8.1. The relation between $(T_n)_{n\in\mathbb{N}} \xrightarrow{s} T$ and $(T_n)_{n\in\mathbb{N}} \xrightarrow{\| \|} T$ 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 $\|T\| \le \liminf_{n\to\infty} \|T_n\|$ such that

$$T_n \xrightarrow{s} T (n \to \infty).$$

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 $m, n \to \infty$ 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

$$T\colon X \to Y, x \mapsto \lim_n T_n x,$$

which is linear since both $T_n(\forall n)$ and $\lim_n$ is linear (i.e. T is a composition of two linear maps, $f_1\colon X \to \mathcal{E}, x \mapsto (T_n x)_{n\in\mathbb{N}}$ and $f_2\colon \mathcal{E} \to Y, (y_n)_{n\in\mathbb{N}} \mapsto \lim_n y_n$, 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 $\|\ \|$ is continuous. And $\forall n \in \mathbb{N}, \|T_n x\| \leq \|T_n\|\|x\|$, take lim inf on both sides and we get

$$\liminf\|T_n x\| \leq \liminf\|T_n\|\|x\|. \tag{18}$$

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 $(T_n)_{n\in\mathbb{N}} \xrightarrow{s} T$. □

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

$$T_n - T_m \xrightarrow{s} 0(m,n \to \infty), \tag{19}$$

we have

$$T_n - T \xrightarrow{s} 0(n \to \infty)$$

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 $\forall x \in X$, $(\|T_n x\|)_{n\in\mathbb{N}}$ in 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. □

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

$$(\forall k \in \mathbb{N}) f_k \colon \ell_p \to \mathbb{R}, (x_n)_{n\in\mathbb{N}} \mapsto \sum_{j=1}^{k} x_j \alpha_j.$$

Then f_k is linear, and bounded since Hölder's inequality implies

$$\begin{aligned} |f_k x| &\le \sum_{j=1}^{k} |x_j \alpha_j| \\ &\le \Big(\sum_{j=1}^{k} |x_j|^p\Big)^{1/p} \Big(\sum_{j=1}^{k} |\alpha_j|^q\Big)^{1/q} \\ &\le \Big(\sum_{j=1}^{k} |\alpha_j|^q\Big)^{1/q} \|x\|_p, \end{aligned}$$

i.e. $\|f_k\| \le \big(\sum_{j=1}^{k} |\alpha_j|^q\big)^{1/q}$. And the reversed inequality holds, to see this, consider the equality condition of Hölder's inequality (and triangle inequality) and hence pick

$$\ell_p \ni x^{(k)}, (x_n^{(k)})_{n\in\mathbb{N}} := (\operatorname{sign}(\alpha_1)|\alpha_1|^{q/p}, \ldots, \operatorname{sign}(\alpha_k)|\alpha_k|^{q/p}, 0, \ldots).$$

And

$$\left|f_k(x^{(k)})\right| = \sum_{j=1}^{k} |\alpha_j|^q, \left\|x^{(k)}\right\|_p = \Big(\sum_{j=1}^{k} |\alpha_j|^q\Big)^{1/p},$$

implies

$$\|f_k\| \ge \Big(\sum_{j=1}^{k} |\alpha_j|^q\Big)^{1-1/p} = \Big(\sum_{j=1}^{k} |\alpha_j|^q\Big)^{1/q}.$$

Above all, $\|f_k\| = \big(\sum_{j=1}^{k} |\alpha_j|^q\big)^{1/q}$.

By assumption, we have $\forall x \in \ell_p, \big(f_n(x)\big)_{n\in\mathbb{N}}$ 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

$$\sup_{n\in\mathbb{N}} \|f_n\| = \lim_{n\to\infty} \|f_n\| = \lim_{n\to\infty} \Big(\sum_{j=1}^{k} |\alpha_j|^q\Big)^{1/q} = \|\alpha\|_q.$$

Therefore, $\|\alpha\|_q < \infty$, i.e. $\alpha \in \ell_q$. □

Remark 8.2. 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

$$\{f\colon \mathbb{R} \to \mathbb{C} \mid f \text{ is continuous and } 2\pi \text{ -periodic}\},$$

with the norm

$$\| \|_\infty\colon C_{2\pi} \to \mathbb{R}, f \mapsto \sup_{x\in\mathbb{R}} |f(x)|.$$

Remark 8.3. The norm max is well-defined since f is 2π-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 $\mathbb{R} \to \mathbb{R}$. A number $T \in \mathbb{R}$ is called a **period** of f, if

$$f = \tau_T f,$$

where

$$\tau_T f\colon \mathbb{R} \to \mathbb{C}, x \mapsto f(x - T).$$

Function f has a period T is called a **T-periodic function**, or a **periodic function** for short.

Remark 8.4. Let

$$\operatorname{per}(f) := \{T \in \mathbb{R} : \tau_T f = f\},$$

then $\operatorname{per}(f)$ is a subgroup of the additive group $\mathbb{R}$. The structure of $\operatorname{per}(f)$ has only 3 possibilities:

1. $\operatorname{per}(f)=\{0\}$, i.e. f is not a periodic function.

2. $\operatorname{per}(f) = T_0\mathbb{Z} = \{T_0 k : k \in \mathbb{Z}\}$ for some $T_0 > 0$. And such T_0 is usually called the fundamental period or the minimum period.

3. $\operatorname{per}(f)$ is a dense subgroup of $\mathbb{R}$, equivalently f has no fundamental period. For example, $\operatorname{per}(\chi_{\mathbb{Q}}) = \mathbb{Q}$.

The Fourier series of a 2π-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

$$\widehat{f}\colon \mathbb{Z} \to \mathbb{C}, n \mapsto \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\mathrm{e}^{-in\theta}\,\mathrm{d}\theta.$$

We use the notation

$$f(x) \sim \sum_{n\in\mathbb{Z}} a_n \mathrm{e}^{inx}$$

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

$$\sum_{k=-n}^{n} \widehat{f}(k)\mathrm{e}^{-ikx} = \sum_{k=-n}^{n} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)\mathrm{e}^{ik(\theta-x)}\,\mathrm{d}\theta.$$

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

$$f * g\colon [-\pi,\pi] \to \mathbb{C}, t \mapsto \begin{cases} \int_{-\pi}^{\pi} f(t-x)g(x)\,\mathrm{d}\mu(x), & \text{if integral is finite} \\ 0. & \text{else} \end{cases}$$

Here μ 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

$$\int_{[-\pi,\pi]} f(t-x)g(x)\,\mathrm{d}\frac{m}{2\pi}(x) < \infty$$

for almost every $t \in [-\pi,\pi]$.

Definition (Dirichlet kernel)**.** Given $n \in \mathbb{N}$, the function

$$D_n\colon \mathbb{R} \to \mathbb{C}, x \mapsto \sum_{k=-n}^{n} \mathrm{e}^{ikx}$$

is called the n-th Dirichlet kernel.

By this definitions, we have a convolution formula for the n-th partial sum of Fourier series

$$S_n(f)(x) = \int_{-\pi}^{\pi} f(\theta) \sum_{k=-n}^{n} \mathrm{e}^{ik(x-\theta)} \frac{\mathrm{d}m}{2\pi}(x) = D_n * f(x),$$

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_δ 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

$$u_n : C_{2\pi} \to \mathbb{C}, f \mapsto S_n(f)(0). \tag{20}$$

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_δ 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

$$\begin{aligned}|u_n(f)| &= |S_n(f)(0)| \\ &\leq \int_{-\pi}^{\pi} \left| f(t) \sum_{k=-n}^{n} \mathrm{e}^{ikt} \right| \frac{\mathrm{d}m}{2\pi}(t) \\ &\leq \|f\|_\infty \|D_n\|_1 \text{(Hölder's inequality)}\end{aligned}$$

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

$$f\colon [0, 2\pi] \to \mathbb{C}, x \mapsto \begin{cases} 1, & x \notin I \wedge D_n(x) > 0; \\ -1, & x \notin I \wedge D_n(x) < 0; \\ l(x), & x \in I. \end{cases}$$

Here l is the affine mapping on each subinterval of I such that f is continuous.

$$\begin{aligned}|u_n(f)| &\geq \left| \int_{[-\pi,\pi]\setminus I} f(t) D_n(t) \frac{\mathrm{d}m}{2\pi}(t) \right| - \left| \int_I f(t) D_n(t) \frac{\mathrm{d}m}{2\pi}(t) \right| \\ &= \int_{[-\pi,\pi]\setminus I} |D_n(t)| \frac{\mathrm{d}m}{2\pi}(t) - \left| \int_I f(t) D_n(t) \frac{\mathrm{d}m}{2\pi}(t) \right| \\ &= \int_I |D_n(t)| \frac{\mathrm{d}m}{2\pi}(t) - 2\left| \int_I f(t) D_n(t) \frac{\mathrm{d}m}{2\pi}(t) \right| \\ &\geq \|D_n\|_1 - \|f\|_\infty \varepsilon.\end{aligned}$$

The first inequality is just $|x+y| \geq |x| - |y|$, and the last inequality follows from

$$\begin{aligned}\left| \int_I f(t) D_n(t) \frac{\mathrm{d}m}{2\pi}(t) \right| &\leq \int_I |f \cdot D_n| \frac{\mathrm{d}m}{2\pi} \\ &\leq \|f\|_\infty \int_I |D_n| \frac{\mathrm{d}m}{2\pi} \\ &\leq \|f\|_\infty (2n+1) \cdot \frac{m}{2\pi}(I) \\ &\leq \|f\|_\infty (2n+1) \cdot \frac{1}{2\pi} \frac{\pi\varepsilon}{2n+1}.\end{aligned}$$

since $\|D_n\|_\infty \leq 2n+1$. From $\|f\|_\infty = 1$, we get

$$\|u_n\| \geq \|D_n\|_1 - \varepsilon \ (\forall \varepsilon > 0),$$

which implies $\|u_n\| \geq \|D_n\|$.

Above all, $\|u_n\| = \|D_n\|$. Now we show that $(\|D_n\|)_{n\in\mathbb{N}}$ is unbounded. Let $\mathrm{d}x$ denote $\mathrm{d}m(x)$ for short

$$\begin{aligned}
\|D_n\|_1 &= \int_{-\pi}^{\pi} \frac{|\sin(n+1/2)x|}{|\sin(x/2)|} \frac{\mathrm{d}x}{2\pi} \\
&\geq \frac{1}{2\pi} \int_0^{\pi} \frac{|\sin(n+1/2)x|}{x/2} \,\mathrm{d}x \\
&= \frac{2}{\pi} \int_0^{\frac{2n+1}{2}\pi} \frac{|\sin x|}{x} \,\mathrm{d}x \\
&> \frac{2}{\pi} \int_0^{\pi} \frac{|\sin x|}{x} \,\mathrm{d}x \\
&= \frac{2}{\pi} \sum_{k=0}^{n-1} \int_0^{\pi} \frac{|\sin x|}{x + k\pi} \,\mathrm{d}x,
\end{aligned}$$

and

$$\int_0^{\pi} \frac{|\sin x|}{x+k\pi} \,\mathrm{d}x \geq \frac{1}{(k+1)\pi} \int_0^{\pi} |\sin x| \,\mathrm{d}x = \frac{2}{(k+1)\pi}.$$

Therefore

$$\|D_n\|_1 \geq \frac{4}{\pi^2} \sum_{k=1}^{n} \frac{1}{k} \sim \frac{4}{\pi^2} \log n \to \infty (n \to \infty).$$

Then, Theorem 8.5 implies the result.

□

Here is the proof of Theorem 8.5.

Proof of Theorem 8.5. Let

$$M \colon X \to [0,\infty], x \mapsto \sup_{i\in I} \|u_i(x)\|.$$

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 $\mathring{F}_m = \varnothing$ 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_δ-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 Ω_m is not dense in X, since X is a Banach space (hence a Baire space). But $\overline{\Omega_m} \neq X$ implies that $\mathring{F}_m \neq \varnothing$, 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 Lecture 8-2

The aim of this lecture is

Aim. Prove the open mapping theorem.

We haven't 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 $B \subseteq \mathring{F}$, where F is a F_σ set such that $\mathring{F} = \varnothing$.
- 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 $(O^c)^\circ = \varnothing$.
    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 $\mathring{F}_n = \varnothing$, we have $\left(\bigcup_{n\geq 1} F_n\right)^\circ = \varnothing$.

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 second 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 $A, B \subseteq X$,

$$kA := \{kx \in X : x \in A\}, A + B := \{x + y : x \in A, y \in B\}.$$

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_σ-set in Y. So $\left(\bigcup_{n\geq 1} \overline{T(nB_X)}\right)^\circ \neq \varnothing$ (else, $\bigcup_{n\geq 1} \overline{T(nB_X)}$ is a F_σ-set containing $T(X)$

while having empty interior, contradiction with $T(X)$ is a set of second category). While Y is a Banach space, hence a Baire space, and there is some $m \in \mathbb{N}$ such that $\left(\overline{T(mB_X)}\right)^\circ \neq \varnothing$. Thus, there is $y_0 \in \left(\overline{T(mB_X)}\right)^\circ$ and $r > 0$ such that

$$y_0 + rB_Y \subseteq \left(\overline{T(mB_X)}\right)^\circ \subseteq \overline{T(mB_X)}.$$

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, $\forall n \in \mathbb{N}$

$$\begin{aligned} y_1 &= y_2/2 + T(x_1) \\ &= y_3/2^2 + T(x_2)/2 + T(x_1) \\ &= \cdots \\ &= y_{n+1}/2^n + \sum_{j=1}^{n} T(x_j)/2^j \\ &= y_{n+1}/2^n + T\Big(\sum_{j=1}^{n} x_j/2^j\Big), \end{aligned}$$

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

$$\left\|\sum_{j=1}^n x_j/2^j\right\| \le \sum_{j=1}^n \|x_j/2^j\| < \sum_{j=1}^n \frac{c}{2^j} < c(\forall n \in \mathbb{N}).$$

and is convergent to some $x_0 \in cB_X$ because X is a Banach space. Therefore, by the continuity of T

$$\begin{aligned}
y_1 &= \lim_{n\to\infty} y_{n+1}/2^n + T\Big(\sum_{j=1}^n x_j/2^j\Big) \\
&= \lim_{n\to\infty} T\Big(\sum_{j=1}^n x_j/2^j\Big) \\
&= T\Big(\lim_{n\to\infty} \sum_{j=1}^n x_j/2^j\Big) \\
&= T(x_0).
\end{aligned}$$

Above all, $B_Y \subseteq T(cB_X) = cT(B_X)$. Then

$$Y = \bigcup_{n\ge 1} nB_Y \subseteq \bigcup_{n\ge 1} nT(cB_X) = T(\bigcup_{n\ge 1} ncB_X) = TX,$$

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 8.5. 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 second 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.

$$\begin{tikzcd}
\text{Thm } 8.6 \arrow[r, Rightarrow] & \text{Thm } 7.4 \arrow[d, Rightarrow] \arrow[dl, Rightarrow] & \\
\text{Thm } 8.8 & \text{Cor } 7.5 \arrow[l, Rightarrow] \arrow[r, Rightarrow] & \text{Thm } 7.6
\end{tikzcd}$$

Now we talk about the Closed Graph Theorem. Here is the natural definition of the graph of a operator.

Definition (Graph)**.** Let X, Y be two sets (allowed to have structures such as topology, norm and so on) and $T\colon X \to Y$ 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 8.6. In this lecture, we assume $X \times Y$ is a linear normed space, the norm of $X \times Y$ is

$$\| \,\| \colon X \times Y \to \mathbb{R}, (x, y) \mapsto \|x\|_X + \|y\|_Y,$$

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, $T\colon X \to Y$ is a map. Then T is said to be closed, if $G(T)$ is closed.

Recall that, a sequence $\big((x_n, y_n)\big)_{n\in\mathbb{N}} \subseteq X \times Y$ converges to $(x, y) \in (X \times Y, \| \,\|_p)$, if and only if

$$\lim_n \|(x_n, y_n) - (x, y)\|_p = \lim_n \Big(\|x_n - x\|^p + \|y_n - y\|^p\Big)^{1/p} = 0,$$

i.e. $(x_n)_{n\in\mathbb{N}} \xrightarrow{\| \,\|_X} x \wedge (y_n)_{n\in\mathbb{N}} \xrightarrow{\| \,\|_Y} y$.

Proposition 8.7. Let $X, Y \in \mathrm{Ob}(\mathsf{Nor})$ 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 $\lim_n x_n = x \wedge \lim_n Tx_n = y$, we have $\big((x_n, Tx_n)\big)_{n\in\mathbb{N}} \subseteq G(T)$ converges to $(x, y) \in X \times Y$. While $G(T)$ is closed, we get $(x, y) \in G(T)$ thus $y = Tx$.

For sufficiency of 1: suppose T satisfies the latter condition in 1. Given an arbitrary convergent sequence $\big((x_n, y_n)\big)_{n\in\mathbb{N}} \subseteq G(T)$, i.e. $(x_n)_{n\in\mathbb{N}} \subseteq X$, $y_n = T(x_n)$ for all $n \in \mathbb{N}$ and $(x_n)_{n\in\mathbb{N}} \to x \in X$. Then the latter condition implies $\lim_n y_n = \lim_n Tx_n$. And continuity of T implies $T(x) = \lim_n y_n$. Thus the limit of $\big((x_n, y_n)\big)_{n\in\mathbb{N}}$, 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 $X, Y \in \mathrm{Ob}(\mathsf{Ban})$. If $T\colon X \to Y$ 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

$$\begin{aligned} p\colon G(T) &\to X \\ (x, Tx) &\mapsto x, \end{aligned}$$

which is a linear bijection and $\|p\| \leq 1$ since

$$\forall x \in X, \left\|p\big((x, Tx)\big)\right\|_X = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\|_{X \times Y}.$$

Theorem 7.4 implies $p^{-1} \in \mathcal{B}(X, G(T))$. Therefore,

$$\begin{aligned} \forall x \in X, \|Tx\| &\leq \|x\| + \|T_x\| \\ &= \|(x, Tx)\|_{X \times Y} \\ &= \left\|p^{-1}(x)\right\|_{X \times Y} \\ &\leq \left\|p^{-1}\right\| \|x\|_X, \end{aligned}$$

i.e. $\|T\| \leq \left\|p^{-1}\right\| < \infty$.

Above all, $T \in \mathcal{B}(X, Y)$. □

Remark 8.7. This can also be proved by Corollary 7.5: since $(X, \|\ \|_X)$ is a Banach space, and

$$(X, \|\ \|_X + \|\ \|_Y \circ T) \cong (G(T), \|\ \|_{X \times Y})$$

is also a Banach space, where the isometry is just $T\colon X \to G(T)$. Then Corollary 7.5 ensures Theorem 8.8.

9 Week 9

9.1 Lecture 9-1

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 $p\colon X \to \mathbb{R}$ 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 $X \in \mathrm{Ob}(\mathsf{Lin}_{\mathbb{K}})$, the dual space, denoted by $X^\sharp$, is

$$X^\sharp := \{f\colon X \to \mathbb{K} \text{ that is } \mathbb{K}\text{-linear}\}.$$

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 $p\colon X \to \mathbb{R}$ 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

$$\mathcal{F} := \bigcup_{X_0 \hookrightarrow D \hookrightarrow X} \{g \in D^\sharp\colon\ g|_{X_0} = f_0 \wedge g \leq p|_D\},$$

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

$$g_1 \le g_2 \iff \mathrm{dom}(g_1) \hookrightarrow \mathrm{dom}(g_2) \wedge g_2|_{\mathrm{dom}(g_1)} = g_1,$$

i.e. $g_1 \le g_2$ iff g_2 is an extension of g_1 in the sense above.

Exercise 9.1. Check that $(\mathcal{F}, \le)$ 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 $\bigcup_{g\in\mathcal{F}_0} \mathrm{dom}(g) = \mathrm{dom}\left(\bigcup \mathcal{F}_0\right)$. Define the linear structure as follows

$$\begin{aligned} +\colon \mathrm{dom}\left(\bigcup \mathcal{F}_0\right) \times \mathrm{dom}\left(\bigcup \mathcal{F}_0\right) &\to \mathrm{dom}\left(\bigcup \mathcal{F}_0\right) \\ (v_1, v_2) &\mapsto v_1 +_{V_1\cup V_2} v_2, \end{aligned}$$

where $V_j \in \mathrm{dom}\left(\bigcup \mathcal{F}_0\right)$ is a space containing v_j, $V_1 \cup V_2$ is a subspace of $\mathrm{dom}\left(\bigcup \mathcal{F}_0\right)$ since $\mathrm{dom}\left(\bigcup \mathcal{F}_0\right)$ 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

$$\cdot\ \colon \mathrm{dom}\left(\bigcup \mathcal{F}_0\right) \times \mathbb{R} \to \mathrm{dom}\left(\bigcup \mathcal{F}_0\right), (v, k) \mapsto k \cdot_V v,$$

where $V \in \mathcal{F}_0$ is a subspace of X containing v. We can prove that $\cdot$ is well-defined similarly. Above all, $\mathrm{dom}\left(\bigcup \mathcal{F}_0\right)$ is a vector space of X and contains X_0. Now we define a linear functional on $\mathrm{dom}\left(\bigcup \mathcal{F}_0\right)$ as follows

$$h\colon \mathrm{dom}\left(\bigcup \mathcal{F}_0\right) \to \mathbb{R}, v \mapsto g(v),$$

whenever g is an element of $\mathcal{F}_0$ such that $v \in \mathrm{dom}(g)$. 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. $\mathrm{dom}(f) = X$;

2. $f|_{X_0} = f_0$;
3. $f \le p$.

Since $f \in \mathcal{F}$, we get $f|_{X_0} = f_0$ and $f \le p|_{\mathrm{dom}(f)}$. Thus it suffices to show that $\mathrm{dom}(f) = X$. Suppose there is an element $x_0 \in X \setminus \mathrm{dom}(f)$, then

$$\mathrm{dom}(f) + \mathbb{R}x_0 = \{y + kx_0 \colon y \in \mathrm{dom}(f) \wedge k \in \mathbb{R}\}$$

is a subspace of X strictly bigger than $\mathrm{dom}(f)$. We prove that f can be extended to a linear functional $\widetilde{f}$ on $\mathrm{dom}(f)+\mathbb{R}x_0$ such that $\widetilde{f} \in \mathcal{F}$, which is a contradiction with f being a maximal element. In order to define $\widetilde{f}$, it suffices to check that $\widetilde{f}(x_0)$ can be defined, since

$$\widetilde{f}(y + kx_0) = \widetilde{f}(y) + k\widetilde{f}(x_0) = f(y) + k\widetilde{f}(x_0)$$

is determined by $\widetilde{f}(x_0)$. The only thing restricts the value of $\widetilde{f}(x_0)$ is $\widetilde{f} \le p|_{\mathrm{dom}(\widetilde{f})}$, i.e. $\forall y \in \mathrm{dom}(f), \forall k \in \mathbb{R} \setminus \{0\}:$

$$\begin{cases} \widetilde{f}(y + kx_0) \le p|_{\mathrm{dom}(\widetilde{f})}(y + kx_0), & k > 0; \\ \widetilde{f}(y + kx_0) \le p|_{\mathrm{dom}(\widetilde{f})}(y + kx_0), & k < 0; \end{cases}$$

i.e. $\forall y \in \mathrm{dom}(f), \forall k \in \mathbb{R} \setminus \{0\}:$

$$\begin{cases} \widetilde{f}(x_0) \le p(y + kx_0)/k - \widetilde{f}(y)/k, & k > 0; \\ \widetilde{f}(x_0) \ge p(y + kx_0)/k - \widetilde{f}(y)/k, & k < 0; \end{cases}$$

i.e. $\forall y, z \in \mathrm{dom}(f), \forall k > 0, k' < 0:$

$$\begin{cases} \widetilde{f}(x_0) \le p(y/k + x_0) - f(y/k); \\ \widetilde{f}(x_0) \ge -p(-z/k' - x_0) - f(z/k'). \end{cases}$$

Here I don't care the case $k = 0$ since $\widetilde{f}(y) \le p|_{\mathrm{dom}(\widetilde{f})}(y)$ can be deduced from $f \le p|_{\mathrm{dom}(f)}$. Therefore, it suffices to show that $\forall y, z \in \mathrm{dom}(f), \forall k > 0, k' < 0:$

$$-p(-z/k' - x_0) - f(z/k') \le p(y/k + x_0) - f(y/k), \tag{21}$$

And (21) holds. To see this, we can set $k = -k' = 1$ and then

$$p(z - x_0) + p(y + x_0) \ge p(y + z) \ge f(y + z) = f(y) + f(z).$$

Then (21) implies

$$\sup S^- \le \inf S^+.$$

where

$$\begin{aligned} S^- &:= \{-p(-z/k' - x_0) - f(z/k') \in \mathbb{R} \colon z \in \mathrm{dom}(f), k' < 0\}, \\ S^+ &:= \{p(y/k + x_0) - f(y/k) \in \mathbb{R} \colon y \in \mathrm{dom}(f), k > 0\}. \end{aligned}$$

Then $\widetilde{f}(x_0)$ can be taken as an arbitrary number in the interval

$$[\sup S^-, \inf S^+].$$

We're done. $\square$

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 9.1. 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

$$p \colon X \to \mathbb{R}, x \mapsto \|f_0\|_{X_0^*}\|x\|.$$

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 $\operatorname{Re} \circ h = \operatorname{Re}(h) \in \mathcal{B}(X, \mathbb{R})$. For all $x \in X$,

$$h(ix) = i\big(\operatorname{Re} h(x) + i \operatorname{Im} h(x)\big) = -\operatorname{Im} h(x) + i \operatorname{Re} h(x),$$

take real parts for both sides and get $-\operatorname{Im} h(x) = \operatorname{Re} h(ix)$. Thus

$$\forall x \in X : h(x) = \operatorname{Re} h(x) - i \operatorname{Re} h(ix).$$

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

$$p\colon X \to \mathbb{R}, x \mapsto \|g_0\|_{\mathcal{B}(X,\mathbb{R})} \|x\|.$$

Then apply the result in 9.1.1 and we get $\exists g \in \mathcal{B}(X, \mathbb{R})$ such that

$$\big(\, g|_{X_0} = g_0\big) \wedge \big(g \leq p\big) \wedge \big(\|g\|_{\mathcal{B}(X,\mathbb{R})} = \|g_0\|_{\mathcal{B}(X_0,\mathbb{R})}\big).$$

Then $f\colon X \to \mathbb{C}, x \mapsto g(x) - ig(ix)$ satisfies

$$\big(\, f|_{X_0} = f_0\big) \wedge \big(\|f\|_{X^*} = \|f_0\|_{X_0^*}\big).$$

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

$$\theta\colon X \to \mathbb{C}, x \mapsto (\operatorname{sign} \circ f)(x).$$

Notice that $|\ | \circ \theta\colon X \to \mathbb{C}$ is a constant function. Then $\forall x \in X$

$$\begin{aligned}
|f(x)| &= f(x) \cdot \theta(x) \\
&= f\big(\theta(x) \cdot x\big) \\
&= g\big(\theta(x) \cdot x\big) && \big(f(y) \in \mathbb{R} \iff f(y) = g(y)\big) \\
&\leq \|g\|_{\mathcal{B}(X,\mathbb{R})} \cdot \|\theta(x) \cdot x\|_X \\
&= \|g\|_{\mathcal{B}(X,\mathbb{R})} \cdot |\theta(x)| \cdot \|x\|_X \\
&= \|g\|_{\mathcal{B}(X,\mathbb{R})} \cdot \|x\|_X.
\end{aligned}$$

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. $\|f\|_{X^*} \leq \|f_0\|_{X_0^*}$. □

Theorem 9.2 is of great importance in the theory of "dual space of Banach space", which can be seen later.

9.2 Lecture 9-2

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 $T\colon X \to Y$. the graph of T is $G(T) := \{(x, Tx) : x \in X\} \hookrightarrow X \times Y$. 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 9.2. 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 $\sup_{\lambda\in\Lambda}\|T_\lambda y\|_Y = \infty$.

9.2.1 Review

Recall the definition of semi-norm

Definition (Semi-norm)**.** Let X be a linear normed space. A function $p\colon X \to \mathbb{R}$ 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 $f\colon X \to \mathbb{R}$ is said to be sub-additive if

$$\forall x, y \in X\colon f(x) + f(y) \leq f(x) + f(y).$$

Definition (Positive-homogeneity)**.** Let X be a linear normed space. A function $f\colon X \to \mathbb{R}$ is said to be positive-homogeneous if

$$\forall x \in X, \alpha \in [0, \infty)\colon f(\alpha x) = \alpha f(x).$$

Thus, a function $p\colon X \to \mathbb{R}$ is a sub-linear functional if and only if (p is sub additive $\wedge$ p is positive-homogeneous). Let

$$X'_+ := \{f\colon X \to \mathbb{R} \mid f \text{ is sub-additive} \wedge \text{ homogeneous}\}.$$

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

$$\begin{tikzcd} X_0 \arrow[rr, "f_0"] \arrow[dr, hook, "\iota"'] & & \mathbb{R} \\ & X \arrow[ur, "\exists f"'] & \end{tikzcd}$$

Remark 9.3. 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 $M := \operatorname{span}(\{x_0\} \cup X_0) \hookrightarrow X$. Notice that $\dim\big(M \setminus X_0\big) = 1$. Thus, bu Mathematical Induction, we can prove the case $\dim\big(X \setminus X_0\big) < \infty$. To get rid of the assumption $\dim\big(X \setminus X_0\big) < \infty$, 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 $f|_{X_0} = f$, i.e. the following diagram commutes

$$\begin{tikzcd} X_0 \arrow[rr, "f_0"] \arrow[dr, hook, "\iota"'] & & \mathbb{R} \\ & X \arrow[ur, "\exists f"'] & \end{tikzcd}$$

2. their norms coincide: $\|f\|_{X^*} = \|f_0\|_{X_0^*}$.

Remark 9.4. 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 some functional $f \in X^*$ such that $|f(x_0)| = \|x_0\|$, since we can multiply f by a constant sign$\big(f(x_0)\big)$. Let $X_0 := \operatorname{span}\{x_0\}$, a subspace of X and

$$f_0 \colon X_0 \to \mathbb{K}, k \cdot x_0 \mapsto k \cdot \|x_0\|.$$

Clearly $\forall k \in \mathbb{K}, |f_0(kx_0)| \leq |k|\|x_0\|$ and $\|kx_0\| = |k|\|x_0\|$, thus $\|f_0\| \leq 1$. And the inverse inequality holds, since

$$f\big(x_0/\|x_0\|\big) = 1, \|x_0/\|x_0\|\| = 1.$$

Then apply Theorem 9.4, there is a function $f \colon X \to \mathbb{C}$ such that $f|_{X_0} = f_0 \wedge \|f\|_{X^*} = \|f_0\|_{X_0^*}$. And

$$f(x_0) = f|_{X_0}(x_0) = f_0(x_0) = \|x_0\|. \qquad \square$$

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 $f|_{X_0} = f_0$ and $\|f\| = \|f_0\|$. In other words, the following diagram commutes.

$$\begin{tikzcd} X_0 \arrow[rr, "f_0"] \arrow[dr, hook, "\iota"'] & & \mathbb{K} \\ & X \arrow[ur, dashed, "\exists f"'] & \end{tikzcd}$$

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. □

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. □

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 (set $x_1 = x_0$, $x_2 = 0$) and get a contradiction. □

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 $\mathrm{span}(\{x_0\} \cup X_0)$).

10 Week 10

10.1 Lecture 10-1

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 $f \in X^*$ such that $\|f\|_{X^*} \leq 1$.

Corollary 10.2. There is another representation of $\| \ \|_X$, 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 10.1. 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 functional 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 $(\mathbb{R}^2, \| \|_1)$, where

$$\| \|_1 \colon (x, y) \mapsto |x| + |y|$$

is equivalent to $\| \|_2$ since $\dim(\mathbb{R}^2) = 2 < \infty$. In fact, $(\mathbb{R}^2, \| \|_1)$ is just $(\mathbb{R}, \| \|_1) \bigoplus (\mathbb{R}, \| \|_1)$. Let $p_2 \colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto y$ 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:

$$f \colon G \to \mathbb{R}, (x, 0) \mapsto x.$$

It's easy to see that $\|f\|_{G^*} = 1$. We can directly construct lots of extensions of f. To see this, define

$$f_\beta \colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x + \beta y$$

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 $|x + \beta y| \le |x| + |\beta||y|$. 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

$$\operatorname{card}[-1, 1] = \operatorname{card} \mathbb{R} = \operatorname{card}({}^{\mathbb{N}}2) > \operatorname{card} \mathbb{N}. \tag{22}$$

The inequality $\operatorname{card}({}^{\mathbb{N}}2) > \operatorname{card} \mathbb{N}$ is known as a special case of Cantor's Theorem.

Here is a proof of "$\mathbb{R}$ is uncountable" using (22).

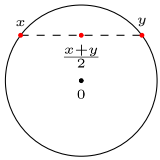
Proof. It suffices to show that ${}^{\mathbb{N}}2$ is uncountable, where 2 means a set of exactly 2 elements, $\{\varnothing, \{\varnothing\}\}$ or $\{-1, 1\}$ for example. WLOG, let $2 = \{-1, 1\}$ then an element of ${}^{\mathbb{N}}2$ is just a sequence with values being 1 or -1. Suppose ${}^{\mathbb{N}}2$ is countable, i.e. ${}^{\mathbb{N}}2 = \{\alpha(k) \colon k \in \mathbb{N}\}$ where α_k is a sequence whose values lie in $\{-1, 1\}$. Then consider the sequence

$$\beta \colon \mathbb{N} \to \{-1, 1\}, n \mapsto -\alpha(n)_n.$$

In other words, β is a sequence whose values lie in $\{-1, 1\}$ and satisfies $\forall n \in \mathbb{N}\colon \beta_n \neq \alpha(n)_n$. Hence $\forall k \in \mathbb{N}\colon \beta \neq \alpha(k)$. While $\beta \in {}^{\mathbb{N}}2$, which contradicts with ${}^{\mathbb{N}}2 = \{\alpha(k)\colon k \in \mathbb{N}\}$. $\square$

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 ℓ_p space and hence it's natural to ask:

Question 10.1. What's ℓ_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 $\varphi\colon \ell_\infty \to \ell_1^*$ and its inverse that is also a contraction. Then $\ell_\infty \cong \ell_1^*$.

Let $\varphi\colon \ell_\infty \to \ell_1^*, \alpha \mapsto \varphi_\alpha$, where

$$\varphi_\alpha\colon \ell_1 \to \mathbb{K}, x \mapsto \sum_{n \geq 1} x_n \overline{\alpha_n}.$$

Hölder's inequality implies that φ_α is well-defined and bounded, with $\|\varphi_\alpha\|_{\ell_1^*} \leq \|\alpha\|_\infty$. In other words, φ_α is a bounded linear functional on ℓ_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 φ. Let

$$e_n := (\underbrace{0, \ldots, 0}_{n-1 \text{ terms}}, 1, 0, \ldots) \text{ for all } n \in \mathbb{N}$$

and $\psi\colon \ell_1^* \to \ell_\infty, f \mapsto \psi^f$, where $\psi_n^f = f(e_n), \forall n \in \mathbb{N}$. Then ψ is well-defined. What should be checked is just: $\forall f \in \ell_1^*, \left\|\psi^f\right\|_\infty < \infty$, which is right since

$$\forall n \in \mathbb{N}\colon \left|\psi_n^f\right| = |f(e_n)| \leq \|f\|_{\ell_1^*} \|e_n\|_{\ell_1} = \|f\|_{\ell_1^*}$$

implies that $\left\|\psi^f\right\|_\infty \leq \|f\|_{\ell_1^*}$ and hence $\|\psi\|_{\mathcal{B}(\ell_1^*, \ell_\infty)} \leq 1$. Now we prove that $\varphi\colon \ell_\infty \to \ell_1^*$ is an isometry, i.e. $(\varphi \circ \psi = \mathrm{id}_{\ell_1^*}) \wedge (\psi \circ \varphi = \mathrm{id}_{\ell_\infty})$.

- For $\varphi \circ \psi = \mathrm{id}_{\ell_1^*}$: let $f \in \ell_1^*$ be an arbitrary element. Then ψ^f is a sequence in ℓ_∞, such that $\forall n \in \mathbb{N}, \psi_n^f = f(e_n)$,

  $$\varphi \circ \psi(f) = \varphi(\psi^f)$$

  and hence for all $x \in \ell_1$:

  $$\begin{aligned}
  [\varphi \circ \psi(f)](x) &= [\varphi(\psi^f)](x) \\
  &= \sum_{n \geq 1} x_n \overline{\psi_n^f} \\
  &= \sum_{n \geq 1} x_n \overline{f(e_n)} \\
  &= \sum_{n \geq 1} f(x_n e_n) \\
  &= \lim_{N \to \infty} \sum_{n=1}^{N} f(x_n e_n) \\
  &= \lim_{N \to \infty} f\Big(\sum_{n=1}^{N} x_n e_n\Big)
  \end{aligned}$$

  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 ℓ_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 ℓ_1. It's natural to guess the answer:

just x. Since

$$\begin{aligned}\left\|\sum_{n=1}^{N} x_n e_n - x\right\|_1 &= \|(0,\ldots,0,x_{N+1},x_{N+2},\ldots)\|_1\\ &= \sum_{n\geq N+1} |x_n|\\ &= \|x\|_1 - \sum_{n=1}^{N}|x_n|\end{aligned}$$

converges to 0 as $N \to \infty$, and by continuity of f, we get

$$\lim_{N\to\infty} f\Big(\sum_{n=1}^{N} x_n e_n\Big) = f\Big(\lim_{N\to\infty}\sum_{n=1}^{N} x_n e_n\Big) = f(x).$$

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 $\varphi \circ \psi = \mathrm{id}_{\ell_1^*}$.

- For $\psi \circ \varphi = \mathrm{id}_{\ell_\infty}$: 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 φ_α and e_n. Therefore, $[\psi \circ \varphi](\alpha) = \alpha$ holds for all $\alpha \in \ell_\infty$, i.e. $\psi \circ \varphi = \mathrm{id}_{\ell_\infty}$.

Above all, φ is an isometry from ℓ_∞ to ℓ_1^* and $\ell_\infty \cong \ell_1^*$. □

Remark 10.2. Verification of "φ 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 ℓ_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 $\|\ \|_2$.

Proof. Consider the orthonormal basis of $\mathbb{K}^n$, i.e.

$$e_k := (\underbrace{0,\ldots,0}_{k-1 \text{ terms}}, 1, 0, \ldots), \forall k \in [n].$$

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

$$\forall j, k \in [n] \colon \varphi_j(e_k) = \delta_k^j.$$

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 $(\mathbb{K}^n)^*$

$$\| \ \| \colon (\mathbb{K}^n)^* \to \mathbb{R}, f \mapsto \Big(\sum_{j=1}^{n} |f(e_j)|^2\Big)^{1/2}.$$

It's easy to verify that $\| \ \|$ is a norm. Let

$$F \colon (\mathbb{K}^n)^* \to \mathbb{K}^n, f \mapsto \big(f(e_j)\big)_{j\in[n]} = \big(f(e_1), \dots, f(e_n)\big),$$

which will be proved to be an isometry. Clearly F is surjective and $\ker F = 0(\in (K^n)^*)$, i.e. F is injective. What is left to prove is just

$$\forall f \in (\mathbb{K}^n)^* \colon \|F(f)\|_2 = \|f\|.$$

Let $f \in (\mathbb{K}^n)^*$ be given. Then

$$\begin{aligned}
\|F(f)\|_2 &= \Big\| \big(f(e_j)\big)_{j\in[n]} \Big\|_2 \\
&= \Big(\sum_{j=1}^{n} |f(e_j)|^2\Big) \\
&= \|f\|,
\end{aligned}$$

as we wanted. $\square$

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 10.3. We have proved that $\ell_p^* \cong \ell_q$ for all $p \in [1, \infty)$. However, it's not very easy to study ℓ_∞^*.

10.2 Lecture 10-2

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 $\varphi\colon \ell_1 \to c_0^*$ and $\psi\colon c_0^* \to \ell_1$ such that $(\varphi \circ \psi = \mathrm{id}_{c_0^*}) \wedge (\psi \circ \varphi = \mathrm{id}_{\ell_1})$.

- Construction of φ: consider

  $$\varphi\colon \ell_1 \to c_0^*, \alpha \mapsto \varphi_\alpha,$$

  where $(c_0^* \ni)\varphi_\alpha\colon c_0 \to \mathbb{K}, x \mapsto \sum_{n=1}^{\infty} x_n\alpha_n$. For all $\alpha \in \ell_1$, the functional φ_α 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 φ is trivial. Thus $\|\varphi\|_{\mathcal{B}(\ell_1, c_0^*)} \leq 1$. Above all, φ is a contraction.

- Construction of ψ: consider

  $$\psi\colon c_0^* \to \ell_1, f \mapsto \psi^f,$$

  where ψ^f is the sequence $\big(f(e_n)\big)_{n\in\mathbb{N}}$. Here $\{e_n \in c_0 : n \in \mathbb{N}\}$ is still the sequence

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

  Now we check that $\psi^f \in \ell_1$ and $\left\|\psi^f\right\|_1 \leq \|f\|_{c_0^*}$. For all $n \in \mathbb{N}$, let

  $$A^{(n)} = \sum_{j=1}^{n} \operatorname{sign}\big(f(e_n)\big)e_n.$$

  Then $A^{(n)} \in c_0$ satisfies $\left\|A^{(n)}\right\|_\infty = 1$ for all $n \in \mathbb{N}$. And

  $$\forall n \in \mathbb{N}\colon f\big(A^n\big) = \sum_{j=1}^{n} \operatorname{sign}\big(f(e_n)\big)f(e_n) = \sum_{j=1}^{n}|f(e_n)|. \tag{23}$$

Then (23) implies $\|f\|_{c_0^*} \geq \sum_{j=1}^{n} |f((e_n)|$ holds for all $n \in \mathbb{N}$. Let $n \to \infty$ 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. ψ is a contraction.

Now we check that both φ, ψ 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 $\forall n \in \mathbb{N}$,

$$\big[(\psi \circ \varphi)(\alpha)\big]_n = \psi_n^{\varphi_\alpha} = \varphi_\alpha(e_n) = \alpha_n,$$

which is easy to check. And hence $(\psi \circ \varphi)(\alpha) = \alpha$. Since $\alpha \in \ell_1$ is arbitrary, we get $\psi \circ \varphi = \mathrm{id}_{\ell_1}$ 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

$$\big[(\varphi \circ \psi)(f)\big](x) = \varphi_{\psi^f}(x) = \sum_{n=1}^{\infty} x_n \psi_n^f.$$

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

$$\begin{aligned}
\sum_{n=1}^{\infty} x_n \psi_n^f &= \sum_{n=1}^{\infty} x_n f(e_n) \\
&= \lim_N \sum_{n=1}^{N} x_n f(e_n) \\
&= \lim_N f\Big(\sum_{n=1}^{N} x_n e_n\Big) \\
&= f\Big(\lim_N \sum_{n=1}^{N} x_n e_n\Big).
\end{aligned}$$

And

$$\left\|\sum_{n=1}^{N} x_n e_n - x\right\|_\infty = \|(0,0,\ldots,x_{N+1},x_{N+2},\ldots)\|_\infty \to \limsup_n |x_n| = 0$$

as $N \to \infty$, since $x \in c_0 \implies \limsup_n |x_n| = \lim_n |x_n| = 0$. Thus

$$f\Big(\lim_N \sum_{n=1}^{N} x_n e_n\Big) = f(x),$$

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 $\varphi \circ \psi = \mathrm{id}_{c_0^*}$.

Above all, both φ, ψ are isometries and $c_0^* \cong \ell_1$. $\square$

Example 27. Set $\mathbb{K} = \mathbb{R}$ in this example. Let

$$\begin{aligned} BV[a,b] &:= \{f \in {}^{\mathbb{R}}[a,b] \colon f \text{ is of bounded variance}\}, \\ BV_0[a,b] &:= \{f \in BV[a,b] \colon f(a) = 0\}, \\ BV_0^+[a,b] &= \{f \in BV_0 \colon f \text{ is right continuous}\}. \end{aligned}$$

Then the dual space of $C[a,b]$ is isometric to $BV_0^+[a,b]$:

$$\big(C[a,b]\big)^* \cong BV_0^+[a,b],$$

where the functional induced by $\rho \in C[a,b]$ is

$$\varphi_\rho \colon C[a,b] \to \mathbb{C}, f \mapsto \int_a^b f(x)\,\mathrm{d}\rho(x),$$

the Riemann-Lebesgue integral of f with respect to ρ.

Example 28. For all σ-finite measure space $(\Omega, \mathcal{F}, \mu)$ we have

$$\forall p \in [1,\infty) \colon \big(L_p(\Omega, \mathcal{F}, \mu)\big)^* \cong L_q(\Omega, \mathcal{F}, \mu),$$

where $q = p'$.

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 $(x_n)_{n\in\mathbb{N}} \xrightarrow{\|\ \|} x$. This is the classical convergence mode (hence we usually omit "in norm" and "$\|\ \|$", writing "$(x_n)_{n\in\mathbb{N}}$ converges to x" and "$(x_n)_{n\in\mathbb{N}} \to x$") 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 $(f_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} f$ (read as "weak star convergence").

Remark 10.4. Weak convergence is a property of sequence in X while weak* convergence is a property of sequence in X^*.

Here are some examples.

Example 29. Consider the linear normed space ℓ_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

$$\ell_q \ni \psi^f = \big(f(e_n)\big)_{n\in\mathbb{N}}$$

  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

$$(x_n)_{n\in\mathbb{N}} \xrightarrow{\|\ \|} x \implies (x_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} x.$$

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

$$\hat{x} \colon X^* \to \mathbb{K}, f \mapsto f(x).$$

From this definition, we have

Claim. For all $x \in X$, the functional $\hat{x}$ is bounded, i.e. $\hat{x} \in \big(X^*\big)^*$.

Proof. For all $f \in X^*$, we have

$$|\hat{x}(f)| = |f(x)| \leq \|f\|_{X^*}\|x\|,$$

which implies $\|\hat{x}\|_{X^{**}} \leq \|x\|$. $\square$

For convenience, we define

Definition (Bidual Space)**.** For a linear normed space, the space $\left(X^*\right)^*$ 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 $\widehat{X} := \{\hat{x} \colon x \in X\}$. The mapping ι defined as follows is called the natural embedding of X

$$\iota \colon X \to \widehat{X}, x \mapsto \hat{x}.$$

For $x \in X$, $\hat{x}$ is also denoted by ι_x.

From definition, we see that ι is surjective. Moreover, we have proved that ι is contractive. In fact, ι 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 ι.

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 $(f_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} f$.

Proof. Suppose $(f_n)_{n\in\mathbb{N}} \rightharpoonup f$, i.e. $\forall F \in X^{**} \colon \lim_n F(f_n) = F(f)$. 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 $(f_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} f$. $\square$

Remark 10.5. 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

$$\left((f_n)_{n\in\mathbb{N}} \xrightharpoonup{*} f\right) \wedge \left((f_n)_{n\in\mathbb{N}} \not\rightharpoonup f\right).$$

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

$$\begin{tikzcd}
0 & & X \arrow[r, equal] & c_0 \\
f_n \arrow[u, harpoon, "*"] \arrow[d, harpoon, "/" marking] & \in & X^* \arrow[r, equal] & \ell_1 \\
0 & & X^{**} \arrow[r, equal] & \ell_\infty
\end{tikzcd}$$

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

$$(e_n)_{n\in\mathbb{N}} \subseteq X^* = \ell_1,$$

where the m-th term of e_n is δ_n^m for all $m, n \in \mathbb{N}$.

Verification. For convenience, we denote φ_a by a, i.e. $a(x) = \varphi_a(x) = \sum_{n\geq 1} a_n x_n$ for all x. Then

- $(e_n)_{n\in\mathbb{N}} \xrightharpoonup{*} 0$: for all $a \in X = c_0$, 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 $(e_n)_{n\in\mathbb{N}} \xrightharpoonup{*} 0$.

- $(e_n)_{n\in\mathbb{N}} \not\rightharpoonup 0$: consider the sequence $\mathbf{1} = (1)_{n\in\mathbb{N}} \in X^{**} = \ell_\infty$, 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 $(e_n)_{n\in\mathbb{N}} \not\rightharpoonup 0$. □

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 $\big(f(x_n)\big)_{n\in\mathbb{N}}$ 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 $\overline{\operatorname{span} G} = X^*$ such that

$$\forall g \in G\colon \big(g(x_n)\big)_{n\in\mathbb{N}} \to g(x).$$

Proof. Sufficiency: for all $f \in \operatorname{span} G$, 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 $\big(g(x_n)\big)_{n\in\mathbb{N}} \to g(x)$, there is $N \in \mathbb{N}$ such that

$$\forall n > N\colon |g(x_n) - g(x)| < \varepsilon.$$

Then $\forall n > N$

$$|f(x_n) - f(x)| \leq \big(M + \|x\|_X + 1\big)\varepsilon,$$

which implies $\lim_n f(x_n) = f(x)$ since $\varepsilon > 0$ is arbitrary. Therefore, $(x_n)_{n\in\mathbb{N}} \rightharpoonup x$ since $f \in X^*$ is arbitrary.

Necessity: For the first condition, consider the natural embedding $\iota \colon : X \to X^{**}$. Then for all $f \in X^*$, the sequence $\big(f(x_n)\big)_{n\in\mathbb{N}}$ in $\mathbb{K}$ is convergent and hence bounded. Thus

$$\forall f \in X^* \colon \sup_{n\in\mathbb{N}} |f(x_n)| = \sup_{n\in\mathbb{N}} |\widehat{x_n}(f)| < \infty.$$

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 $G = X$. ☐

Here are some examples of application of Theorem 10.7.

Example 31. For $p \in (1,\infty)$, consider the space ℓ_p. Then Theorem 10.7 means that: suppose $\big(X^{(n)}\big)_{n\in\mathbb{N}} \subseteq \ell_p$ and $X \in \ell_p$, then $\big(X^{(n)}\big)_{n\in\mathbb{N}} \rightharpoonup X$ if and only if the following 2 conditions hold.

1. $\sup_n \big\|X^{(n)}\big\|_p < \infty$;
2. $\forall i \in \mathbb{N} \colon \lim_n X_i^{(n)} = X_i$.

Proof. Sufficiency: take $G = \{e_n \colon n \in \mathbb{N}\}$ and the second point is done, if $\overline{\operatorname{span} G} = \ell_q$ 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.

$$\left\| x - \sum_{j=1}^{N} x_j e_j \right\|_q = \Big(\sum_{j>N} |x_j|^q \Big)^{1/q} < \varepsilon.$$

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

$$X_i = e_i(X) = \lim_n e_i\big(X^{(n)}\big) = \lim_n \sum_{j=1}^{\infty} \delta_j^i X_j^{(n)} = \lim_n X_i^{(n)}. \qquad ☐$$

Remark 10.6. Here we restrict $p \notin \{1,\infty\}$. Since

- For $p = 1$, then $q = \infty$ and $\operatorname{span} G$ is not dense in ℓ_q;
- For $p = \infty$, the dual space ℓ_p^* is almost unknown.

Example 32. Let a σ-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

$$\lim_n \int_\Omega f_n \chi_E \,\mathrm{d}\mu = \int_\Omega f \chi_E \,\mathrm{d}\mu.$$

Proof. Sufficiency: take $G = \{\chi_E \colon (E \in \mathcal{F}) \wedge (\mu(E) < \infty)\}$, then $\operatorname{span} G$ is the set of all integrable simple functions on Ω, 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

$$\lim_n \int_\Omega f \chi_E \,\mathrm{d}\mu = \lim_n \chi_E(f_n) = \chi_E(f) = \int_\Omega f \chi_E \,\mathrm{d}\mu,$$

where χ_E means the functional

$$\varphi_{\chi_E} : L_p \to \mathbb{K}, f \mapsto \int_\Omega f \chi_E \,\mathrm{d}\mu. \qquad \square$$

Remark 10.7. 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 σ-finite.

11 Week 11

11.1 Lecture 11-1

Something about biduality: let X be a linear normed space, then the space X^{**} is complete and the natural embedding $\iota\colon$ is a injection keeping norms. Thus, the pair (X^{**}, ι) 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:

$$\begin{cases} (x_n)_{n\in\mathbb{N}} \to x \iff \lim_n \|x_n - x\| = 0, \\ (x_n)_{n\in\mathbb{N}} \rightharpoonup x \iff \forall f \in X^*\colon \lim_n f(x_n) = f(x), \\ (x_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} x \iff \forall y \in X_*\colon \lim_n x_n(y) = \lim_n x(y), \end{cases}$$

where X_* is a linear normed space such that $(X_*)^* = X$, called the pre-dual space of X,if exists (so we don't talk the weak-$*$ convergence on X is not the dual space of any linear normed space).The definition of $(x_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} x$ 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 ℓ_1 and we can discuss the 3 modes of convergence on ℓ_1: $\to, \rightharpoonup, \overset{*}{\rightharpoonup}$.

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

$$\exists M > 0\colon \forall x \in A, \|x\| < M.$$

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, Chen 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 for all $f \in X^*, (f_n)_{n\in\mathbb{N}} \subseteq X^*$ 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, denoted by $(x_n)_{n\in\mathbb{N}} \xrightarrow{l} x$.

Claim. This convergence mode is not 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 $(x_n)_{n\in\mathbb{N}} \rightharpoonup x$.

Let $X = H$ be a Hilbert space with a countable orthonormal basis $(x_n)_{n\in\mathbb{N}}$. Corollary 12.9 implies that $(x_n)_{n\in\mathbb{N}} \rightharpoonup 0$. Consider the Riesz map φ^H, then $(\varphi^H_{x_n})_{n\in\mathbb{N}} \xrightarrow{*} 0$ as you should verify. But $\lim_n \varphi^H_{x_n}(x_n) = 1 \neq 0$. □

Weak topology is a topological notion:

Definition (Weak Topology)**.** Let X be a topological space and $\mathcal{F}$ be a family of functions $f\colon X \to \operatorname{cod}(f)$, where $\operatorname{cod}(f)$ 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 $\{\pi_i : X \to X_i \mid i \in I\}$.

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 $\overline{\cdot}^{\|\ \|}, \overline{\cdot}^{\omega}, \overline{\cdot}^{\omega*}$ 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: $\overline{X_0}^{\|\ \|}, \overline{X_0}^{\omega}, \overline{X_0}^{\omega *}$?

Answer and Proof. When $X = (X_*)^*$ exists, the relation is

$$\overline{X_0}^{\|\ \|} \subseteq \overline{X_0}^{\omega} \subseteq \overline{X_0}^{\omega *}.$$

It suffices to see that $\mathcal{T}_{\|\ \|} \subseteq \mathcal{T}_{\omega} \subseteq \mathcal{T}_{\omega *}$, where the $\mathcal{T}$ means the topology induced by its subscript. We know $\mathcal{T}_{\|\ \|} \supseteq \mathcal{T}_{\omega}$ since $\forall f \in X^*$, we know f is continuous when X is equipped with $\mathcal{T}_{\|\ \|}$. 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 $\overline{X_0}^{\|\ \|} \subseteq \overline{X_0}^{\omega}$. □

11.1.2 Conjugate Operators

For 2 linear normed spaces X, Y and a map $T\colon \mathcal{L}(X, Y)$, 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 $(T^* f)(x) = f\big(T(x)\big)$. 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

$$\begin{array}{ccc} x \in X & \xrightarrow{T} & Y \ni Tx \\ T^* f \in X^* & \xleftarrow[T^*]{} & Y^* \ni f \end{array} \qquad \langle f, Tx\rangle = \langle T^* f, x\rangle .$$

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

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[dr, "T^* f"'] & Y \arrow[d, "f"] \\ & \mathbb{K} \end{tikzcd}$$

i.e. $T^*\colon Y^* \to X^*, f \mapsto f \circ T$. For all $f \in Y^*$ we have $f \circ T \in X^*$ since $\|f \circ T\| \leq \|f\| \cdot \|T\|$. This definition satisfies:

$$\forall x \in X, \forall f \in Y^*\colon \langle f, Tx\rangle = \langle T^* f, x\rangle ,$$

which can be easily verified. Existence is ensured. Furthermore, if another map $F\colon Y^* \to X^*$ satisfies:

$$\forall x \in X, \forall f \in Y^*\colon \langle f, Tx \rangle = \langle Ff, x \rangle\,,$$

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 11.1. 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 $T\colon \mathbb{K}^n \to \mathbb{K}^m$ be given, then it must be bounded. Thus there is $T^*\colon (\mathbb{K}^m)^* \to (\mathbb{K}^n)^*$. 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

$$\forall j \in [n]\colon e_j = (\underbrace{0, \ldots, 0}_{j-1 \text{ terms}}, 1, \ldots).$$

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

$$\forall j \in [n]\colon Te_j = \sum_{k=1}^{m} a_{j,k} \mu_k,$$

whose corresponding matrix is $M_T = (a_{j,k})_{k\in[m]}^{j\in[n]}$. 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

$$Y_k := \operatorname{span}\{e_j : j \in [n] \setminus \{k\}\}, \forall k \in [n]$$

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 $M_{T^*} = (b_{j,k})_{k\in[n]}^{j\in[m]}$. From

$$\begin{array}{cccc} T^* & : & (\mathbb{K}^m)^* & \longrightarrow & (\mathbb{K}^n)^* \\ & & \cup & & \cup \\ & & \{\mu_j^* : j \in [m]\} & & \{e_k^* : k \in [n]\} \end{array}$$

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, $M_{T^*} = {}^tM_T$.

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 ι_X, ι_Y is the natural embedding of X, Y respectively;
2. $(\mathrm{id}_X)^* = \mathrm{id}_{X^*}$;
3. $(T \circ S)^* = S^* \circ T^*$.

Proof. What we want is: the following diagram commutes

$$\begin{tikzcd} X \arrow[r, "\iota_X"] \arrow[d, "T"'] & X^{**} \arrow[d, "T^{**}"] \\ Y \arrow[r, "\iota_Y"'] & Y^{**} \end{tikzcd}$$

Remember that $Y^{**} = \left(Y^*\right)^*$. For all $x \in X, f \in Y^*$, we have

$$\begin{aligned} \left[(\iota_Y \circ T)(x)\right](f) &= \left[\iota_Y(Tx)\right](f) \\ &= \left[\widehat{Tx}\right](f) \\ &= f(Tx) \\ &= (f \circ T)(x), \end{aligned}$$

and

$$\begin{aligned} \Big[\big(T^{**} \circ \iota_X\big)(x)\Big](f) &= \left[T^{**}(\hat{x})\right](f) \\ &= \langle T^{**}(\hat{x}), f\rangle \\ &= \langle \hat{x}, T^* f\rangle \\ &= \langle T^* f, x\rangle \\ &= \langle f \circ T. x\rangle \\ &= (f \circ T)(x). \end{aligned}$$

Therefore,

$$\left[(\iota_Y \circ T)(x)\right](f) = \Big[\big(T^{**} \circ \iota_X\big)(x)\Big](f)$$

holds for all $f \in Y^*, x \in X$, i.e.

$$\left[(\iota_Y \circ T)(x)\right] = \big(T^{**} \circ \iota_X\big)(x)$$

holds for all $x \in X$. Hence

$$\iota_Y \circ T = T^{**} \circ \iota_X.$$

The rest is easy. □

11.2 Lecture 11-2

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 11.2. 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 11.3. 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 $\dim(\operatorname{Im} T) < \infty$, 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 11.4. 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

$$\mathcal{C}(X, Y) \subseteq B(X, Y).$$

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 $\big((T+S)x_{n_k}\big)_{k\in\mathbb{N}}$ 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 $\big(\lambda T(x_{n_k})\big)_{k\in\mathbb{N}}$ also converges. Therefore, $\lambda T \in \mathcal{C}(X, Y)$. □

Remark 11.5. 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 $\left(x_{n_k}\right)_{k\in\mathbb{N}} \subseteq (x_n)_{n\in\mathbb{N}}$ be a convergent subsequence, then $\left(y_{n_k}\right)_{k\in\mathbb{N}}$ is a bounded sequence in $\mathbb{R}$. Take a convergent subsequence of $\left(y_{n_k}\right)_{k\in\mathbb{N}}$ 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 11.6. 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 4ε-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 4ε-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 $\dim(\operatorname{Im} T) < \infty$, we know that $\operatorname{Im} T$ is complete. Then for all bounded set $A \subseteq X$, the set $T(A) \subseteq \operatorname{Im} T$ is bounded and hence $\overline{T(A)}$ is bounded. While $\dim(\operatorname{Im} T) < \infty$, 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 cast 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 11.7. The space X is also needed to be finite-dimensional. To see this, set $Y = \mathbb{K}$ and find some space X that $X^\sharp \neq X^*$.

Proof. It suffices to prove that $\forall T \in \mathcal{L}(X,Y)$, we have $T \in \mathcal{F}_r(X,Y)$. Since $\dim(\operatorname{Im} T) \leq \dim(Y) < \infty$, 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

$$\begin{tikzcd}
X \arrow[r, "T"] & Y \\
\mathbb{K}^m \arrow[u, "\varphi"] \arrow[r, dashed, "\exists! L_T"'] & \mathbb{K}^n \arrow[u, "\psi"']
\end{tikzcd}$$

where φ, ψ 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 $T\colon X \to Y, S\colon Y \to Z$ 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 S is compact, we know that $S\big(T(B_X)\big)$ 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

$$\big(S(Tx_{n_k})\big)_{k\in\mathbb{N}} = \big((S \circ T)x_{n_k}\big)_{k\in\mathbb{N}}$$

converges in Z for some subsequence $(x_{n_k})_{k\in\mathbb{N}}$. Thus $S \circ T$ is compact. □

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

$$T\colon \ell_2 \to \ell_2, x \mapsto Tx := \Big(\sum_{j=1}^{\infty} a_{n,j}x_j\Big)_{n\in\mathbb{N}}$$

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

$$\forall n \in \mathbb{N}\colon |Tx_n| = \left|\sum_{j=1}^{\infty} a_{n,j}x_j\right| \leq \|x\|_2 \Big(\sum_{j=1}^{\infty} |a_{n,j}|^2\Big)^{1/2},$$

and hence

$$\|Tx\|_2^2 \leq \|x\|_2^2 \Big(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} |a_{i,j}|^2\Big).$$

Above all

$$\|T\| \leq \Big(\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}|a_{i,j}|^2\Big)^{1/2}. \tag{24}$$

Secondly, let $(T_n)_{n\in\mathbb{N}}$ be defined as follows

$$\forall n\in\mathbb{N}, T_n\colon \ell_2\to\ell_2, x\mapsto \Big(\sum_{j=1}^{n}a_{1,j}x_j,\dots,\sum_{j=1}^{n}a_{n,j}x_j,0,\dots\Big).$$

Then $\dim(\operatorname{Im}T_n)\leq n$ 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

$$\begin{aligned}\|T-T_n\| &\leq \Big(\sum_{i=n+1}^{\infty}\sum_{j=n+1}^{\infty}|a_{i,j}|^2\Big)^{1/2}\\ &= \Big(\sum_{i,j\geq 1}|a_{i,j}|^2-\sum_{1\leq i,j\leq n}|a_{i,j}|^2\Big)^{1/2}\\ &\to 0 (n\to\infty).\end{aligned}$$

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 $(Tx_n)_{n\in\mathbb{N}}\to Tx$.

Remark 11.8. In general: for $T\in\mathcal{B}(X,Y)$, we have $(Tx_n)_{n\in\mathbb{N}}\rightharpoonup Tx$.

Proof of Remark 11.8. 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 $(Tx_n)_{n\in\mathbb{N}}\rightharpoonup Tx$. □

Examples 37 shows that there is an operator T and a sequence $(x_n)_{n\in\mathbb{N}}$ such that $(x_n)_{n\in\mathbb{N}}\rightharpoonup x$ and $(Tx_n)_{n\in\mathbb{N}}\not\to Tx$, where T is a bounded operator. Thus, compact operators are special.

Example 37. 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 expect for the n-th element being 1. Then consider $\mathrm{id}_{c_0} \colon c_0 \to c_0$, we have $(e_n)_{n\in\mathbb{N}} \rightharpoonup 0$ but $(e_n)_{n\in\mathbb{N}} \not\to 0$. See Example 29.

Proof of Proposition 11.11. Suppose $(Tx_n)_{n\in\mathbb{N}} \not\to Tx$, then there is some $\varepsilon_0 > 0$ and a subsequence $(y_k)_{k\in\mathbb{N}} = (x_{n_k})_{k\in\mathbb{N}}$ such that $\|Ty_k - Tx\| > \varepsilon_0$ for all $k \in \mathbb{N}$. Weak convergence ensures that $(x_n)_{n\in\mathbb{N}}$ is weakly bounded and hence bounded, see Exercise 11.1. Therefore, $(y_k)_{k\in\mathbb{N}}$ is bounded and has a subsequence $(z_m)_{m\in\mathbb{N}}$ such that $(Tz_m)_{m\in\mathbb{N}}$ converges to $y \in Y$.

Claim. We have $y = Tx$.

The claim follows from Corollary 9.6 and Remark 11.8. Therefore, there is some N (depending on ε_0) such that for all $m > N$, we have $\|Tz_m - y\| < \varepsilon_0$, contradiction with $\|Ty_k - Tx\| > \varepsilon_0$ for all $k \in \mathbb{N}$. □

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 $T^* \in \mathcal{C}(Y^*, X^*)$.

Remark 11.9. Furthermore, if both of X, Y are complete, then $T^* \in \mathcal{C}(Y^*, X^*)$ implies $T \in \mathcal{C}(X, Y)$ (whose proof can be found in `https://arxiv.org/pdf/1010.1298v4.pdf`).

Proof. Let X, Y be 2 given linear normed spaces and $T \in \mathcal{C}(X, Y)$. Now we prove $T^* \in \mathcal{C}(Y^*, X^*)$. We want to prove that $T^*(B_{Y^*}) \subseteq X^*$ is relatively compact, i.e. it is totally bounded (since X^* is complete). Since T is compact, we have $T(B_X)$ is relatively compact and hence totally bounded, see Corollary 6.4. Given arbitrary $\varepsilon > 0$, there is some $n \in \mathbb{N}$ and $\{x_j \in X : j \in [n]\}$ such that

$$T(B_X) = \bigcup_{j=1}^{n} B_Y(y_j, \varepsilon),$$

where $y_j := Tx_j$ for all $j \in [n]$. Define

$$\begin{aligned} S \colon Y^* &\longrightarrow \mathbb{K}^n, \\ f &\longmapsto \big(f(y_1), f(y_2), \ldots, f(y_n)\big). \end{aligned}$$

Now S is bounded (suppose $\mathbb{K}^n$ is equipped with 1-norm) and compact (since it is of finite-rank), then $S(B_{Y^*})$ is relatively compact, i.e. totally bounded. Let $N := \{f_k : k \in [m]\} \subseteq B_{Y^*}$ be a finite set such that $S(N)$ is an ε-net. For all $f \in B_{Y^*}$:

$$\exists k \in [m] \colon \|S(f) - S(f_k)\| < \varepsilon. \tag{25}$$

And for all $x \in B_X$:

$$\exists j \in [n] \colon \|Tx - y_j\| < \varepsilon. \tag{26}$$

Now for all $x \in B_X$ and $f \in B_{Y^*}$, from (25) and(26), we have

$$\begin{aligned}
&|(T^* f - T^* f_k)(x)| \\
&= |f(Tx) - f_k(Tx)| \\
&\leq |f(Tx) - f(Tx_j)| + |f(Tx_j) - f_k(Tx_j)| + |f_k(Tx_j) - f_k(Tx)| \\
&\leq \|f\| \|Tx - Tx_j\| + \|S(f) - S(f_k)\| + \|f_k\| \|Tx - Tx_j\| \\
&\leq 1 \cdot \varepsilon + \varepsilon + 1 \cdot \varepsilon \\
&= 3\varepsilon,
\end{aligned}$$

where the red part is just (25) and(26). Since $x \in B_X$ is arbitrary, we have

$$\|T^* f - T^* f_k\| = \sup_{\|x\|=1} |(T^* f - T^* f_k)(x)| \leq 3\varepsilon < 4\varepsilon.$$

Therefore, the set $T^*(B_{Y^*})$ has a finite 4ε-net for all $\varepsilon > 0$, as we wanted. □

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 $T \colon X \to Y, S \colon Y \to Z$ be given. Then $S \circ T$ is finite-rank whenever one of S, T is of finite rank.

Proof. It's easy. □

Corollary 11.14. $\mathcal{F}_r(X)$ is a two-sided ideal of the ring $\mathcal{B}(X)$.

12 Week 12

12.1 Lecture 12-1

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 $u\colon X \times X \to \mathbb{K}$ 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: $u(x, y) = u(\bar{y}, x)$.

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 12.1. 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 $u\colon X \times X \to \mathbb{K}$, then $u(x, y)$ is also denoted by $\langle x, y\rangle$, (x, y), $(x \mid y)$ or $\langle x \mid y\rangle$. 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 ℓ_2 is an inner product space, whose inner product is

$$\langle\ ,\ \rangle : \ell_2 \times \ell_2 \to \mathbb{K}, (x, y) \mapsto \sum_{n\geq 1} x_n \bar{y_n}. \tag{27}$$

Hölder's inequality implies that (27) is well-defined.

Example 39. Let (Ω, Σ, μ) 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

$$\langle\ ,\ \rangle : L_2 \times L_2 \to \mathbb{K}, (x, y) \mapsto \int_{\Omega} f\bar{g}\, \mathrm{d}\mu.$$

Hölder's inequality implies that $\langle\ ,\ \rangle$ is well-defined. Example 38 is a special case of this example.

Remark 12.2. 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 12.3. The Cauchy-Schwartz 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 $\langle\ ,\ \rangle$ on a vectors space X induces a norm on X as follows:

$$\|\ \| : X \to \mathbb{R}, x \mapsto \sqrt{\langle x, x\rangle}.$$

Proof. Just verify the axioms:

-
-

□

Corollary 12.3. The inner product of an inner product space is a continuous function on $(X \times X, \|\ \|_1)$. 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 $\langle\ ,\ \rangle$ keeps limit of sequences. Let $x, y \in X$ be given and $(x_n)_{n\in\mathbb{N}} \to x$, $(y_n)_{n\in\mathbb{N}} \to y$. Then Cauchy-Schwarz inequality implies

$$\begin{aligned} |\langle x_n, y_n\rangle - \langle x, y\rangle| &\leq |\langle x_n, y_n\rangle - \langle x_n, y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle| \\ &= |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\| \\ &\leq \sup_{j\in\mathbb{N}}\|x_j\|\|y_n - y\| + \|x_n - x\|\|y\| \\ &\to 0 (n \to \infty). \end{aligned}$$

Since $(x_n)_{n\in\mathbb{N}}$ converges in X implies that $(x_n)_{n\in\mathbb{N}}$ is bounded in X. □

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 $\mathbb{K} = \mathbb{R}$:
$$\langle x, y\rangle = \frac{\langle x+y, x+y\rangle - \langle x, x\rangle - \langle y, y\rangle}{4}.$$
- If $\mathbb{K} = \mathbb{C}$:
$$\langle x, y\rangle = \sum_{k=0}^{3} \frac{i^k}{4}\left\langle x + i^k y, x + i^k y\right\rangle.$$

Remark 12.4. 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

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{28}$$

Proof. Write the norm in the form of inner product and apply linearity. □

Remark 12.5. 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 (28).

Theorem 12.6. Let X be a linear normed space over $\mathbb{K}$ whose norm satisfies (28), then there is an inner product $\langle\ ,\ \rangle$ on X such that $\|\ \|$ is induced by $\langle\ ,\ \rangle$.

Proof.

□

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 Lecture 12-2

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 12.6. We have $A \perp \varnothing$ 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

$$\left\| \sum_{j=1}^{n} k_j x_j \right\|^2 = \sum_{j=1}^{n} |k_j|^2 \|x_j\|^2. \tag{29}$$

Proof. Prove by mathematical induction. For $n = 1$, it's trivial. For $n = 2$, we have . Then suppose (29) holds for n, now we prove (29) 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:

$$\left\|\sum_{j=1}^{n+1} k_j x_j\right\|^2 = \left\|\sum_{j=1}^{n} k_j x_j\right\|^2 + \|k_{n+1}x_{n+1}\|^2 = \sum_{j=1}^{n+1}\|k_j x_j\|. \qquad \square$$

Theorem 12.8 (Bessel's Inequality)**.** Let $\{e_n \colon n \in \mathbb{N}\}$ 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 $k \in [n]\colon h_n \perp e_k$, 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 29 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 $n \to \infty$ and we're done. $\square$

We can generalize this theorem to the case where $\{e_n \colon n \in \mathbb{N}\}$ is replaced by an orthonormal set $\{e_\alpha \colon \alpha \in I\}$ 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 $\{e_\alpha \colon \alpha \in I\}$ 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 $\operatorname{card}\mathcal{E} \leq \operatorname{card}\mathbb{N}$, where $\mathcal{E} := \{\alpha \in I \colon \langle h, e_\alpha\rangle \neq 0\}$.

Proof. Consider $\mathcal{E}_n := \{\alpha \in I : |\langle h, e_\alpha \rangle| > 1/n\}$. Then $\mathcal{E} = \bigcup_{n \geq 1} E_n$. 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 29: for all $\alpha_1, \ldots, \alpha_p \in \mathcal{E}_n$:

$$\frac{p}{n^2} < \sum_{J=1}^{p} \langle h, e_{\alpha_p} \rangle \leq \|h\|^2,$$

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 $n \in \mathbb{N}$:

$$\operatorname{span}\{h_j : j \in [n]\} = \operatorname{span}\{e_j : j \in [n]\}.$$

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

$$\operatorname{span}\{h_j : j \in [n]\} = \operatorname{span}\{e_j : j \in [n]\}.$$

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 $e_{n+1} := \widehat{e_{n+1}} / \left\|\widehat{e_{n+1}}\right\|$. 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 $h_{n+1} \in \operatorname{span}\{e_j : j \in [n]\}$, i.e. $h_{n+1} \in \operatorname{span}\{h_j : j \in [n]\}$ by assumption. Contradiction with $\{h_n : n \in \mathbb{N}\}$ being a linearly independent. Thus we can define $e_{n+1} := \widehat{e_{n+1}} / \left\|\widehat{e_{n+1}}\right\|$ and $\{e_j : j \in [n+1]\}$ is an orthonormal set in H.

Finally, we prove that

$$\operatorname{span}\{h_j : j \in [n+1]\} = \operatorname{span}\{e_j : j \in [n+1]\}. \tag{30}$$

Since

$$\forall k \in [n] \colon e_k \in \operatorname{span}\{h_j \colon j \in [n]\} \subseteq \operatorname{span}\{h_j \colon j \in [n+1]\},$$

and hence

$$e_{n+1} = \frac{h_{n+1} - \sum_{j=1}^{n} \langle h_{n+1}, e_j \rangle e_j}{\left\|\widehat{e_{n+1}}\right\|} \in \operatorname{span}\{h_j \colon j \in [n+1]\}.$$

Therefore, we proved $\supseteq$ part of (30). The $\subseteq$ part follows from

$$h_{n+1} = \left\|\widehat{e_{n+1}}\right\| e_{n+1} + \sum_{j=1}^{n} \langle h_{n+1}, e_j \rangle e_j$$

and the assumption. □

Theorem 12.11. Let $\{e_n \colon n \in \mathbb{N}\}$ 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 $\overline{\operatorname{span}\{e_n \colon n \in \mathbb{N}\}}$ contains h;
2. we have the representation (called the Fourier expansion of h): $h = \sum_{n\geq 1} \langle h, e_n \rangle e_n$;
3. Parseval's Identity holds: $\|h\|^2 = \sum_{n\geq 1} |\langle h, e_n \rangle|^2$.

Proof. We prove in the following order:

$$1 \implies 3 \implies 2 \implies 1.$$

1 $\implies$ 3: from Theorem 12.8, we have

$$\sum_{n\geq 1} |\langle h, e_n \rangle|^2 \leq \|h\|^2.$$

Suppose the other inequality is not true, i.e. $\exists a > 0$ such that

$$\|h\|^2 - \sum_{n\geq 1} |\langle h, e_n \rangle|^2 = a^2 > 0.$$

Since $h \in \overline{\operatorname{span}\{e_n \colon n \in \mathbb{N}\}}$, for the fixed $a > 0$, there is some numbers $\{\alpha_j \colon j \in [N]\}$ such that

$$\left\| h - \sum_{j=1}^{N} \alpha_j e_j \right\| < a.$$

Then apply the first 2 equivalent conditions of Theorem 13.11 (notice that $\sum_{j=1}^{N} \langle h, e_j \rangle e_j$ is the projection on $\operatorname{span}\{e_j : j \in [N]\}$) and we get

$$\begin{aligned} a^2 &> \left\| h - \sum_{j=1}^{N} \alpha_j e_j \right\|^2 \\ &\geq \left\| h - \sum_{j=1}^{N} \langle h, e_j \rangle e_j \right\|^2 . \end{aligned}$$

Clearly, h is orthogonal to $\sum_{j=1}^{N} \langle h, e_j \rangle e_j$ and we can apply Theorem 29:

$$\left\| h - \sum_{j=1}^{N} \langle h, e_j \rangle e_j \right\|^2 = \|h\|^2 - \sum_{j=1}^{N} |\langle h, e_j \rangle|^2.$$

Thus we have

$$a^2 > \|h\|^2 - \sum_{j=1}^{N} |\langle h, e_j \rangle|^2 \geq \|h\|^2 - \sum_{n=1}^{\infty} |\langle h, e_n \rangle|^2 = a^2,$$

which means $a^2 > a^2$, contradiction. Therefore, it is impossible that $\|h\|^2 > \sum_{n \geq 1} |\langle h, e_n \rangle|^2$.

3 $\implies$ 2: for all $n \in \mathbb{N}$, apply Theorem 29 to $h - \sum_{j=1}^{n} \langle h, e_j \rangle e_j$ and h, we have

$$\left\| h - \sum_{j=1}^{n} \langle h, e_j \rangle e_j \right\|^2 = \|h\|^2 - \sum_{j=1}^{n} |\langle h, e_j \rangle| \to 0 (n \to \infty).$$

Therefore,

$$h = \sum_{n=1}^{\infty} \langle h, e_n \rangle e_n.$$

2 $\implies$ 1: the sequence $(h_n)_{n \in \mathbb{N}}$, where $h_n := \sum_{j=1}^{n} \langle h, e_j \rangle e_j$ converges to h and $(h_n)_{n \in \mathbb{N}} \subseteq \operatorname{span}\{e_n : n \in \mathbb{N}\}$. $\square$

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 Lecture 13-1

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 $M \subseteq H$: $M \cap M^\perp = \varnothing$;
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 $M \supseteq (x_n)_{n\in\mathbb{N}} \to x$ by the density. Since $\langle\ ,\ \rangle$ is continuous:
$$\langle y, x\rangle = \lim_{n\to\infty} \langle y, x_n\rangle = \lim_n 0 = 0.$$
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
$$M^{\perp} = \bigcap_{y \in M} \{x \in H \colon \langle x, y \rangle = 0\}.$$
It suffices to show that $\forall y \in M$, the set
$$\{y\}^{\perp} = \{x \in H \colon \langle x, y \rangle = 0\}$$
is a closed subspace of H. That's trivial since it's just the kernel of $f_y \colon H \to \mathbb{K}, x \mapsto \langle x, y \rangle$. □

13.1.1 Orthonormal Basis

Theorem 13.2 (Riesz–Frèchet). 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 $x \in \overline{\operatorname{span} E}$ such that
$$x = \sum_{n \geq 1} \alpha_n e_n.$$

Proof. Define a sequence $(x_n)_{n \in \mathbb{N}}$ by $x_n := \sum_{j=1}^{n} \alpha_j e_j$, then $(x_n)_{n \in \mathbb{N}}$ lies in E. Theorem 29 implies that $\forall m \geq n$:
$$\|x_m - x_n\|^2 = \sum_{j=n}^{m} |\alpha_j|^2 \to 0 (m, n \to \infty).$$
Therefore, the sequence $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence. Since $\mathcal{H}$ is complete, we know $\exists! x \in \mathcal{H}$ such that $x = \lim_n x_n$. □

Remark 13.1. Furthermore, the sequence α is determined by the limit x since $\alpha_n = \langle x, e_n \rangle$.

Definition. Let $\mathcal{H}$ and an orthonormal subset $E \subseteq \mathcal{H}$ be given. For all $x \in \mathcal{H}$, we define the set E_x by
$$E_x := \{y \in \mathcal{H} \colon \langle x, y \rangle \neq 0\}.$$

Definition. Let $\mathcal{H}$ and an orthonormal subset $E \subseteq \mathcal{H}$ be given. If $\overline{\operatorname{span} E} = \mathcal{H}$, then E is called an **orthonormal basis** of $\mathcal{H}H$.

Question 13.1. Does othonormal 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. $\overline{\operatorname{span} E} = H$;
2. $\forall x \in \mathcal{H} \colon x = \sum_{e \in E_x} \langle x, e \rangle e$;

3. $\forall x \in \mathcal{H}$: $x = \sum_{e \in E_x} \langle x, e \rangle e$;

4. $\forall x, y \in \mathcal{H}$: $\langle x, y \rangle = \sum_{e \in E_x \cap E_y} \langle x, e \rangle \overline{\langle y, e \rangle}$;

5. $E^\perp = \{0\}$.

Proof. From Theorem 12.11, we know 1 $\iff$ 2 $\iff$ 3. And we will prove

$$2 \implies 4 \implies 5 \implies 1.$$

2 $\implies$ 4: in the limit sense, we have

$$x = \sum_{e \in E_x} \langle x, e \rangle e.$$

For all $y \in H$, the map $\langle \ , y \rangle$ is a continuous functional and hence

$$\langle x, y \rangle = \sum_{e \in E_x} \langle x, e \rangle \langle e, y \rangle .$$

4 $\implies$ 5: given an arbitrary element $x \in E^\perp$. Then

$$\langle x, x \rangle = \sum_{e \in E_x} |\langle x, e \rangle|^2,$$

and the sum must be 0 anyway. Therefore, $\langle x, x \rangle = 0 \implies x = 0$.

5 $\implies$ 1: if there is an element $x \in \mathcal{H} \setminus \overline{\operatorname{span} E}$, from Corollary 12.9: $\|x\|^2 \geq \sum_{e \in E_x} |\langle x, e \rangle|^2$. From Theorem 13.2, we have

$$\exists! y \in \overline{\operatorname{span} E} \text{ such that } y = \sum_{e \in E_x} \langle x, e \rangle e.$$

Let $z := y - x$, then $z \neq 0$. For all $e \in E \setminus E_x$: $\langle x, e \rangle = 0$, $\langle y, e \rangle = 0$. For all $e \in E_x$: $\langle x, e \rangle = \langle y, e \rangle$. Therefore, for all $e \in E$: $\langle z, e \rangle = \langle y, e \rangle - \langle x, e \rangle = 0$, i.e. $z \perp E$. Above all $E^\perp \ni z$, $E^\perp \neq \{0\}$. □

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. Let

$$\mathcal{F} := \{E \in \mathcal{P}(\mathcal{H}) \colon E \text{ is orthonormal}\}.$$

Then $(\mathcal{F}, \subseteq)$ is an partially-ordered set. For all totally-ordered subset $\mathcal{A} \subseteq \mathcal{F}$, we prove that $\mathcal{A}$ has an upper bound. It suffices to show that $\bigcup \mathcal{A}$ lies in $\mathcal{F}$, i.e. $\bigcup \mathcal{A}$ is orthonormal.For all $e_i, e_j \in \bigcup \mathcal{A}$, there is some $G_i, G_j \in \mathcal{A}$ such that $e_i \in G_i$, $e_j \in G_j$. WLOG, suppose $G_i \subseteq G_j$ then $e_i, e_j \in G_j$, which means $\langle e_i, e_j \rangle = \delta_i^j$. Thus $\bigcup \mathcal{A} \in \mathcal{F}$ is an upper bound of $\mathcal{A}$. Zorn's lemma ensures that there is a maximal element in $\mathcal{F}$, and we denote it by M.

Claim. We have $\overline{\operatorname{span} M} = \mathcal{H}$. In other words, M is an orthonormal basis.

Clearly $\overline{\operatorname{span} M}$ is a closed subspace of $\mathcal{H}$, and Theorem 13.3 ensures that $\overline{\operatorname{span} M} = \mathcal{H}$ is equivalent to $M^\perp = \{0\}$. If we have $M^\perp \neq \{0\}$, i.e. there is an element $y \in M^\perp$ such that $\langle y, y \rangle = 0$. Then $M \cup \{y\} \in \mathcal{F}$ is strictly larger than M, which is a contradiction. □

Example 42. The space ℓ_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

$$\langle\ ,\ \rangle : L_2 \times L_2 \to \mathbb{K}, (f, g) \mapsto \frac{1}{2\pi} \int_{[-\pi,\pi]} f \bar{g} \, \mathrm{d}m.$$

has the basis

$$E = (e_n)_{n \in \mathbb{Z}}, \text{ where we define } e_n : [-\pi, \pi] \to \mathbb{C}, x \mapsto e^{inx}.$$

Proof. We need some lemmas to finish this proof. Let $C[-\pi, \pi]$ denotes 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 $T[-\pi, \pi] := \operatorname{span} E$ 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 simple functions can be approximated, which can be ensured by the outer regularity of Lebesgue measurable sets. See [1, Theorem 3.48]. □

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 $g \in C[-\pi, \pi]$ be given. For all $\delta > 0$, consider

$$g_\delta : [-\pi, \pi] \longrightarrow \mathbb{C},$$

defined by

$$x \longmapsto \begin{cases} g(x), & x \in [-\pi, \pi - \delta]; \\ h_\delta(x), & x \in [\pi - \delta, \pi]; \\ g(-\pi), & x = \pi; \end{cases}$$

where h_δ is the affine map is selected to make that g_δ is continuous. Then Lemma 13.7 ensures that $\lim_{\delta \to 0} \|g_\delta - g\|_2 = 0$ (remember, Dominated Convergence Theorem ensures that convergence in norm, not only changing the order of lim and $\int$).

Lemma 13.7. Let $(X, \mathcal{A}, \mu)$ be a measure space and Ω be a metric space. Fix a point $t_0 \in \Omega$. Suppose there is a function

$$f\colon X \times \Omega \to \mathbb{C}, x \mapsto f(x,t)$$

satisfying the following conditions:

1. for all $x \in X$, the function $t \mapsto f(x,t)$ is measurable;
2. for almost every $x \in X$, the map $t \mapsto f(x,t)$ is continuous at t_0 (i.e. there is a null set N such that for all $x \in N^c$, the map $t \mapsto f(x,t)$ is continuous);
3. there is a function $h \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ such that for all $t \in \Omega$, we have
$$|f(x,t)| \leq h(x)$$
for almost every $x \in X$ (i.e. for each t there is a null set N_t such that $\forall x \in N_t^c$, we have $|f(x,t)| \leq h(x)$).

Then the function

$$F\colon \Omega \to \mathbb{C}, t \mapsto F(t) = \int_X f(x,t)\,\mathrm{d}\mu(x)$$

is well-defined and continuous at t_0.

Therefore, we proved Lemma 13.6. □

Proof of Lemma 13.7. For all $t \in \Omega$, we have $|f(x,t)| \leq h(x)$ for almost every $x \in X$, and hence $f(\ ,t) \in \mathcal{L}^1(X, \mathcal{A}, \mu)$. Then f is well-defined. To see that F is continuous at t_0, let an arbitrary sequence $(t_n)_{n\in\mathbb{N}}$ that converges to t_0 be given, then we want to see

$$\lim_n \int_X f(x,t_n)\,\mathrm{d}\mu(x) = \int_X f(x,t)\,\mathrm{d}\mu(x).$$

This can be proved by Dominated Convergence Theorem. Just let h be the dominating function and

$$N_{(t_n)_{n\in\mathbb{N}}} := \bigcup_{n\geq 1} N_{t_n}$$

is a null set such that $\forall n \in \mathbb{N}$, we have $f_n := f(\ ,t_n)$ is dominated by h for all $x \notin N_{(t_n)_{n\in\mathbb{N}}}$. $\square$

Theorem 13.8 (Weierstrass)**.** The space $T[-\pi,\pi]$ is dense in $C_{2\pi}$, with respect to the infinity norm $\|\ \|_\infty$.

Theorem 13.8 implies that $T[-\pi,\pi]$ is dense in $C_{2\pi}$, with respect to the 2-norm $\|\ \|_2$. Proof of Theorem 13.8 can be found in many books, such as [2, Chapter 2, Corollary 5.4].

Above all, we proved that span $E = T[-\pi,\pi]$ is dense in $L_2[-\pi,\pi]$. $\square$

Remark 13.2. The space $L_2[-\pi,\pi]$ has interesting properties.

1. for all $f \in L_2[-\pi,\pi]$, we have

$$f = \sum_{k\in\mathbb{Z}} \langle f, e_k\rangle\, e_k,$$

where the $=$ means the limit with respect to L_2-norm (and hence the Fourier series convergent to f in measure, which ensures that there is a subsequence convergent to f almost everywhere).

2. in fact, in 1966, Lennart Carleson proved that: for $f \in L_2[-\pi,\pi]$, we have

$$f(x) = \sum_{k\in\mathbb{Z}} \langle f, e_k\rangle\, e_k(x)$$

for almost every $x \in [-\pi,\pi]$. Here is a relevant post.

3. abovel all, $L_2[-\pi,\pi]$ has the orthonormal basis $(e_k)_{k\in\mathbb{Z}}$. For the space $L_2[-1,1]$, it has an orthonormal basis: Legendre Polynomial, defined as

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n\,,$$

for all $n \in \mathbb{Z}_{\geq 0}$. This also means that an orthonormal basis in a Hilbert space may be not unique.

Theorem 13.9. 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 $\operatorname{card} E = \operatorname{card} \mathbb{N}$, then $\mathcal{H} \cong \ell_2$;

3. If $\operatorname{card} E = n \in \mathbb{N}$, then $\mathcal{H} \cong \mathbb{K}^n$.

Proof. 1: if $\mathcal{H}$ is separable, let $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}$ be a dense subset. Pick a maximal linearly independent subset of $(x_n)_{n\in\mathbb{N}}$, denoted by $(y_n)_{n\in\mathbb{N}}$. Apply Theorem 12.10 to $(y_n)_{n\in\mathbb{N}}$ and we get an orthonormal sequence $(e_n)_{n\in\mathbb{N}}$. It's an orthonormal basis, since

$$\mathcal{H} = \overline{\operatorname{span}}(x_n)_{n\in\mathbb{N}} = \overline{\operatorname{span}}(y_n)_{n\in\mathbb{N}} = \overline{\operatorname{span}}(e_n)_{n\in\mathbb{N}}.$$

If $\mathcal{H}$ has an orthonormal basis $(e_n)_{n\in\mathbb{N}}$, i.e.

$$\mathcal{H} = \overline{\operatorname{span}}(e_n)_{n\in\mathbb{N}}.$$

Consider the set

$$A_n := \operatorname{span}_{\mathbb{Q}}\{e_j : j \in [n]\}$$

for all $n \in \mathbb{N}$, where $\operatorname{span}_{\mathbb{Q}}$ means the vector space generated over the field $\mathbb{Q}$. Then A_n is linearly isomorphic to a n-dim vector space over $\mathbb{Q}$, thus

$$\operatorname{card} A_n = \operatorname{card} \mathbb{Q}^n = \operatorname{card} \mathbb{Q},$$

since $\operatorname{card} \mathbb{Q} = \operatorname{card} \mathbb{N}$ implies that $\operatorname{card} \mathbb{Q}^n = \operatorname{card} \mathbb{N}^n = \operatorname{card} \mathbb{N}$, which can be proved by induction. Now

$$\operatorname{span}_{\mathbb{Q}}(e_n)_{n\in\mathbb{N}} = \bigcup_{n\in\mathbb{N}} A_n$$

is a countable (since it is the countable union of countable sets) subset of $\mathcal{H}$, and it is dense in $\mathcal{H}$. For all $x \in \mathcal{H}$, Theorem 13.2 ensures that there is some $\alpha \in \ell_2$ such that $x = \sum_{n\geq 1} \alpha_n e_n$. Then x can be approximated by elements of $\operatorname{span}\{e_j : k \in [N]\}$ for big N, and hence can be approximated by elements of $\operatorname{span}_{\mathbb{Q}}\{e_j : k \in [N]\}$. Above all, we proved that $\mathcal{H} = \overline{\operatorname{span}_{\mathbb{Q}}(e_n)_{n\in\mathbb{N}}}$.

2: let $E := (e_n)_{n\in\mathbb{N}}$ be an orthonormal basis. Consider

$$\Phi : \mathcal{H} \to \ell_2, x \mapsto \big(\langle x, e_n\rangle\big)_{n\in\mathbb{N}}.$$

Then $\mathcal{H}$ is an isometry. Linearity follows from the linearity of the inner product on $\mathcal{H}$. It is surjective as Theorem 13.2. It is injective since $\Phi(x) = 0 \iff x \in E^\perp = \{0\}$, as Theorem 13.3.

3: similarly. □

13.2 Lecture 13-2

This lecture was given by Zhiyao Chen.

We assume that H is an inner product space and $\mathcal{H}$ is a Hilbert space (over $\mathbb{K}$) in this lecture.

13.2.1 Projection

Definition. Let $E \hookrightarrow \mathcal{H}$ be a subspace. For all $x \in \mathcal{H}$, if there are two elements $x_0 \in E, x_1 \in E^\perp$ such that $x = x_0 + x_1$, then we say: x_0 is the projection of x onto E. The projection of x onto E is usually denoted by x_E.

There are 2 natural questions:

Question 13.2. Does the projection of x onto E exist? Is it unique?

Question 13.3. Are there more properties of x_E?

We answer Question 13.2 in this lecture, and some of Question 13.3.

Theorem 13.10. Let $E \hookrightarrow \mathcal{H}, x \in \mathcal{H}$ and $x_0 \in E$ be given. The following statements are equivalent:

1. the projection of x onto E is just x_0;
2. the distance as a infimum is reached at x_0: $d(x, E) = \|x - x_0\|$;
3. for all $z \in E$, the function
$$f\colon \mathbb{R} \to \mathbb{R}, \lambda \mapsto \|x - x_0 - \lambda z\|^2$$
reaches the minimum at 0.

Proof. 1 $\implies$ 2: suppose $x_0 = x_E$. Then $x - x_0 \perp E$. On the one hand: $d(x, x_0) \geq d(x, E)$. On the other hand: for all $y \in E$:

$$\begin{aligned}
\|x - y\|^2 &= \left\| (x - x_0) + \underbrace{(x_0 + y)}_{\in E} \right\|^2 \\
&= \|x - x_0\|^2 + \|x_0 - y\|^2 \\
&\geq \|x - x_0\|^2.
\end{aligned}$$

Therefore $\|x - y\| \geq \|x - x_0\|$. Since y is arbitrary, we have $d(x, E) \geq \|x - x_0\|$.

2 $\implies$ 3: since for all $z \in E$, we have $x_0 - \lambda z \in E$, for all $\lambda \in \mathbb{R}$. Then the definition of infimum implies the result.

3 $\implies$ 1: let $z \in E$ be given. The function f is differentiable at 0:

$$\begin{aligned}
\lim_{t\to 0} \frac{f(t)-f(0)}{t-0} &= \lim_{t\to 0} \frac{\langle x-x_0-tz, x-x_0-tz\rangle - \langle x-x_0, x-x_0\rangle}{t} \\
&= \lim_{t\to 0} \frac{-2t\operatorname{Re}\langle x-x_0, z\rangle + t^2\|z\|^2}{t} \\
&= -2\operatorname{Re}\langle x-x_0, z\rangle .
\end{aligned}$$

And f reaches the minimum at 0 implies that $f'(0) = 0$. In other words, $\operatorname{Re}\langle x-x_0, z\rangle = 0$. Replace z by iz and we get $\operatorname{Im}\langle x-x_0, z\rangle = 0$ and hence $\langle x-x_0, z\rangle = 0$ holds for all $z \in E$. Therefore, $x - x_0 \perp E$ and x_0 is the projection of x onto E. ☐

Theorem 13.11 (Projection)**.** Let $E \hookrightarrow \mathcal{H}$ be a closed subspace. Then for all $x \in H$:

$$\exists! x_E \in E\colon \|x - x_E\| = d(x, E).$$

In other words, the projection of x onto E exists and is unique.

Proof. Existence: let $a := d(x, E)$. By the definition of distance, we can pick a sequence $(y_n)_{n\in\mathbb{N}} \subseteq E$ such that

$$\forall n \in \mathbb{N}\colon a \le \|x - y_n\| < a + 1/n. \tag{31}$$

Now we prove that $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. For all $m, n \in \mathbb{N}$, apply Parallelogram Law:

$$\begin{aligned}
\left(a + 1/n\right)^2 + \left(a + 1/m^2\right) &\ge \|x - y_n\|^2 + \|x - y_m\|^2 \\
&= 2\Big(\|x - (y_n + y_m)/2\|^2 + \|y_n - y_m\|^2\Big) \\
&\ge 2a^2 + 2\|y_n - y_m\|^2.
\end{aligned}$$

Thus $\lim_{m,n}\|y_n - y_m\| = 0$. Since $E \hookrightarrow \mathcal{H}$ is closed and hence complete, there is $\lim_n y_n = y \in E$. Let $n \to \infty$ in (31) and we get $\|x - y\| = a$, by the continuity of norm. Existence has been proved.

Uniqueness: if there is another $\tilde{y} \in E$ such that $d(x, E) = a = \|x - \tilde{y}\|$, then

$$\begin{aligned}
2a^2 &= \|x - \tilde{y}\|^2 + \|x - y\|^2 \\
&= 2\Big(\|x - (y + \tilde{y})/2\|^2 + \|y - \tilde{y}\|^2\Big) \\
&\ge 2a^2 + 2\|y - \tilde{y}\|^2.
\end{aligned}$$

Thus $\|y - \tilde{y}\| = 0$, i.e. $y = \tilde{y}$. ☐

Remark 13.3. Another way to prove uniqueness:

$$\begin{aligned} a^2 &= \|x-y\|^2 \\ &= \Bigg\| \underbrace{(x-\tilde{y})}_{\in E^\perp} + \underbrace{(\tilde{y}-y)}_{\in E} \Bigg\|^2 \\ &= \|x-\tilde{y}\|^2 + \|\tilde{y}-y\|^2 \\ &\geq a^2 + \|x-\tilde{y}\|^2. \end{aligned}$$

Then $y = \tilde{y}$.

In fact, we can change E to be a closed convex subset in Theorem 13.11.

Exercise 13.1. Change E to be a closed convex subset in Theorem 13.11 and prove the same result.

Theorem 13.12. Given $\mathcal{H}$ and $E \hookrightarrow \mathcal{H}$ be a closed subspace, then

1. $\mathcal{H} = E \oplus E^\perp$;

2. $(E^\perp)^\perp = E$.

Proof. The projection operator ensures that for all $x \in \mathcal{H}$, we have $x = x_E + (x - x_E)$ where $x_E \in E$ and $x - x_E \in E^\perp$, thus $\mathcal{H} = E + E^\perp$. Then $E \cap E^\perp = \{0\}$ ensures that $\mathcal{H} = E \oplus E^\perp$.

It is easy to verify $E \subseteq (E^\perp)^\perp$. For another direction, taking an arbitrary $x \in (E^\perp)^\perp$, i.e. $x \perp E^\perp$, consider the projection P_E: we have $x = x_1 + x_2$ where $x_1 = P_E x \in E$, $x_2 = x - P_E x \in E^\perp$. Now we prove $x_2 = 0$ and hence $x = x_1 \in E$. Look at the inner product:

$$\langle x_2, x_2 \rangle = \langle x, x_2 \rangle - \langle x_1, x_2 \rangle = 0 - 0 = 0$$

because $x \in (E^\perp)^\perp$, $X_2 \in E^\perp$ ensures the first 0 and $x_1 \in E$, $x_2 \in E^\perp$ implies the second 0, as we wanted. □

Remark 13.4. More than $\mathcal{H} = E \oplus E^\perp$: for $H \ni x = y + z$, where $y \in E$, $z \in E^\perp$, we have $\|x\|^2 = \|y\|^2 + \|z\|^2$ as Theorem 29. ■

Corollary 13.13. For all $E \hookrightarrow \mathcal{H}$, we have $(E^\perp)^\perp = \overline{E}$.

Proof. Because $E^\perp = \left(\overline{E}\right)^\perp$. □

Consider a simple and familiar example.

Example 44. We have $\mathbb{R}^2 = E \oplus E^\perp$, where $E := \{(x, y) \in \mathbb{R}^2 : y = 0\}$ and $E^\perp = \{(x, y) \in \mathbb{R}^2 : x = 0\}$. ■

Remark 13.5. Above all, given a Hilbert space $\mathcal{H}$ and a closed subspace $E \hookrightarrow \mathcal{H}$, we have studied:

1. the existence of projection: for all $x \in \mathcal{H}$, $\exists! x_E \in E$ such that $x - x_E \in E^\perp$;

2. the orthogonal decomposition $\mathcal{H} = E \oplus E^\perp$;

Finally, we prove that $P_E \in \mathcal{B}(\mathcal{H}, E)$.

Proposition 13.14. Given $\mathcal{H}$ and a closed subspace $E \hookrightarrow \mathcal{H}$. Then $P_E \in \mathcal{B}(\mathcal{H}, E)$.

Proof. Linearity: for arbitrary $x, y \in \mathcal{H}$, we have

$$x = x_E + x_1, y = y_E + y_1$$

where $x_E = P_E(x)$, $y_E = P_E(y)$ and $x_1, y_1 \in E^\perp$. Then

$$x + y = (x_E + y_E) + (x_1 + y_1)$$

where $x_E + y_E \in E$ and $x_1 + y_1 \in E^\perp$. Since $\mathcal{H} = E \oplus E^\perp$ is a direct sum, we know $P_E(x + y) = x_E + y_E$. Similarly we know $P_E(kx) = kP_E(x)$ for all $k \in \mathbb{K}$.

Boundedness: from Remark 13.4, we have

$$\|x\|^2 = \|P_E x\|^2 + \|x - P_E x\|^2 \geq \|P_E x\|^2,$$

which implies $\|P_E\| \leq 1$. And in fact, $\|P_E\| = 1$ if $E \neq \{0\}$, $\|P_E\| = 0$ when $E = \{0\}$. To see this, take an arbitrary $0 \neq x \in E$ (if possible) and we have $P_E(x) = x$. □

For more properties of projection operator, see Lecture 14-2 .

13.2.2 Further Topics

Reference: Characterizing compact sets in L_p-spaces and its application.

Throughout this part, X is a Borel-regular Borel metric measure space such that every open ball with positive radius has a positive and finite measure (definitions can be found below).

We studied compactness and relative compactness in Lecture 6-1 . Recall: in a metric space (X, d), a subset $F \subseteq X$ is

- relatively compact, if and only if the closure of it, i.e. $\overline{F}$ is compact;
- compact, if and only if it is closed and relatively compact.

Notation. We use the operator τ_a for $a \in \mathbb{R}^n$, whose definition can be found in Fourier series's divergence. For convenience, let $B(y, r)$ be the closed ball for a metric space Y, centered at y with radius r.

Theorem (Kolmogorov-Riesz)**.** For $p \in [1, \infty)$, given a subset $F \subseteq L_p(\mathbb{R}^n)$, where $\mathbb{R}^n$ is equipped with the Euclidean metric, Lebesgue measurable sets $\mathcal{M}$ and Lebesgue measure m. Then F is relatively compact, if and only if the following conditions are satisfied.

1. it is bounded, i.e. $\sup_{f \in F} \|f\|_p < \infty$;
2. the two limits hold

$$\lim_{r \to 0} \sup_{f \in F} \|\tau_r f - f\|_p = 0,$$

$$\lim_{R \to 0} \sup_{f \in F} \left\|f \cdot \chi_{\mathbb{R}^n \setminus B(0,R)}\right\|_p = 0.$$

The result is generalized as following:

Definition. Let $(X, \mathcal{M}, \mu, d)$ be a metric measure space. It is said to be **Borel**, if $\mathcal{M}$ contains the Borel σ-algebra, the σ-algebra generated by all open sets. It is said to be **Borel-Regular**, if each $E \subseteq X$ is contained in a Borel set B such that $\mu(B) = \mu(E)$. It is said to be **doubling**, if there is some $\lambda \geq 1$ such that

$$\mu(B(x, 2r)) \leq \lambda \mu(B(x, r))$$

for all $x \in X$, $r > 0$. The above λ is called the **doubling constant**.

Theorem. Let (X, d, μ) be a doubling metric measure space and $p > 1$. Suppose that

$$\inf\{\mu(B(x, r)) \colon x \in X\} > 0$$

for any $r > 0$. Then $F \subseteq L_p := L_p(X, d, \mu)$ is relatively compact if and only if the following conditions are satisfied.

1. it is bounded, i.e. $\sup_{f \in F} \|f\|_p < \infty$;
2. the two limits hold

$$\lim_{r \to 0} \sup_{f \in F} \|A_r f - f\|_p = 0,$$

$$\inf_{\operatorname{diam} E < \infty} \sup_{f \in F} \left\|f \cdot \chi_{X \setminus E}\right\|_p = 0,$$

where the operator A_r is defined as following.

Definition (Average Operator)**.** Let (X, d, μ) be a metric measure space . For $p \geq 1$, we define an operator A_r on $L_p := L_p(X, d, \mu)$ for all $r > 0$. The **average operator**

$$A_r \colon L_p \to L_p, f \mapsto (A_r f),$$

where $A_r f$ is defined by

$$A_r f(x) := \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f \,\mathrm{d}\mu,$$

i.e. the average value of f with respect to the ball centered at x with radius r.

And in 2022, Katsuhisa Koshino proved the following result in Characterizing compact sets in L_p-spaces and its application .

Theorem. Let (X, d, μ) be a doubling metric measure space and $p \geq 1$. Suppose that for any $x \in X$ and any $r > 0$,

$$\mu(B(x,r) \Delta B(y,r)) \to 0$$

as $y \to x$. Then $F \subseteq L_p := L_p(X, d, \mu)$ is relatively compact if and only if the following conditions are satisfied.

1. it is bounded, i.e. $\sup_{f \in F} \|f\|_p < \infty$;
2. the two limits hold

$$\lim_{r \to 0} \sup_{f \in F} \|A_r f - f\|_p = 0,$$
$$\inf_{\operatorname{diam} E < \infty} \sup_{f \in F} \left\| f \cdot \chi_{X \setminus E} \right\|_p = 0.$$

And it is natural to ask, is it possible to generalize the result to $L_{p(\cdot)}$, the variable Lebesgue space?

Reference about variable Lebesgue spaces: Variable Lebesgue Spaces.

14 Week 14

14.1 Lecture 14-1

This Lecture was given by me.

Recall

Last Thursday, we studied "projection on Hilbert space": let $\mathcal{H}$ be a Hilbert space and $E \hookrightarrow \mathcal{H}$ be a subspace of $\mathcal{H}$. Then for all $x \in \mathcal{H}$, there is a unique $P_E(x) \in \mathcal{H}$ such that $x - P_E(x) \perp E$. As a corollary: we have a direct sum decomposition:

$$\mathcal{H} = E \oplus E^{\perp}.$$

Furthermore, there is a non-trivial duality:

$$(E^{\perp})^{\perp} = E. \tag{32}$$

14.1.1 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 (more precisely, by **fixing y at the second position** of the inner product):

$$H^* \ni \varphi_y \colon H \to \mathbb{K}, x \mapsto \langle x, y \rangle .$$

Then $\varphi\colon H \to H^*, y \mapsto \varphi_y$ is conjugate-linear and keeps norms. The map φ depends on H (also denoted by φ^H for clarity), is a conjugate-linear injection satisfying $\| \ \|_{H^*} \circ \varphi = \| \ \|_H$.

Theorem 14.1 (Riesz Representation Theorem)**.** Let H be a Hilbert space, then $\forall f \in H^*$ $\exists! y_f \in H$ such that

$$\forall x \in X \colon f(x) = \langle x, y_f \rangle .$$

Furthermore, we have $\|f\|_{H^*} = \|y_f\|_H$.

Proof. We prove the existence and uniqueness of y_f for all $f \in H^*$. Let an arbitrary functional $f \in H^*$ be given.

1. Existence: if $f = 0$, it suffices to take $y_0 = 0$; suppose $f \neq 0$, i.e. $N := \ker f \neq \{0\}$. As we know, $\ker f$ is a closed subspace of H and there is a direct sum decomposition of H:

$$H = N \bigoplus N^{\perp}$$

Taking an element $0 \neq y \in H^{\perp}$. Then $f(y) = 0$, since $f(y) = 0 \implies y \in N$, and at the same time $N^{\perp} \ni y \neq 0$, thus $y \notin N$.

Claim. We have

$$f(x) = \left\langle x, \frac{\overline{f(y)}}{\|y\|^2} y \right\rangle, \forall x \in H.$$

To see this, notice that for all $x \in H$:

$$x - \frac{f(x)}{f(y)} y \in N,$$

thus it's perpendicular with $y \in N^{\perp}$, i.e.

$$\left\langle x - \frac{f(x)}{f(y)} y, y \right\rangle = 0.$$

Simple calculations imply that

$$f(x) = \left\langle x, \frac{\overline{f(y)}}{\|y\|^2} y \right\rangle, \forall x \in H.$$

Let $y_f := \frac{\overline{f(y)}}{\|y\|^2} \cdot y$ and we're done.

2. Uniqueness: suppose both of $y_f, \tilde{y}$ satisfy that for all $x \in H$:

$$\langle x, y_f \rangle = f(x) = \langle x, \tilde{y} \rangle .$$

Then take $x = y_f - \tilde{y}$, and we see $\langle y_f - \tilde{y}, y_f - \tilde{y} \rangle = 0$, which implies $y_f - \tilde{y} = 0$. □

Remark 14.1. Here we can't change the "a Hilbert space" to "an inner product space" in Theorem 14.1. In other words, completeness is necessary in some sense. See Example 45. ■

Remark 14.2. This proof can be decomposed to 2 ideas:

1. Two linear functionals (not needed to be continuous) on a vector space have the same kernel if and only if they are scalar multiplication of each other.

Proof. Consider a vector space V over the field $\mathbb{K}$. Suppose $f, g \in V^{\sharp}$ has the same kernel $N = \ker f = \ker g$. If $N = V$, then $f = g = 0$ and we're done. Else, if $N \neq 0$, consider the quotient space V/N, which is isomorphic to $\mathbb{K}$ since both of f, g are isomorphisms. Then $\dim V/N = 1$, and so is $\mathcal{L}(V/N, \mathbb{K})$. Consider the quotient map $\pi \colon V \to V/N$ and it induces $\tilde{f}, \tilde{g}$ as the following commutative diagram

$$\begin{tikzcd} V \arrow[r, "f,g"] \arrow[d, "\pi"'] & \mathbb{K} \\ V/N \arrow[ur, "\tilde{f},\tilde{g}"'] & \end{tikzcd}$$

Then we know $\tilde{f}, \tilde{g} \neq 0$ and the diminsion implies that there is $0 \neq c \in \mathbb{K}$ such that $\tilde{f} = c\tilde{g}$. Now for all $x \in V$, we have

$$f(x) = \tilde{f}(\pi(x)) = c\tilde{g}(\pi(x)) = cg(x),$$

since the diagram commutes. □

2. Then consider a Hilbert space H. To find a vector y such the functional $f \in H^*$ is just φ_y as defined above, it suffices to find y such that $\ker(\varphi_y) = \ker f$. What's $\ker(\varphi_y)$? The set of vectors that are orthogonal to y. If $\ker f = \{v\}^\perp$, then the duality (32) implies that
$$\operatorname{span}\{v\} = (\ker f)^\perp.$$
Thus if a non-zero vector of $(\ker f)^\perp$ was found, then we were done. ■

Example 45 (Not Complete Space)**.** Consider the inner product space (the space of all finitely supported sequences) $c_c \hookrightarrow \ell_2$, where

$$c_c := \{x = (x_n)_{n\in\mathbb{N}} \in {}^{\mathbb{N}}\mathbb{K} \colon x_n = 0 \text{ for all but finitely many } n\}$$

and ℓ_2 is equipped with the usual inner product. Then the functional

$$f \colon c_c \to \mathbb{K}, x \mapsto \sum_{n=1}^{\infty} x_n/n$$

is bounded, with norm $\pi/\sqrt{6}$ but there is no element $y \in c_c$ such that

$$\forall x \in c_c \colon f(x) = \langle x, y \rangle .$$

Definition (Conjugate-Isometry)**.** Let H, K be 2 linear normed spaces. If $f \colon X \to Y$ be a conjugate-linear bijection such that

$$\forall x \in X \colon \|f(x)\|_Y = \|x\|_X,$$

then f is said to be a **conjugate-isometry**. We write $H \underset{\text{conj}}{\cong} K$ if there is an conjugate isometry from H to K.

Corollary 14.2. For a Hilbert space H, there is a conjugate-isometry:

$$\varphi\colon H \to H^*, y \mapsto \varphi(y) = \varphi_y,$$

where the map φ_y is defined as follows

$$\varphi_y\colon H \to \mathbb{K}, x \mapsto \langle x, y \rangle,$$

for all $y \in H$.

Proof. Theorem 14.1 implies that φ is bijective. Furthermore, φ is

- conjugate-linear: let arbitrary $y, z \in H$ and $k, l \in \mathbb{K}$ be given. Then

$$\begin{aligned} \forall x \in H\colon \varphi_{ky+lz}(x) &= \langle x, ky + lz \rangle \\ &= \bar{k} \langle x, y \rangle + \bar{l} \langle x, z \rangle \\ &= \bar{k}\varphi_y(x) + \bar{l}\varphi_z(x) \\ &= (\bar{k}\varphi_y + \bar{l}\varphi_z)(x), \end{aligned}$$

  and hence

$$\varphi_{ky+lz} = \bar{k}\varphi_y + \bar{l}\varphi_z.$$

- isometry: it suffices to show that it keeps norms. Let $y \in H$ be an arbitrary fixed element. Then

$$\forall x \in H\colon |\varphi_y(x)| = |\langle x, y \rangle| \leq \|x\| \cdot \|y\|,$$

  which means $\|\varphi_y\| \leq \|y\|$. And $\varphi(y) = \|y\|^2$ implies that $\|\varphi_y\| \geq \|y\|$, thus $\|\varphi_y\| = \|y\|$.

Above all, φ is a conjugate-isometry. □

Remark 14.3. The conjugate isometry given by this corollary is called the Riesz map (with respect to H), also denoted by φ^H.

Example 46. Consider a measure space (Ω, Σ, μ). Then $L_2(\Omega, \Sigma, \mu)$ is a Hilbert space (denoted by L_2 for short from now on). This implies

$$L_2 \underset{\text{conj}}{\cong} L_2^*.$$

On the other hand, we have $L_2 \underset{\text{conj}}{\cong} L_2$, since

$$\overline{\mathrm{id}_{L_2}}\colon L_2 \to L_2, f \mapsto (\bar{f}\colon \omega \mapsto \overline{f(\omega)})$$

is a conjugate isometry. Then $\varphi^{L_2} \circ \overline{\mathrm{id}_{L_2}}$ is an isometry from L_2 to L_2^*.

Question 14.1. 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 there is an inner product space X and there is a conjugate-isometry from X to X^*, then X is a Hilbert space.

We will answer this later.

Proposition 14.3. Let H be a Hilbert space and φ be the bijection of above corollary. Then the norm (as a operator norm) of H^* is induced by the following inner product:

$$\langle\ ,\ \rangle_{H^*} : H^* \times H^* \to \mathbb{K}, (\varphi(y), \varphi(x)) \mapsto \langle x, y\rangle_H\,, \tag{33}$$

and H^* is a Hilbert space.

Proof. It suffices to prove that $\|\ \|_{H^*}$ 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 $\|\ \|_{H^*}$ is induced by an inner product $\langle\ ,\ \rangle_{H^*}$. Now we prove (33). 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:

$$\begin{aligned}
\langle \varphi(y), \varphi(x) \rangle_{H^*} &= \frac{1}{4} \sum_{\alpha^4=1} \alpha \|\varphi(y) + \alpha\varphi(x)\|^2 \\
&= \frac{1}{4} \sum_{\alpha^4=1} \alpha \|\varphi(y + \overline{\alpha} x)\|^2 \\
&= \frac{1}{4} \sum_{\alpha^4=1} \alpha \|y + \overline{\alpha} x\|^2 \\
&= \overline{\frac{1}{4} \sum_{\beta^4=1} \beta \|y + \beta x\|^2} \\
&= \overline{\langle y, x \rangle} \\
&= \langle x, y \rangle .
\end{aligned}$$

The space H^* is the dual space of H and must be complete. $\square$

Theorem 14.4. Every Hilbert space is reflexive. In other words, for all Hilbert space H, the canonical embedding $\iota_H : H \to H^{**}$ is an isometry. And $\iota_H = \varphi^{H^*} \circ \varphi^H$.

Proof. Theorem 14.1 gives an conjugate-isometry $\varphi^H : H \to H^*$. We have proved that H^* is a Hilbert space, thus there is an conjugate-isometry φ^{H^*} from H^* to $H^{**} = \big(H^*\big)^*$.

Claim. The natural embedding is just the composition of 2 conjugate-isometries: $\iota_H = \varphi^{H^*} \circ \varphi^H$. In other words, the following diagram commutes

$$\begin{tikzcd}
H \arrow[rr, hook, "\iota_H"] \arrow[dr, "\varphi^H"'] & & H^{**} \\
& H^* \arrow[ur, "\varphi^{H^*}"'] &
\end{tikzcd}$$

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

$$\begin{aligned}
\big[(\varphi^{H^*} \circ \varphi^H)(x)\big](f) &= \big[\varphi^{H^*}\big(\varphi^H(x)\big)\big](f) \\
&= \big\langle f, \varphi^H(x) \big\rangle_{H^*} \\
&= \big\langle \varphi^H_y, \varphi^H_x \big\rangle_{H^*} \\
&= \langle x, y \rangle_H .
\end{aligned}$$

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 14.1.

Answer of Question 14.1. That's true (and this proof is almost the same as the proof of Theorem 14.4). 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 $\varphi\colon X \to X^*$ be the conjugate-isometry, Let $Y := X^*$, and Y^* be the dual space of Y, then both of Y, Y^* are Hilbert spaces. There is an conjugate-isometry

$$\varphi^Y : Y \to Y^*, f \mapsto \varphi^Y_f = \langle\ , f\rangle_{H^*}$$

given by Theorem 14.1. Just like what we do in the proof of Theorem 14.1, we can define an inner product on Y^* such by

$$\forall g, f \in Y : \left\langle \varphi^Y(g), \varphi^Y(f) \right\rangle_{Y^*} := \langle f, g\rangle_Y ,$$

which induces the norm on Y^*. Then consider $\mu := \varphi^Y \circ \varphi\colon X \to Y^*$, which is an isometry. Thus $X \cong Y^*$ and hence X is a Hilbert space. $\square$

Remark 14.4. Let H be a Hilbert space, the conjugate-isometry from H to H^* doesn't need to be unique. In fact, it suffices to show that there are some Hilbert spaces that have at least one non-trivial "auto-isometry". Consider the following examples:

Example 47. The Hilbert space $\ell_2(\mathbb{Z})$, whose elements are all square summable sequence from $\mathbb{Z}$ to $\mathbb{K}$:

$$\ell_2(\mathbb{Z}) := \{x \in {}^{\mathbb{Z}}\mathbb{K}\colon \sum_{n\in\mathbb{Z}} |x_n|^2 < \infty.\}$$

Then the k-shift operator is an "auto-isometry":

$$\tau_k\colon \ell_2(\mathbb{Z}) \to \ell_2(\mathbb{Z}), x \mapsto (\tau_k x\colon n \mapsto x_{n-k}).$$

Therefore, $F := \varphi \circ \lambda_k$ is a conjugate-isometry from $\ell_2(\mathbb{Z})$ to $\big(\ell_2(\mathbb{Z})\big)^*$

$$\begin{tikzcd}
\ell_2(\mathbb{Z}) \arrow[d, "\tau_k"'] \arrow[dr, "F"] & \\
\ell_2(\mathbb{Z}) \arrow[r, "\varphi"] & \big(\ell_2(\mathbb{Z})\big)^*
\end{tikzcd}$$

This example can be generalized to $L_2(G)$, where G is an Abelian group with a suitable measure. $\blacksquare$

14.1.2 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 $\psi\colon H \times K \to \mathbb{K}$ 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

$$\begin{aligned} \psi(kx + ly, z) &= k\psi(x, z) + l\psi(y, z), \\ \psi(x, kz + lw) &= \bar{k}\psi(x, z) + \bar{k}\psi(x, w). \end{aligned}$$

3. a **bounded functional** on $H \times K$, if ψ is bilinear or sesquilinear and the supermum $\sup_{\substack{\|y\| \le 1 \\ \|x\| \le 1}} |\psi(x, y)|$ is finite. Equivalently, there is $M > 0$ such that

$$\forall x \in X, y \in Y\colon |\psi(x, y)| \le M\|x\| \cdot \|y\|.$$

   And denote $\sup_{\substack{\|y\| \le 1 \\ \|x\| \le 1}} |\psi(x, y)|$ by $\|\psi\|$ ($\| \ \|$ is a norm).

Remark 14.5. Inner product is a special sesquilinear functional.

Remark 14.6. For the boundedness of ψ, 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 ψ is bilinear: consider the tensor product of H and K, i.e. a pair $(X \bigotimes Y, \theta)$ satisfying the universal property of tensor product. Then the following diagram commutes:

$$\begin{tikzcd} H \times K \arrow[d, "\theta"'] \arrow[dr, "\psi"] & & (x, y) \arrow[r, mapsto, "\psi"] \arrow[d, mapsto, "\theta"] & \psi(x, y) \arrow[d, equal] \\ H \bigotimes K \arrow[r, dashed, "\Psi"'] & K & x \otimes y \arrow[r, mapsto, "\Psi"'] & \Psi(x \otimes y) \end{tikzcd}$$

Here $H \bigotimes K$ is a linear normed space whose norm satisfies

$$\forall x \in H, y \in K\colon \|x \otimes y\| = \|x\| \cdot \|y\|.$$

Let

$$B_{E\bigotimes F} \supseteq S := \Big\{\theta(x,y) = x\otimes y \in X \bigotimes Y : x \in B_E, y \in B_F\Big\},$$

then

$$\sup_{\substack{\|x\|\le 1\\ \|y\|\le 1}} |\psi(x,y)| = \sup_{\substack{\|x\|\le 1\\ \|y\|\le 1}} \Big|\big(\Psi\circ\theta\big)(x,y)\Big| = \sup_{x\otimes y\in S} |\Psi(x\otimes y)|.$$

And in fact, the convex hull of S is just $B_{E\bigotimes F}$. Thus

$$\sup_{u\in S}|\Psi(u)| = \sup_{u\in B_{E\bigotimes F}} |\Psi(u)| = \|\Psi\|_{(E\bigotimes F)^*}.$$

In other words, the norm ψ is just the norm of the bounded linear functional Ψ. ■

Theorem 14.5 (Representation of Sesquilinear Functional)**.** Let H, K be 2 Hilbert spaces. Then $\psi\colon H\times K \to \mathbb{K}$ 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 $f_x\colon K \to \mathbb{K}, y \mapsto \overline{\psi(x,y)}$. 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. □

Remark 14.7. 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 $\psi(x,\ )$. ■

Theorem 14.6. 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

$$\forall x\in H, y\in K\colon\ \langle Tx, y\rangle_K = \langle x, Sy\rangle_H .$$

Proof. Let
$$\psi\colon K \times H \to \mathbb{K}, (y, x) \mapsto \overline{\langle Tx, y\rangle}.$$
Then ψ is a sesquilinear bounded functional. Apply Theorem 14.5. $\square$

Definition (Adjoint Operator)**.** Let H, K be 2 Hilbert spaces and $T \in \mathcal{B}(H, K)$. The operator S uniquely determined by

$$\forall x \in H, y \in K\colon \langle Tx, y\rangle_K = \langle x, Sy\rangle_H .$$

is called the adjoint operator of T, which will be denoted by T^{h*}.

Remark 14.8. For convenience, we will denote the conjugate operator of T by T^{b*} and the adjoint operator of T by T^*. ■

There is an interesting relation between adjoint and conjugate operators:

Theorem 14.7. 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:

$$\begin{tikzcd} H & K \arrow[l, dashed, "T^*"'] \\ H^* \arrow[from=u, "\varphi^H"'] & K^* \arrow[from=u, "\varphi^K"] \arrow[l, "T^{b*}"] \end{tikzcd}$$

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:

$$\varphi^H \circ T^*(x) = \varphi^H_{T^*x}$$

and hence for all $y \in H$

$$\left[\varphi^H \circ T^*(x)\right](y) = \varphi^H_{T^*x}(y) = \langle y, T^*x\rangle_H = \langle Ty, x\rangle_K .$$

On the other hand:

$$(T^{b*} \circ \varphi^K)(x) = T^{b*}(\varphi^K_x) = \varphi^K_x \circ T$$

and hence

$$[(T^{b*} \circ \varphi^K)(x)](y) = \left[\varphi^K_x \circ T\right](y) = \varphi^K_x(Ty) = \langle Ty, x\rangle_K .$$

Since y is arbitrary, we have $(T^{b*} \circ \varphi^K)(x) = (\varphi^K \circ T^*)(x)$. And x is arbitrary, therefore $T^{b*} \circ \varphi^K = \varphi^H \circ T^*$. $\square$

Remark 14.9. This theorem means that $\{\varphi^H : H \in \mathrm{Ob}(\mathsf{Hil})\}$ is a natural transformation from the Banach functor to the Hilbert functor.

Remark 14.10. We can also define the adjoint operator by this. In other words: both φ^H, φ^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. ■

14.2 Lecture 14-2

This lecture was given by Chen Li.

Notation. In this lecture: We write H, K to be 2 Hilbert spaces unless otherwise specified and the composition of two operators is written TS for short sometimes.

Recall

In the last lecture, we studied the conjugate operator:

Definition. Let H, K be 2 Hilbert spaces and $T \in \mathcal{B}(H, K)$. There is a unique operator $T^* \in \mathcal{B}(K, H)$ satisfies

$$\forall x \in H, y \in K \colon \langle Tx, y\rangle_K = \langle x, T^*y\rangle_H ,$$

which is called the adjoint operator of T.

Naturally we consider the special case: $K = H$ and $T = T^*$:

Definition. An operator $T \in \mathcal{B}(H)$ is said to be a **self-adjoint operator**, if $T^* = T$.

14.2.1 More about adjoint operators

Then we introduce some useful equivalent conditions for an operator to be a self-adjoint operator.

Proposition 14.8. Let $T \in \mathcal{B}(H)$. The following statements are equivalent:

1. it's a self-adjoint operator;
2. the bounded sesquilinear functional $\Psi_T \colon (x, y) \mapsto \langle Tx, y\rangle$ induced by T is Hermitian.

Here a **Hermitian** functional means:

Definition. A sesquilinear functional $\Psi\colon H \times H \to \mathbb{K}$ is said to be Hermitian, if

$$\forall x, y \in H\colon \Psi(x, y) = \overline{\Psi(y, x)}.$$

Furthermore, if $\mathbb{K} = \mathbb{C}$, they are both equivalent to

3. for all $x \in H$, we have $\langle Tx, x\rangle \in \mathbb{R}$.

Proof of Proposition 14.8. We prove in the following order:

$$1 \iff 2 \implies 3 \xRightarrow{\mathbb{K}=\mathbb{C}} 2.$$

$1 \iff 2$: $T = T^*$ if and only if for all $x, y \in H$,

$$\langle Tx, y\rangle = \langle x, Ty\rangle\,,$$

i.e.

$$\Psi_T(x, y) = \overline{\Psi_T(y, x)},$$

which means Ψ_T is Hermitian.

$2 \implies 3$: this doesn't need $\mathbb{K} = \mathbb{C}$. Suppose Ψ_T is Hermitian, then

$$\forall x \in H\colon\ \langle Tx, x\rangle = \Psi_T(x, x) = \overline{\Psi(x, x)} = \overline{\langle x, Tx\rangle},$$

thus $\langle Tx, x\rangle \in \mathbb{R}$

$3 \implies 2$: suppose $\mathbb{K} = \mathbb{C}$. We apply the polar identity (for a seseuilinear functional, since the proof of polar identity only needs the sesquilinearity of inner product): let

$$\mathcal{Q}\colon H \to \mathbb{K}, x \mapsto \Psi_T(x, x) = \langle Tx, x\rangle\,.$$

Notice that, we have $\operatorname{Im}\mathcal{Q} \subseteq \mathbb{R}$ since Ψ_T is Hermitian. Then we have: for all $x, y \in H$:

$$4\Psi_T(x, y) = \sum_{k=0}^{3} i^k \mathcal{Q}(x + i^k y),$$
$$4\Psi_T(y, x) = \sum_{k=0}^{3} i^k \mathcal{Q}(y + i^k x).$$

Checking that they are conjugate to each other is boring and hence omitted. Therefore, we know Ψ_T is Hermitian. $\square$

Now we introduce some examples about adjoint operators.

Example 48. Consider $H = \mathbb{K}^n$ and $K = \mathbb{K}^m$ and $T \in \mathcal{B}(H, K) = \mathcal{B}(\mathbb{K}^n, \mathbb{K}^m)$. Then there is a matrix representation of T with respect to the standard bases:

$$\{\alpha_j \in \mathbb{K}^n \ : \alpha_j = (\underbrace{0, \dots, 0}_{j-1 \text{ items}}, 1, 0, \dots, 0)^t, j \in [n]\},$$
$$\{\beta_k \in \mathbb{K}^n \ : \beta_k = (\underbrace{0, \dots, 0}_{k-1 \text{ items}}, 1, 0, \dots, 0)^t, k \in [m]\}.$$

Let M_T be the corresponding $m \times n$ matrix of T. Similarly T^* corresponds to a matrix M_{T^*}. Then $T\alpha = M_T\alpha$ for all $\alpha \in H$.

Claim. We have $M_{T^*} = M_T^H$, i.e. the matrix of T^* is just the conjugate-transpose of M_T.

Proof. Let $M_T = (a_{j,k})_{k\in[m]}^{j\in[n]}$ and $M_{T^*} = (b_{k,j})_{j\in[n]}^{k\in[m]}$. Then for all $j \in [n]$:

$$T\alpha_j = M_T\alpha_j = \sum_{l=1}^{n} a_{l,j}\beta_j.$$

For all $k \in [m]$:

$$\langle T\alpha_j, \beta_k\rangle = \langle \alpha_j, T^*\beta_k\rangle .$$

Replace T, T^* with M_T, M_{T^*} and we get

$$\sum_{l=1}^{m} a_{l,j} \langle \beta_l, \beta_k\rangle = \sum_{l=1}^{n} \bar{b}_{l,k} \langle \alpha_j, \alpha_k\rangle .$$

Thus $a_{k,j} = \overline{b_{j,k}}$. □

Example 49. Let $E \hookrightarrow H$ be a closed subspace. Consider the inclusion mapping

$$\iota_E \colon E \to H, x \mapsto x.$$

Clearly it's a bounded linear map.

Question 14.2. What's the adjoint operator of ι_E ?

Answer and Proof. The adjoint operator of ι_E is the projection onto E, i.e. $\iota_E^* = P_E \colon H \to E, x \mapsto x_E$ where x_E is the unique element in E such that $x - x_E \perp E$, as defined before.

Now we prove this. For all $x \in E, y \in H$:

$$\langle \iota_E x, y\rangle_H = \langle x, \iota_E^*(y)\rangle_E \implies \langle x, y - \iota_E^*(y)\rangle_E = 0.$$

Thus $E \perp y - \iota_E^*(y)$, i.e. $\iota_E^*(y) = P_E(y)$ and hence $P_E = \iota_E^*$. □

Remark 14.11. This also proves that $P_E^* = \iota_E$. ■

Now we back to the topic about projection and look for some equivalent descriptions of projection.

Theorem 14.9. Let $P \in \mathcal{B}(H)$, then the following statements are equivalent:

1. it's a projection operator composed with the embedding $\iota_E \colon E \to H$, i.e. $P = \iota_E \circ P_E$ where $E \hookrightarrow H$ is a closed subspace and P_E is the projection onto E;
2. it's idempotent and self-adjoint, i.e. $P^2 = P$ and $P^* = P$;
3. it's idempotent and satisfies $\ker P \perp \operatorname{Im} P$;
4. (in the case $\mathbb{K} = \mathbb{C}$) for all $x \in H$: $\langle Px, x \rangle = \|Px\|^2$;
5. there is a closed subspace $E \hookrightarrow H$ such that $\iota_E^* = P \circ \iota_E$.

Proof. Example 49 implies that 1 is equivalent to 5.

1 $\Longrightarrow$ 2. Let $E \hookrightarrow H$ and $P = \iota_E \circ P_E$ be given. Then for all $x \in H$: $P(x) = \iota_E \circ P_E(x) = P_E(x) \in E$, and hence

$$P^2(x) = P(P(x)) = \iota_E \circ P_E(P(x)) = \iota_E(P(x)) = P(x).$$

We proved $P^2 = P$. For all $x, y \in H$: we have $x = P_E x + x_1, y = P_E y + y_1$ where $x_1 \perp E, y_1 \perp E$. Thus

$$\begin{aligned}
\langle Px, y \rangle &= \langle x, y \rangle - \langle x_1, y \rangle \\
&= \langle x, Py \rangle + \langle x, y_1 \rangle - \langle x_1, Py \rangle - \langle x_1, y_1 \rangle \\
&= \langle x, Py \rangle + \langle x_1 + y_1 \rangle + \langle Px, y_1 \rangle - \langle x_1, y_1 \rangle \\
&= \langle x, Py \rangle .
\end{aligned}$$

Therefore, P is self-adjoint.

2 $\Longrightarrow$ 3. Let $x \in \ker P$ and $y \in \operatorname{Im} P$ be given. Suppose $z \in H$ satisfies $y = Pz$, then

$$\langle x, y \rangle = \langle x, Pz \rangle = \langle Px, z \rangle = 0$$

since $Px = 0$. (In fact, P is self adjoint implies $\ker P \perp \operatorname{Im} P$.)

3 $\Longrightarrow$ 1. What we need to do is to find a closed subspace $E \hookrightarrow H$ such that $P = \iota_E \circ P_E$. If $P = P_E$, then $\operatorname{Im} P = E = \ker(\mathrm{id}_H - P)$ and $E^\perp = \ker P$. It's natural to consider a kernel (that is naturally closed). Thus, let $E := \ker(\mathrm{id}_H - P)$, then we want to prove $P = P_E$.

- First, we have $E = \operatorname{Im} P$. Let $y \in \operatorname{Im} P$, then there is some $x \in H$ such that $y = Px = P^2x$ and hence

$$0 = Px - P^2x = P(x - P(x)) = (\mathrm{id}_H - P)(Px) = (\mathrm{id}_H - P)(y),$$

which means $y \in \ker(\mathrm{id}_H - P)$. And we proved $\operatorname{Im} P \subseteq E$. Let $z \in E$, i.e. $z \in \ker(\mathrm{id}_H - P)$, i.e. $z = Pz$. Then $z \in \operatorname{Im} P$. Thus $E \subseteq P$.

- Secondly, we have $P = P_E$. It suffices to prove that for all $x \in H\colon x - Px \perp E$. For all $y \in E$, we know $y \in \operatorname{Im} P$ and hence

$$\langle y, x - Px \rangle = 0,$$

where the equality follows from: $\ker P \perp \operatorname{Im} P$ and

$$P(x - Px) = (P - P^2)(x) = 0 \implies x - Px \in \ker P.$$

2 $\implies$ 4. This doesn't need $\mathbb{K} = \mathbb{C}$. For all $x \in H$, we have

$$\langle Px, x \rangle = \left\langle P^2x, x \right\rangle = \langle Px, P^*x \rangle = \langle Px, Px \rangle = \|Px\|^2.$$

4 $\implies$ 2. Suppose $\mathbb{K} = \mathbb{C}$. Consider the sesquilinear functional

$$\Psi_P\colon H \times H \to \mathbb{C}, (x, y) \mapsto \langle Px, y \rangle\,.$$

Then Proposition 14.8 implies that T is self-adjoint. Since P is self-adjoint, we have

$$\left\langle P^2x, x \right\rangle = \langle Px, Px \rangle = \|Px\|^2,$$

thus $\left\langle (P^2 - P)x, x \right\rangle = 0$. Then consider the **polar identity for a sesquilinear functional** (whose proof is the same as the proof for the case that a sesquilinear functional is an inner product, which only needs the sesquilinearity of an inner product). We have: for all $x, y \in H$:

$$\left\langle (P^2 - P)x, y \right\rangle = \sum_{n=0}^{3} i^n \left\langle (P^2 - P)(x + i^n y), x + i^n y \right\rangle = 0,$$

and hence $P^2 - P = 0$.

Everything has been proved as the following diagram.

$$\begin{tikzcd}
(4) \arrow[r, Leftrightarrow] & (2) \arrow[d, Rightarrow] & (1) \arrow[l, Rightarrow] \arrow[r, Leftrightarrow] & (5) \\
& (3) \arrow[ur, Rightarrow] & &
\end{tikzcd}$$

□

In fact, we have also proved the following result

Corollary 14.10. If $\mathbb{K} = \mathbb{C}$ then $T \in \mathcal{B}(H)$ satisfies that $\forall x \in H\colon \langle Tx, x\rangle = 0$ if and only if $T = 0$.

Notice that: let $\mathrm{pr}(H)$ be the set of all projection operators in H and $\mathrm{cs}(H)$ be the set of all closed subspaces of H, we have proved that there is a bijection

$$\Pi\colon \mathrm{cs}(H) \to \mathrm{pr}(H) \to, E \mapsto P_E,$$

whose inverse can be interpreted as ker, i.e. the inverse of Π maps a projection operator to the kernel of the projection operator. Since there is a natural partial order on $\mathrm{cs}(H)$: $\subseteq$. Then we define a partial order $\leq$ on $\mathrm{pr}(H)$ as follows:

$$P_M \leq P_E \iff M \subseteq E,$$

i.e. let Π keeps the order. The following theorem give some equivalent descriptions of this order.

Theorem 14.11. Let $M, E \in \mathrm{cs}(H)$ and P_M, P_E is their corresponding projection operators (regarded as operators in $\mathcal{B}(H)$). Then the following statements are equivalent:

1. the order relation holds $P_M \leq P_E$;
2. they commute and satisfy: $P_E P_M = P_M P_E = P_M$;
3. the difference is a projection: $P_E - P_M \in \mathrm{pr}(H)$;
4. for all $x \in H$, $\langle P_M x, x\rangle \leq \langle P_E x, x\rangle$;
5. for all $x \in H$, $\|P_M x\| \leq \|P_E x\|$.

Remark 14.12. For the forth, the numbers are real, see Theorem 14.9.

Proof. 1 $\Longrightarrow$ 2: from definition, we know $M \subseteq E$. For all $x \in H$, we have

$$x = P_M x + x_1, x = P_E x + x_2,$$

then $P_E P_M(x) = P_E(P_M x) = P_M x$, since $P_M x \in M \subseteq E$. Then $P_M(P_E x) = P_M(x) + P_M(x_2)$ where $x_2 = x - P_E x \perp E$. Then $x_2 \perp M$ and hence $P_M(x_2) = 0$, which ensures that

$$P_M(P_E x) = P_M(x) + P_M(x_2) = P_M(x).$$

Above all, $P_M P_E = P_E P_M = P_M$.

2 $\implies$ 3: it suffices to prove that $P_E - P_M$ is idempotent and self-adjoint. They are both true, since

$$(P_E - P_M)^2 = P_E^2 - P_E P_M - P_E P_M + P_M^2 = P_E - 2P_M + P_M = P_E - P_M.,$$

and $(P_E - P_M)^* = P_E^* - P_M^* = P_E - P_M$, as we wanted.

3 $\implies$ 4: suppose $P_E - P_M = P_N$, where $N \in \text{cs}(H)$, then for all $x \in H$:

$$\langle P_E x, x\rangle - \langle P_M x, x\rangle = \langle P_E x - P_M x, x\rangle = \langle P_N x, x\rangle = \|P_N x\|^2 \geq 0.$$

4 $\implies$ 5: this follows from 4 of Theorem 14.9

5 $\implies$ 1: consider

$$P_M \leq P_E \iff M \subseteq E \iff M^\perp \supseteq E^\perp.$$

For all $x \in E^\perp$, we have $P_E x = 0$ and hence $\|P_M x\| \leq \|P_E x\| = 0$, thus $P_M x = 0$, i.e. $x \in E^\perp$. Therefore, we know $M^\perp \supseteq E^\perp$. $\square$

From the proof above, we find that there is some closed suspace such that $P_E - P_M = P_N$, and it's natural to ask the relation between E, M, N.

Theorem 14.12. Then the following statements are equivalent:

1. two operators are orthogonal: $P_M P_N = 0$;

2. two operators are orthogonal: $P_N P_M = 0$;

3. the addition of two operators is a projection, i.e. $P_M + P_N$ is a projection;

4. the two subspaces are orthogonal, i.e. $M \perp N$.

Proof. 1 $\iff$ 2: suppose 1 holds. We start from inner product since it enjoys good properties; for all $x \in H$:

$$\begin{aligned} \langle P_N P_M x, P_N P_M x\rangle &= \left\langle P_M x, P_N^2 P_M x\right\rangle \\ &= \langle P_M x, P_N P_M x\rangle \\ &= \langle x, (P_M P_N) P_M x\rangle \\ &= \langle x, 0\rangle \\ &= 0 \end{aligned}$$

thus $P_N P_M x = 0 (\forall x \in H)$, i.e. $P_N P_M = 0$. By the symmetry, we have also $P_N P_M = 0 \implies P_M P_N = 0$.

2 $\implies$ 3: clearly $P_M + P_N$ is self-adjoint. And

$$(P_M + P_N)^2 = P_M^2 + P_M P_N + P_N P_M + P_N^2 = P_M + P_N,$$

then $P_M + P_N$ is a projection.

3 $\implies$ 4: let $x \in M$, $y \in N$ be given. Suppose $P_E = P_M + P_N$, then $P_E - P_M = P_N$, and Theorem 14.11 implies $M \subseteq E$. For all $x \in M$, then $P_E x = P_M x$; for all $y \in N$, $P_N y = y$. Then for all $x \in M$, $y \in N$:

$$\begin{aligned} 2\langle x, y\rangle &= \langle P_M x, y\rangle + \langle x, P_N y\rangle \\ &= \langle P_M x, y\rangle + \langle P_N x, y\rangle \\ &= \langle P_E x, y\rangle \\ &= \langle x, y\rangle, \end{aligned}$$

thus $\langle x, y\rangle = 0$. Since x, y are arbitrary, we have $M \perp N$.

4 $\implies$ 1: for all $x, y \in H$, we have

$$\langle P_M P_N x, y\rangle = \langle P_N x, P_M y\rangle = 0$$

since $M \perp N$. Thus $P_M P_N = 0$. □

Here we talk something about algebra. From now on, view H as a $\mathbb{K}$-module.

Definition (Module)**.** Let R be a commutative ring. An ***R*-module** is an abelian group M with addition operation $+$ and a map $R \times M \to M, (r, m) \mapsto rm$ that satisfies the following axioms; for all $m, m_1, m_2 \in M$ and $r, r_1, r_2 \in R$:

1. Identity Law (if R has an identity 1): $1m = m$;
2. Distributive Law $\sharp$1: $(r_1 + r_2)m = r_1 m + r_2 m$;
3. Distributive Law $\sharp$2: $r(m_1 + m_2) = rm_1 + rm_2$;
4. Associative Law: $(r_1 r_2)m = r_1(r_2 m)$.

Remark 14.13. An R-module is a vector space over the ring R. ■

Then a vector space over $\mathbb{K}$ is naturally an $\mathbb{K}$-module.

Definition. A **chain** of R-modules is a sequence of R-modules and R-module homomorphisms

$$M_0 \xrightarrow{f_1} M_1 \xrightarrow{f_2} M_2 \xrightarrow{f_3} \cdots \xrightarrow{f_n} M_n. \tag{34}$$

An **exact sequence** is a chain (34) such that

$$\forall k \in [n-1]\colon \ \ker f_{k+1} = \operatorname{Im} f_k.$$

A **short exact sequence** is a five-term exact sequence in which the edge modules are 0; i.e. an exact sequence of the form, where the arrow with domain/codomain being 0 is just 0:

$$0 \longrightarrow M_1 \xrightarrow{\ f\ } M_2 \xrightarrow{\ g\ } M_3 \longrightarrow 0. \tag{35}$$

Thus a chain of the form (35) is a short exact sequence if and only if the four conditions are satisfied:

1. f is injective;
2. g is surjective;
3. $g \circ f = 0$;
4. $g(n) = 0 \implies \exists m\colon f(m) = n$.

Claim. Suppose $M, N \hookrightarrow H$ satisfy $M \perp N$. Then the following chain is a short exact sequence

$$0 \longrightarrow M \xrightarrow{\ \iota_M\ } M \oplus N \xrightarrow{\ P_N\ } N \longrightarrow 0. \tag{36}$$

Proof. By definitions, we know ι_M is injective and P_N is surjective, thus it suffices to show $P_N \circ \iota_M = 0$ and $P_N(y) = 0 \implies \exists x\colon \iota_M(x) = y$.

1. $P_N \circ \iota_M = 0$: for all $x \in M$, we have

   $$P_N \circ \iota_M(x) = P_N\big(\iota_M(x)\big) = P_N(x) = 0,$$

   since $x \in M$ and $M \perp N$ implies $P_N(x) = 0$.

2. $P_N(y) = 0 \implies \exists x\colon \iota_M(x) = y$: suppose $P_N(y) = 0$ for some $y \in M \oplus N$, then $y \in M \hookrightarrow M \oplus N$ and hence $M \ni y = \iota_M(y)$. To see that $y \in M$: suppose $y \in M \oplus N$ satisfies $P_N(y) = 0$, then $y \in M \oplus N$ implies

   $$y = y' + z \text{ for some } y' \in M, z \in N.$$

   Now $z \in N$ implies $P_N(z) = z$, while $P_N(y) = P_N(y') + P_N(z)$, thus

   $$0 = P_N(y') + z, \text{ i.e. } -z = P_N(y).$$

   But $M \cap N = \{0\}$ since $M \perp N$, then $z = 0$, $P_N(y) = 0$ and hence $y = y' \in M$. □

Furthermore, we can apply the covariant functor $\mathcal{B}(H, \ )$ on the chain (35) to get another chain:

$$0 \longrightarrow \mathcal{B}(H,M) \xrightarrow{(\iota_M)_*} \mathcal{B}(H, M \oplus N) \xrightarrow{(P_N)_*} \mathcal{B}(H,N) \longrightarrow 0$$

where $f_* := \mathcal{B}(H,f)$ is defined as follows for each $f \in \mathcal{B}(A,B)$:

$$f_* \colon \mathcal{B}(H,A) \to \mathcal{B}(H,B), g \mapsto f \circ g.$$

Theorem 14.13. The chain

$$0 \longrightarrow \mathcal{B}(H,M) \xrightarrow{(\iota_M)_*} \mathcal{B}(H, M \oplus N) \xrightarrow{(P_N)_*} \mathcal{B}(H,N) \tag{37}$$

is exact.

Proof. We should prove 3 points: $(\iota_M)_*$ is injective, $(P_N)_* \circ (\iota_M)_* = 0$ and $(P_N)_*(g) = 0 \implies \exists f \in \mathcal{B}(H,M) \colon (\iota_M)_*(f) = g$.

1. $(\iota_M)_*$ is injective; i.e. $\ker(\iota_M)_* = 0$. Let $f \in \ker(\iota_M)_*$, equivalently, $(\iota_M)_*(f) = \iota \circ f = 0$. This implies $f = 0$ and thus $\ker(\iota_M)_* = 0$.

2. $(P_N)_* \circ (\iota_M)_* = 0$: this is implied by the functor property, but we also check it: let $f \in \mathcal{B}(H,M)$, then

$$\begin{aligned}
\big[(P_N)_* \circ (\iota_M)_*\big](f) &= (P_N)_*\big[(\iota_M)_*(f)\big] \\
&= (P_N)_*(\iota_M \circ f) \\
&= P_N \circ \iota_M \circ f \\
&= 0 \circ f \\
&= 0,
\end{aligned}$$

   since $P_N \circ \iota_M = 0$, which follows from the fact that (36) is exact.

3. $(P_N)_*(g) = 0 \implies \exists f \in \mathcal{B}(H,M) \colon (\iota_M)_*(f) = g$: from $(P_N)_*(g) = 0$, we know that $\operatorname{Im} g \perp N$, while $\operatorname{Im} g \subseteq M \oplus N$, where $M \perp N$, thus $\operatorname{Im} g \subseteq M$. Then we have an element

$$f := g|^M \colon H \to M, x \mapsto g(x)$$

   that lies in $\mathcal{B}(H,M)$ satisfying $g = (\iota_M)_*(f)$. □

Remark 14.14. This theorem means that the functor $\mathcal{B}(H, \ )$ is left-exact. ■

Definition. Two projection operators P_M, P_E are said to be **orthogonal**, if $P_M P_N = 0$ or $P_N P_M = 0$.

Finally, we prove that the family $\mathrm{pr}(H) \subseteq \mathcal{B}(H)$ is closed. In general, this family is not closed for addition, thus it has no algebraic structure.

Lemma 14.14. If $(P_n)_{n\in\mathbb{N}} \subseteq \mathrm{pr}(H)$ and $P_n \xrightarrow{s} P(n \to \infty)$, then $P \in \mathrm{pr}(H)$.

Remark 14.15. Since convergence in norm implies strong convergence, this implies that $\mathrm{pr}(H)$ is closed in $\mathcal{B}(H)$.

Proof of Lemma 14.14. We prove that P is bounded, satisfying $P^2 = P$ and $P^* = P$.

It is bounded: see Theorem 8.1 and apply the fact that $\|P_E\| \leq 1$ for all $E \in \mathrm{cs}(H)$.

It is self-adjoint: for all $x, y \in H$, we have

$$\begin{aligned} \langle Px, y\rangle &= \left\langle \lim_n P_n x, y\right\rangle \\ &= \lim_n \langle P_n x, y\rangle \\ &= \lim_n \langle x, P_n y\rangle \\ &= \left\langle x, \lim_n P_n y\right\rangle \\ &= \langle x, Py\rangle . \end{aligned}$$

It is idempotent: for all $x \in H$, from $P_n^2 = P_n$, we have

$$\begin{aligned} \left\|(P^2 - P)x\right\| &= \left\|(P^2 - P_nP + P_nP - P_n^2 + P_n - P)x\right\| \\ &\leq \|(P - P_n)(Px)\| + \|P_n\|\|(P - P_n)x\| + \|(P_n - P)x\| \\ &\leq \|(P - P_n)(Px)\| + \|(P - P_n)x\| + \|(P_n - P)x\| \\ &\to 0(n \to \infty), \end{aligned}$$

since $P_n \xrightarrow{s} P$. Thus $(P^2 - P)x = 0$ for all x, i.e. P is idempotent. $\square$

Theorem 14.15. Given H,

1. if $(Q_n)_{n\in\mathbb{N}} \subseteq \Pi(H)$ is orthogonal, then $\exists! P \in \Pi(H)$ such that $\sum_{j=1}^n Q_j \xrightarrow{s} P$ as $n \to \infty$.
2. if $(E_n)_{n\in\mathbb{N}} \subseteq \mathrm{cs}(H)$ is non-decreasing and $E = \overline{\bigcup_{n\geq 1} E_n}$, then $P_{E_n} \xrightarrow{s} P_E$.

Proof. 1: define $P_n = \sum_{j=1}^{n} Q_j$ for all $n \in \mathbb{N}$. Then Theorem 14.12 ensures that P_n is a projection operator, as you can prove by induction. For all $x \in H$, for all $n \in \mathbb{N}$, apply Theorem 29, we have

$$\|x\|^2 \geq \|P_n x\|^2 = \sum_{j=1}^{n} \|Q_j x\|^2.$$

Thus $\sum_{j=1}^{\infty} \|Q_j x\|^2 < \infty$ and hence for all $m \geq n$

$$\|(P_m - P_n)x\|^2 = \sum_{j=n+1}^{n} \|Q_j x\|^2 \to 0(m, n \to \infty).$$

Now H is complete, we can define

$$P \colon H \to H, x \mapsto \lim_n \sum_{j=1}^{n} Q_j x$$

and Lemma 14.14 ensures that $P \in \Pi(H)$. The uniqueness of P is just the uniqueness of strong limit.

2: Theorem 14.11 implies that $P_{E_k} - P_{E_j} \in \Pi(H)$ and $P_{E_k} P_{E_j} = P_{E_k} P_{E_j} = P_{E_j}$ for all $k \geq j$. For all $j < k$:

$$\begin{aligned}
&(P_{E_{j+1}} - P_{E_j})(P_{E_{k+1}} - P_{E_k}) \\
&= P_{E_{j+1}} P_{E_{k+1}} - P_{E_j} P_{E_{k+1}} - P_{E_{j+1}} P_{E_k} + P_{E_j} P_{E_k} \\
&= P_{E_{j+1}} - P_{E_j} - P_{E_j} + P_{E_j} \\
&= 0,
\end{aligned}$$

and

$$P_{E_1}(P_{E_k} - P_{E_j}) = P_{E_1} - P_{E_1} = 0.$$

Then we have an orthogonal projection sequence as following

$$P_{E_1}, P_{E_2} - P_{E_1}, \ldots, P_{E_{j+1}} - P_{E_j}, \ldots$$

Apply 1 to this sequence and we get an operator P such that $P_{E_n} \to P$. Now we prove $P = P_E$. Suppose $P = P_F$ for some $F \in \mathrm{cs}(H)$ and we need to prove $F = E$. For all $n \geq 1$, $E_n \subseteq E$, Theorem 14.11 implies that

$$\forall x \in H \colon \|P_{E_n} x\| \leq \|P_E x\|.$$

Then

$$\forall x \in H \colon \|P_F x\| = \lim_n \|P_{E_n} x\| \leq \|P_E x\|,$$

i.e. $F \subseteq E$. On the other hand, taking $x \in E_n$, when $k \geq n$, we have $E_k \supseteq E_n$ and $P_{E_k}x = x$. Thus

$$P_F(x) = \lim_k P_k(x) = x \in F.$$

Since $n \in \mathbb{N}$ is arbitrary, we have $\bigcup_{n\geq 1} E_n \subseteq F$. Since F is closed, we have $F \supseteq E$. $\square$

15 Week 15

15.1 Lecture 15-1

This lecture was given by Deyu Yu.

In this lecture, we assume that H, K are Hilbert spaces over $\mathbb{C}$.

In this lecture, an isometry means an operator that satisfies

$$\forall x\colon \|ux\| = \|x\|.$$

It's not needed to be a bijection! If an operator is isometric and bijective, we will say it's an isometric isomorphism in this lecture.

15.1.1 Normal Operator

Definition. An operator is said to be normal, if it satisfies the conditions in Theorem 15.1.

Theorem 15.1. Let $u \in \mathcal{B}(H)$. The following statements are equivalent:

1. they commute: $u^*u = uu^*$;
2. for all $x \in H$, we have $\|ux\| = \|u^*x\|$, i.e. $\| \ \| \circ u = \| \ \| \circ u^*$.

Proof of Theorem 15.1. For all $x \in H$:

$$\|ux\|^2 = \langle ux, ux \rangle = \langle x, u^*ux \rangle,$$

and

$$\|u^*x\|^2 = \langle u^*x, u^*x \rangle = \langle x, uu^*x \rangle,$$

since $(u^*)^* = u$. Thus $\|ux\| = \|u^*x\|$ for all $x \in H$ if and only if

$$\langle (u^*u - uu^*)x, x \rangle = 0$$

holds for all $x \in H$. Corollary 14.10 implies the result. □

Example 50. Consider $H = K = \mathbb{C}^n$ and $T \in \mathcal{B}(\mathbb{C}^n)$. Then T is normal if and only if the matrix M_T satisfies $M_T M_T^H = M_T^H M_T$, i.e. M_T is a normal matrix.

Example 51. Let $I \subseteq \mathbb{R}$ be an interval (we don't care the boundedness of this interval) and an essentially bounded (with respect to Lebesgue measure) map $f\colon \to \mathbb{C}$. Let $L_2 = L_2(I)$ and $L_\infty = L_\infty(I)$. The map

$$u\colon L_\infty \to \mathcal{B}(L_2), f \mapsto u_f,$$

is linear and bounded, where

$$u_f \colon L_2 \to L_2, g \mapsto fg.$$

Clearly u_f is linear and satisfies $\|u_f\| \leq \|f\|_\infty$ for all $f \in L_\infty$. If $0 < a < \|f\|$, there is a subset $A \subseteq I$ such that $\infty > m(A) > 0$ and for all $x \in A$: $|f(x)| > a$. Then

$$u_f(\chi_A) = f \cdot \chi_A$$

and hence

$$\|u_f(\chi_A)\|^2 = \int_I |f\chi_A|^2 \, \mathrm{d}m = \int_A |f|^2 \, \mathrm{d}m > a^2 \|\chi_A\|^2.$$

Therefore, $\|u_f\| > a$ for all $a < \|f\|_\infty$ and hence $\|u_f\| = \|f\|_\infty$. Thus u is an isometry from L_∞ to a subspace of $\mathcal{B}(L_2)$.

Now we go back to talk about normal operators.

Claim. For all $f \in L_\infty$, the operator induced by f, i.e. u_f is normal.

Proof. For all $g, h \in L_2$:

$$\langle u_f g, h \rangle = \int_I u_f(g)\bar{h} \, \mathrm{d}m = \int_I fg\bar{h} \, \mathrm{d}m.$$

Similarly, for $\bar{f} \in L_\infty$ we have

$$\langle g, u_{\bar{f}} h \rangle = \int_I g\overline{u_{\bar{f}}(h)} \, \mathrm{d}m = \int_I gf\bar{h} \, \mathrm{d}m.$$

Thus we know $u_f^* = u_{\bar{f}}$. It's easy to see that $|f|^2 \in L_\infty$ and

$$u_f u_{\bar{f}} = u_{\bar{f}} u_f = u_{|f|^2},$$

i.e. $u_f u_f^* = u_f^* u_f$. □

In fact, for all $f_1, f_2 \in L_\infty$, we have $u_{f_1} u_{f_2} = u_{f_2} u_{f_1} = u_{f_1 f_2}$(algebra homomorphism). ■

Remark 15.1. Normal operators have the following properties.

1. An operator u is normal if and only if it's adjoint operator is normal, since $u^{**} = u$;
2. If an operator u is normal, then for all polynomial $p(t) \in \mathbb{C}[t]$, the operator $p(u)$ defined by (39) is also normal. ■

Theorem 15.2. Let $u \in \mathcal{B}(H)$ be a normal operator. For all $\lambda \in \mathbb{C}$, $x \in H$, we have

$$u(x) = \lambda x \iff u^*(x) = \bar{\lambda}x.$$

Proof. From the definition of kernel: $u(x) = \lambda x$ if and only if $x \in \ker(u - \lambda\,\mathrm{id}_H)$. Since $(u - \lambda\,\mathrm{id}_H)^* = u - \bar{\lambda}\,\mathrm{id}_H$ and $x \in \ker(u - \lambda\,\mathrm{id}_H)$ if and only if $\|(u - \lambda\,\mathrm{id}_H)x\| = 0$, Theorem 15.1 implies that $\|(u - \lambda\,\mathrm{id}_H)x\| = 0$ if and only if $\left\|(u^* - \bar{\lambda}\,\mathrm{id}_H)x\right\| = 0$. □

Definition (Characteristic Subspace)**.** Let $u \in \mathcal{B}(H)$ and $\lambda \in \mathbb{C}$. The subspace

$$E_\lambda := \ker(u - \lambda\,\mathrm{id}_H)$$

is called the **characteristic subspace** of u with respect to λ.

Theorem 15.3. Let $u \in \mathcal{B}(H)$ be a normal operator and $\lambda \in \mathbb{C}$, $\mu \in \mathbb{C}$ such that $\lambda \neq \mu$. Then we have

$$E_\lambda \perp E_\mu.$$

Proof. Let $x \in E_\lambda$ and $y \in E_\mu$ be given. Then

$$\lambda \langle x, y\rangle = \langle \lambda x, y\rangle = \langle ux, y\rangle = \langle x, u^* y\rangle.$$

And Theorem 15.2 implies that $u^* y = \bar{\mu} y$, and hence

$$\lambda \langle x, y\rangle = \mu \langle x, y\rangle \implies \langle x, y\rangle = 0.$$ □

Corollary 15.4. Let $E := \sum_{\lambda\in\mathbb{C}} E_\lambda \hookrightarrow H$. Then $K := \overline{E}$ is a Hilbert space and we have

1. direct sum decomposition: $K \cong \bigoplus_{\lambda\in\mathbb{C}} E_\lambda$, where the direct sum means

$$\bigoplus_{\lambda\in\mathbb{C}} E_\lambda := \{(x^\lambda)_{\lambda\in\mathbb{C}} \in \prod_{\lambda\in\mathbb{C}} E_\lambda \mid x^\lambda \in E_\lambda, \sum_{\lambda\in\mathbb{C}} \left\|x^\lambda\right\|^2 < \infty\};$$

2. the operator $u|_{K^\perp}$ has no eignevalues.

Proof. The second is easy: if $u|_{K^\perp}$ has an eigenvector $v \in K^\perp$ with respect to the eigenvalue λ, then $v \in K \bigcap K^\perp$, thus $v = 0$. That's a contradiction.

There is a natural map

$$\psi\colon \bigoplus_{\lambda\in\mathbb{C}} E_\lambda \to K, (x^\lambda)_{\lambda\in\mathbb{C}} \mapsto \sum_{\lambda\in\mathbb{C}} x^\lambda,$$

where $\sum_{\lambda\in\mathbb{C}} x_\lambda$ is defined as a limit of net. The existence of limit is guaranteed by

$$\sum_{\lambda\in\mathbb{C}} \left\|x^\lambda\right\|^2 < \infty \tag{38}$$

and Question 16.1. Equation (38) should be proved by taking all finite sets of $\mathbb{C}$ and apply Theorem 29. Clearly ψ is an isometric isomorphism. $\square$

Theorem 15.5. Let $u \in \mathcal{B}(H)$ be a normal operator. Then we have

$$\ker u = (\operatorname{Im} u)^\perp, \overline{\operatorname{Im} u} = (\ker u)^\perp.$$

This follows from the following Theorem.

Theorem 15.6. Let $u \in \mathcal{B}(H, K)$. We have

$$\ker u = (\operatorname{Im} u^*)^\perp, \overline{\operatorname{Im} u} = (\ker u^*)^\perp.$$

Proof of Theorem 15.6. Since

$$\begin{aligned} x \in \ker u &\iff u(x) = 0 \in K \\ &\iff \langle u(x), y\rangle_K = 0 \forall y \in K \\ &\iff \langle x, u^*(y)\rangle_H = 0 \forall y \in K \\ &\iff x \in (\operatorname{Im} u^*)^\perp. \end{aligned}$$

Replace u by u^*, then the duality (32) and $u^{**} = u$ implies the second equality. $\square$

Here we introduce some equivalent conditions of isometry.

Theorem 15.7. Let $u\colon H \to K$ be a bounded operator. The following statements are equivalent.

1. it's an isometry;
2. it satisfies: $\langle ux, uy\rangle_K = \langle x, y\rangle_H$ for all $x, y \in H$;
3. it has a left inverse u^*, i.e. $u^*u = \mathrm{id}_H$.

Proof. $1 \implies 2$: apply the polar identity.

$2 \implies 3$: for all $x, y \in H$, we have

$$\langle x, y\rangle = \langle ux, uy\rangle = \langle x, u^*uy\rangle.$$

This implies $y = u^*uy$ for all $y \in H$, i.e. $u^*u = \mathrm{id}_H$.

$3 \implies 1$: for all $x \in H$, we have $u^*ux = x$ and then

$$\|x\|^2 = \langle x, x\rangle = \langle x, u^*ux\rangle = \langle ux, ux\rangle = \|ux\|^2. \qquad \square$$

Remark 15.2. If $\dim H = n \in \mathbb{N}$, then $u \in \mathcal{B}(H)$ is an isometry implies that u is unitary. If $\dim H = \infty$, we will see that things are different , see Example 52. ■

Definition. An operator $u\colon H \to K$ is called a **unitary** operator if it's a bijection with inverse $u^{-1} = u^*$. Here we say an operator v is an **inverse** of u, if $vu = \mathrm{id}_H$ and $uv = \mathrm{id}_K$.

Theorem 15.8. Let $u\colon H \to K$ be a bounded operator. The following statements are equivalent.

1. the operator u is unitary;
2. it's a surjective isometry;
3. it satisfies $uu^* = \mathrm{id}_K$, and $u^*u = \mathrm{id}_H$;
4. it's an isometric isomorphism.

Proof. We will prove in the following order:

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

$1 \Longrightarrow 4$: suppose $u^{-1} = u^*$ and it suffices to show that u is isometric. For all $x \in H$:

$$\|x\|^2 = \langle x, x\rangle = \langle x, u^*ux\rangle = \langle ux, ux\rangle = \|ux\|^2,$$

since $u^{-1} = u^*$ implies $u^*ux = x$.

$4 \Longrightarrow 3$: suppose u is an isometric isomorphism. Then Theorem 15.7 implies $u^*u = \mathrm{id}_H$. Therefore, we know $uu^* = \mathrm{id}_K$ and $u^*u = \mathrm{id}_H$ since the inverse map of a bijection is unique, and its left inverse must be the inverse.

$3 \Longrightarrow 2$: clearly u is surjective. It's also isometric by Theorem 15.7.

$2 \Longrightarrow 1$: this also follows from Theorem 15.7. □

Example 52. Consider the space of square-summable sequence ℓ_2 and an operator u is defined as

$$u\colon \ell_2 \to \ell_2, (x_n)_{n\in\mathbb{N}} \mapsto (0, x_1, x_2, \ldots),$$

i.e. the right-shifting operator. Then u is isometric but not unitary, since $\operatorname{Im} u \neq \ell_2$. ■

The following example describes some unitary operators.

Example 53. Consider the measure space $(I, \mathcal{P}(I), \mu)$ where μ is the counting measure on I. Let $\alpha = (\alpha_i)_{i\in I} \in \ell_\infty(I)$. There is an operator

$$u_\alpha \colon \ell_2(I) \to \ell_2(I), x = (x_i)_{i\in I} \mapsto \alpha x = (\alpha_i x_i)_{i\in I}.$$

We will prove some properties of u_α.

Proposition. The following statements are equivalent.

1. the operator u_α is unitary;
2. the operator u_α is isometric;
3. for all $i \in I$, we have $|\alpha_i| = 1$.

Proof. $1 \implies 2$: a unitary operator must be isometric.

$2 \implies 3$: for all $i \in I$, take the characteristic function $\chi_{\{i\}} \in \ell_2(I)$ and we see $|\alpha_i| = 1$.

$3 \implies 2$: it follows from the fact that $\forall \beta_1, \beta_2 \in \ell_\infty(I)$, $u_{\beta_1\beta_2} = u_{\beta_1} u_{\beta_2}$ and the claim:

Claim. For all $\alpha \in \ell_\infty(I)$, the adjoint operator of u_α is just $u_{\bar{\alpha}}$.

Taking an arbitrary element $i, j \in I$, we see

$$\left\langle u_\alpha \chi_{\{i\}}, \chi_{\{j\}} \right\rangle = \alpha_i \delta_j^i = \left\langle \chi_{\{i\}}, u_{\bar{\alpha}} \chi_{\{j\}} \right\rangle,$$

and by definition of adjoint operators

$$\left\langle u_\alpha \chi_{\{i\}}, \chi_{\{j\}} \right\rangle = \left\langle \chi_{\{i\}}, u_\alpha^* \chi_{\{j\}}, \right\rangle.$$

We proved that $\left\langle \chi_{\{i\}}, u_{\bar{\alpha}} \chi_{\{j\}} \right\rangle = \left\langle \chi_{\{i\}}, u_\alpha^* \chi_{\{j\}}, \right\rangle$. Now the density of $\operatorname{span}\{\chi_{\{i\}} : i \in I\}$ implies that

$$\langle f, u_\alpha^* g \rangle = \langle f, u_{\bar{\alpha}} g \rangle$$

for all $f, g \in \ell_2(I)$, and hence $u_\alpha^* = u_{\bar{\alpha}}$. □

15.1.2 Partial Isometric Operator

Definition. Let $u \in \mathcal{B}(H, K)$. The space $(\ker u)^\perp$ is called the **support subspace** of u, denoted by $\operatorname{supp} u$.

Remark 15.3. This is a generalization of the support of a function $f \colon X \to \mathbb{C}$:

$$\operatorname{supp} f := \overline{\{x \in X \mid f(x) \neq 0\}},$$

where X is a topological space.

Definition. An operator $u \in \mathcal{B}(H, K)$ is called a **partial isometric operator**, if $v := u|_{\operatorname{supp} u}$ is isometric.

Remark 15.4. We have the following properties.

1. for all $u \in \mathcal{B}(H, K)$: the support subspace of u^* is $\operatorname{supp} u^* = (\ker u^*)^\perp = \overline{\operatorname{Im} u}$;
2. for all $u \in \mathcal{B}(H, K)$: $v := u|_{\operatorname{supp} u}$ is injective, since $\ker v = \ker u \cap \operatorname{supp} u = \{0\}$;
3. If u be a partial isometric operator, then the operator $v = u|_{\operatorname{supp} u}$ is an isometric isomorphism from $\operatorname{supp} u$ to $\operatorname{supp} u^*$;

Proof of 3. It suffices to show that $\operatorname{Im} u = \operatorname{Im} v$. For $y \in \operatorname{Im} u$, there is some $x \in H$ such that $u(x) = y$. Then the projection of x upon $\operatorname{supp} u$, denoted by $x_{\operatorname{supp} u}$ satisfies $v(x_{\operatorname{supp} u}) = u(x_{\operatorname{supp} u}) = u(x)$. Thus $\operatorname{Im} v = \operatorname{Im} u$. Then v is an isometric isomorphism from $\operatorname{supp} u$ to $\operatorname{Im} u$, which implies that $\operatorname{Im} u$ is closed and hence $\operatorname{Im} u = \operatorname{supp} u^*$. □

Theorem 15.9. Let u be a partially isometric operator. Then

1. the composition is $u^* u = P_{\operatorname{supp} u}$;
2. the inverse of the unitary operator $u|_{\operatorname{supp} u} : \operatorname{supp} u \to \operatorname{Im} u$ is $u^*|_{\operatorname{Im} u} : \operatorname{Im} u \to \operatorname{supp} u$;
3. the adjoint operator u^* is a partially isometric operator.

Exercise 15.1. Let H be a Hilbert space over $\mathbb{C}$. Then $\|T\| = 0$ if and only if $\langle Tx, x \rangle = 0$ holds for all $x \in H$.

Exercise 15.2. Let $T \in \mathcal{B}(H)$, prove that: $A + A^* = 0$ if and only if $\operatorname{Re} \langle Tx, x \rangle = 0$ holds for all $x \in H$.

Hint: apply Corollary 14.10.

15.1.3 Look Back

Here we overview what we have learnt about Hilbert spaces.

First, we studied "**inner product space**". **Cauchy-Schwarz inequality** ensures that an inner product induces a norm and hence we can consider the dual space of an inner product space. Theorem 12.4 is of great importance and you should also know the version for a sesquilinear functional. **Parallelogram Law** was used to prove the projection theorem. Theorem 12.6 was used to prove that the norm of H^* is induced by an inner product (Theorem 12.4 was also used).

Secondly, a **Hilbert space** is an inner product space that is complete. Things about **orthogonality** were discussed, such as Bessel's Inequality, Parseval's Identity and Schmidt orthogonalization progress. We defined the **projection operator** for a closed subspace of a Hilbert space. Theorem 13.11 implies the **decomposition**: $H = E \oplus E^{\perp}$ where $E \hookrightarrow H$ is closed.

Finally, **Riesz Representation Theorem** came and so was **adjoint operator**.

15.2 Lecture 15-2

Final exam doesn't need this lecture. Let H be a Hilbert space and X be a Banach space, in this lecture.

15.2.1 Introduction to Spectrum Theory

Given a Banach space $\mathcal{A}$, there is some operations: addition $+$, scalar-multiplication $\cdot$ and norm $\| \ \|$.

Example 54. Consider the Banach space $\mathcal{A} = L_p(\Omega)$, where we have a measure space (Ω, Σ, μ) and $p \in [1, \infty]$.

What if we consider more operations on $\mathcal{A}$? Consider a map, called the multiplication on $\mathcal{A}$, defined as follows:

$$\cdot\colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}, (x, y) \mapsto x \cdot y,$$

which is bilinear and associative, i.e. for all $x, y, z \in \mathcal{A}$ and $\alpha \in \mathbb{K}$:

$$\begin{aligned}
(x+y) \cdot z &= x \cdot z + y \cdot z, \\
z \cdot (x+y) &= z \cdot x + z \cdot y, \\
\alpha(x \cdot y) &= (\alpha x) \cdot y = x \cdot (\alpha y), \\
(x \cdot y) \cdot z &= x \cdot (y \cdot z).
\end{aligned}$$

We also write $x \cdot y = xy$ for simplicity. It's natural to ask the multiplication is continuous (for more good results, we want it to be contractive), i.e.

$$\forall x, y \in \mathcal{A}\colon \|x \cdot y\| \leq \|x\| \|y\|.$$

The definition of a unity is natural. That's how we define an algebra.

Definition (Banach Algebra)**.** A **Banach Algebra** is a Banach space $\mathcal{A}$ with a continuous multiplication. For simplicity, a Banach algebra is also called an algebra.

Definition (Unity)**.** Let $\mathcal{A}$ be an algebra. If there is an element $\mathbb{1}_{\mathcal{A}} \in \mathcal{A}$ such that $\mathbb{1}_{\mathcal{A}} x = x \mathbb{1}_{\mathcal{A}} = x$ holds for all $x \in \mathcal{A}$, then $\mathbb{1}_{\mathcal{A}}$ is called the **unity** of $\mathcal{A}$. An algebra that has a unity is called a **unital algebra**.

Remark 15.5. Clearly, if an algebra has a unity, then it's unique.

Definition. An algebra is said to be **commutative**, if for all $x, y \in \mathcal{A}$: $xy = yx$.

What if there is one more map on this algebra?

Definition (Involution)**.** Let $\mathcal{A}$ be an algebra. An **involution** on $\mathcal{A}$ is a map

$$*\colon \mathcal{A} \to \mathcal{A}, x \mapsto x^*$$

that is conjugate-linear and satisfies

$$\forall x, y \in \mathcal{A}\colon (xy)^* = y^* x^*, \|xx^*\| = \|x\|^2, x^{**} = x,$$

where $x^{**} := (x^*)^*$. For $x \in \mathcal{A}$, x^* is also called the **adjoint** to x.

Remark 15.6. This is a generalization of adjoint in Hilbert spaces. We can define self-adjoint element in $\mathcal{A}$ as what we did for Hilbert spaces.

Definition. An algebra with an involution is called a C^***-algebra**.

Example 55. The Banach space $\mathcal{B}(H)$ is a C^*-algebra.

This follows from the following lemma.

Lemma 15.10. For all $T \in \mathcal{B}(H)$, we have $\|T\|^2 = \|T^*T\| = \|TT^*\|$.

Proof. We know $\|T^*T\| \leq \|T\|\|T^*\| = \|T\|^2$. For all $x \in H$, we have

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, T^*Tx \rangle \leq \|T^*T\|\|x\|^2.$$

Then we know $\|T\|^2 \leq \|T^*T\|$. The another equality follows from $(T^*T)^* = T^*T$ and $\|T\| = \|T^*\|$. □

Consider a Banach space X, then there is an algebra $\mathcal{A} = \mathcal{B}(X)$. Suppose $T \in \mathcal{A}$ is invertible, then Theorem 7.4 implies that $T^{-1} \in \mathcal{A}$. Thus we define:

Definition. An element $x \in \mathcal{A}$ is said to be **invertible**, if there is an element $y \in \mathcal{A}$ such that

$$xy = yx = \mathbb{1}_{\mathcal{A}}.$$

And such y is called the inverse of x, denoted by x^{-1}.

Remark 15.7. Things shoul be noticed:

1. the inverse of x is unique;
2. if $y = x^{-1}$, then $x = y^{-1}$;
3. for a square matrix $A \in \mathbb{K}^{n\times n}$, if there is a square matrix B such that $AB = I$, then we have $BA = I$ and hence $B = A^{-1}$. But in a unital algebra, it's possible $xy = \mathbb{1}_{\mathcal{A}}$ while $yx \neq \mathbb{1}_{\mathcal{A}}$.

Example 56. Consider $X = \ell_2$, the unital algebra $\mathcal{A} = \mathcal{B}(\ell_2)$. Consider $T, S \in \mathcal{A}$, defined by

$$\begin{aligned} T\colon\ & \ell_2 \to \ell_2, (x_n)_{n\in\mathbb{N}} \mapsto (x_{n+1})_{n\in\mathbb{N}}, \\ S\colon\ & \ell_2 \to \ell_2, (x_n)_{n\in\mathbb{N}} \mapsto (0, x_1, x_2, \ldots). \end{aligned}$$

Then $T \circ S = \mathbb{1}_{\mathcal{A}}$ while it's impossible that $S \circ T = \mathbb{1}_{\mathcal{A}}$.

From now on, we assume all algebras mentioned are unital. The following proposition shows some relations in a C^*-algebra.

Proposition 15.11. Let $\mathcal{A}$ be a C^*-algebra and $x, y \in \mathcal{A}$, then

1. the unity is self-adjont, i.e. $\mathbb{1}_{\mathcal{A}}^* = \mathbb{1}_{\mathcal{A}}$;
2. if x is invertible, then $(x^{-1})^{-1} = x$;
3. if both of x, y are invertible then so is xy and $(xy)^{-1} = y^{-1}x^{-1}$;
4. if x is invertible, then so is x^* and $(x^*)^{-1} = (x^{-1})^*$.

Proof. We prove the first and the last only, since the second and the third follow from the uniqueness of inverse. For the first, apply the uniqueness of $\mathbb{1}_{\mathcal{A}}$ and

$$\forall x \in \mathcal{A}\colon \mathbb{1}_{\mathcal{A}} x = x \mathbb{1}_{\mathcal{A}} = x,$$

which implies

$$\forall x \in \mathcal{A}\colon x^* \mathbb{1}_{\mathcal{A}}^* = \mathbb{1}_{\mathcal{A}}^* x^* = x^*.$$

We're done, since for all $x \in \mathcal{A}$, we have $x = (x^*)^*$. The last follows from

$$\begin{aligned} x^*(x^{-1})^* &= (x^{-1}x)^* = \mathbb{1}_{\mathcal{A}}^* = \mathbb{1}_{\mathcal{A}}, \\ (x^{-1})^* x^* &= (xx^{-1})^* = \mathbb{1}_{\mathcal{A}}^* = \mathbb{1}_{\mathcal{A}}. \end{aligned}$$

and the uniqueness of inverse. □

We can define the polynomial for an element in a unital algebra $\mathcal{A}$.

Example 57. Let $\mathcal{A}$ be an algebra and $x \in \mathcal{A}$. For a complex polynomial $p(t) = \sum_{k=0}^{n} a_k t^k \in \mathbb{C}[t]$, define

$$p(x) := \sum_{k=0}^{n} a_k x^k, \tag{39}$$

where $x^0 := \mathbb{1}_{\mathcal{A}}$. Then $p(x) \in \mathcal{A}$.

Similarly, we can define the power series for an element in a unital algebra $\mathcal{A}$.

Example 58. Let $(a_n)_{n\geq 0}$ be a complex sequence and $x \in \mathcal{A}$, where $\mathcal{A}$ is an algebra. We show that $\sum_{n\geq 0} a_n x^n$ is meaningful in some sense. For convenience, consider $a_n = 1/n! (\forall n \geq 0)$.

Claim. There is a unique $y \in \mathcal{A}$ such that

$$\sum_{n\geq 0} \frac{x^n}{n!} = y$$

in the meaning that $y = \lim_n P_n(x)$, where $P_n(t) \in \mathbb{C}[t]$ is defined as

$$P_n(t) = \sum_{k=0}^{n} \frac{t^n}{n!}.$$

Proof. Since $\mathcal{A}$ is a Banach space, we can apply Proposition 3.1. It suffices to show that $(P(x)_n)_{n\in\mathbb{N}}$ is Cauchy. For all $m \geq n \geq 1$:

$$\begin{aligned} \|P_m(x) - P_n(x)\| &= \left\| \sum_{k=n+1}^{m} P_n(x) \right\| \\ &\leq \sum_{k=n+1}^{m} \frac{1}{k!} \|x^n\| \\ &\leq \sum_{k=n+1}^{m} \frac{1}{k!} \|x\|^n \\ &\to 0 (m, n \to \infty) \end{aligned}$$

since for all $y \in \mathbb{R}$, $\sum_{n\geq 1} y^n/n!$ converges (ratio test works). □

Theorem 15.12. Let X be a Banach space, $\mathcal{A} := \mathcal{B}(X)$ and $x \in \mathcal{A}$ satisfying $\|x\| < 1$. Then $\mathbb{1}_{\mathcal{A}} - x$ is invertible.

Aim. We want to find $y \in \mathcal{A}$ such that $yx = xy = \mathbb{1}_{\mathcal{A}}$. Recall that, for $x \in \mathbb{R}$ such that $x < 1$, we have

$$(\mathbb{1} - x)^{-1} = \sum_{n \geq 0} x^n.$$

Proof. Clearly $\sum_{n \geq 0} x^n$ converges in $\mathcal{A}$. Suppose $\mathcal{A} \supseteq (a_n)_{n \in \mathbb{N}} \to a \in \mathcal{A}$ and $b \in \mathcal{A}$, then

$$\forall n \in \mathbb{N} \colon \|ba_n - ba\| \leq \|b\| \cdot \|a_n - a\|$$

implies $(ba_n)_{n \in \mathbb{N}} \to ba$. Thus

$$\begin{aligned}
(\mathbb{1}_{\mathcal{A}} - x) \sum_{n \geq 0} x^n &= (\mathbb{1}_{\mathcal{A}} - x) \lim_N \sum_{n=0}^{N} x^n \\
&= \lim_N (\mathbb{1}_{\mathcal{A}} - x) \sum_{n=0}^{N} x^n \\
&= \lim_N (\mathbb{1}_{\mathcal{A}} - x^{N+1}) \\
&= \mathbb{1}_{\mathcal{A}}.
\end{aligned}$$

The another equality $\sum_{n \geq 0} x^n (\mathbb{1}_{\mathcal{A}} - x) = \mathbb{1}_{\mathcal{A}}$ can be checked similarly and we're done. □

Question 15.1. Let $\lambda \in \mathbb{C}$ and $T \in \mathcal{B}(X)$, where X is a Banach space. When will $\lambda I - T$ be invertible? Here $I = \mathbb{1}_{\mathcal{B}(X)}$.

Partial Answer. A sufficient condition is: $|\lambda| > \|T\|$. If $|\lambda| > \|T\| > 0$, then $\lambda^{-1} \in \mathbb{C}$ and $\lambda I - T$ is invertible if and only $I - \lambda^{-1} T$ is invertible. Then Theorem 15.12 implies:

$$(\lambda I - T)^{-1} = \lambda^{-1} \sum_{n \geq 0} (T/\lambda)^n.$$

And we can also show that

$$\left\| (\lambda I - T)^{-1} \right\| \leq \frac{1}{|\lambda| - \|T\|}.$$

If T is an invertible operator, then $\lambda = 0$ works. □

15.2.2 Recall Linear Algebra

We studied the eigenvalue of matrices: let A be a $n \times n$ complex matrix, a complex number is called an **eigenvalue** of A, if there is a (column) vector $x \in \mathbb{C}^n$ such that

$$(\lambda I - A)x = 0,$$

where I is the identity matrix. Equivalently, $\ker(\lambda I - A) \neq \{0\}$. In linear algebra, a square matrix (also viewed as a linear transformation) is invertible if and only if it has the trivial kernel.

Definition. Let X be a Banach space and $T \in \mathcal{B}(X)$. If $\lambda I - T$ is not invertible, then λ is said to be a **spectrum point** of T. The set of all spectrum points of T is denoted by $\sigma(T)$, called the resolvent set of T.

Definition. Let X be a Banach space and $T \in \mathcal{B}(X)$. If $\lambda I - T$ is invertible, then λ is said to be a **regular point** of T. The set of all regular points of T is denoted by $\rho(T)$, called the spectrum set of T.

Clearly we have $\mathbb{C} = \sigma(T) \bigsqcup \rho(T)$.

Question 15.2. Let $T \in \mathcal{B}(X)$ and $r > 0$ satisfy that $\sigma(T) \subseteq \{z \in \mathbb{C} \colon |z| < r\}$. Can we give an upper bound of r?

Question 15.1 shows that $r \leq \|T\|$.

Example 59. Consider the shift operator

$$T \colon \ell_2 \to \ell_2, (x_n)_{n \in \mathbb{N}} \mapsto (0, x_1, x_2, \ldots).$$

Clearly, $\|T\| = 1$ and hence $\sigma(T) \subseteq \{z \in \mathbb{C} \colon |z| \leq 1\}$.

15.2.3 Classification of Spectrum and so on

For an algebra $\mathcal{A} = \mathcal{B}(X)$, where X is a Banach space, the spectrum of $T \in \mathcal{A}$ is

$$\sigma(T) = \{\lambda \in \mathbb{C} \colon \lambda I - T \text{ is not invertible}\}.$$

Let $I := \mathbb{1}_{\mathcal{A}}$. If $\lambda \in \sigma(T)$, then there are 3 cases:

Case 1. the operator $\lambda I - T$ has a non-trivial kernel, i.e. $\ker(\lambda I - T) \neq \{0\}$

Case 2. the operator $\lambda I - T$ has a trivial kernel but it's image, being smaller than X, is dense in X, i.e.

$$\ker(\lambda I - T) = \{0\}, \overline{\operatorname{Im}(\lambda I - T)} = X, \operatorname{Im}(\lambda I - T) \neq X.$$

Case 3. the operator $\lambda I - T$ has a trivial kernel but it's image is not even dense in X, i.e.

$$\ker(\lambda I - T) = \{0\}, \overline{\mathrm{Im}(\lambda I - T)} \neq X.$$

Then we know

$$\sigma(T) = \sigma_p(T) \bigsqcup \sigma_c(T) \bigsqcup \sigma_r(T).$$

Remark 15.8. A complex number $\lambda \in \mathbb{C}$ is a spectrum of T, if and only if the following chain

$$\cdots \longrightarrow 0 \longrightarrow X \xrightarrow{\lambda I - T} X \longrightarrow 0 \longrightarrow \cdots$$

fails to be exact.

Question 15.3. Let $T \in \mathcal{B}(X)$ be a compact operator. Can we say about $0 \in \sigma(T)$?

Answer and Proof. If $\dim X = \infty$, it holds. Suppose $\dim X = \infty$ and $0 \notin \sigma(T)$, i.e. T is invertible. Then $T^{-1} \in \mathcal{B}(X)$. Since $T \in \mathcal{C}(X)$ and $\mathcal{C}(X)$ is an ideal of $\mathcal{B}(X)$, we find

$$I = TT^{-1} \in \mathcal{C}(X).$$

The identity operator I is compact if and only if $I(B_X) = B_X$ is compact, equivalently, $\dim X < \infty$ (see Theorem 6.8). Thus, it's sure that $0 \in \sigma(T)$ when $\dim(X) = \infty$.

If $\dim X < \infty$, then all operators are bounded and hence of finite-rank. Thus all operators are compact. But clearly the identity operator I is invertible and hence $0 \notin \sigma(I)$. □

Special operators have special spectrum:

Theorem 15.13. Let H be a Hilbert space and $T \in \mathcal{B}(H)$.

1. if $T = T^*$, then $\sigma(T) \subseteq \mathbb{R}$;
2. if $T \geq 0$, then $\sigma(T) \subseteq [0, \infty)$.

Definition. An operator $T \in \mathcal{B}(H)$ is said to be positive-semidefinite, denoted by $T \geq 0$, if for all $x \in H$, we have $\langle Tx, x\rangle \in [0, \infty)$.

Remark 15.9. A positive-semidefinite operator is self-adjoint.

16 Week 16

16.1 Lecture 16-1

16.1.1 Exercise Course: Question Part

Question 16.1. Let H be a Hilbert space and $(e_n)_{n\in\mathbb{N}}$ be a sequence of orthogonal vector, i.e. $(e_n, e_m) = 0$ when $n \neq m$. Prove that the following statements are equivalent:

1. the sequence $(\sum_{j=1}^{n} e_j)_{n\in\mathbb{N}}$ converges;
2. the sequence $(\sum_{j=1}^{n} e_j)_{n\in\mathbb{N}}$ converges weakly;
3. the sum $\sum_{n\geq 1}\|e_n\|^2$ is finite.

Question 16.2. Let X be a Banach space and $T \in \mathcal{L}(X)$. Prove that: $T \in \mathcal{B}(X)$ if and only if for all $x \in X$: $(x_n)_{n\in\mathbb{N}} \rightharpoonup x$ implies that $(Tx_n)_{n\in\mathbb{N}} \rightharpoonup Tx$.

Question 16.3. Let X be a reflexive linear normed space. Prove that: for all $f \in X^{**}$, $(f_n)_{n\in\mathbb{N}} \subseteq X^{**}$ such that $(f_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} f$, we have $(f_n)_{n\in\mathbb{N}} \rightharpoonup f$.

Remark 16.1. For a linear normed space X with a pre-dual space X_*, we know that $(f_n)_{n\in\mathbb{N}} \rightharpoonup f$ implies $(f_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} f$.

Question 16.4. Let X be a linear normed space X with a closed subspace $M \hookrightarrow X$. Show that: if $(x_n)_{n\in\mathbb{N}} \subseteq M$ and $x_0 \in X$ satisfy that $(x_n)_{n\in\mathbb{N}} \rightharpoonup x)$, then $x_0 \in M$.

Question 16.5. Let X be a reflexive linear normed space. Prove that: if $(x_n)_{n\in\mathbb{N}} \subseteq X$ satisfies that $\forall f \in X^*$:

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

then there is some $x \in X$ such that $(x_n)_{n\in\mathbb{N}} \rightharpoonup x$.

16.1.2 Exercise Course: Solution Part

Proof of Question 16.1. Let $(x_n)_{n\in\mathbb{N}} := (\sum_{j=1}^{n} e_j)_{n\in\mathbb{N}}$. We will prove in the following order

$$1 \implies 2 \implies 3 \implies 1.$$

$1 \implies 2$: suppose $(x_n)_{n\in\mathbb{N}} \to x \in H$. For all $f \in H^*$, we have

$$|f(x_n) - f(x)| \leq \|f\|\|x_n - x\| \to 0(n \to \infty).$$

Thus $(x_n)_{n\in\mathbb{N}} \rightharpoonup x$.

2 $\Longrightarrow$ 3: weak convergence implies that $(x_n)_{n\in\mathbb{N}}$ is weakly bounded, and then Exercise 11.1 implies that $(x_n)_{n\in\mathbb{N}}$ is bounded. Suppose M is a bound of $(x_n)_{n\in\mathbb{N}}$. Then for all $n \in \mathbb{N}$: Theorem 29 implies that

$$\sum_{j=1}^{n} \|e_j\|^2 = \|x_n\|^2 \leq M^2,$$

therefore

$$\sum_{n\geq 1} \|e_n\|^2 = \sup_{n\in\mathbb{N}} \|x_n\|^2 \leq M^2 < \infty.$$

3 $\Longrightarrow$ 1: it suffices to show that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, since H is complete. For all $n > m$:

$$\|x_n - x_m\|^2 = \sum_{j=m+1}^{n} \|e_j\|^2 \to 0 (n \to \infty).$$

Therefore, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and hence $(x_n)_{n\in\mathbb{N}}$ converges. $\square$

Proof of Question 16.2. Necessity: for all $(x_n)_{n\in\mathbb{N}} \subseteq$ and $x \in X$ such that $(x_n)_{n\in\mathbb{N}} \rightharpoonup x$, we have

$$\lim_n f(Tx_n) = \lim_n (T^* f)(x_n) = (T^* f)(x) = f(Tx),$$

since $T^* f \in X^*$ and $(x_n)_{n\in\mathbb{N}} \rightharpoonup x$, where $T^* \colon X^* \to X^*, f \mapsto f \circ T$ is the conjugate operator of T.

Sufficiency (method 1): consider proof by contradiction. If T is not bounded, then

$$\sup_{\|x\|=1} \|Tx\| = \infty,$$

and we can pick a sequence $(x_n)_{n\in\mathbb{N}}$ such that $\forall n \in \mathbb{N}$, $\|x_n\| = 1$ and $\|Tx_n\| \geq n$.

Claim. The sequence $(Ty_n)_{n\in\mathbb{N}}$ can't converge to 0 weakly.

To see the claim, if suffices to see that: weak convergence implies that $(Ty_n)_{n\in\mathbb{N}}$ is weakly bounded, and hence bounded, which is a contradiction. Now consider the sequence $(y_n)_{n\in\mathbb{N}}$, where $y_n := (\sqrt{n}/\|Tx_n\|)x_n$, then we have $\|y_n\| \leq 1/\sqrt{n} \to 0 (n \to \infty)$ and $\|Ty_n\| = \sqrt{n} \to \infty (n \to \infty)$. Now we have a sequence $(y_n)_{n\in\mathbb{N}}$ such that $(y_n)_{n\in\mathbb{N}} \rightharpoonup 0$ and $(Ty_n)_{n\in\mathbb{N}} \not\rightharpoonup 0$, contradiction with the assumption.

Sufficiency (method 2): consider **Closed Graph Theorem**, see Theorem 8.8. It suffices to show that the graph $G(T)$ is closed in $X \times X$, i.e. $\forall (x_n, Tx_n)_{n\in\mathbb{N}}$ such that $(x_n)_{n\in\mathbb{N}} \to x \in X$, $(Tx_n)_{n\in\mathbb{N}} \to y \in X$ then we have $y = Tx$. Suppose $(x_n)_{n\in\mathbb{N}} \to x \in X$, $(Tx_n)_{n\in\mathbb{N}} \to y \in X$, then $(x_n)_{n\in\mathbb{N}} \rightharpoonup x$ and $(Tx_n)_{n\in\mathbb{N}} \rightharpoonup y$. And the assumption means that $(Tx_n)_{n\in\mathbb{N}} \rightharpoonup Tx$. The uniqueness of weak limit implies $y = Tx$. □

Remark 16.2. Both proofs of sufficiency need the Banach Steinhaus Theorem.

Proof of Question 16.3. Suppose that $(f_n)_{n\in\mathbb{N}} \subseteq X^{**}$, $f \in X^{**}$ such that $(f_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} f$, i.e. $\forall x \in X^*$: $f_n(x) \to f(x)$. Thus: $\forall x \in X^*$: $\hat{x}(f_n) \to \hat{x}(f)$, where $\hat{x} = \iota_{X^*}(x)$ and ι_{X^*} is the natural embedding of X^*. Then $(f_n)_{n\in\mathbb{N}} \rightharpoonup f$ follows from the following claim:

Claim. For all $F \in X^{***}$, there is some $x \in X^*$ such that $F = \iota_{X^*}(x)$.

In other words, the natural embedding ι_{X^*} is also an isometry, i.e. X^* is also reflexive. That's true, since

$$\iota_X \circ \iota_X^{-1} = \mathrm{id}_{X^{**}} \wedge \iota_X^{-1} \circ \iota_X = \mathrm{id}_X$$

implies that (by the property of the functor $*$)

$$\iota_X^* \circ (\iota_X^{-1})^* = \mathrm{id}_{X^*} \wedge (\iota_X^{-1})^* \circ \iota_X^* = \mathrm{id}_{X^{***}}.$$

And Theorem 11.1 implies that $\|\iota_X^*\| = \left\|(\iota_X^*)^{-1}\right\| = 1$, which means ι_X^* is an isometry. Then that X^* is reflexive follows from the following claim:

Claim. We have

$$\iota_X^* \circ \iota_{X^*} = \mathrm{id}_{X^*},$$

which implies $\iota_{X^*}^{-1} = \iota_X^*$ (the left inverse of an invertible map is the inverse).

Given arbitrary $f \in X^*$ and $x \in X$. On the one hand

$$\iota_X^* \circ \iota_{X^*}(f) = \iota_X^*(\hat{f}) = \hat{f} \circ \iota_X \in X^*,$$

where $X^{***} \ni \hat{f} \colon X^{**} \to \mathbb{K}, A \mapsto A(f)$. Then

$$\left[\iota_X^* \circ \iota_{X^*}(f)\right](x) = \hat{f} \circ \iota_X(x) = \hat{f}(\hat{x}) = \hat{x}(f) = f(x).$$

On the other hand

$$\left[\mathrm{id}_{X^*}(f)\right](x) = f(x).$$

Then we proved $\iota_X^* \circ \iota_{X^*}(f) = \mathrm{id}_{X^*}(f)$ since $x \in X$ is arbitrary. Therefore, we proved $\iota_X^* \circ \iota_{X^*} = \mathrm{id}_{X^*}$ since $f \in X^*$ is arbitrary. □

Proof of Question 16.4. See Exercise 9.4. If $x_0 \notin M$, we have $d(x_0, M) > 0$ since M is closed and there is a linear functional such that $f(x_0) \neq 0$ while $f|_M = 0$ and hence it's impossible that $(x_n)_{n\in\mathbb{N}} \rightharpoonup x_0$, which is a contradiction. □

Remark 16.3. To prove Exercise 9.4, consider the functional

$$f_0\colon \operatorname{span}\left(\{x_0\} \bigcup M\right) \to \mathbb{K}$$

defined by

$$f_0(x_0) = d(x_0, M) \wedge f_0|_M = 0.$$

Then apply Theorem 9.2. ■

Proof of Question 16.5. Given a sequence $(x_n)_{n\in\mathbb{N}}$ satisfies the property stated in this Question, consider the natural embedding $\iota\colon X \to X^{**}$. Then for all $f \in X^*$:

$$\lim_{m,n} |\widehat{x_m}(f) - \widehat{x_n}(f)| = 0.$$

Since X is a reflexive space, we know X is a Banach space. Theorem 8.3 implies that $(\hat{x}_n)_{n\in\mathbb{N}} \xrightarrow{s} \hat{x}$ for some $\hat{x} \in X^{**}$. Clearly: in the space X^{**}, strong convergence is equivalent to weak star convergence, and hence we proved $(\hat{x}_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} \hat{x}$ for some $x \in X$. Question 16.3 implies that it suffices to show $(x_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} x$ for some $x \in X$, and

$$\begin{aligned} (x_n)_{n\in\mathbb{N}} \rightharpoonup x &\iff \forall f \in X^*\colon f(x_n) \to f(x) \\ &\iff \forall f \in X^*\colon \hat{x}_n(f) \to \hat{x}(f) \\ &\iff (\hat{x}_n)_{n\in\mathbb{N}} \overset{*}{\rightharpoonup} \hat{x}. \end{aligned}$$ □

16.2 Lecture 16-2

16.2.1 Exercise Course: Question Part

Question 16.6. Let H be a Hilbert space and $(x_n)_{n\in\mathbb{N}}$ be a sequence of orthogonal vector, i.e. $(x_n, x_m) = 0$ when $n \neq m$. Prove that the following statements are equivalent:

1. the sequence $(\sum_{j=1}^{n} x_j)_{n\in\mathbb{N}}$ converges;
2. the sequence $(\sum_{j=1}^{n} \langle x_j, y\rangle)_{n\in\mathbb{N}}$ converges in $\mathbb{K}$;
3. the sum $\sum_{n\geq 1} \|x_n\|^2$ is finite.

Question 16.7. Give an example that a sequence is weakly* convergent but not weakly convergent.

Question 16.8. Let $1 < p < \infty$ and $\alpha = (\alpha_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{K}$. If $\forall x \in \ell_q$, we have $\sum_{n\geq 1} x_n\alpha_n$ exists, show that $\alpha \in \ell_p$.

Question 16.9. Given $\alpha \in \ell_\infty$ and $1 \leq p < \infty$. Define

$$T_\alpha \colon \ell_p \to \ell_p, x \mapsto \alpha \cdot x.$$

Find $\|T_\alpha\|$.

Question 16.10. Let r, p, q be positive real numbers such that $1/r = 1/p + 1/q$ and (Ω, Σ, μ) be a measure space. Let $\alpha \in L_q$ and defined an operator

$$T_\alpha \colon L_p \to L_p, x \mapsto \alpha \cdot x.$$

Show that $\|T_\alpha\| = \|\alpha\|_q$.

16.2.2 Exercise Course: Solution Part

Proof of Question 16.6. We have proved 1 $\iff$ 3. It suffices to prove

$$1 \Longrightarrow 2 \Longrightarrow 3.$$

Define a sequence $(s_n)_{n\in\mathbb{N}}$ by $s_n := \sum_{j=1}^n x_j$ for all $n \in \mathbb{N}$.

First, 1 $\iff$ 2. Let φ^H be the Riesz map. Suppose $(s_n)_{n\in\mathbb{N}}$ converges, then

$$(\sum_{j=1}^n \langle x_j, y\rangle)_{n\in\mathbb{N}} = (\langle s_n, y\rangle)_{n\in\mathbb{N}}$$

converges, since $\langle\ , y\rangle = \varphi_y^H$ is a continuous functional.

Secondly, 2 $\Longrightarrow$ 3. The convergence implies that $(s_n)_{n\in\mathbb{N}}$ is weakly bounded. Then the proof of Question 16.1, 2 $\Longrightarrow$ 3 works.

Another method for 2 $\Longrightarrow$ 3. This idea is similar to Question 3 Define a functional sequence $(f_n)_{n\in\mathbb{N}}$ as follows: for all $n \in \mathbb{N}$,

$$f_n \colon H \to \mathbb{K}, x \mapsto \langle x, s_n\rangle .$$

Then $f_n = \varphi^H_{s_n}$ and hence $\|f_n\| = \|s_n\| = \sqrt{\sum_{j=1}^n \|x_j\|^2}$, by Theorem 29. The sequence $(\sum_{j=1}^n \langle x_j, y\rangle)_{n\in\mathbb{N}}$ converges in $\mathbb{K}$, and hence is bounded, i.e. $(f_n)_{n\in\mathbb{N}}$ is bounded pointwisely. Thus Theorem 7.6 ensures that $(f_n)_{n\in\mathbb{N}}$ is uniformly bounded, i.e.

$$\infty > \sup_{n\in\mathbb{N}} \|f_n\| = \Big(\sum_{n\geq 1} \|x_n\|^2\Big)^{1/2}. \qquad \square$$

Remark 16.4. Banach-Steinhaus Theorem 7.6 is important for 2 $\Longrightarrow$ 3. Both of 2 methods need it. ■

Proof of Question 16.7. See Example 29. □

Proof of Question 16.8. See this. □

Proof of Question 16.9. Clearly $T_\alpha \in \mathcal{L}(\ell_p)$, and Lemma 16.1 implies that $\|T_\alpha\| \leq \|\alpha\|_\infty$. Consider $(e_n)_{n\in\mathbb{N}} \subseteq \ell_p$, and we can see that

$$\|T_\alpha(e_n)\| = |\alpha_n|$$

for all $n \in \mathbb{N}$. Since $\|e_n\|_p = 1$, we have

$$\forall n \in \mathbb{N}\colon \|T_\alpha\| \geq |\alpha_n|,$$

i.e. $\|T_\alpha\| \geq \|\alpha\|_\infty$. □

Proof of Question 16.10. The boundedness of T_α is implied by the following lemma.

Lemma 16.1. Given a measure space (Ω, Σ, μ), and $p, q, r \in (0, \infty]$ such that $1/r = 1/p + 1/q$. For all $f \in L_p$ and $g \in L_q$, we have $fg \in L_r$ and

$$\|fg\|_r \leq \|f\|_p \|g\|_q.$$

The inequality comes to be a equality if and only if $|f|^p = |g|^q$ almost everywhere.

Lemma 16.1 can be proved by applying Hölder's inequality to $|f|^r$ and $|g|^r$.

Consider $\|\alpha\|_q > 0$, i.e. $\alpha \neq 0 \in L_q$. Notice the condition for equality, and we consider the function $\tilde{\alpha} := \operatorname{sign}(\alpha)|\alpha|^{q/p} \in L_p$, then

$$T_\alpha(\tilde{\alpha}) = |\alpha|^{p/q+1}.$$

Thus

$$\|T_\alpha(\tilde{\alpha})\|_r = \|\alpha\|_q^{q/r},$$

as you should verify. Furthermore, $\|\tilde{\alpha}\| = \|\alpha\|_q^{q/p}$. Therefore

$$\|T_\alpha\| \geq \|T_\alpha(\tilde{\alpha})\|_r / \|\tilde{\alpha}\|_p = \|\alpha\|_q.$$ □

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 A.1. 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 $e_1, e_2, \ldots, e_n \in \bigcup \mathcal{A}$ such that $e_j \in X_j \in \mathcal{A}$ for all $j = 1, 2, \ldots, n$. Since $\mathcal{A}$ is totally ordered, we can suppose $X_1 \subseteq X_2 \subseteq \cdots \subseteq X_n$ and hence $e_j \in X_n (\forall j = 1, 2, \ldots, n)$. Since $X_n \in \mathcal{A} \subseteq \mathcal{D}$, X_n is linear independent and hence $e_1, e_2, \ldots, e_n$ 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 $\operatorname{span}(B)$. If there is an element $v \in V$ such that $v \notin \operatorname{span}(B)$, i.e. $B \cup \{v\}$ is linearly independent. This is impossible by the definition of the maximal element. □

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 $\varphi \colon V \to W$.

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 $\varphi \colon V \to W$ by

$$\forall \alpha \in I, \forall z \in \mathbb{C} : \varphi(z v_\alpha) := \big(\operatorname{Re}(z) v_\alpha, \operatorname{Im}(z) v_\alpha\big).$$

Extend φ to V keeping $\mathbb{R}$-linear. Then $\varphi|^{\operatorname{Im}(\varphi)}$ is what we wanted. □

B Banach Functor

This appendix comes from [4].

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

$$\begin{aligned} \|u_n x - u_m x\|_Y &\leq \|(u_n - u_m)x\|_Y \\ &\leq \|u_n - u_m\|_{\mathcal{B}(X,Y)} \|x\|_X \\ &\to 0 (m, n \to \infty). \end{aligned} \tag{40}$$

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

$$u\colon X \to Y, x \mapsto \lim_n u_n(x).$$

And now we prove that $(u_n)_{n\in\mathbb{N}} \to u$ 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 $m \to \infty$, by the continuity of $\| \ \|_Y$, we have

$$\|(u_n - u)x\| = \|u_n x - ux\|_Y \leq \varepsilon \|x\|, \forall x \in X.$$

Therefore, $\|u_n - u\| \leq \varepsilon$ holds for all $n > N$. That is $(u_n)_{n\in\mathbb{N}} \to u$. $\square$

Now we can define

Definition. The contravariant functor

$$\begin{aligned} *\colon \mathsf{Nor} &\longrightarrow \mathsf{Ban}, \\ \mathrm{Ob}(\mathsf{Nor}) \ni X &\longmapsto X^*, \\ \mathrm{Mor}(\mathsf{Nor}) \ni \varphi\colon X_1 \to X_2 &\longmapsto \varphi^*\colon X_2^* \to X_1^*. \end{aligned} \tag{41}$$

Where $\varphi^*\colon X_2^* \to X_1^*, f \mapsto f \circ \varphi$.

Remark B.1. Banach Functor is a special case of the functor $\mathcal{B}(\ ,Y)$ where Y is a Banach space, defined as

$$\begin{aligned} \mathcal{B}(\ ,Y)\colon \mathsf{Nor} &\longrightarrow \mathsf{Ban}, \\ \mathrm{Ob}(\mathsf{Nor}) \ni X &\longmapsto \mathcal{B}(X,Y), \\ \mathrm{Mor}(\mathsf{Nor}) \ni \varphi\colon X_1 \to X_2 &\longmapsto \mathcal{B}(\varphi,Y)\colon \mathcal{B}(X_1,Y) \to \mathcal{B}(X_2,Y). \end{aligned}$$

Here $\mathcal{B}(\varphi,Y)\colon \mathcal{B}(X_1,Y) \to \mathcal{B}(X_2,Y), f \mapsto f \circ \varphi$.

Banach functor is surely a functor.

Proof. It suffices to prove that $(\mathrm{id}_X)^* = \mathrm{id}_{X^*}$ and $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$. To prove two maps are the same, we should prove that they coincide at every point.

- Given $X \in \mathrm{Ob}(\mathsf{Nor})$, we have

  $$\forall f \in X^* : (\mathrm{id}_X)^*(f) = f \circ \mathrm{id}_X = f = \mathrm{id}_{X^*}(f).$$

  Thus $(\mathrm{id}_X)^* = \mathrm{id}_{X^*}$, since id_{X^*} is uniquely determined by this property.

- Given $X_1 \xleftarrow{\varphi} X_2 \xleftarrow{\psi} X_3$. Notice that $\mathrm{dom}((\varphi \circ \psi)^*) = X_1^*$ and for any $f \in X_1^*$, we have

  $$\begin{aligned} (\varphi \circ \psi)^*(f) &= f \circ (\varphi \circ \psi) \\ &= (f \circ \varphi) \circ \psi \\ &= \big(\varphi^*(f)\big) \circ \psi \\ &= \psi^*\big(\varphi^*(f)\big) \\ &= (\psi^* \circ \varphi^*)(f). \end{aligned}$$

  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 $\forall f \in X_2^*$

$$\|\varphi^*(f)\|_{X_1^*} = \|f \circ \varphi\|_{\mathcal{B}(X_1,\mathbb{K})} \le \|f\|_{X_2^*}\|\varphi\|_{\mathcal{B}(X_1,X_2)},$$

hence $\|\varphi^*\|_{\mathcal{B}(X_2^*,X_1^*)} \le \|\varphi\|_{\mathcal{B}(X_1,X_2)}$. □

If we restrict $*\colon \mathsf{Nor} \to \mathsf{Ban}$ to the full subcategory Ban, and call it **Banach adjointness functor**, then we can consider the composition of $*\colon \mathsf{Ban} \to \mathsf{Ban}$, i.e. $**\colon \mathsf{Ban} \to \mathsf{Ban}$, 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

$$\begin{aligned} \mathcal{B}(\mathbb{K},\ )\colon \mathsf{Ban} &\longrightarrow \mathsf{Ban} \\ X &\longmapsto \mathcal{B}(\mathbb{K}, X), \\ (f\colon X \to Y) &\longmapsto (\mathcal{B}(\mathbb{K}, f)\colon \mathcal{B}(\mathbb{K}, X) \to \mathcal{B}(\mathbb{K}, Y)), \end{aligned}$$

where

$$\mathcal{B}(\mathbb{K}, f)\colon \mathcal{B}(\mathbb{K}, X) \to \mathcal{B}(\mathbb{K}, Y), \psi \mapsto f \circ \psi.$$

Proposition B.1. Show that $\mathcal{B}(\mathbb{K},\ )$ is naturally equivalent to $\mathrm{id}_{\mathsf{Ban}}$, the identity functor of Ban.

Proof. Define the natural transformation $\theta = \{\theta_X \colon X \in \mathrm{Ob}(\mathsf{Ban})\}$ as follows

$$\theta_X \colon \mathcal{B}(\mathbb{K}, X) \to X, \varphi \mapsto \varphi(1),$$

where 1 is just the multiplicative identity of $\mathbb{K}$. We check that θ_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: $\mathrm{Im}\, \theta_X = X$. For arbitrary $x \in X$, define $\varphi_x \colon \mathbb{K} \to X, z \mapsto zx$. Then $\|\varphi_x\| = \|x\|$ and $\theta_X(\varphi_x) = x$ as we want. Furthermore, we know that θ_X keeps norms.

Thus θ_X is an isomorphism for each $X \in \mathrm{Ob}(\mathsf{Ban})$.

Then we check that the following diagram commutes

$$\begin{tikzcd} \mathcal{B}(\mathbb{K}, X) \arrow[r] \arrow[d, "\mathcal{B}(\mathbb{K}, f)"'] & \mathrm{id}_{\mathsf{Ban}}(X) \arrow[r, equal] \arrow[d, "\mathrm{id}_{\mathsf{Ban}}(f)"'] & X \arrow[d, "f"] \\ \mathcal{B}(\mathbb{K}, Y) \arrow[r] & \mathrm{id}_{\mathsf{Ban}}(Y) \arrow[r, equal] & Y \end{tikzcd}$$

Given $\varphi \in \mathcal{B}(X, C)$, on the one hand

$$(f \circ \theta_X)(\varphi) = f\big(\theta_X(\varphi)\big) = f\big(\varphi(1)\big).$$

On the other hand

$$\big(\theta_Y \circ \mathcal{B}(f, \mathbb{K})\big)(\varphi) = \theta_Y\Big(\big(\mathcal{B}(f, \mathbb{K})\big)(\varphi)\Big) = \theta_Y(f \circ \varphi) = (f \circ \varphi)(1).$$

We're done since $f(\varphi(1)) = (f \circ \varphi)(1)$. $\square$

Remark B.2. The definition of θ_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 $F\colon \mathcal{C} \to \mathcal{D}, G\colon \mathcal{D} \to \mathcal{E}$ are two functors. The composition functor of F and G, denoted by $G \circ F$ is defined as follows.

$$\begin{aligned} G \circ F\colon \mathcal{C}_1 &\longrightarrow \mathcal{C}_3, \\ \mathrm{Ob}(\mathcal{C}_1) \ni A &\mapsto G\big(F(A)\big), \\ \mathrm{Mor}(\mathcal{C}_1) \ni \varphi\colon X \to Y &\mapsto G\big(F(\varphi)\big)\colon G\big(F(X)\big) \to G\big(F(Y)\big). \end{aligned}$$

Thus the bidual functor $**$ is just the composition of $*\colon \mathsf{Nor} \to \mathsf{Ban}$ and itself (more precisely, $** := * \circ \iota \circ *$, where $\iota\colon \mathsf{Ban} \to \mathsf{Nor}$ is the identity functor that keeps everything the same). Here is a easy exercise that ensures something like $\big(X^{**}\big)^* = \big(X^*\big)^{**}$.

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 Nor to itself.

Proposition B.2. Consider the category Nor and two functors

$$\begin{aligned} **\colon \mathsf{Nor} &\to \mathsf{Nor}, \\ \mathrm{id}\colon \mathsf{Nor} &\to \mathsf{Nor}, \end{aligned}$$

where id is $\mathrm{id}_{\mathsf{Nor}}$ for short. Then the family of natural embedding

$$\{\iota_X\colon X \to X^{**}, x \mapsto \hat{x} \mid X \in \mathrm{Ob}(\mathsf{Nor})\}$$

is a natural transformation from id to $**$.

Proof. See Corollary 11.2. $\square$

C Uniqueness of Completion

To ensure the uniqueness of completion, we need to look at 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 $\iota\colon X \hookrightarrow \widehat{X}$ keeps norms with a dense image.

Theorem C.1 (Uniqueness of completion)**.** The completion of a linear normed space X is unique up to an unique isometry (that conincides with the two inclusions). That is, if $\widehat{X}, Y$ with isometric inclusion map ι, ι' respectively are completions of X, then the following diagram commutes

$$\begin{tikzcd} & X \arrow[dl, hook', "\iota"'] \arrow[dr, hook, "\iota'"] & \\ \widehat{X} \arrow[rr, dashed, "\exists\varphi"'] & & Y \end{tikzcd}$$

Proof. Consider the corestriction of ι, that is $\iota_0 := \iota|^{\iota(X)}$. Clearly ι_0 is an isometry from X to $\iota(X)$(which is dense in $\widehat{X}$). Now we define a map φ_0 by the following diagram (i.e. $\varphi_0 := \iota' \circ (\iota_0)^{-1}$)

$$\begin{tikzcd} & X \arrow[dr, "\iota'"] & \\ \iota(X) \arrow[ur, "(\iota_0)^{-1}"] \arrow[rr, "\varphi_0"] & & Y \end{tikzcd}$$

Now φ_0 is linear and keeps norm. Since $\iota(X)$ is dense in $\widehat{X}$, Y is complete and φ_0 is continuous (φ_0 keeps norm and hence is continuous), we can extend φ_0 to a continuous map $\varphi\colon \widehat{X} \to Y$ (see this exercise, to prove which, it suffices to displace $X_0, \mathbb{K}$ with $\iota(X), Y$ respectively).

To show that φ is an isometry, we should show that:

1. φ is linear;
2. φ keeps norm.
3. φ is surjective;

First, φ is linear. We have proved this in the exercise.

Second, φ 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 $(x_n)_{n\in\mathbb{N}} \to x$. Then

$$\begin{aligned}
\|x\| &= \left\|\lim_n x_n\right\| \\
&= \lim_n \|x_n\| && \text{(continuity of } \| \ \| \text{)} \\
&= \lim_n \|\varphi_0(x_n)\| && (\varphi_0 \text{ is an isometry}) \\
&= \lim_n \|\varphi(x_n)\| && (\varphi|_{\iota(X)} = \varphi_0) \\
&= \left\|\lim_n \varphi(x_n)\right\| && \text{(continuity of } \| \ \| \text{)} \\
&= \|\varphi(x)\| && \text{(continuity of } \varphi\text{)}.
\end{aligned}$$

Thirdly, φ is surjective. $\forall y \in Y$, by the density of $\iota'(X)$, $\exists (y_n)_{n\in\mathbb{N}} \subseteq \iota'(X)$ such that $(y_n)_{n\in\mathbb{N}} \to y$. 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 φ 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, φ is surjective.

Above all, φ is an isometry. If there is another isometry $\phi\colon \widehat{X} \to Y$ such that the diagram commutes, then $\varphi|_{\iota(X)} = \phi|_{\iota(X)} = \varphi_0$. Then φ and ϕ conincide on a dense subset of $\widehat{X}$ and hence $\varphi = \phi$. □

In fact, a completion of a linear normed space is a final object of some category and hence is unique up to a unique isomorphism, see [4, Theorem 2.6.1].

References

[1] S. Axler. *Measure, Integration & Real Analysis*. Graduate Texts in Mathematics. Springer International Publishing, 2019. ISBN: 9783030331429.

[2] Rami Shakarchi Elias M. Stein. *Fourier analysis: an introduction*. PLA01, PUP. Princeton Lectures in Analysis, Volume 1. Princeton University Press, 2003. ISBN: 9780691113845.

[3] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.

[4] A. Ya. Helemskii. *Lectures and exercises on functional analysis*. Translations of mathematical monographs 233. AMS Bookstore, 2006. ISBN: 9780821840986.

[5] J.L. Kelley. *General Topology*. Graduate Texts in Mathematics. Springer New York, 1975.

[6] 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.