Functional analysis class note
(泛函分析笔记)

Huarui ZHOU

December 14, 2024

Contents

Chapter 1

Foundations

1.1 Norm and normed vector space

Definition 1.1.1. 1. A normed vector space $(X, \|\cdot\|)$ is a pair consisting a vector space X over $\mathbb{F}$ and a function called norm $\|\cdot\| : X \to \mathbb{R}_{\geq 0}$ s.t.

(1) $\|v\| = 0$ if and only if $v = 0$

(2) for any $c \in \mathbb{F}$ and $v \in X$, we have $\|cv\| = |c| \cdot \|v\|$

(3) for any $u, v \in X$, we have $\|u + v\| \leq \|x\| + \|v\|$.

2. A normed vector space $(X, \|\cdot\|)$ defines a metric $d : X \times X \to \mathbb{R}_{\geq 0}$ by

$$d(u, v) := \|u - v\|,$$

which then induces a metric topology.

3. A complete normed vector space is called a Banach space.

Example 1.1.2. 1. For any $1 \leq p < \infty$, define

$$\ell^p := \{\{x_i\}_{i=1}^{\infty} \subseteq \mathbb{R} : \sum_{i=1}^{\infty} |x_i|^p < \infty\}.$$

Then ℓ^p is a Banach space with the norm defined by

$$\|\{x_i\}_{i=1}^{\infty}\|_p := \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}.$$

2. Define

$$\ell^\infty := \{\{x_i\}_{i=1}^{\infty} \subseteq \mathbb{R} : \sup_{i \geq 1} |x_i| < \infty\}.$$

Then ℓ^∞ is a Banach space with the norm defined by

$$\|\{x_i\}_{i=1}^{\infty}\|_\infty := \sup_{i \geq 1} |x_i|.$$

3. Let X be a topological space, define

$$C_b(X) := \{f : X \to \mathbb{R} \text{ continuous and bounded}\},$$

then $C_b(X)$ is a Banach space with the supremum norm

$$\|f\|_\infty := \sup_{x\in X} |f(x)|, \quad \forall f \in C_b(X).$$

Proposition 1.1.3. $\|\cdot\| : X \to \mathbb{R}$ *is continuous.*

Proof. For $u, v \in X$,

$$\|u\| = \|v + (u - v)\| \le \|v\| + \|u - v\|,$$

thus $\|u\| - \|v\| \le \|u - v\|$, similarly, $\|v\| - \|u\| \le \|u - v\|$, then we have

$$|\, \|u\| - \|v\| \,| \le \|u - v\|,$$

which means $\|\cdot\|$ is Lipschitz continuous and hence continuous. □

Definition 1.1.4. Suppose $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are two normed vector spaces.

1. A function $A : X \to Y$ is called bounded if there is a $c \ge 0$, s.t. for any $w \in X$,

$$\|Aw\|_Y \le c \|w\|_X.$$

2. The smallest such c is called the operator norm of A, denoted as $\|A\|$, i.e.

$$\|A\| = \sup_{w\in X\setminus\{0\}} \frac{\|Aw\|_Y}{\|w\|_X}.$$

3. Define $\mathcal{L}(X, Y) = \{A : X \to Y \text{ s.t. } A \text{ is linear and bounded}\}$.

Proposition 1.1.5. *Suppose $\|\cdot\|$ is the operator norm defined on $\mathcal{L}(X, Y)$, then*

1. *For any $A \in \mathcal{L}(X, Y)$ and $w \in X$,*

$$\|Aw\|_Y \le \|A\| \cdot \|w\|_X.$$

2. *The operator norm is a norm. Thus $(\mathcal{L}(X, Y), \|\cdot\|)$ is a normed vector space. The resulting metric topology is called the uniform operator topology.*

3. *For any $A \in \mathcal{L}(X, Y)$,*

$$\begin{aligned}\|A\| &= \sup_{w\in X\setminus\{0\}} \frac{\|Aw\|_Y}{\|w\|_X} \\ &= \sup_{w\in X, \|w\|_X \le 1} \|Aw\|_Y \\ &= \sup_{w\in X, \|w\|_X = 1} \|Aw\|_Y.\end{aligned}$$

Theorem 1.1.6. *Suppose $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are two normed vector spaces, $A : X \to Y$ is linear, then TFAE*

1. *A is bounded*

2. *A is continuous*

3. *A is continuous at 0.*

Proof. $1 \implies 2$. Suppose A is bounded, then for any $u, v \in X$,

$$d_Y(Au, Av) = \|Au - Av\|_Y = \|A(u-y)\|_Y \leq \|A\| \cdot \|u-y\|_X = \|A\| \, d_X(u, y),$$

thus A is Lipschitz continuous, hence continuous.
$2 \implies 3$: Obvious.
$3 \implies 1$: Suppose A is continuous at 0. Take $\varepsilon = 1$, then there is a $\delta > 0$ s.t. for any $\|u\|_X < \delta$, we have $\|Au\|_Y < 1$. Let $w \in X \setminus \{0\}$,

$$\|Aw\|_Y = \frac{2\|w\|_X}{\delta} \left\| A(\frac{\delta w}{2\|w\|_X}) \right\|_Y \leq \frac{2}{\delta} \cdot \|w\|_X \,,$$

where $\delta w / (2\|w\|_X)$ has the norm $\dfrac{\delta}{2} < \delta$. □

Corollary 1.1.7. *Suppose X, Y are normed vector spaces, and $A \in \mathcal{L}(X, Y)$. Then* $\mathrm{Ker}\,(A)$ *is a closed subspace of X.*

Proof. Let $\{x_k\}_{k=1}^{\infty}$ is a convergent sequence in $\mathrm{Ker}\,(A)$ and $x \in X$ is the limit, then by the continuity of A,

$$Ax = A(\lim_{k\to\infty} x_k) = \lim_{k\to\infty} Ax_k = 0,$$

thus $x \in \mathrm{Ker}\,(A)$, and hence $\mathrm{Ker}\,(A)$ is closed. □

1.2 Finite-dimensional normed vector space

Definition 1.2.1. Suppose $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms on X. They are called equivalent if there is $0 < c_1 \leq c_2$ s.t. for any $w \in X$,

$$c_1 \|w\|_1 \leq \|w\|_2 \leq c_2 \|w\|_1 \,.$$

Theorem 1.2.2. *Suppose X is a finite-dimensional vector space over $\mathbb{F}(= \mathbb{R}$ or $\mathbb{C})$. Then any two norms on X are equivalent.*

Proof. 1. Choose the standard basis $\{e_j\}_{j=1}^n$ for X, then for any $x = \sum_{j=1}^n x_j e_j$, define the norm $\|\cdot\|_2$ on X by

$$\|x\|_2 = \sqrt{\sum_{j=1}^n x_j^2}.$$

Suppose $\|\cdot\|$ is any other norm on X, then

$$\|x\| = \left\|\sum_{j=1}^{n} x_j e_j\right\| \leq \sum_{j=1}^{n} \|x_j e_j\| = \sum_{j=1}^{n} |x_j| \|e_j\| \leq \sqrt{\sum_{j=1}^{n} x_j^2}\sqrt{\sum_{j=1}^{n} \|e_j\|^2} = c_2 \|x\|_2 ,$$

where $c_2 = \sqrt{\sum_{j=1}^{n} \|e_j\|^2}$ is a constant.
2. For the other inequality, consider

$$S = \{v \in X : \|v\|_2 = 1\}.$$

Let $n = \dim X$, then $S \subseteq \mathbb{R}^n$, easy to check S is closed and bounded, therefore compact by Heine-Borel theorem. And $\|\cdot\| : S \to \mathbb{R}$ is continuous by Proposition 1.1.3. Thus $\|\cdot\|$ achieves its minimum c_1 on S. For any non-zero $w \in X$, $w/\|w\|_2 \in S$, then

$$\left\|\frac{w}{\|w\|_2}\right\| \geq c_1,$$

i.e. $\|w\| \geq c_2 \|w\|_2$. □

Corollary 1.2.3. *Every finite-dimensional normed vector space is Banach.*

Corollary 1.2.4. *Every finite-dimensional subspace of a normed vector space is closed.*

Proof. By Corollary 1.2.3, any finite-dimensional normed vector space is Banach, thus complete, and every complete subspace is closed. □

Corollary 1.2.5. *Suppose $(X, \|\cdot\|)$ is a finite-dimensional normed vector space. Then $K \subseteq X$ is compact if and only if K is closed and bounded.*

Corollary 1.2.6. *Suppose $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are two normed vector spaces and X is finite-dimensional. Then every linear map $A : X \to Y$ is continuous.*

Proof. Suppose $A : X \to Y$ is linear, define a new norm $\|\cdot\|_A$ on X by

$$\|x\|_A = \|x\|_X + \|Ax\|_Y , \quad \forall x \in X.$$

Since X is finite-dimensional, $\|\cdot\|_A$ is equivalent to $\|\cdot\|_X$, then there is $c \geq 0$ s.t.

$$\|Ax\|_Y \leq \|x\|_A \leq c \|x\|_X ,$$

i.e. A is bounded and hence continuous. □

Theorem 1.2.7. *Suppose $(X, \|\cdot\|)$ is a normed vector space, let*

$$\bar{B} = \{v \in X : \|v\| \leq 1\}, \quad S = \{v \in X : \|v\| = 1\}.$$

Then TFAE:

1. $\dim X < \infty$.
2. $\bar{B}$ *is compact.*
3. S *is compact.*

To prove Theorem 1.2.7, we need the following Lemma.

Lemma 1.2.8 (Riesz's Lemma)**.** *Suppose* $(X, \|\cdot\|)$ *is a normed vector space,* $Y \subsetneq X$ *is a closed subspace, then for any* $\delta \in (0,1)$*, there exists* $w \in X$ *(actually* $w \in X \setminus Y$*) s.t.* $\|w\| = 1$ *and*

$$\|w - y\| \geq 1 - \delta, \quad \forall y \in Y.$$

Proof. 1. Since $Y \subsetneq X$, we can find $w_0 \in X \setminus Y$.
2. Since Y is closed, $X \setminus Y$ is open, then there is an open ball $B_r(w_0)$ with radius $r > 0$ s.t. $B_r(x) \subseteq X \setminus Y$ i.e. $B_r(w_0) \cap Y = \varnothing$. Then

$$d = \inf_{y \in Y} \|w_0 - y\| > r > 0.$$

3. By the definition of inf (d is the maximal lower bound of $\|w_0 - y\|$ for $y \in Y$, any number greater than d is no longer a lower bound), there is $y_0 \in Y$ s.t.

$$\|w_0 - y_0\| \leq \frac{d}{1-\delta}.$$

4. Let $w = \frac{w_0 - y_0}{\|w_0 - y_0\|}$, then $\|w\| = 1$ and for any $y \in Y$,

$$\begin{aligned}
\|w - y\| &= \left\| \frac{w_0 - y_0}{\|w_0 - y_0\|} - y \right\| \\
&= \frac{1}{\|w_0 - y_0\|} \cdot \|w_0 - (y_0 + y\,\|w_0 - y_0\|)\| \\
&\geq \frac{d}{\|w_0 - y_0\|} \qquad \text{(since } y_0 + y\,\|w_0 - y_0\| \in Y \text{ and } d \text{ is a lower bound)} \\
&\geq \frac{d}{d/(1-\delta)} \\
&= 1 - \delta.
\end{aligned}$$

□

Proof (Theorem 1.2.7). 1 $\implies$ 2 $\implies$ 3 is clear. Left to show 3 $\implies$ 1. We will prove by contradiction.
1. Assume S is compact and X is infinite-dimensional. Let $x_1 \in S$ and $Y_1 = \mathrm{Span}(\{x_1\})$. Y_1 is closed by Corollary 1.2.4. By Riesz's Lemma, we can find $x_2 \in X \setminus Y_1$ s.t.

$$\|x_2\| = 1 \quad \text{and} \quad \|x_1 - x_2\| \geq \frac{1}{2},$$

it's clear $x_2 \in S$.
2. Inductively, suppose we have found $\{x_1, \cdots, x_n\} \subseteq S$ s.t.

$$\|x_i - x_j\| \geq \frac{1}{2}, \quad \forall 1 \leq i \neq j \leq n.$$

Let $Y_n = \text{Span}(\{x_1, x_2, \cdots, x_n\})$, Y_n is closed thus by Riesz's Lemma, we can find $x_{n+1} \in X \setminus Y_n$ s.t.

$$\|x_{n+1}\| = 1 \quad \text{and} \quad \|x_{n+1} - x_i\| \geq \frac{1}{2} \quad \forall 1 \leq i \leq n.$$

3. In this way, we can construct a sequence $\{x_j\}_{j=1}^{\infty} \subseteq S$ (infinite-dimension guarantees that we can find infinitely many x_j) s.t.

$$\|x_i - x_j\| \geq \frac{1}{2}, \quad \forall i \neq j,$$

which obviously has no Cauchy subsequences thus no convergent subsequences. This contradicts that S is compact. □

Remark. This theorem tells us compactness may be lost in infinite-dimensional normed vector spaces.

1.3 Quotient space

Definition 1.3.1. Let $(X, \|\cdot\|)$ be a normed vector space, and $Y \subseteq X$ is a closed subspace.

1. Define equivalence relation $\sim$ on X by

$$x \sim y \quad \Longleftrightarrow x - y \in Y,$$

   and denote $[x] := \{y \in X : y \sim x\}$ to be the equivalence class containing x.

2. Define the quotient space by $X/Y := \{[x] : x \in X\}$.

3. Define a norm $\|\cdot\|_{X/Y}$ on X/Y by

$$\|[x]\|_{X/Y} = \inf_{y \in Y} \|x + y\|_X .$$

Lemma 1.3.2. *$\|\cdot\|_{X/Y}$ defined above is a norm on X/Y.*

1.4 Dual space

Theorem 1.4.1. *Suppose X is a normed vector space and Y is a Banach space. Then $\mathcal{L}(X, Y)$ is a Banach space.*

Proof. 1. Suppose $\{A_n\}_{n=1}^\infty \subseteq \mathcal{L}(X,Y)$ is a Cauchy sequence, our goal is to show its limit is also contained in $\mathcal{L}(X,Y)$.
2. For any $w \in X$, $\{A_n w\}_{n=1}^\infty \subseteq Y$ is also a Cauchy sequence because

$$\|A_n w - A_m w\|_Y \le \|A_n - A_m\| \, \|w\|_X .$$

Since Y is Banach, $\{A_n w\}_{n=1}^\infty$ has a limit in Y, denoted as

$$A_\infty w := \lim_{n\to\infty} A_n w.$$

3. We can pointwise define $A_\infty : X \to Y$ by $w \mapsto A_\infty w$. Then A_∞ is linear.
4. $A_n \to A_\infty$ w.r.t. operator topology.
Since $\{A_n\}_{n=1}^\infty$ is Cauchy, for any $\varepsilon > 0$, there is $N_\varepsilon \in \mathbb{Z}_+$ s.t. for any $m, n \ge N_\varepsilon$,

$$\|A_n - A_m\| < \varepsilon.$$

Thus

$$\|A_\infty w - A_n w\|_Y \le \|A_\infty w - A_m w\|_Y + \|A_m w - A_n w\|_Y < \|A_\infty w - A_m w\|_Y + \varepsilon \|w\|_X ,$$

let $m \to \infty$, we have

$$\|A_\infty w - A_n w\|_Y < \varepsilon \|w\|_X ,$$

i.e. $\|A_\infty - A_n\| < \varepsilon$.
5. A_∞ is bounded. From Step 4 and $A_{N_\varepsilon} \in \mathcal{L}(X,Y)$ hence bounded, we have

$$\|A_\infty w\|_Y \le \|A_\infty w - A_{N_\varepsilon} w\|_Y + \|A_{N_\varepsilon} w\|_Y \le \varepsilon \|w\|_X + \|A_{N_\varepsilon}\| \cdot \|w\|_X = (\varepsilon + \|A_{N_\varepsilon}\|) \|w\|_X . \quad \square$$

Definition 1.4.2. Suppose X is a normed vector space, $X^* := \mathcal{L}(X, \mathbb{R})$ is called the dual space of X.

Remark. From Theorem 1.4.1, although X may not be Banach, X^* is always a Banach space.

Definition 1.4.3. Suppose $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are two normed vector spaces.

1. A map $A : X \to Y$ is called isometric (or an isometric embedding) if it preserves the norm, i.e. $\|Aw\|_Y = \|w\|_X$ for all $w \in X$.
2. An isometric map $A : X \to Y$ is called an isometry if it is bijective.
3. A linear isometry $A : X \to Y$ is called an isometric isomorphism.

Remark. 1. Any linear isometric map is continuous because it is 1-Lipschitz.

2. Any linear isometric map is injective. Suppose there is $w_1, w_2 \in X$ s.t. $Aw_1 = Aw_2$, then $A(w_1 - w_2) = 0$. Since A is isometric, we have $0 = \|A(w_1 - w_2)\|_Y = \|w_1 - w_2\|_X$, then $w_1 = w_2$.

3. An isometric isomorphism gives an equivalence relation.
4. An isometric isomorphism preserves all necessary properties between two normed vector spaces, so we can regard two normed vector spaces as the same if there is an isometric isomorphism between them.

Example 1.4.4. $(\mathbb{R}^n)^* = \mathbb{R}^n$

Example 1.4.5. For any $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, we have $[L^p(\mu)]^* = L^q(\mu)$.

Example 1.4.6. $[L^1(\mu)]^* = L^\infty(\mu)$.

Example 1.4.7. For any $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, we have $(\ell^p)^* = \ell^q$.

Proof. Define $\phi : \ell^q \to (\ell^p)^*$ as follows, for any $x = \{x_i\}_{i=1}^\infty \in \ell^q$, $y = \{y_i\}_{i=1}^\infty \in \ell^p$,

$$\phi(x)(y) := \sum_{i=1}^\infty x_i y_i.$$

Then ϕ is well-defined ($\phi(x) \in (\ell^p)^*$), linear, isometric and bijective, i.e. it is an isometric isomorphism. □

Example 1.4.8. $(\ell^1)^* = \ell^\infty$.

Proof. Define a map $\phi : \ell^\infty \to (\ell^1)^*$ as follows, for any $x = \{x_i\}_{i=1}^\infty \in \ell^\infty$, $y = \{y_i\}_{i=1}^\infty \in \ell^1$,

$$\phi(x)(y) := \sum_{i=1}^\infty x_i y_i.$$

Example 1.4.9. let $c_0 := \{(x_1, x_2, \cdots) \in \mathbb{R}^\mathbb{N} : x_i \to 0\} \subseteq \ell^\infty$. Then $(c_0)^* = \ell^1$.

1.5 Hilbert space

Definition 1.5.1. Let H be a real vector space.

1. A bilinear map $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{R}$ is called an inner product if it is
   (1) symmetric: $\langle x, y \rangle = \langle y, x \rangle$ for all $x, y \in H$;
   (2) and positive definite: $\langle x, x \rangle > 0$ for all $x \in H \setminus \{0\}$.
2. $(H, \langle \cdot, \cdot \rangle)$ is called an inner product space (often denoted as H for short).
3. Define a norm on an inner product space by $\|x\| = \sqrt{\langle x, x \rangle}$.
4. An inner product space $(H, \langle \cdot, \cdot \rangle)$ is called a Hilbert space if $(H, \|\cdot\|)$ is a Banach space.

Proposition 1.5.2. *Suppose H is an inner product space. Then for any $x, y \in H$,*

1. *Polarization identity:* $2\langle x, y\rangle = \|x+y\|^2 - \|x\|^2 - \|y\|^2$.
2. *Parallelogram law:* $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + \|y\|^2$.
3. *Cauchy-Schwarz inequality:* $|\langle x, y\rangle| \leq \|x\|\,\|y\|$.
4. *Triangle inequality:* $\|x+y\| \leq \|x\| + \|y\|$.

Remark. From 4, $\|\cdot\|$ is indeed a norm on H.

Theorem 1.5.3 (Riesz)**.** *Suppose H is a Hilbert space. The map $\Lambda : H \to H^*$ defined by $\Lambda(y) = \langle \cdot, y\rangle$ is an isometric isomorphism.*

Definition 1.5.4. Suppose Ω is a subset of a vector space. Ω is called convex if for any $u, v \in \Omega$ and $t \in [0,1]$, we have

$$tu + (1-t)v \in \Omega.$$

Example 1.5.5. Suppose X is a vector space, then

1. X and $\varnothing$ are convex.
2. Suppose $\|\cdot\|$ is a norm defined on X, then for any $x_0 \in X$, the unit ball

$$B_1(x_0) := \{x \in X : \|x - x_0\| < 1\}$$

is convex.

3. Any subspace of X is convex.

Proof. For any $u, v \in B_1(x_0)$ and $t \in [0,1]$,

$$\|tu + (1-t)v - x_0\| \leq \|tu - tx_0\| + \|(1-t)v - (1-t)x_0\| = t\|u - x_0\| + (1-t)\|v - x_0\| < 1.$$

□

Lemma 1.5.6. *Suppose H is a Hilbert space and $K \subseteq H$ is a non-empty, closed and convex subset. Then there is a unique element $x_0 \in K$ s.t. $\|x_0\| \leq \|x\|$ for all $x \in K$.*

Proof. 1. Since $\{\|x\| : x \in K\}$ has a lower bound 0, it must have an inf, let

$$\delta = \inf\{\|x\| : x \in K\} \geq 0.$$

By the definition of inf, for any $n \in \mathbb{Z}_+$, there is $x_n \in K$ s.t.

$$\delta \leq \|x_n\| < \delta + \frac{1}{n},$$

then there exists a sequence $\{x_n\}_{n=1}^{\infty} \subseteq K$ s.t.

$$\lim_{n\to\infty} \|x_n\| = \delta.$$

2. Claim: $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence.
Given $\varepsilon > 0$, there is $N \in \mathbb{Z}_+$, s.t. for any $n \geq N$,

$$\|x_n\|^2 < \delta^2 + \frac{\varepsilon^2}{4}.$$

(solving the inequality $\delta + \frac{1}{N} < \sqrt{\delta^2 + \frac{\varepsilon^2}{4}}$ gives the value of N). For any $m, n \geq N$, since K is convex, we have

$$\frac{x_m + x_n}{2} \in K,$$

thus

$$\left\| \frac{x_m + x_n}{2} \right\| \geq \delta.$$

Then

$$\|x_m - x_n\|^2 = 2\|x_m\|^2 + 2\|x_n\|^2 - \|x_m + x_n\|^2 < 4(\delta^2 + \frac{\varepsilon^2}{4}) - 4\delta^2 = \varepsilon^2,$$

i.e. $\|x_m - x_n\| < \varepsilon$ and hence $\{x_n\}_{n=1}^{\infty}$ is Cauchy.
3. Since H is Hilbert and K is closed, K is complete. Then there is $x_\infty \in K$ s.t. $x_n \to x_\infty$. Therefore by the continuity of norm,

$$\|x_\infty\| = \left\| \lim_{n\to\infty} x_n \right\| = \lim_{n\to\infty} \|x_n\| = \delta.$$

4. Uniqueness. Suppose there is $y \in K$ with $\|y\| = \delta$. Then $(y + x_\infty)/2 \in K$, thus

$$\|y + x_\infty\| \geq 2\delta.$$

Then

$$\|y - x_\infty\|^2 = 2\|y\|^2 + 2\|x_\infty\|^2 - \|y + x_\infty\|^2 \leq 2\delta^2 + 2\delta^2 - 4\delta^2 = 0,$$

which implies $y = x_\infty$. □

Proof of Theorem 1.5.3. 1. It's clear $\Lambda : H \to H^*$ is linear, we will show it is also isometric. For any $y \in H$, recall $\Lambda(y)(\cdot) = \langle \cdot, y \rangle$, then

$$\|\Lambda(y)\| = \sup_{x \in H\setminus\{0\}} \frac{|\langle x, y\rangle|}{\|x\|} \leq \sup_{x \in H\setminus\{0\}} \frac{\|x\|\,\|y\|}{\|x\|} = \|y\|.$$

Since $\Lambda(y) \in H^*$ is bounded, we have

$$\|y\|^2 = |\Lambda(y)(y)| \leq \|\Lambda(y)\|\,\|y\|,$$

thus $\|y\| \leq \|\Lambda(y)\|$. Therefore $\|\Lambda(y)\| = \|y\|$ for any $y \in H$, i.e. Λ is isometric.
2. Since the linear isometric map is already injective, we only need to show Λ is surjective. If $\phi = 0$, then $\phi = \Lambda(0)$. Let $\phi \in H^* \setminus \{0\}$ and $K = \{x \in H : \phi(x) = 1\}$. Then K is:

(1) Non-empty. Since $\phi \neq 0$, there is $\xi \in H$ s.t. $\phi(\xi) \neq 0$. Let $x = \xi/\phi(\xi)$, then $\phi(x) = 1$.

(2) Closed. Because ϕ is continuous and $\{1\}$ is closed.

(3) Convex. For any $u, v \in K$, and $t \in [0, 1]$,

$$\phi(tu + (1-t)v) = t\phi(u) + (1-t)\phi(v) = t + (1-t) = 1,$$

thus $tu + (1-t)v \in K$.

By Lemma 1.5.6, there is $x_0 \in K$ s.t. $\|x_0\| \leq \|x\|$ for any $x \in K$.
3. Claim: $x_0 \perp \operatorname{Ker}\phi$.
Let $y \in \operatorname{Ker}\phi$, we need to show $\langle x_0, y\rangle = 0$. For any $t \in \mathbb{R}$,

$$x_0 + ty \in K,$$

from Step 2,

$$\|x_0\|^2 \leq \|x_0 + ty\|^2 = \|x_0\|^2 + t^2 \|y\|^2 + 2t\langle x_0, y\rangle.$$

Let $f(t) = \|x_0 + ty\|^2$, then $f(t)$ attains its minimum at $t = 0$, we have

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} f(0) = 2\langle x_0, y\rangle,$$

thus $\langle x_0, y\rangle = 0$.
4. For any $x \in H$, our goal is to find $y \in H$, s.t. $\phi(x) = \Lambda(y)(x)$, i.e. Λ is surjective. Recall $\phi(x_0) = 1$, so

$$\phi(x - \phi(x)x_0) = \phi(x) - \phi(x)\phi(x_0) = 0,$$

which means $x - \phi(x)x_0 \in \operatorname{Ker}\phi$. By Step 3,

$$\langle x_0, x - \phi(x)x_0\rangle = 0,$$

i.e.

$$\langle x_0, x\rangle = \langle x_0, \phi(x)x_0\rangle = \phi(x)\|x_0\|^2,$$

thus

$$\phi(x) = \langle x, \frac{x_0}{\|x_0\|^2}\rangle = \Lambda(\frac{x_0}{\|x_0\|^2})(x).$$

□

Corollary 1.5.7. *Suppose H is a Hilbert space, then for any $\phi \in H^*$, there is a unique $y \in H$, s.t.*

$$\phi(x) = \langle x, y\rangle, \quad \forall x \in H.$$

Definition 1.5.8. Let $S \subseteq H$ be a subset of a Hilbert space. Define

$$S^\perp = \{x \in H : \langle x, y\rangle = 0, \ \forall y \in S\}.$$

Remark. By Theorem 1.5.3, $S^\perp$ is the same (in the sense of isometric isomorphism) as the annihilator of S, i.e. $\{\phi \in H^* : \phi(y) = 0, \ \forall y \in S\}$.

Corollary 1.5.9. *Suppose H is a Hilbert space and $E \subseteq H$ is a closed subspace, then*

$$H = E \oplus E^{\perp}.$$

Proof. It suffices to prove $H = E + E^{\perp}$ and $E \cap E^{\perp} = \{0\}$. It's clear $0 \in E \cap E^{\perp}$. On the other hand, for any $x \in E \cap E^{\perp}$, $\langle x, x \rangle = 0$, which implies $x = 0$, thus $E \cap E^{\perp} = \{0\}$.
For any $y \in E$, $y + E$ is non-empty, closed and convex, then by Lemma 1.5.6, there is a unique element $\xi \in E$, s.t.

$$\|y + \xi\| \leq \|x\| \quad \forall x \in y + E.$$

For any $e \in E$ and $t \in \mathbb{R}$, $y + \xi + te \in y + E$. Let

$$f(t) = \|y + \xi + te\| = \|y + \xi\|^2 + t^2 \|e\| + t\langle y + \xi, e \rangle,$$

since f attains its minimum at $t = 0$, we have

$$0 = f'(0) = \langle y + \xi, e \rangle,$$

i.e. $y + \xi \in E^{\perp}$, therefore $y = (-\xi) + (y + \xi) \in E + E^{\perp}$. □

Corollary 1.5.10. *Suppose H is a Hilbert space and $E \subseteq H$ is a closed subspace, then*

$$E = (E^{\perp})^{\perp}.$$

Proof. 1. Let $x \in (E^{\perp})^{\perp} \subseteq H$, $x = x_1 + x_2$ for some $x_1 \in E$ and $x_2 \in E^{\perp}$, which implies

$$\langle x_1, x_2 \rangle = 0, \quad \langle x, x_2 \rangle = 0,$$

then $\langle x_2, x_2 \rangle = \langle x - x_1, x_2 \rangle = 0$, i.e. $x_2 = 0$ and hence $x = x_1 \in E$. For the other direction, since $E^{\perp}$ is always closed (See Definition 2.6.1), $H = E^{\perp} \oplus (E^{\perp})^{\perp}$. Then for any $x \in E$, $x = x_1 + x_2$ for some $x_1 \in E^{\perp}$ and $x_2 \in (E^{\perp})^{\perp}$, then

$$\langle x, x_1 \rangle = 0, \quad \langle x_1, x_2 \rangle = 0,$$

so $x_1 = 0$ and $x = x_2 \in (E^{\perp})^{\perp}$. □

The following is an application of Riesz representation theorem.

Theorem 1.5.11 (Lax-Milgram)**.** *Suppose H is a real Hilbert space. Let $B : H \times H \to \mathbb{R}$ be a bilinear and there exists $\alpha, \beta > 0$ s.t. for any $u, v \in H$,*

1. *$|B(u, v)| \leq \alpha \|u\| \|v\|$;*

2. *$B(u, u) \geq \beta \|u\|^2$.*

Then for any $\phi \in H^*$, there is a unique $u \in H$ s.t. for any $v \in H$,

$$B(u, v) = \varphi(v).$$

Proof. 1. For a fixed $u \in H$, define $T_u : H \to \mathbb{R}$ by $T_u(\cdot) = B(u, \cdot)$, then T_u is linear and bounded, i.e. $T_u \in H^*$. By Riesz representation theorem (1.5.3), there is a unique $w \in H$, s.t.

$$T_u(\cdot) = \langle w, \cdot \rangle.$$

Define a map $A : H \to H$ by taking u as input and returning w as output, i.e.

$$B(u, \cdot) = \langle Au, \cdot \rangle.$$

2. $A \in \mathcal{L}(H)$.
Let $u_1, u_2 \in H$ and $c \in \mathbb{R}$, then

$$\langle A(cu_1 + u_2), \cdot \rangle = B(cu_1 + u_2, \cdot) = cB(u_1, \cdot) + B(u_2, \cdot) = \langle cAu_1, \cdot \rangle + \langle Au_2, \cdot \rangle = \langle cAu_1 + Au_2, \cdot \rangle,$$

therefore A is linear. For any $u \in H$ s.t. $Au \neq 0$ ($Au = 0$ is trivially bounded), we have

$$\|Au\|^2 = \langle Au, Au \rangle = B(u, Au) \leq \alpha \|u\| \|Au\|,$$

divided by $\|Au\|$, we have $\|Au\| \leq \alpha \|u\|$, thus A is bounded.
3. A is injective.
For any $u \in H$,

$$\beta \|u\|^2 \leq B(u, u) = \langle Au, u \rangle \leq \|Au\| \cdot \|u\|,$$

assume $u \neq 0$, then

$$\beta \|u\| \leq \|Au\|,$$

which is also true for $u = 0$, thus is true for all $u \in H$. For any $u_1, u_2 \in H$ with $u_1 \neq u_2$, we have

$$\|A(u_1 - u_2)\| \geq \beta \|u_1 - u_2\| > 0,$$

then $Au_1 \neq Au_2$, i.e. A is injective.
4. $\mathrm{Im}\,(A)$ is closed in H.
Let $\{y_i\}_{i=1}^{\infty} \subseteq \mathrm{Im}\,(A)$ be a convergent sequence (i.e. Cauchy) in H, we want to show $y := \lim_{i\to\infty} y_i \in \mathrm{Im}\,(A)$. Since A is injective, for any $i \geq 1$, we can find a unique $x_i \in H$ s.t. $Ax_i = y_i$. By Step 3,

$$0 \leq \|x_m - x_n\| \leq \frac{1}{\beta} \|Ax_m - Ax_n\|,$$

thus $\{x_i\}_{i=1}^{\infty}$ is Cauchy. Then there is $x \in H$ s.t. $x_i \to x$ by completeness of H. Moreover, by the boundedness of A,

$$\|Ax - y\| = \lim_{i\to\infty} \|A(x - x_i)\| \leq \lim_{i\to\infty} \alpha \|x - x_i\| = 0,$$

therefore $Ax = y$ and $y \in \mathrm{Im}\,(A)$, i.e. $\mathrm{Im}\,(A)$ is closed.
5. $\mathrm{Im}\,(A) = H$ and hence A is bijective.
Since $\mathrm{Im}\,(A)$ is closed, by Corollary 1.5.9,

$$H = \mathrm{Im}\,(A) \oplus \mathrm{Im}\,(A)^{\perp}.$$

If $\mathrm{Im}\,(A) \subsetneq H$, $\mathrm{Im}\,(A)^{\perp} \neq \{0\}$, then there exists $u \in \mathrm{Im}\,(A)^{\perp} \setminus \{0\}$, then we have

$$\|u\|^2 \leq \frac{1}{\beta} B(u,u) = \langle Au, u\rangle = 0,$$

which contradicts $u \neq 0$.
4. For any $\phi \in H^*$, by Riesz representation theorem (1.5.3), there is a unique $w \in H$ s.t.

$$\phi(\cdot) = \langle w, \cdot\rangle.$$

To find $u \in H$ s.t.

$$\langle Au, \cdot\rangle = B(u,\cdot) = \phi(\cdot) = \langle w, \cdot\rangle,$$

i.e. we need to find $u \in H$ s.t. $Au = w$. Since $A : H \to H$ is bijective, there is a unique solution $u = A^{-1}w$. □

Remark. From Step 2 we have $\|A\| \leq \alpha$. From Step 3, for any $w \in H$,

$$\left\|A^{-1}w\right\| \leq \frac{1}{\beta}\left\|A(A^{-1}w)\right\| = \frac{1}{\beta}\|w\|,$$

thus $\left\|A^{-1}\right\| \leq 1/\beta$. This implies $A^{-1} \in \mathcal{L}(H)$. We can also apply the Inverse operator theorem to show $A^{-1} \in \mathcal{L}(H)$.

1.6 Banach algebra

Definition 1.6.1. 1. A Banach algebra is a Banach space $\mathcal{A}$ equipped with a bilinear map called product $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$: $(x,y) \mapsto xy$ s.t. it is

   (i) associative, i.e. for any $a,b,c \in \mathcal{A}$, $(ab)c = a(bc)$;

   (ii) and for any $a,b \in \mathcal{A}$, $\|ab\| \leq \|a\| \cdot \|b\|$.

2. A Banach algebra is commutative if for any $a,b \in \mathcal{A}$, $ab = ba$.

3. A Banach algebra is unital if there exists a unit $\mathbb{1}_{\mathcal{A}} \in \mathcal{A}$ s.t. for any $b \in \mathcal{A}$, $\mathbb{1}_{\mathcal{A}} \cdot b = b$ and $b \cdot \mathbb{1}_{\mathcal{A}} = b$.

4. If $\mathcal{A}$ is unital, then $b \in \mathcal{A}$ is called invertible if there is $x \in \mathcal{A}$ s.t. $bx = xb = \mathbb{1}_{\mathcal{A}}$. x is called the inverse of b, denoted by b^{-1}.

Remark. 1. If they exist, both the unit and the inverse are unique.

2. If $\mathcal{A}$ is unital, the set of invertible elements forms a group.

Example 1.6.2. Suppose X is a Banach space. Let $\mathcal{L}(X) := \mathcal{L}(X,X)$ be the set of endomorphisms of X. Then $\mathcal{A}$ is a Banach algebra with the product given by composition. $\mathcal{L}(A)$ is also unital because the identity map is the unit.

Proof. Only need to show (ii). Let $a, b \in \mathcal{L}(X)$, then

$$\|ab\| = \sup_{x \in X \setminus \{0\}} \frac{\|a \circ b(x)\|_X}{\|x\|_X} \leq \sup_{x \in X \setminus \{0\}} \frac{\|a\| \, \|b(x)\|_X}{\|x\|_X} \leq \|a\| \cdot \|b\| \, . \qquad \square$$

Example 1.6.3. Suppose Ω is a complete metric space. Let $C(\Omega)$ be the set of all continuous functions: $\Omega \to \mathbb{R}$. Then $C(\Omega)$ with the supremum norm is a Banach space. And $C(\Omega)$ with the function product is a commutative and unital Banach algebra.

Example 1.6.4. Let

$$C_0(\mathbb{R}^n) = \{f \in C(\mathbb{R}^n) : \lim_{|x| \to \infty} f(x) = 0, \}$$

then $C_0(\mathbb{R}^n)$ is a commutative Banach algebra but not unital.

Theorem 1.6.5. *Suppose $\mathcal{A}$ is a Banach algebra. Then*

1. *For any $a \in \mathcal{A}$,*

$$r_a := \lim_{n \to \infty} \|a^n\|^{1/n}$$

exists and $r_a \leq \|a\|$. We call r_a the spectral radius of a.

2. *If $\mathcal{A}$ is unital and $a \in \mathcal{A}$ satisfies $r_a < 1$, then $\mathbb{1} - a$ is invertible and*

$$(\mathbb{1} - a)^{-1} = \sum_{k=0}^{\infty} a^k.$$

3. *If $\mathcal{A}$ is unital, let G be the set of all invertible elements in $\mathcal{A}$. Then G forms a group and is an open subset of $\mathcal{A}$. The function $G \to G : b \mapsto b^{-1}$ is continuous.*

Proof. 1. Since $\left\|a^{n+1}\right\| = \|a \cdot a^n\| \leq \|a\| \cdot \|a^n\|$, inductively, we have

$$\|a^n\| \leq \|a\|^n \, , \quad \forall n \geq 1.$$

Then

$$\|a^n\|^{1/n} \leq \|a\| \, ,$$

i.e. $\{\|a^n\|^{1/n}\}_{n=1}^{\infty}$ is a bounded real sequence and $\lim_{n \to \infty} \|a^n\|^{1/n} \leq \|a\|$ if the limit exists. Let $r = \inf_{n \geq 1} \|a^n\|^{1/n}$, by definition of inf, for any $\varepsilon > 0$, there is $m \in \mathbb{Z}_+$ s.t.

$$\|a^m\|^{1/m} < r + \varepsilon.$$

For any $n \in \mathbb{Z}_+$, there is $k, l \geq 0$ s.t. $l < m$ and $n = km + l$. Then

$$\|a^n\|^{1/n} = \left\|a^{km+l}\right\|^{1/n} \leq \|a^m\|^{k/n} \|a\|^{l/n} \leq (r + \varepsilon)^{km/n} \|a\|^{l/n} \, ,$$

then

$$r \leq \liminf_{n \to \infty} \|a^n\|^{1/n} \leq \limsup_{n \to \infty} \|a^n\|^{1/n} = r + \varepsilon,$$

therefore $r = \lim_{n\to\infty} \|a^n\|^{1/n}$.
2.
3. For any $a \in G$, we will show $B_{1/\|a^{-1}\|}(a) \subseteq G$, G is therefore open. Let $b \in B_{1/\|a^{-1}\|}(a)$, i.e. $\|a-b\| < 1/\left\|a^{-1}\right\|$, and $\left\|\mathbb{1} - a^{-1}b\right\| \leq \|a-b\| \left\|a^{-1}\right\| < 1$. Let $c = \mathbb{1} - a^{-1}b$, we have $r_c \leq \|c\| < 1$, then by (2), $a^{-1}b = \mathbb{1} - c \in G$ and hence $b = a(a^{-1}b) \in G$, i.e. G is open.
Next, we will show $G \to G : b \mapsto b^{-1}$ is continuous. It suffice to show for any $a \in G$, it is continuous at a, i.e. for any $\varepsilon > 0$, there is $B_\delta(a) \subseteq G$ s.t. for all $b \in B_\delta(a)$,

$$\left\|a^{-1} - b^{-1}\right\| < \varepsilon.$$

Let $\delta < 1/\left\|a^{-1}\right\|$, then by previous step, $b \in G$, then

$$\left\|b^{-1} - a^{-1}\right\| \leq \frac{\|a-b\| \left\|a^{-1}\right\|^2}{1 - \|a-b\| \left\|a^{-1}\right\|}.$$

If we let

$$\|a-b\| < \frac{\varepsilon}{\left\|a^{-1}\right\|^2 + \varepsilon \left\|a^{-1}\right\|},$$

then we have

$$\left\|b^{-1} - a^{-1}\right\| < \varepsilon,$$

therefore we find

$$\delta = \min\{\frac{\varepsilon}{\left\|a^{-1}\right\|^2 + \varepsilon \left\|a^{-1}\right\|}, 1/\left\|a^{-1}\right\|\}.$$

□

1.7 Baire category theorem

Definition 1.7.1. Suppose (X, d) is a metric space.

1. $A \subseteq X$ is called nowhere dense if
$$(\overline{A})^\circ = \varnothing.$$
2. $A \subseteq X$ is called meager if it is a countable union of nowhere dense sets.
3. A is called non-meager if it is not meager.
4. $A \subseteq X$ is called residual if A^c is meager.

Remark. 1. If X is non-empty, then $\varnothing$ is meager and X is residual.

2. If $X = \varnothing$, then X is both meager and residual.

Lemma 1.7.2. *Suppose (X, d) is a metric space.*

1. *$A \subseteq X$ is nowhere dense if and only if A^c contains a dense open subset.*

2. *If $B \subseteq X$ is meager and $A \subseteq B$, then A is meager.*

3. *If $A \subseteq X$ is non-meager and $A \subseteq B \subseteq X$, then B is non-meager.*

4. *Any countable union of meager sets is meager.*

5. *Any countable intersection of residual sets is residual.*

6. *$A \subseteq X$ is residual if and only if A contains a countable intersection of open subsets of X.*

Proof. 1. Suppose A is nowhere dense, using the fact $(\overline{A})^c = (A^c)^\circ$ and $(A^\circ)^c = \overline{A^c}$, we have

$$X = ((\overline{A})^\circ)^c = \overline{(\overline{A})^c} = \overline{(A^c)^\circ},$$

i.e. $(A^c)^\circ$ is dense in X, we are done since $(A^c)^\circ$ is an open set contained in A^c. For the other direction, suppose A^c contains a dense open subset B. Since $B \subseteq (A^c)^\circ$, $X = \overline{B} \subseteq \overline{(A^c)^\circ}$.
2. 3. 4. Clear.
5.

Lemma 1.7.3. *Suppose (X, d) is a metric space, TFAE*

1. *Every residual subset is dense*

2. *Every non-empty open subset is non-meager*

3. *If $\{A_i\}_{i=1}^\infty$ are subsets of X with $A_i^\circ = \varnothing$, then*

$$\left(\bigcup_{i=1}^\infty U_i\right)^\circ = \varnothing$$

4. *If $\{U_i\}_{i=1}^\infty$ are dense open subsets of X, then $\bigcap_{i=1}^\infty U_i$ is dense.*

Proof. $1 \implies 2$. Let $U \subseteq X$ be a non-empty open set. For any $x_0 \in U$, by definition of openness, there is $\delta > 0$ s.t. $B_\delta(x_0) \subseteq U$, then $B_\delta(x_0) \cap U^c = \varnothing$, which implies U^c is not dense. By statement 1, U^c is not residual, thus not meager.
$2 \implies 3$
$3 \implies 4$
$4 \implies 1$ □

Theorem 1.7.4 (Baire category theorem)**.** *Let (X, d) be a complete metric space. Then $1 \sim 4$ in Lemma 1.7.3 hold and*
5. *Every residual subset is non-meager*
holds.

Proof. $2 \implies 5$. Let $R \subseteq X$ be residual, then by definition, R^c is meager. If R is meager too, then by Lemma 1.7.2, $R \cup R^c = X$ is also meager, however by Statement 2, X is a non-empty open set and thus non-meager.

Completeness $\Longrightarrow$ 4. Step (1). Let $\{U_i\}_{i=1}^{\infty}$ be a sequence of dense open subsets of X. By definition of "dense", we can take $x_0 \in X$ and $\varepsilon_0 > 0$ s.t.

$$B_{\varepsilon_0}(x_0) \cap U_1 \neq \varnothing. \tag{1.1}$$

Then take $x_1 \in B_{\varepsilon_0}(x_0) \cap U_1$, since $B_{\varepsilon_0}(x_0) \cap U_1$ is open, there is $0 < \varepsilon_1 < \frac{1}{2}$ [1] s.t.

$$B_{\varepsilon_1}(x_1) \subseteq B_{\varepsilon_0}(x_0) \cap U_1.$$

By definition of "dense" again, we have $B_{\varepsilon_1}(x_1) \cap U_2$ is non-empty, which is also open, thus we can find $x_2 \in B_{\varepsilon_1}(x_1) \cap U_2$ and ε_2 with $0 < \varepsilon_1 < \frac{1}{4}$ s.t.

$$B_{\varepsilon_2}(x_2) \subseteq B_{\varepsilon_1}(x_1) \cap U_2 \subseteq U_1 \cap U_2.$$

Inductively, for any $k \in Z_+$, we can find $x_k \in X$ and $\varepsilon_k \in \mathbb{R}$ s.t.

$$0 < \varepsilon_k < \frac{1}{2^k}, \quad B_{\varepsilon_k}(x_k) \subseteq \bigcap_{j=1}^{k} U_j, \quad \text{and} \quad B_{\varepsilon_k}(x_k) \subseteq B_{\varepsilon_j}(x_j) \quad \forall 0 \leq j \leq k.$$

(The choice of the sequence $\{x_k\}_{k=1}^{\infty}$ requires the axiom of dependent choice.)
Step (2). $\{x_k\}_{k=1}^{\infty}$ is Cauchy. (why?)
Step (3). By Step (2) and completeness of X, there is $x_\infty \in X$ s.t. $x_k \to x_\infty$. We have (why?)

$$x_\infty \in B_{\varepsilon_k}(x_k) \quad \text{and} \quad x_\infty \in U_k, \quad \forall k \in \mathbb{Z}_{\geq 0}.$$

Therefore

$$x_\infty \in \left(\bigcap_{k=1}^{\infty} U_k \right) \cap B_{\varepsilon_0}(x_0).$$

Since x_0 and ε_0 are chosen arbitrarily, we conclude that $\bigcap_{k=1}^{\infty} U_k$ is dense. $\square$

[1] If $\varepsilon_1 \geq \frac{1}{2}$, any ball with radius $\leq \varepsilon_1$ also satisfies the following condition.

Chapter 2

Principles of functional analysis

2.1 Uniform boundedness principle

Definition 2.1.1. Suppose X is a set, $\{Y_i\}_{i\in I}$ is a sequence of normed vector spaces with the index set I, the function sequence $\{f_i\}_{i\in I}$ with $f_i : X \to Y_i$ is called pointwise bounded if for any $x \in X$,

$$\sup_{i\in I} \|f_i(x)\|_{Y_i} < \infty.$$

Theorem 2.1.2 (Uniform boundedness principle)**.** *Suppose X is a Banach space, $\{Y_i\}_{i\in I}$ is a sequence of normed vector spaces, $A_i \in \mathcal{L}(X, Y_i)$ for each $i \in I$. Suppose $\{A_i\}_{i\in I}$ is pointwise bounded, then*

$$\sup_{i\in I} \|A_i\| < \infty.$$

The proof of Theorem 2.1.2 needs the following lemma.

Lemma 2.1.3. *Suppose (X, d) is a complete, non-empty metric space. For each $i \in I$, $f_i : X \to \mathbb{R}$ is continuous. If $\{f_i\}_{i=1}^{\infty}$ is pointwise bounded, then there is $x_0 \in X$ and $\varepsilon > 0$ s.t.*

$$\sup_{i\in I} \sup_{x\in B_\varepsilon(x_0)} |f_i(x)| < \infty.$$

Proof. 1. For any $i \in I$ and $n \in \mathbb{Z}_+$, define

$$F_{n,i} := \{x \in X : |f_i(x)| \le n\},$$

which is a closed set, thus

$$F_n := \bigcap_{i\in I} F_{n,i} = \{x \in X : \sup_{i\in I} |f_i(x)| \le n\}$$

is also closed. For any $x \in X$, by pointwise boundedness, there is $N \in \mathbb{Z}_+$ s.t.

$$\sup_{i\in I} |f_i(x)| \le N,$$

thus $x \in F_N$ and then

$$X \subseteq \bigcup_{n=1}^{\infty} F_n \subseteq X,$$

i.e. $X = \bigcup_{n=1}^{\infty} F_n$.

2. Since X is non-empty, X is residual, then by Baire category theorem (1.7.4) and completeness, X is non-meager. By the definition of "non-meager", not all F_n is nowhere dense. Suppose F_m is not a nowhere dense set, since F_m is closed,

$$\varnothing \neq (\overline{F_m})^\circ = (F_m)^\circ,$$

which means F_m contains a non-empty open set. Then there is $x_0 \in F_m$ and $\varepsilon_0 > 0$ s.t. $B_{\varepsilon_0}(x_0) \subseteq F_m$. Therefore, for any $x \in B_{\varepsilon_0}(x_0)$,

$$\sup_{i\in I} |f_i(x)| \leq m,$$

then change the order of two sup, we have

$$\sup_{i\in I} \sup_{x\in B_{\varepsilon_0}(x_0)} |f_i(x)| = \sup_{x\in B_{\varepsilon_0}(x_0)} \sup_{i\in I} |f_i(x)| \leq m < \infty.$$

□

Proof of Theorem 2.1.2. Step 1. For each $i \in I$, define $f_i : X \to \mathbb{R}$ by $f_i(x) = \|A_i(x)\|_{Y_i}$, $\forall x \in X$. Then f_i is continuous since A_i and the norm are both continuous. Since $\{A_i\}_{i\in I}$ is pointwise bounded, $\{f_i\}_{i\in I}$ is also pointwise bounded. By Lemma 2.1.3, there is $x_0 \in X$ and $\varepsilon > 0$ s.t.

$$M := \sup_{i\in I} \sup_{x\in B_\varepsilon(x_0)} \|A_i(x)\|_{Y_i} = \sup_{i\in I} \sup_{x\in B_\varepsilon(x_0)} |f_i(x)| < \infty.$$

Step 2. Let $w \in X$ with $\|w\|_X = 1$, then

$$\begin{aligned}
\|A_i(w)\|_{Y_i} &= \frac{1}{\varepsilon} \left\| A[(x_0 + \frac{\varepsilon}{2}w) - (x_0 - \frac{\varepsilon}{2}w)] \right\|_{Y_i} \\
&\leq \frac{1}{\varepsilon} \left\| A(x_0 + \frac{\varepsilon}{2}w) \right\|_{Y_i} + \frac{1}{\varepsilon} \left\| A(x_0 - \frac{\varepsilon}{2}w) \right\|_{Y_i} \\
&\leq \frac{2M}{\varepsilon} = \frac{2M}{\varepsilon} \|w\|_X,
\end{aligned}$$

thus $\|A_i\| \leq 2M/\varepsilon$ for all $i \in I$. □

Example 2.1.4. If a real sequence $a := \{a_n\}_{n=1}^{\infty}$ satisfies

$$\{a_n x_n\}_{n=1}^{\infty} \in c_0$$

for all $\{x_n\}_{n=1}^{\infty} \in c_0$, then $a \in \ell^\infty$.

Proof. Define $\phi : c_0 \to c_0$ by
$$\{x_n\}_{n=1}^\infty \mapsto \{a_n x_n\}_{n=1}^\infty,$$
then $\phi \in \mathcal{L}(c_0, c_0)$. Define $\phi_m : c_0 \to c_0$ by
$$\{x_n\}_{n=1}^\infty \mapsto (a_1 x_1, \cdots, a_m x_m, 0, 0, \cdots).$$

Definition 2.1.5. Suppose X, Y are normed vector spaces and $\{A_i\}_{i=1}^\infty \subseteq \mathcal{L}(X,Y)$. $\{A_i\}_{i=1}^\infty$ is said to converge strongly to $A \in \mathcal{L}(X,Y)$ if for any $x \in X$,
$$Ax = \lim_{i\to\infty} A_i x.$$

Theorem 2.1.6 (Banach-Steinhaus)**.** *Suppose X, Y are Banach spaces and $\{A_i\}_{i=1}^\infty \subseteq \mathcal{L}(X,Y)$. TFAE*

1. *For any $x \in X$, $\{A_i x\}_{i=1}^\infty$ is a convergent sequence.*
2. $\sup_{i\geq 1} \|A_i\| < \infty$ *and there is a dense subset $D \subseteq X$ s.t. for any $x \in D$, $\{A_i x\}_{i=1}^\infty$ is Cauchy.*
3. $\sup_{i\geq 1} \|A_i\| < \infty$ *and there is $A \in \mathcal{L}(X,Y)$ s.t. $A_i \to A$ strongly and*
$$\|A\| \leq \liminf_{i\to\infty} \|A_i\|.$$

Proof. $1 \implies 3$. Suppose for any $x \in X$, $\{A_i x\}_{i=1}^\infty$ is convergent, then $\{\|A_i x\|_Y\}_{i=1}^\infty$ is bounded, i.e. $\{A_i\}_{i=1}^\infty$ is pointwise bounded, then by the Uniform boundedness principle,
$$\sup_{i\geq 1} \|A_i\| < \infty.$$
Define $Ax = \lim_{i\to\infty} A_i x$, then by the continuity of $\|\cdot\|$,
$$\|Ax\|_Y = \lim_{i\to\infty} \|A_i x\|_Y = \liminf_{i\to\infty} \|A_i x\|_Y \leq \liminf_{i\to\infty} \|A_i\| \cdot \|x\|_X \leq \sup_{i\geq 1} \|A_i\| \cdot \|x\|_X,$$
i.e. A is bounded. And A is also linear, thus $A \in \mathcal{L}(X,Y)$, $A_i \to A$ strongly and
$$\|A\| \leq \liminf_{i\to\infty} \|A_i\|.$$
$2 \implies 1$. Fix $x \in X$, we want to show $\{A_i x\}_{i=1}^\infty$ is Cauchy in Y. Since D is dense in X, there is $\{x_k\}_{k=1}^\infty \subseteq D$ s.t. $x_k \to x$. Notice that
$$\begin{aligned}
\|A_m x - A_n x\|_Y &= \|A_m(x - x_k) - A_n(x - x_k) + A_m x_k - A_n x_k\|_Y \\
&\leq \|(A_m - A_n)(x - x_k)\|_Y + \|A_m x_k - A_n x_k\|_Y \\
&\leq \|A_m - A_n\| \|x - x_k\|_X + \|A_m x_k - A_n x_k\|_Y \\
&\leq (\|A_m\| + \|A_n\|) \|x - x_k\|_X + \|A_m x_k - A_n x_k\|_Y \\
&\leq 2(\sup_{i\geq 1} \|A_i\|) \|x - x_k\|_X + \|A_m x_k - A_n x_k\|_Y.
\end{aligned}$$

Since $x_k \to x$, for any $\varepsilon > 0$, there is $N_1 \in \mathbb{Z}_+$ s.t. for all $k \geq N_1$, we have

$$\|x - x_k\|_X < \frac{\varepsilon}{4(\sup_{i\geq 1}\|A_i\|)}.$$

Then choose a $k \geq N$, since $\{A_i x_k\}_{i=1}^{\infty}$ is Cauchy, there is $N_2 \in Z_+$ s.t. for all $m, n \geq N_2$,

$$\|A_m x_k - A_n x_k\|_Y < \frac{\varepsilon}{2}.$$

Therefore $\|A_m x - A_n x\|_Y < \varepsilon$, i.e. $\{A_i x\}_{i=1}^{\infty}$ is Cauchy and hence convergent. □

Corollary 2.1.7 (Bilinear map)**.** *Suppose X, Y, Z are Banach spaces and $B : X \times Y \to Z$ is a bilinear map. TFAE*

1. *B is bounded, i.e. there is $c \geq 0$ s.t.*

$$\|B(x,y)\|_Z \leq c\,\|x\|_X\,\|y\|_Y\,, \quad \forall x \in X, y \in Y.$$

2. *B is continuous.*

3. *For any $x \in X$, $B(x,\cdot) : Y \to Z$ is continuous, and for any $y \in Y$, $B(\cdot,y) : X \to Z$ is continuous.*

Proof. $1 \implies 2$. Locally Lipschitz.
$2 \implies 3$. Clear.
$3 \implies 1$. Assume (3) holds. For any $y \in Y$, define $F_y : X \to Z$ by $F_y(\cdot) = B(\cdot, y)$. For any $x \in X$, define $G_x : Y \to Z$ by $G_x(\cdot) = B(x,\cdot)$. Since G_x is continuous, there exist $c_x \geq 0$ s.t.

$$\|G_x(\cdot)\|_Z \leq c_x\,\|\cdot\|_Y\,.$$

Let $I = \{y \in Y : \|y\|_Y = 1\}$, then for any $x \in X$,

$$\sup_{y\in I}\|F_y(x)\|_Z = \sup_{y\in I}\|B(x,y)\|_Z = \sup_{y\in I}\|G_x(y)\|_Z \leq c_x < \infty,$$

applying Uniform boundedness principle (2.1.2), we have

$$c := \sup_{y\in I}\|F_y\| < \infty.$$

Then for all $y \in I$,

$$\|F_y(x)\|_Z \leq c\,\|x\|_X\,, \quad \forall x \in X.$$

For any $x \in X$, $y \in Y$,

$$\|B(x,y)\|_Z = \|y\|_Y\,B(x, \frac{y}{\|y\|_Y}) \leq c\,\|x\|_X\,\|y\|_Y\,.$$ □

2.2 Open mapping theorem

Definition 2.2.1. Suppose (X, d_X) and (Y, d_Y) are two metric spaces. $f : X \to Y$ is called open if for any open set $U \subseteq X$, $f(U)$ is open in Y.

Theorem 2.2.2 (open mapping theorem)**.** *Suppose X and Y are Banach spaces. Let $A \in \mathcal{L}(X, Y)$ be surjective, then A is open.*

We need two lemmas to prove Theorem 2.2.2.

Lemma 2.2.3. *Suppose X and Y are Banach spaces. Let $A \in \mathcal{L}(X, Y)$ be surjective, then there is $\delta > 0$ s.t.*

$$B_\delta^Y(0) \subseteq \overline{A[B_1^X(0)]}.$$

Proof. Step 0. For any subset W of a vector space and $\lambda \in \mathbb{R}$, define the scaled set as

$$\lambda W := \{\lambda y : y \in W\}.$$

Step 1. Then

$$X = \bigcup_{n=1}^{\infty} B_n^X(0) = \bigcup_{n=1}^{\infty} nB_1^X(0).$$

Surjectivity of A implies

$$Y = A(X) = A\left(\bigcup_{n=1}^{\infty} nB_1^X(0)\right) = \bigcup_{n=1}^{\infty} A[nB_1^X(0)] = \bigcup_{n=1}^{\infty} nA[B_1^X(0)].$$

Step 2. Since Y is residual and complete, Baire category theorem (1.7.4) implies Y is non-meager. Then there is $n_0 \in \mathbb{Z}_+$ s.t. $n_0A[B_1^X(0)]$ is not nowhere dense, i.e.

$$\left(\overline{n_0A[B_1^X(0)]}\right)^\circ \neq \varnothing,$$

scaling the set, we have

$$\left(\overline{\frac{1}{2}A[B_1^X(0)]}\right)^\circ \neq \varnothing.$$

Then there is $y_0 \in Y$ and $\delta > 0$ s.t.

$$B_\delta^Y(y_0) \subseteq \left(\overline{\frac{1}{2}A[B_1^X(0)]}\right)^\circ \subseteq \overline{\frac{1}{2}A[B_1^X(0)]}.$$

Step 3. $B_\delta^Y(0) \subseteq \overline{A[B_1^X(0)]}$.
Fix $y \in B_\delta^Y(0)$ i.e. $\|y\|_Y < \delta$. Then

$$y_0 + y \subseteq B_\delta^Y(y_0) \subseteq \overline{\frac{1}{2}A[B_1^X(0)]}, \quad y_0 \subseteq \overline{\frac{1}{2}A[B_1^X(0)]}.$$

By the definition of closure, there exist sequences $\{x_i\}_{i=1}^{\infty}, \{x_i'\}_{i=1}^{\infty} \subseteq \frac{1}{2}B_1^X(0)$ s.t.

$$x_i' \to y + y_0, \quad x_i \to y_0.$$

Since $\|x_i' - x_i\|_X \leq \|x_i\|_X + \|x_i'\|_X < 1$, i.e. $x_i' - x_i \in B_1^X(0)$, we have

$$A(x_i' - x_i) \in A[B_1^X(0)],$$

then

$$y = \lim_{i\to\infty} A(x_i' - x_i) \in \overline{A[B_1^X(0)]}.$$

The lemma is now proved. □

Lemma 2.2.4. *Suppose X and Y are Banach spaces and $A \in \mathcal{L}(X, Y)$. If*

$$B_\delta^Y(0) \subseteq \overline{A[B_1^X(0)]} \quad \textit{for some } \delta > 0,$$

then

$$B_\delta^Y(0) \subseteq A[B_1^X(0)].$$

Proof. Step 1. For any $y \in Y$,

$$\frac{\delta y}{\|y\|_Y} \in B_\delta^Y(0) \subseteq \overline{A[B_1^X(0)]},$$

thus

$$y \in \overline{A[B_{\|y\|_Y/\delta}^X(0)]}.$$

Step 2. Let $y \in B_\delta^Y(0)$, then $y \in \overline{A[B_{\|y\|_Y/\delta}^X(0)]}$, by the definition of closure[1], let $\varepsilon := \delta - \|y\|_Y > 0$, there is $x_0 \in B_{\|y\|_Y/\delta}^X(0)$ (i.e. $\|x_0\|_X < \|y\|_Y/\delta$) s.t.

$$\|y - Ax_0\|_Y < \frac{\varepsilon}{2},$$

i.e. $y - Ax_0 \in B_{\varepsilon/2}^Y(0)$, rescale the vector, we have

$$\frac{\delta(y - Ax_0)}{\varepsilon/2} \in B_\delta^Y(0).$$

Repeat the previous procedure, there is $\tilde{x}_1 \in B_1^X(0)$, s.t.

$$\left\|\frac{\delta(y - Ax_0)}{\varepsilon/2} - A\tilde{x}_1\right\|_Y < \frac{\delta}{2},$$

rescale the vector, we have

$$\left\|y - Ax_0 - A\frac{\varepsilon\tilde{x}_1}{2\delta}\right\|_Y = \|y - Ax_0 - Ax_1\|_Y < \frac{\varepsilon}{4},$$

[1] If $x \in \overline{S}$, then for any $r > 0$, there is $s \in S$, s.t. $d(x, s) < r$.

where $x_1 = \dfrac{\varepsilon \tilde{x}_1}{2\delta}$ and $\|x_1\|_X < \dfrac{\varepsilon}{2\delta}$.
Step 3. Suppose we have found $\{x_j\}_{j=0}^k \subseteq X$ s.t.

$$\left\| y - A\sum_{j=0}^k x_j \right\|_Y < \frac{\varepsilon}{2^{k+1}}, \quad \text{with} \quad \|x_j\|_X < \frac{\varepsilon}{\delta 2^j}. \tag{2.1}$$

Then

$$\frac{\delta 2^{k+1}(y - A\sum_{j=0}^k x_j)}{\varepsilon} \in B_\delta^Y(0),$$

so there is $\tilde{x}_{k+1} \in B_1^X(0)$ s.t.

$$\left\| \frac{\delta 2^{k+1}(y - A\sum_{j=0}^k x_j)}{\varepsilon} - \tilde{x}_{k+1} \right\|_Y < \frac{\delta}{2},$$

then

$$\left\| y - A\sum_{j=0}^k x_j - \frac{\varepsilon \tilde{x}_{k+1}}{\delta 2^{k+1}} \right\|_Y = \left\| y - A\sum_{j=0}^k x_j - x_{k+1} \right\|_Y < \frac{\varepsilon}{2^{k+2}},$$

where $x_{k+1} = \dfrac{\varepsilon \tilde{x}_{k+1}}{\delta 2^{k+1}}$ and $\|x_{k+1}\|_X < \dfrac{\varepsilon}{\delta 2^{k+1}}$. Therefore (2.1) holds for all $k \in \mathbb{Z}_+$.
Step 4. Let $k \to \infty$, we have

$$y = A\sum_{j=0}^\infty x_j = Ax^*$$

where we let $x^* = \sum_{j=0}^\infty x_j$. Since

$$\|x^*\|_X = \left\| \sum_{j=0}^\infty x_j \right\|_X \le \sum_{j=0}^\infty \|x_j\|_X = \|x_0\|_X + \sum_{j=1}^\infty \|x_j\|_X < \frac{\|y\|_Y}{\delta} + \frac{\varepsilon}{\delta}\sum_{j=1}^\infty \frac{1}{2^j} = \frac{\|y\|_Y}{\delta} + \frac{\delta - \|y\|_Y}{\delta} < 1,$$

we have $x^* \in B_1^X(0)$ and hence $y \in A[B_1^X(0)]$. □

Next, we will prove the Open mapping theorem 2.2.2.

Proof of theorem 2.2.2. Suppose $A \in \mathcal{L}(X, Y)$ and A is surjective, then from Lemma 2.2.3 and 2.2.4, there is $\delta > 0$ s.t.

$$B_\delta^Y(0) \subseteq A[B_1^X(0)].$$

Let $U \subseteq X$ be an open set, we want to show $A(U)$ is also open. For any $y_0 \in A(U)$, there is $x_0 \in U$ s.t. $y_0 = Ax_0$. Since U is open, there is $\varepsilon > 0$ s.t. $B_\varepsilon(x_0) \subseteq U$.

Claim. $B^Y_{\varepsilon\delta}(y_0) \subseteq A(U)$.
Let $y \in B^Y_{\varepsilon\delta}(y_0)$, then

$$\frac{y-y_0}{\varepsilon} \in B^Y_\delta(0) \subseteq A[B^X_1(0)].$$

Then there is $x_1 \in B^X_1(0)$ s.t.

$$Ax_1 = \frac{y-y_0}{\varepsilon},$$

i.e. $y = \varepsilon Ax_1 + y_0 = A(x_0 + \varepsilon x_1)$, where $x_0 + \varepsilon x_1 \in B^X_\varepsilon(x_0)$. Now the claim is proved. Then for any element in $A(U)$, we can find an open ball around that element and contained in $A(U)$, i.e. $A(U)$ is open. □

Corollary 2.2.5. *Let X, Y be Banach spaces, $A \in \mathcal{L}(X,Y)$ be surjective. Let $\delta > 0$ s.t. $B^Y_\delta(0) \subseteq A[B^X_1(0)]$. Then for any $y \in Y$,*

$$\inf_{x\in X, Ax=y} \|x\|_X \le \frac{\|y\|_Y}{\delta}.$$

Proof. Fix $y \in Y$. For any $C \in \mathbb{R}$ s.t. $C > \|y\|_Y/\delta$, we have

$$\frac{\|y\|_Y}{C} < \delta,$$

then $y/C \in B^Y_\delta(0)$, by the hypothesis,

$$\frac{y}{C} \in A[B^X_1(0)].$$

Then there is $\tilde{x} \in B^X_1(0)$ s.t.

$$\frac{y}{C} = A\tilde{x}.$$

Let $x = C\tilde{x}$, $y = Ax$ and $\|x\|_X = c\|\tilde{x}\|_X < C$. We complete the proof since C is arbitrary. □

Corollary 2.2.6 (Inverse operator)**.** *Let X, Y be Banach spaces, $A \in \mathcal{L}(X,Y)$ be bijective. Then A is invertible.*

Proof. Bijectivity implies A^{-1} exists, We need to show A^{-1} is linear and continuous.
A^{-1} is linear: for any $u, v \in Y$ and $c \in \mathbb{R}$,

$$A^{-1}(cu+v) = A^{-1}[cAA^{-1}u + AA^{-1}v] = A^{-1}A(cA^{-1}u + A^{-1}v) = cA^{-1}u + A^{-1}v.$$

$A^{-1}: Y \to X$ is continuous: for any open set $U \in X$, by the Open mapping theorem (2.2.2), $(A^{-1})^{-1}(U) = A(U)$ is also open. Thus A^{-1} is continuous. □

Corollary 2.2.7. *Let X be a Banach space, X_1, X_2 be closed subspaces of X s.t.*

$$X = X_1 \oplus X_2.$$

Then there is $c > 0$ s.t. for any $x_1, x_2 \in X$,

$$\|x_1\| + \|x_2\| \le c\|x_1 + x_2\|.$$

Proof. X_1 and X_2 are both Banach spaces, easy to check $X_1 \times X_2$ is also Banach w.r.t. the product norm defined by

$$\|(x,y)\| = \|x_1\| + \|x_2\|, \quad \forall (x_1,x_2) \in X_1 \times X_2.$$

Define $A : X_1 \times X_2 \to X$ by $A((x_1,x_2)) = x_1 + x_2$, $\forall (x_1,x_2) \in X_1 \times X_2$. A is clearly linear and bounded. Moreover A is bijective: (i) A is injective, because 0 can be uniquely expressed, hence $0 = x_1 + x_2$ with $x_1 \in X_1, x_2 \in X_2$ has the only solution $x_1 = x_2 = 0$, i.e. $\operatorname{Ker} A = \{0\}$. (ii) A is surjective because every element in X can be expressed as $x = x_1 + x_2$ with $x_1 \in X_1$ and $x_2 \in X_2$.
Therefore A^{-1} is also bounded by Corollary 2.2.6. By definition of boundedness, there is $c > 0$, s.t. for any $x \in X$, we have

$$\left\|A^{-1}x\right\| \leq c\|x\|,$$

let $x = x_1 + x_2$ with $x_1 \in X_1, x_2 \in X_2$, then $\|x_1\| + \|x_2\| \leq c\|x_1 + x_2\|$. □

2.3 Closed graph theorem

Definition 2.3.1. Suppose X, Y are Banach spaces and U is a subspace of X. Let $A : U \to Y$ be linear (not necessarily be bounded), and write $\operatorname{dom}(A) = U$.

1. Define the product norm on $X \times Y$ by

$$\|(x,y)\| = \|x\|_X + \|y\|_Y, \quad \forall (x,y) \in X \times Y.$$

   Easy to check $(X \times Y, \|\cdot\|)$ is Banach.

2. We say A is closed if its graph

$$\Gamma_A := \{(x, Ax) : x \in \operatorname{dom}(A)\} \subseteq X \times Y$$

   is closed (w.r.t. the topology induced by the product norm).

3. Γ_A induces a norm on $\operatorname{dom}(A)$ called the graph norm, defined by

$$\|x\|_{\Gamma_A} := \|x\|_X + \|Ax\|_Y, \quad \forall x \in \operatorname{dom}(A).$$

   Then $(\operatorname{dom}(A), \|\cdot\|_{\Gamma_A})$ is a normed vector space.

Remark.

1. $\{(x_k, y_k)\}_{k=1}^\infty$ is a Cauchy sequence in $X \times Y$ if and only if $\{x_k\}_{k=1}^\infty$ is Cauchy in X and $\{y_k\}_{k=1}^\infty$ is Cauchy in Y.

2. If we endow $\operatorname{dom}(A)$ with the graph norm, then $A : (\operatorname{dom}(A), \|\cdot\|_{\Gamma_A}) \to Y$ is bounded.

Theorem 2.3.2 (Closed graph theorem)**.** *Let X, Y be Banach spaces, $A : X \to Y$ be linear. Then A is bounded if and only if Γ_A is closed in $X \times Y$.*

Proof. $\Longrightarrow$: Suppose $A \in \mathcal{L}(X, Y)$. For any convergent sequence $\{(x_n, Ax_n)\}_{n=1}^{\infty} \subseteq \Gamma_A$, let its limit be $(x_\infty, y_\infty) \in X \times Y$. We want to show $(x_\infty, y_\infty) \in \Gamma_A$, i.e. $y_\infty = Ax_\infty$. Easy to check $x_n \to x_\infty$ and $Ax_n \to y_\infty$. Since A is bounded hence continuous, we have

$$y_\infty = \lim_{n\to\infty} Ax_n = A(\lim_{n\to\infty} x_n) = Ax_\infty.$$

$\Longleftarrow$: Suppose Γ_A is closed. Then Γ_A is Banach w.r.t. the product norm. Define the projection map $\pi : \Gamma_A \to X$ by $(x, Ax) \mapsto x$. Easy to check π is bijective and bounded, thus by Corollary 2.2.6, π^{-1} is bounded. Then there is $c \geq 0$, s.t.

$$\|x\|_X + \|Ax\|_Y = \|(x, Ax)\| = \left\|\pi^{-1}(x)\right\| \leq c\,\|x\|_X\,,$$

therefore $\|Ax\|_Y \leq (c-1)\,\|x\|_X$, take $c' = \max\{c-1, 0\} \geq 0$, then $\|Ax\|_Y \leq c'\,\|x\|_X$, i.e. A is bounded. ☐

Corollary 2.3.3 (Hellinger–Toeplitz theorem)**.** *Let H be a $\mathbb{R}$-Hilbert space. Let $A : H \to H$ be linear and symmetric i.e. $\langle Ax, y\rangle = \langle x, Ay\rangle$ for any $x, y \in H$. Then A is bounded.*

Proof. By Theorem 2.3.2, it suffices to show Γ_A is closed in $H \times H$. Suppose $\{(x_n, Ax_n)\}_{n=1}^{\infty} \subseteq \Gamma_A$ converges to (x_∞, y_∞), we want to show $y_\infty = Ax_\infty$. For any $z \in H$, since A is symmetric, we have

$$\langle Ax_\infty, z\rangle = \langle x_\infty, Az\rangle = \lim_{n\to\infty} \langle x_n, Az\rangle = \lim_{n\to\infty} \langle Ax_n, z\rangle = \langle y_\infty, z\rangle,$$

then $\langle Ax_\infty - y_\infty, z\rangle = 0$. Let $z = Ax_\infty - y_\infty$, we have $Ax_\infty = y_\infty$. ☐

Definition 2.3.4 (Closeable operator)**.** Suppose X, Y are Banach spaces, let $\mathrm{dom}(A) \subseteq X$ be a subspace and $A : \mathrm{dom}(A) \to Y$ be linear. A is called closeable if there is a closed linear operator $\tilde{A} : \mathrm{dom}(\tilde{A}) \to Y$ s.t.

$$\mathrm{dom}(A) \subseteq \mathrm{dom}(\tilde{A}), \quad \tilde{A}\big|_{\mathrm{dom}(A)} = A.$$

Remark. The extension $\tilde{A}$ is closed, thus it is continuous by Theorem 2.3.2.

Lemma 2.3.5. *Suppose X, Y are Banach spaces, let $\mathrm{dom}(A) \subseteq X$ be a subspace and $A : \mathrm{dom}(A) \to Y$ be linear. TFAE:*

1. *A is closeable*
2. *The projection $\pi_X : \overline{\Gamma_A} \to X$ is injective*
3. *Suppose there is a sequence $\{x_n\}_{n=1}^{\infty} \subseteq \mathrm{dom}(A)$ and $y \in Y$ s.t.*

$$\lim_{n\to\infty} x_n = 0, \quad \lim_{n\to\infty} Ax_n = y,$$

then $y = 0$.

Proof. $1 \Longrightarrow 3$. Since $\tilde{A}$ is continuous, we have

$$y = \lim_{n\to\infty} Ax_n = \lim_{n\to\infty} \tilde{A}x_n = \tilde{A}(\lim_{n\to\infty} x_n) = \tilde{A}0 = 0.$$

$3 \Longrightarrow 2$. First, $\overline{\Gamma_A}$ is a subspace of X since the closure of a subspace is still a subspace (addition and multiplication are continuous). Second, consider the kernel of π,

$$\mathrm{Ker}\,(\pi) = \{(x, y) \in \overline{\Gamma_A} : x = 0\}.$$

For any $(x, y) \in \overline{\Gamma_A}$ with $x = 0$, there is a sequence $(x_i, Ax_i) \in \Gamma_A$ s.t. $(x_i, Ax_i) \to (0, y)$, i.e. $x_i \to 0$ and $Ax_i \to y$. From 2 we have $y = 0$, therefore $\mathrm{Ker}\,(\pi) = \{(0, 0)\}$ and hence π is injective.
$2 \Longrightarrow 1$.

2.4 Hahn-Banach theorem

Definition 2.4.1. Suppose X is a real vector space. A function $p : X \to \mathbb{R}$ is called a quasi-semi-norm (q-s-norm) if for any $x, y \in X$

1. $p(\lambda x) = \lambda p(x)$, $\forall \lambda \in [0, \infty)$
2. $p(x + y) \leq p(x) + p(y)$.

A function $p : X \to \mathbb{R}$ is called a semi-norm if it is a q-s-norm satisfying

$$p(\lambda x) = |\lambda| p(v), \quad \forall \lambda \in \mathbb{R}.$$

Remark. A q-s-norm is not necessarily non-negative, but a semi-norm is always non-negative since for any $x \in X$, $2p(x) = p(x) + p(-x) \geq p(0) = 0$.

Theorem 2.4.2 (Hahn-Banach)**.** *Suppose X is a vector space and $p : X \to \mathbb{R}$ is a q-s-norm. If there is a subspace $Y \subseteq X$ and a linear map $\phi : Y \to \mathbb{R}$ s.t.*

$$\phi(y) \leq p(y), \quad \forall y \in Y,$$

then there is a linear map $\psi : X \to \mathbb{R}$ s.t. $\psi|_Y = \phi$ on Y and

$$\psi(x) \leq p(x), \quad \forall x \in X.$$

Lemma 2.4.3. *Suppose X is a vector space and $p : X \to \mathbb{R}$ is a q-s-norm. If there is a proper subspace $Y \subsetneq X$ and a linear map $\phi : Y \to \mathbb{R}$ s.t.*

$$\phi(y) \leq p(y), \quad \forall y \in Y,$$

then for any $x_0 \in X \setminus Y$, there is a linear map $\psi : Y \oplus \mathrm{Span}(x_0) \to \mathbb{R}$ s.t. $\psi|_Y = \phi$ and

$$\psi(x) \leq p(x), \quad \forall x \in Y \oplus \mathrm{Span}(x_0).$$

Proof. If such ψ exists, for any $y+\lambda x_0 \in Y \oplus \mathrm{Span}(x_0)$ where $y \in Y, \lambda \in \mathbb{R}$,

$$\psi(y+\lambda x_0) = \psi(y) + \lambda\psi(x_0) = \phi(y) + \lambda\psi(x_0),$$

therefore ψ is determined by $\psi(x_0)$. Our goal is to find suitable $\psi(x_0)$ s.t.

$$\phi(y) + \lambda\psi(x_0) = \psi(y+\lambda x_0) \leq p(y+\lambda x_0), \quad \forall y \in Y, \lambda \in \mathbb{R}. \tag{2.2}$$

Easy to check (2.2) is equivalent to

$$\phi(y) + \psi(x_0) \leq p(y+x_0) \quad \text{and} \quad \phi(y) - \psi(x_0) \leq p(y-x_0), \quad \forall y \in Y,$$

i.e.

$$\phi(y) - p(y-x_0) \leq \psi(x_0) \leq p(y+x_0) - \phi(y), \quad \forall y \in Y.$$

For any $y_1, y_2 \in Y$, we have

$$\begin{aligned}\phi(y_1) + \phi(y_2) &= \phi(y_1+y_2) \\ &\leq p(y_1+y_2) \\ &= p(y_1 - x_0 + y_2 + x_0) \\ &\leq p(y_1 - x_0) + p(y_2 + x_0),\end{aligned}$$

thus

$$\phi(y_1) - p(y_1 - x_0) \leq p(y_2 + x_0) - \phi(y_2), \quad \forall y_1, y_2 \in Y.$$

Let

$$c_1 = \sup_{y \in Y}[\phi(y) - p(y-x_0)], \quad c_2 = \inf_{y \in Y}[p(y+x_0) - \phi(y)],$$

we have $c_1 \leq c_2$, so we can always choose $c \in [c_1, c_2]$ and let $\psi(x_0) = c$. $\square$

Proof of Theorem 2.4.2. Define the set

$$P = \{(Z,\psi) : Y \subseteq Z \subseteq X;\ \psi : Z \to \mathbb{R} \text{ is linear } s.t.\ \phi = \psi\big|_Y,\ \psi(z) \leq p(z),\ \forall z \in Z.\}$$

Define a partial order $\preceq$ on P by $(Z,\psi) \preceq (Z',\psi')$ if

$$Z \subseteq Z', \quad \psi = \psi'\big|_Z.$$

Let $C \subseteq P$ be a chain, define (Z_C, ψ_C) by

$$Z_C = \bigcup_{(Z,\psi)\in C} Z, \quad \psi_C = \psi(x), \quad \forall x \in Z \quad \text{with} \quad (Z,\psi) \in C.$$

Then $(Z_C,\psi_C) \in P$ and for any $(Z,\psi) \in C$, we have $(Z,\psi) \preceq (Z_C,\psi_C)$, i.e. (Z_C,ψ_C) is an upper bound for the chain C. By Zorn's lemma (A.1.3), there is a maximal element $(Z_m,\psi_m) \in P$, i.e. there is no other element $m' \in P$ s.t. $m' \neq m$ and $m \preceq m'$.
Claim: $Z_m = X$.

Otherwise, assume $Z_m \subsetneq X$, by Lemma 2.4.3, there is $x_0 \in X \backslash Z_m$ and $\psi : Z_m \oplus \text{Span}(x_0) \to \mathbb{R}$ s.t. $(Z_m \oplus \text{Span}(x_0), \psi) \in P$ and

$$(Z_m, \psi_m) \preceq (Z_m \oplus \text{Span}(x_0), \psi),$$

which contradicts (Z_m, ψ_m) is a maximal element. □

Corollary 2.4.4. *Let X be a normed vector space, $Y \subseteq X$ be a subspace. If $\phi \in Y^*$, then there is $\psi \in X^*$ s.t. $\psi|_Y = \phi$ and $\|\psi\| = \|\phi\|$.*

Proof. Let $p(x) = \|\phi\| \cdot \|x\|$. Then p is a norm, and

$$\phi(x) \le |\phi(x)| \le \|\phi\| \cdot \|x\| = p(x), \quad \forall x \in Y.$$

By Hahn-Banach theorem, there is a linear map $\psi : X \to \mathbb{R}$ s.t. $\psi|_Y = \phi$ and

$$\psi(x) \le p(x) = \|\phi\| \cdot \|x\|, \quad \forall x \in X.$$

thus $|\psi(x)| \le \|\phi\| \cdot \|x\|$, $\forall x \in X$, which implies ψ is bounded and $\|\psi\| \le \|\phi\|$. Since

$$\|\phi\| = \sup_{x \in Y \backslash \{0\}} \frac{|\phi(x)|}{\|x\|} = \sup_{x \in Y \backslash \{0\}} \frac{|\psi(x)|}{\|x\|} \le \sup_{x \in X \backslash \{0\}} \frac{|\psi(x)|}{\|x\|} = \|\psi\|,$$

we conclude that $\|\phi\| = \|\psi\|$. □

Corollary 2.4.5. *Let X be a normed vector space, for any $x_0 \in X \setminus \{0\}$, there is $\psi \in X^*$ s.t. $\|\psi\| = 1$ and $\psi(x_0) = \|x_0\|$.*

Proof. Let $Y = \text{Span}(x_0)$. Define $\phi : Y \to \mathbb{R}$ by

$$\phi(x) = \phi(tx_0) = t\|x_0\|, \quad \forall x = tx_0 \in Y.$$

Easy to check $\phi \in Y^*$, $\phi(x_0) = \|x_0\|$ and $\|\phi\| = 1$. By Corollary 2.4.4, there is $\psi \in X^*$ s.t. $\psi|_Y = \phi$ and $\|\psi\| = \|\phi\| = 1$. Since $x_0 \in Y$, $\psi(x_0) = \phi(x_0) = \|x_0\|$. □

Corollary 2.4.6. *Let X be a normed vector space, then X^* separate points on X, i.e. for any $x, y \in X$ with $x \neq y$, there is $\phi \in X^*$, s.t. $\phi(x) \neq \phi(y)$.*

Proof. Let $x, y \in X$ with $x \neq y$. Then $x - y \neq 0$, by Corollary 2.4.5, there is $\psi \in X^*$ s.t.

$$\psi(x) - \psi(y) = \psi(x - y) = \|x - y\| \neq 0.$$

□

2.5 Separation of convex sets

In this section, we will introduce another important application of Hahn-Banach theorem.

Theorem 2.5.1 (Separation of convex sets)**.** *Let X be a real normed vector space, A, B be non-empty, disjoint convex subsets of X with $B^\circ \neq \varnothing$. Then there is $\Lambda \in X^* \setminus \{0\}$ and $c \in \mathbb{R}$ s.t.*

$$\Lambda(a) \leq c \leq \Lambda(b), \quad \forall a \in A, b \in B,$$

and

$$\Lambda(a) \leq c < \Lambda(b), \quad \forall a \in A, b \in B^\circ.$$

Before proving Theorem 2.5.1, we will introduce some definitions and lemmas.

Lemma 2.5.2. *Let X be a normed vector space, $u, v \in X$, $r_1, r_2 > 0$, $a, b \geq 0$. Then*

$$aB_{r_1}(u) + bB_{r_2}(v) = B_{ar_1+br_2}(au + bv).$$

Proof. 1. For any $x \in aB_{r_1}(u) + bB_{r_2}(v)$, $x = ax_1 + bx_2$ where $x_1, x_2 \in X$ s.t.

$$\|x_1 - u\| < r_1, \quad \|x_2 - v\| < r_2.$$

Therefore

$$\|x - (au + bv)\| \leq a\|x_1 - u\| + b\|x_2 - v\| < ar_1 + br_2,$$

i.e. $x \in B_{ar_1+br_2}(au + bv)$.
2. For any $x \in B_{ar_1+br_2}(au + bv)$, let $x = au + bv + y$, where $\|y\| < ar_1 + br_2$. Then we can write x as

$$x = au + \frac{ar_1}{ar_1 + br_2}y + bv + \frac{br_2}{ar_1 + br_2}y.$$

Since

$$\left\|u + \frac{r_1}{ar_1 + br_2}y - u\right\| = \frac{r_1}{ar_1 + br_2}\|y\| < r_1,$$

we have

$$u + \frac{r_1}{ar_1 + br_2}y \in B_{r_1}(u).$$

Similarly,

$$v + \frac{r_2}{ar_1 + br_2}y \in B_{r_2}(v),$$

therefore $x \in aB_{r_1}(u) + bB_{r_2}(v)$. □

Lemma 2.5.3. *Suppose X is a normed vector space and $A \subseteq X$ is convex. Then*

1. *A° and $\overline{A}$ are convex.*

2. *If $A^\circ \neq \varnothing$, then $A \subseteq \overline{A^\circ}$.*

Proof. 1. For any $u, v \in A^\circ$, there are open balls $B_\varepsilon(u), B_\varepsilon(v) \subseteq A^\circ$. By convexity of A, for any $t \in [0, 1]$,

$$tB_\varepsilon(u) + (1 - t)B_\varepsilon(v) \subseteq A.$$

On the other hand, by Lemma 2.5.2,

$$tB_\varepsilon(u) + (1-t)B_\varepsilon(v) = B_\varepsilon(tu + (1-t)v),$$

which is open, thus

$$tB_\varepsilon(u) + (1-t)B_\varepsilon(v) \subseteq A^\circ,$$

in particular, $tu + (1-t)v \in A^\circ$, i.e. A° is convex.
Next, we will show $\overline{A}$ is convex. For any $u, v \in \overline{A}$, there is sequences $\{u_i\}_{i=1}^\infty, \{v_i\}_{i=1}^\infty \subseteq A$ s.t.

$$u_i \to u, \quad v_i \to v.$$

Let $t \in [0,1]$, since A is convex, for each i,

$$tu_i + (1-t)v_i \in A,$$

then

$$tu + (1-t)v = \lim_{i\to\infty}[tu_i + (1-t)v_i] \in \overline{A}.$$

2. Fix $x_0 \in A^\circ$, since A° is open, there is $B_\varepsilon(x_0) \subseteq A^\circ$. For any $x \in A$ and $t \in (0,1)$, since A is convex,

$$tx + (1-t)B_\varepsilon(x_0) \subseteq A.$$

Moreover, by Lemma 2.5.2, $tx + (1-t)B_\varepsilon(x_0) = B_{(1-t)\varepsilon}(tx + (1-t)x_0)$ is open, thus

$$tx + (1-t)B_\varepsilon(x_0) \subseteq A^\circ.$$

Define

$$U_x = \bigcup_{0<t<1} [tx + (1-t)B_\varepsilon(x_0)],$$

then $U_x \subseteq A^\circ$. And we can choose $\{y_n\}_{n=1}^\infty \subseteq U_x$ s.t. $y_n \to x$, for example, let

$$y_n = (1 - \frac{1}{n})x + \frac{x_0}{n},$$

which implies $x \in \overline{U_x} \subseteq \overline{A^\circ}$. □

Lemma 2.5.4. *Suppose X is a normed vector space and $A \subseteq X$ is convex s.t. $A^\circ \neq \varnothing$. If there is $\Lambda \in X^* \setminus \{0\}$ and $c \in \mathbb{R}$ s.t.*

$$\Lambda(x) \geq c, \quad \forall x \in A^\circ,$$

then

$$\Lambda(x) \geq c, \quad \forall x \in A,$$

and

$$\Lambda(x) > c, \quad \forall x \in A^\circ.$$

Proof. From Lemma 2.5.3, $A \subseteq \overline{A^\circ}$, thus for any $x \in A$, there is a sequence $\{x_i\}_{i=1}^\infty \subseteq A^\circ$ s.t. $x_i \to x$. By the continuity of Λ,

$$\Lambda(x) = \lim_{i\to\infty} \Lambda(x_i) \geq c.$$

For the second statement, let $x \in A^\circ$. We can find $x_0 \in X$ s.t. $\Lambda(x_0) = 1$, for example, choose $y \in X \setminus \{0\}$ s.t. $\Lambda y \neq 0$, let $x_0 = y/\|\Lambda y\|$. Since A° is open, there is an open ball $B_\delta(x) \subseteq A^\circ$ with $\delta > 0$. And then $x - \dfrac{\delta x_0}{2} \in B_\delta(x) \subseteq A^\circ$ because

$$\left\| x - \frac{\delta x_0}{2} - x \right\| = \frac{\delta}{2}\|x_0\| = \frac{\delta}{2} < \delta.$$

Therefore

$$\Lambda(x) = \Lambda(x - \frac{\delta}{2}x_0) + \frac{\delta}{2}\Lambda(x_0) \geq c + \frac{\delta}{2} > c.$$

□

Definition 2.5.5 (Minkowski function)**.** Suppose X is a real normed vector space and $A \subseteq X$ is convex s.t. $0 \in A$. Define the Minkowski function $\rho_A : X \to [0, \infty]$ by

$$\rho_A(x) = \inf\{t > 0 : \frac{x}{t} \in A\}, \quad \forall x \in X.$$

Define $\inf \varnothing = \infty$.

Remark. 1. $\rho_A(0) = 0$.
2. If $x \in A$, $\rho_A(x) \leq 1$ because $1 \in \{t > 0 : x/t \in A\}$.
3. If $x \notin A$, $\rho_A(x) \geq 1$. Assume $\rho_A(x) < 1$, there is $\delta > 0$ s.t. $\rho_A(x) < \delta < 1$, and by the definition of inf, δ is no longer a lower bounded, i.e. there is $t_0 > 0$ s.t. $t_0 < \delta < 1$ and $x/t_0 \in A$. Since A is convex, $0, x/t_0 \in A$, we have

$$x = (1 - t_0) \cdot 0 + t_0 \cdot \frac{x}{t_0} \in A.$$

Lemma 2.5.6. *ρ_A defined above is a q-s-norm, i.e. for any $x, y \in A$ and $\lambda \in [0, \infty)$,*

1. $\rho_A(\lambda x) = \lambda \rho_A(x)$

2. $\rho_A(x + y) \leq \rho_A(x) + \rho_A(y)$.

Proof. 1. The case $\lambda = 0$ is clear. Assume $\lambda > 0$, by definition,

$$\rho_A(\lambda x) = \inf\{t > 0 : \frac{\lambda x}{t} \in A\} = \inf\{\lambda \cdot (\frac{t}{\lambda}) > 0 : \frac{x}{t/\lambda} \in A\} = \lambda \rho_A(x).$$

2. For any $t, s > 0$ s.t. $t > \rho_A(x), s > \rho_A(y)$, we have

$$\rho_A(\frac{x}{t}) < 1, \quad \rho_A(\frac{y}{s}) < 1,$$

then $x/t, y/s \in A$. Since A is convex, we have

$$\frac{x+y}{s+t} = \frac{t}{s+t} \cdot \frac{x}{t} + \frac{s}{s+t} \cdot \frac{y}{s} \in A,$$

then

$$\rho_A(\frac{x+y}{s+t}) \leq 1, \quad i.e. \quad \rho_A(x+y) \leq s+t.$$

For any $\varepsilon > 0$, let $t = \rho_A(x) + \varepsilon$, $s = \rho_A(y) + \varepsilon$, then

$$\rho_A(x+y) \leq \rho_A(x) + \rho_A(y) + 2\varepsilon,$$

since ε is arbitrary, we have

$$\rho_A(x+y) \leq \rho_A(x) + \rho_A(y).$$

□

Proof of Theorem 2.5.1. 1. Let $A, B \subseteq X$ be non-empty disjoint convex subsets and $B^\circ \neq \varnothing$. Let $M = A - B^\circ := \{a - b : a \in A, b \in B^\circ\}$.
2. Claim: M is convex.
Let $x, y \in M$, i.e. $x = a_1 - b_1, y = a_2 - b_2$ for some $a_1, a_2 \in A$ and $b_1, b_2 \in B^\circ$. Then for any $t \in [0, 1]$,

$$tx + (1-t)y = [ta_1 + (1-t)a_2] - [tb_1 + (1-t)b_2] \in A - B^\circ = M,$$

because A and B° are convex by Lemma 2.5.3.
3. M is non-empty since A and B° are non-empty, then we can choose $x_0 \in M$. Let $M_0 = M - \{x_0\}$, then $0 \in M_0$ and M_0 is also convex. Thus we can define Minkowski function on M_0. Let $p = \rho_{M_0} : X \to [0, \infty]$ be the Minkowski function. Since $A \cap B^\circ = \varnothing$, $0 \notin M$, then $-x_0 \notin M_0$, thus $p(-x_0) \geq 1$.
4. Let $E_0 = \text{Span}(x_0)$. Define $f_0 \in E_0^*$ by

$$f_0(sx_0) = -s, \quad \forall s \in \mathbb{R}.$$

In particular, $f_0(x_0) = -1$.
5. Claim: For any $x \in E_0$, $f_0(x) \leq p(x)$.
For any $s \leq 0$, apply $p(-x_0) \geq 1$ and Lemma 2.5.6, we have

$$f_0(sx_0) = -s \leq -sp(-x_0) = p(sx_0)$$

For any $s > 0$,

$$f_0(sx_0) = -s < 0 \leq p(sx_0),$$

because p is always non-negative.
6. Since p is a q-s-norm, by Step 5 and Hahn-Banach theorem, there is $f \in X^*$ s.t.

$$f\big|_{E_0} = f_0, \quad f(x) \leq p(x) \quad \forall x \in X.$$

In particular, for any $x \in M_0$, $f(x) \leq p(x) \leq 1$. And $f(x_0) = f_0(x_0) = -1$. For any $x \in M$, $x - x_0 \in M_0$, thus

$$f(x) = f(x - x_0) + f(x_0) = f(x - x_0) - 1 \leq 0.$$

7. Therefore, for any $a \in A, b \in B^\circ$, $a - b \in M$, then

$$f(a) - f(b) = f(a - b) \leq 0,$$

i.e. $f(a) \leq f(b)$. Let

$$c_1 = \sup_{a \in A} f(a), \quad c_2 = \inf_{b \in B^\circ} f(b),$$

then $c_1 \leq c_2$, choose $c \in [c_1, c_2]$, we have

$$f(a) \leq c \leq f(b), \quad \forall a \in A, b \in B^\circ.$$

8. From Lemma 2.5.4,

$$f(a) \leq c \leq f(b), \quad \forall a \in A, b \in B,$$

and

$$f(a) \leq c < f(b), \quad \forall a \in A, b \in B^\circ.$$

□

2.6 Dual space and annihilator

Definition 2.6.1. Suppose X be a normed vector space, for any $S \subseteq X$, define the annihilator of S to be

$$S^\perp = \{\varphi \in X^* : \varphi(a) = 0, \forall a \in S\}.$$

Remark. 1. $S^\perp$ is a subspace of X^* no matter whether S is a subspace of X.
2. $S^\perp$ is always closed. Suppose $\{\phi_n\}_{n=1}^\infty \subseteq S^\perp$ is convergent in X^*, i.e. there is $\phi \in X^*$ s.t. $\phi_n \to \phi$. Then for any $x \in X$,

$$|\phi x| = |\phi_n x - \phi x| + |\phi_n x| \leq \|\phi_n - \phi\| \cdot \|x\|_X \to 0,$$

thus $\phi \in S^\perp$, and $S^\perp$ is closed.
3. Since X^* is always Banach (Theorem 1.4.1), any closed subspace of Banach space is also Banach, so $S^\perp$ is Banach.

Theorem 2.6.2. *Suppose X is a normed vector space and $Y \subseteq X$ is a subspace. Then for any $x_0 \in X \setminus \overline{Y}$,*

$$d(x_0, Y) := \inf_{y \in Y} \|x_0 - y\|_X > 0$$

and there is $\phi \in Y^\perp$ s.t.

$$\|\phi\| = 1, \quad \phi(x_0) = d(x_0, Y).$$

Proof. 1. Assume $d(x_0, Y) = 0$, by the definition of inf, there is a sequence $\{y_n\}_{n=1}^{\infty} \subseteq Y$ s.t.

$$d(x_0, y_n) < \frac{1}{n},$$

thus $y_n \to x_0$, which implies $x_0 \in \overline{Y}$, a contradiction!
2. Let $\delta := d(x_0, Y) > 0$. we want to find such ϕ. Denote $Z = Y \oplus \mathrm{Span}(x_0)$, define $\psi \in Z^*$ by $\psi(tx_0) = \delta t$, $\forall t \in \mathbb{R}$, and $\psi(y) = 0$, $\forall y \in Y$. Then

$$\begin{aligned}
\|\psi\| &= \sup_{y\in Y, t\in\mathbb{R}\setminus\{0\}} \frac{|\psi(y+tx_0)|}{\|y+tx_0\|} \\
&= \sup_{y\in Y, t\in\mathbb{R}\setminus\{0\}} \frac{\delta|t|}{\|y+tx_0\|} \\
&= \sup_{y\in Y, t\in\mathbb{R}\setminus\{0\}} \frac{\delta}{\|y/t+x_0\|} \\
&= \sup_{y'\in Y} \frac{\delta}{\|x_0-y'\|} \qquad (\text{let } y' = -y/t) \\
&= \frac{\delta}{\inf_{y'\in Y}\|x_0-y'\|} \\
&= \frac{\delta}{\delta} = 1.
\end{aligned}$$

By Corollary 2.4.4, there is $\phi \in X^*$ s.t.

$$\|\phi\| = \|\psi\| = 1, \quad \phi\big|_Z = \psi.$$

Moreover, $\phi(x_0) = \psi(x_0) = \delta$, and for any $y \in Y$, $\phi(y) = \psi(y) = 0$, i.e. $\phi \in Y^\perp$. □

Corollary 2.6.3. *Suppose X is a normed vector space and $Y \subseteq X$ is a subspace. Let $x \in X$, then $x \in \overline{Y}$ if and only if*

$$\phi(x) = 0, \quad \forall \phi \in Y^\perp.$$

Proof. If $x \in \overline{Y}$, then there is a sequence $\{y_n\}_{n=1}^{\infty} \subseteq Y$ s.t. $y_n \to x$. For any $\phi \in Y^\perp$, by the continuity of ϕ,

$$\phi(x) = \lim_{n\to\infty} \phi(y_n) = 0.$$

On the other hand, if $x \notin \overline{Y}$, by Theorem 2.6.2, there is $\phi \in Y^\perp$, s.t. $\phi(x) > 0$. □

Corollary 2.6.4. *Suppose X is a normed vector space and $Y \subseteq X$ is a subspace. Then Y is dense in X, i.e. $\overline{Y} = X$ if and only if $Y^\perp = \{0\}$.*

Proof. By Corollary 2.6.3, $X = \overline{Y}$ if and only if for any $x \in X$ and any $\phi \in Y^\perp$,

$$\phi(x) = 0,$$

which implies $\phi = 0$, i.e. $Y^\perp = \{0\}$. □

Corollary 2.6.5. *Suppose X is a normed vector space and $Y \subseteq X$ is a subspace. Then*

1. *The map $[\phi] \mapsto \phi|_Y : X^*/Y^\perp \to Y^*$ is an isometric isomorphism.*
2. *If Y is a closed subspace and $\pi : X \to X/Y$ is the canonical projection defined by $x \mapsto x+Y$, then the map $\phi \mapsto \phi \circ \pi : (X/Y)^* \to Y^\perp$ is an isometric isomorphism.*

2.7 Reflexive space

Definition 2.7.1. Suppose X is a normed vector space. Denote the double dual of X as $X^{**} = (X^*)^*$. And define $\iota_X : X \to X^{**}$ by

$$\iota_X(x)(\phi) = \phi(x), \quad \forall x \in X, \phi \in X^*.$$

Lemma 2.7.2. *$\iota_X : X \to X^{**}$ defined above is an isometric embedding.*

Proof. First, for any $x \in X$,

$$\|\iota_X(x)\|_{X^{**}} = \sup_{\phi \in X^* \setminus \{0\}} \frac{|\iota_X(x)(\phi)|}{\|\phi\|_{X^*}} = \sup_{\phi \in X^* \setminus \{0\}} \frac{|\phi(x)|}{\|\phi\|_{X^*}} \leq \sup_{\phi \in X^* \setminus \{0\}} \frac{\|\phi\|_{X^*} \|x\|_X}{\|\phi\|_{X^*}} = \|x\|_X .$$

Second, by Corollary 2.4.5, for any $x \in X$, there is $\phi \in X^*$ s.t.

$$\|\phi\|_{X^*} = 1, \quad \phi(x) = \|x\|_X .$$

Then

$$\|x\|_X = \phi(x) = \iota_X(x)(\phi) \leq \|\iota_X(x)\|_{X^{**}} \|\phi\|_{X^*} = \|\iota_X(x)\|_{X^{**}} .$$

Therefore $\|x\|_X = \|\iota_X(x)\|_{X^{**}}$. □

Remark. ι_X is injective since an isometric embedding is always injective.

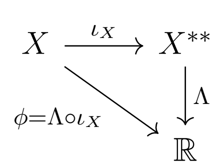
Definition 2.7.3. A normed vector space X is called reflexive if ι_X is bijective.

Remark. 1. If X is reflexive, then X is Banach since X^{**} is always Banach.
2. To show X is reflexive, since ι_X is always injective, we only need to show ι_X is surjective.

Theorem 2.7.4. *Suppose X is Banach, the following holds*

1. *X is reflexive if and only if X^* is reflexive.*
2. *If X is reflexive and Y is a closed subspace, then both Y and X/Y are reflexive.*

Proof. 1. $\Longrightarrow$. Suppose X is reflexive, we want to show $\iota_{X^*} : X^* \to X^{***}$ is surjective. Let $\Lambda \in X^{***}$, we want to find $\phi \in X^*$ s.t. $\iota_{X^*}(\phi) = \Lambda$. Let $\phi = \Lambda \circ \iota_X \in X^*$.



Claim: $\iota_{X^*}(\phi) = \Lambda$.
For any $f \in X^{**}$, since X is reflexive by assumption, we have ι_X is bijective, thus there is $x_f \in X$ s.t. $\iota_X(x_f) = f$. Then

$$\iota_{X^*}(\phi)(f) = f(\phi) = f(\Lambda \circ \iota_X) = \iota_X(x_f)(\Lambda \circ \iota_X) = \Lambda \circ \iota_X(x_f) = \Lambda \circ f.$$

The claim is proved and hence X^* is reflexive.
$\Longleftarrow$. Suppose X^* is reflexive, we want to show $\iota_X : X \to X^{**}$ is surjective. Since ι_X is isometric by Lemma 2.7.2, $\iota_X(X)$ is a closed subspace of X^{**}. Let $\psi \in \iota_X(X)^\perp \subseteq X^{***}$, since X^* is reflexive by the assumption, there is $\phi \in X^*$ s.t. $\psi = \iota_{X^*}(\phi)$. For any $a \in X$, $\iota_X(a) \in \iota_X(X)$, then

$$0 = \psi(\iota_X(a)) = \iota_{X^*}(\phi)(\iota_X(a)) = \iota_X(a)(\phi) = \phi(a),$$

i.e. $\phi = 0$, and then $\psi = \iota_{X^*}(\phi) = 0$, thus $\iota_X(X)^\perp = \{0\}$. By Corollary 2.6.4 and $\iota_X(X)$ is closed, we have

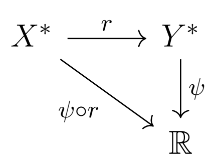
$$\iota_X(X) = \overline{\iota_X(X)} = X^{**},$$

therefore ι_X is surjective.

2. Assume X is reflexive and $Y \subseteq X$ is closed. First, we want to show Y is reflexive, i.e. $\iota_Y : Y \to Y^{**}$ is surjective. Define the restriction map $r : X^* \to Y^*$ by

$$r(\phi) = \phi\big|_Y, \quad \forall \phi \in X^*.$$

Then for any $\psi \in Y^{**}$, $\psi \circ r \in X^{**}$.



Since X is reflexive i.e. $\iota_X : X \to X^{**}$ is bijective, there is $x_\psi \in X$, s.t.

$$\iota_X(x_\psi) = \psi \circ r.$$

Claim: $x_\psi \in Y$.
Let $f \in Y^\perp \subseteq X^*$, then

$$f(x_\psi) = \iota_X(x_\psi)(f) = \psi \circ r(f) = 0,$$

because $f \in Y^\perp$ implies $r(f) = f|_Y = 0$. By Corollary 2.6.3 and Y is closed, $x_\psi \in \overline{Y} = Y$.

We have shown for any $\psi \in Y^{**}$, there is $x_\psi \in Y$ s.t. $\iota_X(x_\psi) = \psi$. Left to show $\iota_Y(x_\psi) = \psi$. For any $g \in Y^*$, by Corollary 2.4.4, there is $h \in X^*$ s.t.

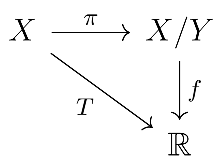
$$g = h\big|_Y = r \circ h.$$

Then

$$\iota_Y(x_\psi)(g) = g(x_\psi) = h(x_\psi) = \iota_X(x_\psi)(h) = \psi \circ r(h) = \psi(g),$$

therefore $\iota_Y(x_\psi) = \psi$, i.e. ι_Y is surjective and hence Y is reflexive.

Second, we want to show X/Y is reflexive, i.e. $\iota_{X/Y}$ is surjective. Let $\pi : X \to X/Y$ be the canonical projection, i.e. $\pi(x) = [x] = x + Y$ for any $x \in X$. Define $T : (X/Y)^* \to X^*$ by $T(f) = f \circ \pi$, for any $f \in (X/Y)^*$.



In fact, $\operatorname{Im} T \subseteq Y^\perp$ because for any $y \in Y$, $\pi(y) = [y] = 0$, thus for any $f \in (X/Y)^*$, $T(f)(y) = f \circ \pi(y) = 0$. Moreover, by Theorem 2.6.5, $T : (X/Y)^* \to Y^\perp$ is an isometric isomorphism.
Fix $\psi \in (X/Y)^{**}$, we want to find $x_\psi \in X/Y$ s.t. $\iota_{X/Y}(x_\psi) = \psi$. Notice that $\psi \circ T^{-1} \in (Y^\perp)^* \subseteq X^{**}$, then by Corollary 2.4.4, there is $\phi \in X^{**}$ s.t.

$$\psi \circ T^{-1} = \phi\big|_{Y^\perp},$$

i.e. for any $g \in Y^\perp$,

$$\psi \circ T^{-1}(g) = \phi(g),$$

and for any $f \in (X/Y)^*$, $Tf \in Y^\perp$, thus

$$\psi(f) = \psi \circ T^{-1}(Tf) = \phi(Tf) = \phi(f \circ \pi).$$

Since X is reflexive, there is $x \in X$ s.t. $\iota_X(x) = \phi$. Let $x_\psi = \pi(x)$.
Claim: $\iota_{X/Y}(x_\psi) = \psi$.
For any $f \in (X/Y)^*$,

$$\begin{aligned}\iota_{X/Y}(x_\psi)(f) &= f(x_\psi) \\ &= f(\pi(x)) \\ &= f \circ \pi(x) \\ &= \iota_X(x)(f \circ \pi) \\ &= \phi(f \circ \pi) \\ &= \psi(f).\end{aligned}$$

The claim is proved and therefore $\iota_{X/Y}$ is surjective. □

Example 2.7.5. Every finite-dimensional normed vector space is reflexive.

Example 2.7.6. Every hilbert space is reflexive.

Example 2.7.7. L^p is reflexive for any $1 < p < \infty$.

Example 2.7.8. c_0 is not reflexive.

2.8 Separable space

Definition 2.8.1. A normed vector space is separable if it has a countable dense subset.

Example 2.8.2. $\mathbb{R}^n$ is separable.

Lemma 2.8.3. *A normed vector space is separable if and only if it has a countable set $\{e_j\}_{j=1}^{\infty}$ s.t. the set of finite linear combination of $\{e_j\}_{j=1}^{\infty}$, i.e.*

$$\left\{ \sum_{j=1}^{n} a_j e_j : a_j \in \mathbb{R}, n \in \mathbb{Z}_+ \right\}$$

is dense.

Theorem 2.8.4. *Suppose X is a normed vector space. Then*

(1) *If X^* is separable, then X is separable.*

(2) *If X is reflexive and separable, then X^* is separable.*

Proof. Claim: (1) implies (2).
Assume (1) holds and assume X is reflexive and separable. By reflexivity, X^{**} has the same topology as X, thus X^{**} is also separable. By (1), X^* is separable.

Left to show (1). Assume X^* is separable. Let $\{\phi_j\}_{j=1}^{\infty} \subseteq X^*$ be dense. Denote the unit sphere in X by $S_1(X)$, and let $\{b_j\}_{j=1}^{\infty} \subseteq S_1(X)$ s.t.

$$\phi_j(b_j) \geq \frac{1}{2} \|\phi_j\| .$$

Define

$$Y = \left\{ \sum_{j=1}^{n} y_j b_j : y_j \in \mathbb{R}, n \in \mathbb{Z}_+ \right\},$$

by Lemma 2.8.3, it suffices to show Y is dense in X.

Claim: Y is dense.
From Corollary 2.6.4, Y is dense if and only if $Y^{\perp} = \{0\}$. Let $\psi \in Y^{\perp} \subseteq X^*$. Since $\{\phi_j\}_{j=1}^{\infty}$ is dense, there is $\{\phi_{j_k}\}_{k=1}^{\infty}$ s.t. $\phi_{j_k} \to \psi$, i.e.

$$\|\phi_{j_k} - \psi\| \to 0.$$

Then

$$\begin{aligned}
\|\psi\| &\le \|\phi_{j_k} - \psi\| + \|\phi_{j_k}\| \\
&\le \|\phi_{j_k} - \psi\| + 2\phi_{j_k}(b_{j_k}) \\
&= \|\phi_{j_k} - \psi\| + 2(\phi_{j_k} - \psi)(b_{j_k}) \quad (\text{since } \psi \in Y^{\perp}) \\
&\le \|\phi_{j_k} - \psi\| + 2\left\|\phi_{j_k} - \psi\right\| \left\|b_{j_k}\right\|_X \\
&= 3\left\|\phi_{j_k} - \psi\right\| \\
&\to 0,
\end{aligned}$$

thus $\psi = 0$. $\square$

Chapter 3

Weak and weak* topologies

3.1 Weak topology

Motivation

The fewer open sets a topology has, the easier it is for sequences to converge. To facilitate convergence, we will introduce a topology that is coarser than the norm topology.

However, we want this new topology to preserve important properties, such as the continuity of continuous functions. Specifically, we aim to construct the smallest, or *weakest*, topology on X such that every linear functional that is continuous with respect to the norm topology remains continuous with respect to this new topology. In other words, for any $\phi \in X^*$ and any open interval $(\phi(x_0) - \varepsilon, \phi(x_0) + \varepsilon) \subseteq \mathbb{R}$, we want the set

$$\phi^{-1}((\phi(x_0) - \varepsilon, \phi(x_0) + \varepsilon)) = \{x \in X : |\phi(x) - \phi(x_0)| < \varepsilon\}$$

to be open in the new topology. We will call these sets "new open" sets.

Recall that a topology is closed under finite intersections, meaning the new topology must include every finite intersection of these "new open" sets. This gives us the sets

$$\begin{aligned} N(x_0, \varepsilon, \phi_1, \phi_2, \ldots, \phi_N) &= \{x \in X : |\phi_1(x) - \phi_1(x_0)| < \varepsilon, \ldots, |\phi_N(x) - \phi_N(x_0)| < \varepsilon\} \\ &= \bigcap_{i=1}^{N} \{x \in X : |\phi_i(x) - \phi_i(x_0)| < \varepsilon\}, \end{aligned}$$

where $x_0 \in X$, $\varepsilon > 0$, and $\phi_1, \phi_2, \ldots, \phi_N \in X^*$. Let

$$\mathcal{B} = \{N(x_0, \varepsilon, \phi_1, \phi_2, \ldots, \phi_N) : x_0 \in X, \varepsilon > 0, \phi_1, \phi_2, \ldots, \phi_N \in X^*\}$$

be the basis for this topology.

Thus, the weakest topology that ensures the continuity of every $\phi \in X^*$ must contain the basis $\mathcal{B}$. Let $\mathcal{U}^{\mathrm{w}}$ denote the topology generated by $\mathcal{B}$. By construction, $\mathcal{U}^{\mathrm{w}}$ is the smallest topology containing $\mathcal{B}$ and, therefore, the weakest topology that makes every $\phi \in X^*$ continuous.

Definition 3.1.1. Suppose X is a normed vector space.

1. For any $x_0 \in X$, define a weak neighborhood around x_0 by
$$N(x_0, A, \varepsilon) = \{x \in X : |\phi(x) - \phi(x_0)| < \varepsilon, \forall \phi \in A\},$$
where $A \subseteq X^*$ is finite and $\varepsilon > 0$.

2. The weak topology $\mathcal{U}^{\mathrm{w}}$ on X is the topology generated by the basis (analog of the open ball in the metric topology) $N(x_0, \varepsilon, A)$ where $x_0 \in X$, $\varepsilon > 0$ and $A \subseteq X^*$ is finite. In other words, we say $U \subseteq X$ is weakly open if for any $p \in U$, there is a finite set $A \subseteq X^*$ and $\varepsilon > 0$ s.t. $N(p, \varepsilon, A) \subseteq U$.

3. We say $\{x_n\}_{n=1}^{\infty} \subseteq X$ converges to $x_\infty \in X$ weakly if x_n converges to x_∞ in $\mathcal{U}^{\mathrm{w}}$, denoted by $x_n \xrightarrow{\mathrm{w}} x_\infty$.

Remark. 1. $\mathcal{U}^{\mathrm{w}}$ is the weakest topology on X to make every $\phi \in X^*$ continuous.
2. Let $\mathcal{U}$ be the norm topology on X. For any $x_0 \in X$, $\varepsilon > 0$ and finite $A \subseteq X^*$, $N(x_0, \varepsilon, A)$ is open in $\mathcal{U}$ because $\phi \in A \subseteq X^*$ is continuous in $\mathcal{U}$ and the finite intersection of open sets is still open in $\mathcal{U}$.
3. From Remark 2, $\mathcal{U}^{\mathrm{w}} \subseteq \mathcal{U}$, Then

- Open in $\mathcal{U}^{\mathrm{w}}$ implies open in $\mathcal{U}$
- Closed in $\mathcal{U}^{\mathrm{w}}$ implies closed in $\mathcal{U}$
- Convergent in $\mathcal{U}$ implies convergent in $\mathcal{U}^{\mathrm{w}}$.

Lemma 3.1.2. *$x_n \xrightarrow{\mathrm{w}} x_\infty$ if and only if for any $\phi \in X^*$,*
$$\phi(x_n) \to \phi(x_\infty).$$

Definition 3.1.3 (Convex hull)**.** Suppose V is a vector space, $A \subseteq V$. Define the convex hull of A by
$$\mathrm{Conv}(A) := \left\{ \sum_{i=1}^{n} t_i a_i : t_i \geq 0, \sum_{i=1}^{n} t_i = 1, a_i \in A, n \in \mathbb{Z}_+ \right\}$$

Lemma 3.1.4. *Suppose K is a convex set and $n \geq 2$. Then for any $v_1, \cdots, v_n \in K$ and any $t_1, \cdots, t_n$ with $t_i \geq 0$ and $\sum_{i=1}^{n} t_i = 1$, we have*
$$\sum_{i=1}^{n} t_i v_i \in K.$$

Proof. We will prove inductively. The case $n = 2$ is true by definition. Assume the statement is true for $n-1$, i.e. for any $v_1, \cdots, v_{n-1} \in K$ and any $t_1, \cdots, t_{n-1}$ with $t_i \geq 0$ and $\sum_{i=1}^{n-1} t_i = 1$, we have
$$\sum_{i=1}^{n-1} t_i v_i \in K.$$

Then for any $v_1, \cdots, v_n \in K$ and any $t_1, \cdots, t_n$ with $t_i \geq 0$ and $\sum_{i=1}^n t_i = 1$, then

$$\sum_{i=1}^{n-1} t_i = 1 - t_n, \quad \sum_{i=1}^{n-1} \frac{t_i}{1-t_n} = 1,$$

by the assumption, we have

$$\sum_{i=1}^{n-1} \frac{t_i}{1-t_n} v_i \in K.$$

Therefore

$$\sum_{i=1}^{n} t_i v_i = (1-t_n) \sum_{i=1}^{n-1} \frac{t_i}{1-t_n} v_i + t_n v_n \in K. \qquad \square$$

Lemma 3.1.5. Conv(A) *is the smallest convex set that contains* A*, i.e.* Conv(A) *is convex and for any convex set* $K \subseteq X$ *s.t.* $A \subseteq K$*, we have*

$$\mathrm{Conv}(A) \subseteq K.$$

Proof. 1. For any $u = \sum_{i=1}^n t_i a_i, v = \sum_{i=1}^m t_i' a_i \in \mathrm{Conv}(A)$ (we can assume $m = n$, otherwise, if $m < n$, we can let $t_{m+1}' = \cdots = t_n' = 0$) and any $t \in [0,1]$,

$$tu + (1-t)v = t\sum_{i=1}^n t_i a_i + (1-t)\sum_{i=1}^n t_i' a_i = \sum_{i=1}^n [tt_i + (1-t)t_i'] a_i \in \mathrm{Conv}(A)$$

because

$$\sum_{i=1}^n [tt_i + (1-t)t_i'] = t\sum_{i=1}^n t_i + (1-t)\sum_{i=1}^n t_i' = t + (1-t) = 1,$$

Therefore Conv(A) is convex.
2. Moreover, suppose $K \subseteq X$ is convex and $A \subseteq K$. For any $u = \sum_{i=1}^n t_i a_i \in \mathrm{Conv}(A)$, from Lemma 3.1.4 we have $u \in K$. $\square$

Theorem 3.1.6 (Mazur)**.** *Suppose* X *is normed vector space. If there is a sequence* $\{x_n\}_{n=1}^\infty \subseteq X$ *and* $x_\infty \in X$ *s.t.* $x_n \xrightarrow{\mathrm{w}} x_\infty$*, then*[1]

$$x_\infty \in \overline{\mathrm{Conv}(\{x_n\}_{n=1}^\infty)}^{\mathrm{norm}}.$$

Proof. Let $K := \overline{\mathrm{Conv}(\{x_n\}_{n=1}^\infty)}^{\mathrm{norm}}$. Then K is clearly closed, and K is also convex by Lemma 2.5.3. Suppose $x_\infty \notin K$. Then $x_\infty \in K^c$ and K^c is open, thus there is an non-empty open ball $B_\varepsilon(x_\infty) \subseteq K^c$, which is also convex. By Theorem 2.5.1, there is $\Lambda \in X^*$ and $c \in \mathbb{R}$ s.t. for any $k \in K$ and $x \in B_\varepsilon(x_\infty)$,

$$\Lambda(k) \leq c < \Lambda(x).$$

[1]The superscript "norm" means the closure is in the norm topology.

Then there is $\delta > 0$ s.t.

$$|\Lambda(x_j - x_\infty)| = \Lambda(x_\infty) - \Lambda(x_j) \geq \Lambda(x_\infty) - c > \delta > 0, \quad \forall j \in \mathbb{Z}_+,$$

then

$$\lim_{j\to\infty} |\Lambda(x_j - x_\infty)| \geq \delta > 0,$$

which contradicts

$$\Lambda(x_j) \to \Lambda(x_\infty),$$

thus x_j does not converge to x_∞ weakly. □

Corollary 3.1.7. *Suppose X is a normed vector space and $K \subseteq X$ is convex. Then K is closed in $\mathcal{U}$ if and only if it is closed in $\mathcal{U}^{\mathrm{w}}$.*

Proof. If K is closed in $\mathcal{U}^{\mathrm{w}}$, then K is also closed in $\mathcal{U}$, since closed in $\mathcal{U}^{\mathrm{w}}$ always implies closed in $\mathcal{U}$. Conversely, if K is closed in $\mathcal{U}$, suppose there is $\{x_n\}_{n=1}^\infty \subseteq K$ and $x_\infty \in X$ s.t. $x_n \xrightarrow{\mathrm{w}} x_\infty$, then by Mazur's theorem (3.1.6) and Lemma 3.1.5,

$$x_\infty \in \overline{\mathrm{Conv}(\{x_n\}_{n=1}^\infty)}^{\mathrm{norm}} \subseteq \overline{K}^{\mathrm{norm}} = K,$$

i.e. K is closed in $\mathcal{U}^{\mathrm{w}}$.

Another proof without using Mazur's theorem. First, any closed half space $H = \{x \in X : \Lambda(x) \geq c\}$ is weakly closed, because Λ is also continuous in $\mathcal{U}^{\mathrm{w}}$, and then

$$H = \Lambda^{-1}([c, \infty)) \in \mathcal{U}^{\mathrm{w}}.$$

Second, by the Corollary **??**,

$$\begin{aligned} K &= \bigcap \{\text{closed half spaces } H \text{ containing } K\} \\ &= \bigcap \{\text{weakly closed half spaces } H \text{ containing } K\}, \end{aligned}$$

then K is also weakly closed since any intersection of closed sets is also closed. □

We will show a fun fact about the weak topology.

Lemma 3.1.8. *Suppose X is a normed vector space with $\dim(X) = \infty$ and $\phi_1, \cdots, \phi_n \in X^*$. Then*

$$\dim\left(\bigcap_{i=1}^n \mathrm{Ker}\,(\phi_i)\right) = \infty,$$

and hence $\bigcap_{i=1}^n \mathrm{Ker}\,(\phi_i) \neq \varnothing$.

Proof. . □

Theorem 3.1.9. *Suppose X is an infinite-dimensional normed vector space. Then the weak closure of the unit sphere is the closed unit ball.*

Proof. 1. Since the closed unit ball $\overline{B_1} := \{x \in X : \|x\| \leq 1\}$ is convex, by Corollary 3.1.7, we have $\overline{B_1} = \overline{B_1}^{\text{norm}} = \overline{B_1}^{\text{w}}$. Then the unit sphere $S_1 \subseteq \overline{B_1} = \overline{B_1}^{\text{w}}$, thus $\overline{S_1}^{\text{w}} \subseteq \overline{B_1}^{\text{w}} = \overline{B_1}$.
2. It suffices to show $\overline{B_1} \subseteq \overline{S_1}^{\text{w}}$. For any $x_0 \in \overline{B_1}$, let U be a weakly open set containing x_0, then there exists

$$N(x_0, \varepsilon, A) = \{x \in X : |\phi_j(x) - \phi_j(x_0)| < \varepsilon, \forall \phi_j \in A\} \subseteq U$$

for some $\varepsilon > 0$ and $A = \{\phi_1, \cdots, \phi_N\} \subseteq X^*$. From Lemma 3.1.8, $\text{Ker}\,(\phi_j) \setminus \{0\} \neq \varnothing$ and we can choose $y \in \cap_{j=1}^N \text{Ker}\,(\phi_j) \setminus \{0\}$.
3. Claim: there is $t_1 \in \mathbb{R}$ s.t. $\|x_0 + t_1 y\| = 1$.
Let $f(t) = \|x_0 + ty\|$, then $f(0) = \|x_0\| \leq 1$, and

$$f(t) = \|x_0 + ty\| \geq |t| \, \|y\| - \|x_0\| \,,$$

so let $t_0 = (2 + \|x_0\|)/ \, \|y\|$, we have $f(t_0) \geq 2$. Since f is continuous and $1 \in [f(0), f(t_0)]$, by the intermediate value theorem, there is $t_1 \in (0, t_0)$ s.t.

$$f(t_1) = \|x_0 + t_1 y\| = 1,$$

i.e. $x_0 + t_1 y \in S_1$.
4. On the other hand, since y is in the kernel of ϕ_j, we have $\phi_j(x_0 + t_1 y) = \phi_j(x_0)$ for each $j = 1, \cdots, N$, thus

$$|\phi_j(x_0 + t_1 y) - \phi_j(x_0)| = 0 < \varepsilon, \quad \forall j = 1, \cdots, N,$$

i.e. $x_0 + t_1 y \in N(x_0, \varepsilon, A)$. Then $x_0 + t_1 y \in S_1 \cap N(x_0, \varepsilon, A) \subseteq S_1 \cap U$. Since U is arbitrary[2], we have $x_0 \in \overline{S_1}^{\text{w}}$. □

Definition 3.1.10. Suppose X is a real normed vector space and $E \subseteq X^*$, then the set

$$^{\perp}E := \{x \in X : \phi(x) = 0, \forall \phi \in E\}$$

is called the pre-annihilator of E.

Lemma 3.1.11. *Let $E \subseteq X^*$, then $^{\perp}E$ is closed in X.*

Proof. Suppose $\{x_n\}_{n=1}^{\infty} \subseteq {^{\perp}E}$ converges to x_∞, then for any $\phi \in E$,

$$\phi(x_\infty) = \phi(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} \phi(x_n) = 0,$$

so $x_\infty \in {^{\perp}E}$. □

[2]Recall that in the general topology, $x \in \overline{S}$ if and only if every neighborhood of x contains a point of S.

Theorem 3.1.12 (weak closure of a subspace)**.** *Suppose X is a normed vector space and $E \subseteq X$ is a subspace.*

(1) $\overline{E} = {}^{\perp}(E^{\perp}) = \overline{E}^{\mathrm{w}}$.

(2) *E is closed if and only if it is weakly closed if and only if $E = {}^{\perp}(E^{\perp})$.*

(3) *E is dense if and only if E is weakly dense if and only if $E^{\perp} = \{0\}$.*

Theorem 3.1.13 (Eberlein-Shmulyan)**.** *A Banach space is reflexive if and only if every bounded sequence has a weakly convergent subsequence.*

Proof. "$\Longrightarrow$" is more interesting, we will only prove this direction.
Step 1. Let X be reflexive, and $\{x_n\}_{n=1}^{\infty} \subseteq X$ be bounded, we want to show it has a weakly convergent subsequence. Let $Y = \overline{\mathrm{Span}(\{x_n\})}$, then Y is a separable, reflexive Banach space. By Theorem 2.8.4, Y^* is also separable, then Y^* has a countable dense subset, denote it to be $\{\varphi_j\}_{j=1}^{\infty}$.
Step 2. By the diagonalization method, we can find a subsequence $\{x_{\sigma(n)}\}_{n=1}^{\infty} \subseteq \{x_n\}_{n=1}^{\infty}$ s.t. $\{\varphi_i(x_{\sigma(n)})\}_{n=1}^{\infty}$ converges for all $i \geq 1$.
Step 3. Consider the sequence $\{\iota_X(x_{\sigma(n)})\}_{n=1}^{\infty} \subseteq X^{**}$, which converges pointwise for every element of the dense set $\{\varphi_j\}_{j=1}^{\infty}$, then by Theorem 2.1.6, $\{\iota_X(x_{\sigma(n)})\}_{n=1}^{\infty}$ converges strongly in Y^{**} to some element $A \in Y^{**}$. Since X is reflexive, there is $x_\infty \in X$ s.t. $A = \iota_X(x_\infty)$. Therefore for every $\varphi \in X^*$,

$$\varphi(x_{\sigma(n)}) = \iota_X(x_{\sigma(n)})(\varphi) \to \iota_X(x_\infty)(\varphi) = \varphi(x_\infty),$$

i.e. $\{x_{\sigma(n)}\}_{n=1}^{\infty} \subseteq \{x_n\}_{n=1}^{\infty}$ weakly convergent. □

3.2 Weak* topology

Definition 3.2.1. The weak* topology on X^* is the weakest topology on X^* s.t. every element of $\iota_X(X) \subseteq X^{**}$ is continuous.

Example 3.2.2. Let X be a normed vector space. Then $\iota : X \to X^{**}$ is continuous w.r.t. the weak topology on X and the weak* topology on $(X^*)^*$.

Proof. For any open set $O \subseteq \mathcal{U}^{\mathrm{w}^*}(X^*)$, we want to show $\iota^{-1}(O) \in \mathcal{U}^{\mathrm{w}}(X)$. Consider the base

$$N = \{\}$$

Theorem 3.2.3 (Alaoglu)**.** *Suppose X is a separable normed vector space. Then every bounded sequence in X^* has a w*-convergent subsequence.*

Proof. Since X is separable, there is $\{x_j\}_{j=1}^{\infty} \subseteq X$ s.t. it is countable and dense in X. Let $\{\phi_k\}_{k=1}^{\infty} \subseteq X^*$ be a bounded sequence, i.e.

$$\sup_k \|\phi_k\| < \infty.$$

Then for any $j \in \mathbb{Z}_+$,

$$\sup_k |\phi_k(x_j)| \leq \sup_k \|\phi_k\| \cdot \|x_j\| < \infty,$$

i.e. for each $j \in \mathbb{Z}_+$, $\{\phi_k(x_j)\}_{k=1}^{\infty} \subseteq \mathbb{R}$ is bounded. For $j = 1$, since $\{\phi_k(x_1)\}_{k=1}^{\infty} \subseteq \mathbb{R}$ is bounded we can find a subsequence

$$\{\phi_{\sigma_1(k)}\} \subseteq \{\phi_k\}_{k=1}^{\infty}$$

s.t. $\{\phi_{\sigma_1(k)}(x_1)\}_{k=1}^{\infty}$ converges. Since $\{\phi_{\sigma_1(k)}(x_2)\}_{k=1}^{\infty} \subseteq \mathbb{R}$ is bounded, we can find a subsequence

$$\{\phi_{\sigma_2(k)}\} \subseteq \{\phi_{\sigma_1(k)}\}_{k=1}^{\infty}$$

s.t. $\{\phi_{\sigma_2(k)}(x_2)\}_{k=1}^{\infty}$ converges. Iteratively, we can construct a subsequence

$$\{\phi_{\sigma_j(k)}\} \subseteq \{\phi_{\sigma_{j-1}(k)}\}_{k=1}^{\infty}$$

s.t. $\{\phi_{\sigma_j(k)}(x_j)\}_{k=1}^{\infty}$ converges. Therefore the subsequence $\{\phi_{\sigma_j(j)}\}_{j=1}^{\infty}$ satisfies $\{\phi_{\sigma_j(j)}(x_i)\}_{j=1}^{\infty}$ converges for every $i \geq 1$. By Banach-Steinhaus Theorem (2.1.6), $\{\phi_{\sigma_j(j)}(x)\}_{j=1}^{\infty}$ converges for all $x \geq X$, i.e. $\{\phi_{\sigma_j(j)}\}_{j=1}^{\infty}$ is w*-convergent. □

Theorem 3.2.4 (weak* closure of a subspace)**.** *Suppose X is a normed vector space and $E \subseteq X^*$ is a subspace.*

(1) $\overline{E} = {}^{\perp}(E^{\perp}) = \overline{E}^{w^*}$.

(2) *E is closed if and only if it is weak* closed if and only if $E = {}^{\perp}(E^{\perp})$.*

(3) *E is dense if and only if E is weak* dense if and only if $E^{\perp} = \{0\}$.*

Chapter 4

Dual operator and compact operator

4.1 Dual operator

Definition 4.1.1. Let X, Y be normed vector spaces and $A : X \to Y$. Define $A^* : Y^* \to X^*$ by

$$A^*\phi = \phi \circ A, \quad \forall \phi \in Y^*.$$

Lemma 4.1.2. *Suppose X, Y are normed vector spaces and $A \in \mathcal{L}(X, Y)$, $B \in \mathcal{L}(Y, Z)$. Then*

1. *$A^* \in \mathcal{L}(Y^*, X^*)$ and $\|A^*\| = \|A\|$.*

2. *$(BA)^* = A^*B^*$, and $\mathbb{1}_X^* = \mathbb{1}_{X^*}$*

3. *$A^{**} \in \mathcal{L}(X^{**}, Y^{**})$ and $A^{**} \circ \iota_X = \iota_Y \circ A$.*

Proof. 1. We have

$$\begin{aligned}
\|A^*\| &= \sup_{\phi \in Y^* \setminus \{0\}} \frac{\|A^*\phi\|_{X^*}}{\|\phi\|_{Y^*}} \\
&= \sup_{\phi \in Y^* \setminus \{0\}} \sup_{x \in X \setminus \{0\}} \frac{1}{\|\phi\|_{Y^*}} \cdot \frac{|A^*\phi(x)|}{\|x\|_X} \\
&= \sup_{\phi \in Y^* \setminus \{0\}} \sup_{x \in X \setminus \{0\}} \frac{|\phi(Ax)|}{\|\phi\|_{Y^*} \|x\|_X} \\
&= \sup_{x \in X \setminus \{0\}} \frac{1}{\|x\|_X} \sup_{\phi \in Y^* \setminus \{0\}} \frac{|\iota_Y(Ax)(\phi)|}{\|\phi\|_{Y^*}} \\
&= \sup_{x \in X \setminus \{0\}} \frac{\|\iota_Y(Ax)\|_{Y^{**}}}{\|x\|_X} \\
&= \sup_{x \in X \setminus \{0\}} \frac{\|Ax\|_Y}{\|x\|_X} \\
&= \|A\| < \infty.
\end{aligned}$$

2. Since $BA \in \mathcal{L}(X, Z)$, $(BA)^* \in \mathcal{L}(Z^*, X^*)$. For any $\phi \in Z^*$,

$$(BA)^*(\phi) = \phi \circ (BA) = (\phi \circ B) \circ A = B^*(\phi) \circ A = A^*(B^*(\phi)) = A^*B^*(\phi).$$

3. For any $x \in X$, $\phi \in Y^*$,

$$\begin{aligned}[A^{**} \circ \iota_X(x)](\phi) &= [A^{**}(\iota_X(x))](\phi) = \iota_X(x) \circ A^*(\phi) \\ &= \iota_X(x)(A^*\phi) = (A^*\phi)(x) = \phi(Ax) = [\iota_Y \circ A(x)](\phi).\end{aligned}$$

□

Example 4.1.3. Let X, Y be a Hilbert space. Let $A \in \mathcal{L}(X, Y)$. $A^\dagger \in \mathcal{L}(Y, X)$ is called the adjoint operator of A if it satisfies

$$\langle v, Au \rangle = \langle A^\dagger v, u \rangle, \quad \forall u \in X, v \in Y.$$

If $X = Y = H$, $A \in \mathcal{L}(H)$, $R : H \to H^*$ is the Riesz representation isometry, then

$$A^\dagger = R^{-1} \circ A^* \circ R.$$

Lemma 4.1.4. *Let X, Y be real Hilbert spaces and $A \in \mathcal{L}(X, Y)$, then*

$$\|A\|^2 = \left\|A^\dagger A\right\|.$$

Proof. By the definition of adjoint operator, we have

$$\begin{aligned}\|A\|^2 &= \sup_{\|x\|_X=1} \|Ax\|_Y^2 = \sup_{\|x\|_X=1} \langle Ax, Ax \rangle \\ &= \sup_{\|x\|_X=1} \langle A^\dagger Ax, x \rangle \\ &\leq \sup_{\|x\|_X=1} \left\|T^\dagger Tx\right\|_X \|x\|_X \\ &= \left\|T^\dagger T\right\| \\ &\leq \left\|T^\dagger\right\| \|T\| \\ &= \|T\|^2.\end{aligned}$$

□

Theorem 4.1.5 (Duality)**.** *Suppose X, Y are normed vector spaces and $A \in \mathcal{L}(X, Y)$, then:*

(1) $\mathrm{Im}\,(A)^\perp = \mathrm{Ker}\,(A^*)$ *and* ${}^\perp\mathrm{Im}\,(A^*) = \mathrm{Ker}\,(A)$.

(2) $\mathrm{Im}\,(A)$ *is dense in Y if and only if A^* is injective.*

(3) *A is injective if and only if A^* has a w^*-dense image.*

Proof. (1) Notice that

$$\begin{aligned}\operatorname{Im}(A)^{\perp} &= \{\phi \in Y^* : \phi(Ab) = 0,\ \forall b \in X\} \\ &= \{\phi \in Y^* : (A^*\phi)(b) = 0,\ \forall b \in X\} \\ &= \{\phi \in Y^* : A^*\phi = 0\} \\ &= \operatorname{Ker}(A^*),\end{aligned}$$

and

$$\begin{aligned}{}^{\perp}\operatorname{Im}(A^*) &= \{x \in X : A^*\phi(x) = 0,\ \forall \phi \in Y^*\} \\ &= \{x \in X : \phi(Ax) = 0,\ \forall \phi \in Y^*\} \\ &= \{x \in X : \iota_Y(Ax)(\phi) = 0,\ \forall \phi \in Y^*\} \\ &= \{x \in X : \iota_Y(Ax) = 0\} \\ &= \{x \in X : Ax = 0\} \\ &= \operatorname{Ker}(A).\end{aligned}$$

(2)A^* is injective if and only if $\{0\} = \operatorname{Ker}(A^*) = \operatorname{Im}(A)^{\perp}$ if and only if $\overline{\operatorname{Im}(A)} = Y$ by Corollary 2.6.4.
(3) □

Theorem 4.1.6 (Closed image theorem)**.** *Suppose X, Y are Banach and $A \in \mathcal{L}(X, Y)$. TFAE*

(1) $\operatorname{Im}(A) =^{\perp} \operatorname{Ker}(A^*)$

(2) $\operatorname{Im}(A)$ *is closed in* Y

(3) *There is $c > 0$ s.t. for any $w \in X$,*

$$\|[w]\|_{X/\operatorname{Ker}(A)} \leq c\,\|Ax\|_Y .$$

(4) $\operatorname{Im}(A^*) = \operatorname{Ker}(A)^{\perp}$

(5) $\operatorname{Im}(A^*)$ *is w^*-closed in X^*.*

(6) $\operatorname{Im}(A^*)$ *is closed in* X^*

(7) *There is $c > 0$ s.t. for any $\phi \in Y^*$,*

$$\|[\phi]\|_{Y^*/\operatorname{Ker}(A^*)} \leq c\,\|A^*\phi\|_{X^*} .$$

Proof. (1)$\Longrightarrow$(2). By Lemma 3.1.11.
(2)$\Longrightarrow$(3) □

Corollary 4.1.7. *Suppose X, Y are Banach spaces, $A \in \mathcal{L}(X, Y)$. Then*

(1) *A is surjective if and only if A^* is injective with a closed image, i.e. there is $c > 0$ s.t. for any $\phi \in Y^*$,*

$$\|\phi\|_{Y^*} \le c \|A^*\phi\|_{X^*}.$$

(2) *A^* is surjective if and only if A is injective with a closed image i.e. there is $c > 0$ s.t. for any $x \in X$,*

$$\|x\|_X \le c \|Ax\|_Y.$$

Corollary 4.1.8. *Suppose X, Y are Banach spaces, $A \in \mathcal{L}(X, Y)$. Then*

(1) *A is bijective if and only if A^* is bijective.*

(2) *If A^* is bijective, then $(A^*)^{-1} = (A^{-1})^*$*

(3) *A is an isometry if and only if A^* is an isometry.*

4.2 Compact operator

Definition 4.2.1. Suppose X, Y are Banach, $K \in \mathcal{L}(X, Y)$. K is

1. compact if for any bounded subset $S \subseteq X$, $\overline{K(S)}$ is compact.
2. completely continuous if the image of every weakly convergent sequence in X is norm convergent in Y
3. finite-rank if $\dim \operatorname{Im}(K) < \infty$.

Lemma 4.2.2. *Suppose X, Y are Banach and $K \in \mathcal{L}(X, Y)$. TFAE*

(1) *K is compact*

(2) *$\overline{K(\overline{B_1(0)})}$ is compact*

(3) *If $\{x_n\}_{n=1}^\infty$ is bounded, then $\{Kx_n\}_{n=1}^\infty$ has a convergent subsequence.*

Lemma 4.2.3. *Suppose X, Y are Banach, $K \in L(X, Y)$.*

(1) *If K is compact, then K is completely continuous.*

(2) *If X is reflexive and K is completely continuous, then K is compact.*

Proof. (1) Step 1. Let $\{x_n\}_{n=1}^\infty$ be weakly convergent.
Claim: $\{x_n\}_{n=1}^\infty$ is bounded.
By the definition of weakly convergence, for any $\phi \in X^*$,

$$\{\iota_X(x_n)(\phi)\}_{n=1}^\infty = \{\phi(x_n)\}_{n=1}^\infty$$

is convergent and hence bounded, i.e. $\{\iota_X(x_n)\}_{n=1}^{\infty} \subseteq X^{**}$ is pointwise bounded. By the Uniformly bounded principle (2.1.2), there is $c > 0$ s.t.

$$\sup_n \|x_n\| = \sup_n \|\iota_X(x_n)\| < c,$$

i.e. $\{x_n\}_{n=1}^{\infty}$ is bounded.
Step 2. Since K is compact and $\{x_n\}_{n=1}^{\infty}$ is bounded, $\{Kx_n\}_{n=1}^{\infty}$ has a convergent subsequence $\{Kx_{\sigma_1(n)}\}_{n=1}^{\infty}$ which converges to $y_\infty \in Y$.
Claim: $Kx_n \to y_\infty$.
Suppose $\{Kx_n\}_{n=1}^{\infty}$ does not converge to y_∞, there is $\varepsilon > 0$ s.t. for all $N \in Z_+$, there is $n > N$ s.t.

$$\|Kx_n - y_\infty\| \geq \varepsilon,$$

so there is a new subsequence $\{Kx_{\sigma_2(n)}\}_{n=1}^{\infty}$ s.t.

$$\left\|Kx_{\sigma_2(n)} - y_\infty\right\| \geq \varepsilon.$$

$\{x_{\sigma_2(n)}\}_{n=1}^{\infty}$ is weakly convergent thus bounded, then there is a convergent subsequence

$$\{Kx_{\sigma_3(n)}\}_{n=1}^{\infty} \subseteq \{Kx_{\sigma_2(n)}\}_{n=1}^{\infty},$$

s.t. $Kx_{\sigma_3(n)} \to z_\infty \in Y$ and $z_\infty \neq y_\infty$. By Corollary 2.4.6, there is $\psi \in Y^*$ s.t.

$$\psi(y_\infty) \neq \psi(z_\infty).$$

However, let x_∞ be the weak limit of $\{x_n\}_{n=1}^{\infty}$, then

$$\psi(y_\infty) = \lim_{n\to\infty} \psi(Kx_{\sigma_1(n)}) = \lim_{n\to\infty} K^*\psi(x_{\sigma_1(n)}) = K^*\psi(x_\infty),$$

similarly,

$$\psi(z_\infty) = K^*\psi(x_\infty),$$

which contradicts $\psi(y_\infty) \neq \psi(z_\infty)$.
(2) Assume X is reflexive and K is completely continuous. Let $\{x_n\}_{n=1}^{\infty} \subseteq X$ be a bounded sequence, then by Theorem 3.1.13, there is a weakly convergent subsequence

$$\{x_{\sigma(n)}\}_{n=1}^{\infty} \subseteq \{x_n\}_{n=1}^{\infty}.$$

By the definition of "completely continuous", $\{Kx_{\sigma(n)}\}_{n=1}^{\infty}$ is convergent, so K is compact. □

Example 4.2.4. Suppose X, Y are Banach, $K \in \mathcal{L}(X, Y)$ is finite-rank, then K is compact.

Example 4.2.5. Consider the space $L^2(S^1)$, any function $f \in L^2(S^1)$ can be spanned in this way:

$$f(t) = \sum_{k\in\mathbb{Z}} f_k e^{ikt},$$

where $f_k \in \mathbb{C}$. Define the norm $\|\cdot\|_{L^2}$ and $\|\cdot\|_{H_1}$ by

$$\|f\|_{L^2} = \sqrt{2\pi \sum_{k\in\mathbb{Z}} |f_k|^2},$$

and

$$\|f\|_{H_1} = \sqrt{\|f\|_{L^2}^2 + \left\|\frac{\mathrm{d}f}{\mathrm{d}t}\right\|_{L^2}^2} = \sqrt{2\pi \sum_{k\in\mathbb{Z}} (1+k^2)|f_k|^2}.$$

Let $H_1 = \{f \in L^2(S^1),\ s.t.\ \|f\|_{H_1} < \infty.\}$

Proposition. The inclusion map $\iota : H_1 \to L^2(S^1)$ is compact.

Proof. Let $\{f^n\}_{n=1}^\infty \subseteq H_1$ s.t. it is bounded, i.e. there is $M > 0$ s.t.

$$\sup_n \|f^n\|_{H_1}^2 < M,$$

i.e. for any $n \geq 1$,

$$\sum_{k\in\mathbb{Z}} (1+k^2)|f_k^n|^2 < M.$$

Then $\{f_k^n\}_{n=1}^\infty \subseteq \mathbb{C}$ is bounded for every $k \in \mathbb{Z}$. By Bolzano–Weierstrass theorem, there is a subsequence $\{f^{\sigma_0(n)}\}_{n=1}^\infty \subseteq \{f^n\}_{n=1}^\infty \subseteq H_1$ s.t. $\{f_0^{\sigma_0(n)}\}_{n=1}^\infty \subseteq \mathbb{C}$ converges. Inductively, we can find subsequences

$$\{f^{\sigma_k(n)}\}_{n=1}^\infty \subseteq \{f^{\sigma_{k-1}(n)}\}_{n=1}^\infty \subseteq \cdots \subseteq \{f^n\}_{n=1}^\infty \subseteq H_1,$$

s.t. $\{f_j^{\sigma_k(n)}\}_{n=1}^\infty \subseteq \mathbb{C}$ converges for all $|j| \leq k$. We can then pass to the diagonal subsequence $\{f^{\sigma_n(n)}\}_{n=1}^\infty$ to get

$$\{f_j^{\sigma_n(n)}\}_{n=1}^\infty$$

converges for all $j \in \mathbb{Z}$.
Claim: $\{f^{\sigma_n(n)}\}_{n=1}^\infty$ converges in $L^2(S^1)$.
It suffices to show it is Cauchy. Given $\varepsilon > 0$, we want to find N s.t. for any $l, m \geq N$,

$$\left\|f^{\sigma_l(l)} - f^{\sigma_m(m)}\right\|_{L^2} < \varepsilon.$$

Notice that there is $k_0 \geq 1$ s.t. for all $|k| \geq k_0$,

$$\frac{4M}{1+k^2} < \frac{\varepsilon^2}{2}.$$

Then

$$\begin{aligned}
\left\|f^{\sigma_l(l)}-f^{\sigma_m(m)}\right\|_{L^2}^2 &= 2\pi\sum_{k\in\mathbb{Z}}\left|f_k^{\sigma_l(l)}-f_k^{\sigma_m(m)}\right|^2 \\
&\le 2\pi\sum_{k:|k|<k_0}\left|f_k^{\sigma_l(l)}-f_k^{\sigma_m(m)}\right|^2+\frac{2\pi}{1+k_0^2}\sum_{k:|k|\ge k_0}\left|f_k^{\sigma_l(l)}-f_k^{\sigma_m(m)}\right|^2(1+k^2) \\
&\le 2\pi\sum_{k:|k|<k_0}\left|f_k^{\sigma_l(l)}-f_k^{\sigma_m(m)}\right|^2+\frac{1}{1+k_0^2}\left\|f^{\sigma_l(l)}-f^{\sigma_m(m)}\right\|_{H_1}^2 \\
&\le 2\pi\sum_{k:|k|<k_0}\left|f_k^{\sigma_l(l)}-f_k^{\sigma_m(m)}\right|^2+\frac{4M}{1+k_0^2} \\
&\le 2\pi\sum_{k:|k|<k_0}\left|f_k^{\sigma_l(l)}-f_k^{\sigma_m(m)}\right|^2+\frac{\varepsilon^2}{2}.
\end{aligned}$$

Since $\{f_k^{\sigma_n(n)}\}_{n=1}^\infty$ converges for all $k\in\mathbb{Z}$, there is $N\ge 1$ s.t. for all $l,m\ge N$ and $|k|<k_0$,

$$2\pi\sum_{k:|k|<k_0}\left|f_k^{\sigma_l(l)}-f_k^{\sigma_m(m)}\right|^2<\frac{\varepsilon^2}{2},$$

therefore

$$\left\|f^{\sigma_l(l)}-f^{\sigma_m(m)}\right\|_{L^2}^2<\varepsilon^2.$$

□

Theorem 4.2.6. *Suppose X, Y, Z are Banach, $A\in\mathcal{L}(X,Y)$, $B\in\mathcal{L}(Y,Z)$.*

(1) *If A or B is compact, then $B\circ A$ is compact.*

(2) *$K(X,Y)$ is a closed subspace of $\mathcal{L}(X,Y)$*

(3) *$A\in K(X,Y)$ if and only if $A^*\in K(Y^*,X^*)$.*

Proof. (1) Clear
(2) Step 1. Let $\{K_j\}_{j=1}^\infty\subseteq K(X,Y)$ s.t.

$$K_j\xrightarrow{\|\cdot\|}K_\infty\in\mathcal{L}(X,Y),$$

we want to show K_∞ is compact. Let $\{x_n\}_{n=1}^\infty\subseteq X$ be a bounded sequence, say $\|x_n\|\le M$ for all n, our goal is to show $\{K_\infty(x_n)\}_{n=1}^\infty$ has a convergent subsequence.
Step 2. Since K_1 is compact, $\{x_n\}_{n=1}^\infty$ is bounded, we can find a subsequence

$$\{x_{\sigma_1(n)}\}_{n=1}^\infty\subseteq\{x_n\}_{n=1}^\infty$$

s.t. $\{K_1 x_{\sigma_1(n)}\}_{n=1}^{\infty}$ converges, iteratively, we can find subsequences

$$\{x_{\sigma_k(n)}\}_{n=1}^{\infty} \subseteq \{x_{\sigma_{k-1}(n)}\}_{n=1}^{\infty} \subseteq \cdots \subseteq \{x_{\sigma_1(n)}\}_{n=1}^{\infty} \subseteq \{x_n\}_{n=1}^{\infty}$$

s.t. $\{K_j x_{\sigma_k(n)}\}_{n=1}^{\infty}$ converges for all $j \le k$. Therefore, we find a subsequence

$$\{x_{\sigma_n(n)}\}_{n=1}^{\infty} \subseteq \{x_n\}_{n=1}^{\infty}$$

s.t. $\{K_j x_{\sigma_n(n)}\}_{n=1}^{\infty}$ converges for all $j \ge 1$.
Step 3. Now we want to show $\{K_\infty x_{\sigma_n(n)}\}_{n=1}^{\infty}$ converges, i.e. it is Cauchy.
Let $\varepsilon > 0$, we want to find $N > 0$ s.t.

$$|K_\infty x_{\sigma_n(n)} - K_\infty x_{\sigma_m(m)}| < \varepsilon, \quad \forall m, n \ge N.$$

Since $K_j \to K_\infty$ in norm, there is $N_1 > 0$ s.t.

$$\|K_j - K_\infty\| < \frac{\varepsilon}{4M}, \quad \forall j \ge N_1.$$

Choose $j_0 = N_1$, then

$$\begin{aligned}
|K_\infty x_{\sigma_n(n)} - K_\infty x_{\sigma_m(m)}| &= |(K_\infty - K_{j_0})(x_{\sigma_n(n)} - x_{\sigma_m(m)}) + K_{j_0}(x_{\sigma_n(n)} - x_{\sigma_m(m)})| \\
&\le |(K_\infty - K_{j_0})(x_{\sigma_n(n)} - x_{\sigma_m(m)})| + |K_{j_0}(x_{\sigma_n(n)} - x_{\sigma_m(m)})| \\
&\le \|K_\infty - K_{j_0}\| \cdot \left\|x_{\sigma_n(n)} - x_{\sigma_m(m)}\right\| + |K_{j_0}(x_{\sigma_n(n)} - x_{\sigma_m(m)})| \\
&\le \frac{\varepsilon}{4M} \cdot 2M + |K_{j_0}(x_{\sigma_n(n)} - x_{\sigma_m(m)})| \\
&= \frac{\varepsilon}{2} + |K_{j_0}(x_{\sigma_n(n)} - x_{\sigma_m(m)})|.
\end{aligned}$$

Since $\{K_{j_0} x_{\sigma_n(n)}\}_{n=1}^{\infty}$ converges, hence is Cauchy, there is $N > 0$ s.t.

$$|K_{j_0}(x_{\sigma_n(n)} - x_{\sigma_m(m)})| < \frac{\varepsilon}{2}, \quad \forall m, n \ge N,$$

thus

$$|K_\infty x_{\sigma_n(n)} - K_\infty x_{\sigma_m(m)}| < \varepsilon, \quad \forall m, n \ge N.$$

Therefore, K_∞ is compact and $K(X, Y)$ is closed.

(3) We only prove $A \in K(X, Y) \Longrightarrow A^* \in K(Y^*, X^*)$. The other direction is similar.
Step 1. Let $A \in K(X, Y)$, then $\overline{A(\overline{B_1(0)})}$ is a compact subset of Y, denoted as M, which is a compact metric space. For $\phi \in Y^*$, define $f_\phi : M \to \mathbb{R}$ by $f_\phi = \phi|_M$. Define

$$\mathcal{F} = \{f_\phi \in C(M) : \|\phi\|_{Y^*} \le 1\} \subseteq C(M).$$

For any $\phi \in Y^*$,

$$\left\|f_\phi\right\|_{C(M)} = \sup x \in X : \|x\|_X \le 1 \phi(Ax) = \sup x \in X : \|x\|_X \le 1 A^*\phi(x) = \|A^*\phi\|_{X^*}. \quad (*)$$

Step 2. $\mathcal{F}$ is a bounded subset of $C(M)$.
For $f_\phi \in \mathcal{F}$,

$$\begin{aligned}\left\|f_\phi\right\|_{C(M)} &= \sup_{y\in M} |f_\phi(y)| \\ &= \sup_{x\in X:\|x\|\leq 1} |\phi(Ax)| \\ &\leq \sup_{x\in X:\|x\|\leq 1} \|\phi\|_{Y^*}\cdot\|A\|\cdot\|x\|_X \\ &\leq \|A\|, \qquad (\text{since } \|\phi\|_{Y^*}\leq 1 \text{ and } \|x\|\leq 1)\end{aligned}$$

Step 3. $\mathcal{F}$ is pre-compact in $C(M)$.
By Arzelà–Ascoli theorem A.2.3, it suffices to show $\mathcal{F}$ is equi-continuous. For any $y_1, y_2 \in M$ and $f_\phi \in \mathcal{F}$,

$$|f_\phi(y_1) - f_\phi(y_2)| = |\phi(y_1) - \phi(y_2)| \leq \|\phi\|_{Y^*}\|y_1 - y_2\|_Y \leq \|y_1 - y_2\|_Y\,,$$

therefore $\mathcal{F}$ is equi-continuous and hence pre-compact.
Step 4. Finally, A^* is compact.
We will show $A^*(\overline{B_1^{Y^*}(0)})$ is pre-compact. Let $\{\phi_n\}_{n=1}^\infty \subseteq \overline{B_1^{Y^*}(0)}$, i.e. $\|\phi_n\|_{Y^*} \leq 1$, so $\{f_{\phi_n}\}_{n=1}^\infty \subseteq \mathcal{F}$. Since $\mathcal{F}$ is pre-compact, by Theorem A.2.1, it has a convergent subsequence $\{f_{\phi_{\sigma(n)}}\}_{n=1}^\infty$. Then by $(*)$,

$$\left\|A^*\phi_{\sigma(m)} - A^*\phi_{\sigma(n)}\right\|_{X^*} = \left\|f_{\phi_{\sigma(m)}} - f_{\phi_{\sigma(n)}}\right\|_{C(M)}\,,$$

i.e. $\{A^*\phi_{\sigma(n)}\}_{n=1}^\infty \subseteq A^*(\overline{B_1^{Y^*}(0)})$ is Cauchy and thus convergent. By Theorem A.2.1 again, $\overline{A^*(B_1^{Y^*}(0))}$ is pre-compact, therefore A^* is compact. □

Chapter 5

Spectral theory

5.1 Spectrum

Definition 5.1.1. Let X be Banach, $A \in \mathcal{L}(X)$. Define the spectrum of A by

$$\sigma(A) = \{\lambda \in \mathbb{C} : \lambda\mathbb{1} - A \text{ is not bijective}\} = P_\sigma(A) \sqcup C_\sigma(A) \sqcup R_\sigma(A),$$

where P_σ is the point spectrum,

$$P_\sigma := \{\lambda \in \sigma(A) : \lambda\mathbb{1} - A \text{ is not injective}\},$$

C_σ is the continuous spectrum

$$C_\sigma := \{\lambda \in \sigma(A) : \lambda\mathbb{1} - A \text{ is injective and } \operatorname{Im}(\lambda\mathbb{1} - A) \text{ is dense in } X\},$$

R_σ is the residual spectrum

$$R_\sigma := \{\lambda \in \sigma(A) : \lambda\mathbb{1} - A \text{ is injective and } \operatorname{Im}(\lambda\mathbb{1} - A) \text{ is not dense in } X\}.$$

Define the resolvent set of A by

$$\rho(A) = \sigma(A)^c = \{\lambda \in \mathbb{C} : \lambda\mathbb{1} - A \text{ is bijective}\}.$$

Example 5.1.2 (Spectrum of left-shift operator)**.** Let $X = \ell^2$, define $A \in \mathcal{L}(\ell^2)$ by

$$A(x_1, x_2, \cdots) = (x_2, x_3, \cdots), \quad \forall (x_1, x_2, \cdots) \in X.$$

For any $(x_1, x_2, \cdots), (y_1, y_2, \cdots) \in X$, we have

$$\langle A(x_1, x_2, \cdots), (y_1, y_2, \cdots)\rangle = x_2 y_1 + x_3 y_2 + \cdots = \langle (x_1, x_2, \cdots), (0, y_1, y_2, \cdots)\rangle,$$

so

$$A^\dagger(y_1, y_2, \cdots) = (0, y_1, y_2, \cdots).$$

Claim: $\sigma(A) = \sigma(A^\dagger) = \overline{B_1(0)}$.

Proof. Let $\lambda \in B_1(0) \setminus \{0\}$, then $(\lambda, \lambda^2, \cdots) \in X$ because

$$\sum_{n=1}^{\infty} |\lambda^n|^2 = \frac{|\lambda|^2}{1-|\lambda|^2} < \infty.$$

And observe that

$$A(\lambda, \lambda^2, \cdots) = (\lambda^2, \lambda^3, \cdots) = \lambda(\lambda, \lambda^2, \cdots),$$

which implies λ is an eigenvalue, i.e. $\lambda \in B_1(0) \setminus \{0\} \subseteq P_\sigma(A)$. And

$$A(1, 0, 0, \cdots) = 0,$$

so $0 \in P_\sigma(A)$ and hence $B_1(0) \subseteq P_\sigma(A)$.

Suppose $\lambda \in P_\sigma(A) \cap B_1^c(0)$, then there is $v = (v_1, v_2, \cdots) \in X$ s.t.

$$(v_2, v_3, \cdots) = \lambda(v_1, v_2, \cdots),$$

i.e. $v_2 = \lambda v_1, v_3 = \lambda v_2, \cdots$, and we have

$$v_n = \lambda^{n-1} v_1,$$

i.e.

$$v = (v_1, \lambda v_1, \lambda^2 v_1, \cdots),$$

however

$$\sum_{n=1}^{\infty} |\lambda^{n-1} v_1|^2 = |v_1|^2 \sum_{n=1}^{\infty} |\lambda|^{2(n-1)}$$

is not convergent because $|\lambda| \geq 1$. Therefore $B_1(0) \setminus \{0\} = \varnothing$, thus $P_\sigma(A) = B_1(0)$.

Consider other spectrum, let $\lambda \in \sigma(A) \setminus P_\sigma(A)$. Then $(\lambda \mathbb{1} - A)$ is injective. If $\operatorname{Im}(\lambda \mathbb{1} - A)$ is closed, then $\operatorname{Im}(\lambda I - A)$ is Banach. Since $(\lambda \mathbb{1} - A) : X \to \operatorname{Im}(\lambda \mathbb{1} - A)$ is bijective, by the inverse operator theorem (2.2.6), there is $c > 0$ s.t. for any $v \in X$,

$$\|v\| \leq c \, \|\lambda \mathbb{1} - A\| . \tag{$*$}$$

Suppose $\lambda \in S^1$, choose $\varepsilon > 0$, consider the sequence $(v_1, v_2, \cdots) \in X$ where

$$v_j = (\frac{\lambda}{1+\varepsilon})^j.$$

We have

$$\|(\lambda \mathbb{1} - A)v\|^2 = \sum_{j=1}^{\infty} |\lambda v_j - v_{j+1}|^2 = \sum_{j=1}^{\infty} \left| \lambda v_j - \frac{\lambda}{1+\varepsilon} v_j \right|^2 = \frac{|\lambda|^2 \varepsilon^2}{(1+\varepsilon)^2} \sum_{j=1}^{\infty} |v_j|^2 = \frac{\varepsilon^2}{(1+\varepsilon)^2} \|v\| ,$$

then $(*)$ implies

$$\|v\|^2 \leq c^2 \|\lambda \mathbb{1} - A\|^2 = \frac{c^2\varepsilon^2}{(1+\varepsilon)^2} \|v\|^2,$$

so

$$1 \leq \frac{c^2\varepsilon^2}{(1+\varepsilon)^2},$$

but since ε is arbitrary, if we choose $\varepsilon = 1/c$, then

$$\frac{c^2\varepsilon^2}{(1+\varepsilon)^2} = \frac{1}{(1+1/c)^2} < 1,$$

which leads to a contradiction! So for any $\lambda \in S^1$, $\mathrm{Im}\,(\lambda\mathbb{1} - A)$ is not closed, hence $\mathrm{Im}\,(\lambda\mathbb{1} - A) \subsetneq X$, and $S^1 \subseteq \sigma(A)$.

For any $\lambda \in S^1$, is $\mathrm{Im}\,(\lambda\mathbb{1} - A)$ dense in X? Notice that

$$X = \overline{\mathrm{Im}\,(\lambda\mathbb{1} - A)} \oplus \mathrm{Ker}\,((\lambda\mathbb{1} - A)^\dagger) = \overline{\mathrm{Im}\,(\lambda\mathbb{1} - A)} \oplus \mathrm{Ker}\,(\overline{\lambda}\mathbb{1} - A^\dagger).$$

If $v = (v_1, v_2, \cdots) \in \mathrm{Ker}\,(\overline{\lambda}\mathbb{1} - A^\dagger)$, we have

$$0 = (\overline{\lambda}\mathbb{1} - A^\dagger)(v_1, v_2, \cdots) = (\overline{\lambda}v_1, \overline{\lambda}v_2 - v_1, \overline{\lambda}v_3 - v_2, \cdots),$$

then $v_1 = 0, v_2 = 0, \cdots$, and hence $v = 0$. Therefore $X = \overline{\mathrm{Im}\,(\lambda\mathbb{1} - A)}$ and $\mathrm{Im}\,(\lambda\mathbb{1} - A)$ is dense in X, i.e. $S^1 \subseteq C_\sigma(A)$. Now we have shown $\overline{B_1(0)} \subseteq P_\sigma(A) \cup C_\sigma(A)$.

If $\lambda \in \overline{B_1(0)}^c$, i.e. $\lambda > 1$. Let $A \in \mathcal{L}(X)$ s.t. $\|A\| = 1$,

$$\lambda\mathbb{1} - A = \lambda(\mathbb{1} - \frac{A}{\lambda}),$$

where the spectrum radius $r_{A/\lambda} \leq \|A/\lambda\| < 1$, thus $\mathbb{1} - A/\lambda$ is invertible by Theorem 1.6.5 and so is $\lambda\mathbb{1} - A$, therefore $\lambda\mathbb{1} - A$ is bijective and $\lambda \notin \sigma(A)$. Together, we have

$$P_\sigma(A) = B_1(0), \quad C_\sigma(A) = S^1, \quad R_\sigma(A) = \varnothing.$$

Similarly, we can show that

$$P_\sigma(A^\dagger) = \varnothing, \quad C_\sigma(A^\dagger) = S^1, \quad R_\sigma(A^\dagger) = B_1(0).$$ □

Theorem 5.1.3. *Let X be Banach, $A \in \mathcal{L}(X)$.*

(1) *$\sigma(A)$ is compact in $\mathbb{C}$.*

(2) *$\sigma(A) = \sigma(A^*)$.*

(3) $P_\sigma(A^*) \subseteq P_\sigma(A) \cup R_\sigma(A)$, $\quad P_\sigma(A) \subseteq P_\sigma(A^*) \cup R_\sigma(A^*)$,
$R_\sigma(A^*) \subseteq P_\sigma(A) \cup C_\sigma(A)$, $\quad R_\sigma(A) \subseteq P_\sigma(A^*)$,
$C_\sigma(A^*) \subseteq C_\sigma(A)$, $\quad C_\sigma(A) \subseteq R_\sigma(A^*) \cup C_\sigma(A^*)$.

(4) *If X is reflexive, then $C_\sigma(A^*) = C_\sigma(A)$ and $R_\sigma(A^*) \subseteq P_\sigma(A)$.*

Proof. (1) If $|\lambda| > \|A\|$, then

$$r_{A/\lambda} \le \left\| \frac{A}{\lambda} \right\| = \frac{\|A\|}{|\lambda|} < 1,$$

by Theorem 1.6.5, we have $\mathbb{1} - A/\lambda$ is invertible, thus $\lambda\mathbb{1} - A$ is invertible and hence bijective. So $\sigma(A) \subseteq \overline{B_{\|A\|}(0)}$, i.e. $\sigma(A)$ is bounded. It suffices to show $\sigma(A)$ is closed, then by Heine-Borel theorem, $\sigma(A)$ is compact. We will then show $\rho(A) = \sigma(A)^c$ is open. For any $\lambda \in \rho(A)$, $\lambda\mathbb{1} - A$ is invertible. The set of all invertible operators denoted G is open by Theorem 1.6.5, then there is $B_\delta(\lambda\mathbb{1} - A) \subseteq G$ for some $\delta > 0$. Let $B_\varepsilon = (\lambda + \varepsilon)\mathbb{1} - A$, where $\varepsilon \in \mathbb{C}$ satisfies $0 \le |\varepsilon| < \delta$, then

$$\|B_\varepsilon - (\lambda\mathbb{1} - A)\| = \|\varepsilon\mathbb{1}\| = |\varepsilon| < \delta,$$

so $B_\varepsilon \in B_\delta(\lambda\mathbb{1} - A) \subseteq G$ is invertible, i.e. $\lambda + \varepsilon \in \rho(A)$ for any $0 \le |\varepsilon| < \delta$, which implies $B_\delta(\lambda) \subseteq \rho(A)$, hence we have shown $\rho(A)$ is open.
(2) By Corollary 4.1.8, $\lambda\mathbb{1} - A$ is bijective if and only if $(\lambda\mathbb{1} - A)^* = \lambda\mathbb{1} - A^*$ is bijective, thus $\rho(A) = \rho(A^*)$ and hence $\sigma(A) = \sigma(A^*)$.
(3)We only prove part of these relations.
(i) Prove $P_\sigma(A^*) \subseteq P_\sigma(A) \cup R_\sigma(A)$.
First $P_\sigma(A^*) \subseteq \sigma(A)$ by (2). Let $\lambda \in P_\sigma(A^*)$, then $\lambda\mathbb{1} - A^*$ is not injective, by Theorem 4.1.5, $\mathrm{Im}\,(\lambda\mathbb{1} - A)$ is not dense, thus $\lambda \notin C_\sigma(A)$.
(ii) Prove $R_\sigma(A^*) \subseteq P_\sigma(A) \cup C_\sigma(A)$.
Let $\lambda \in R_\sigma(A^*)$, then $\lambda\mathbb{1} - A^*$ is injective but has no dense image. By Theorem 4.1.5, injectivity implies that $\lambda\mathbb{1} - A$ is dense, thus $\lambda \notin R_\sigma(A)$.
(iii) Prove $C_\sigma(A^*) \subseteq C_\sigma(A)$.
Let $\lambda \in C_\sigma(A^*)$, then $\lambda\mathbb{1} - A^*$ is injective and $\mathrm{Im}\,(\lambda\mathbb{1} - A^*)$ is dense. By Theorem 3.2.4, $\lambda\mathbb{1} - A^*$ is also weak* dense, then by Theorem 4.1.5, $\lambda\mathbb{1} - A$ is injective, thus $\lambda \notin P_\sigma(A)$. Again by Theorem 4.1.5, $\lambda\mathbb{1} - A^*$ is injective implies $\lambda\mathbb{1} - A$ has a dense image. Therefore $\lambda \in C_\sigma(A)$. □

Theorem 5.1.4. *Let X be Banach, $A \in \mathcal{L}(A)$, then $\sigma(A) \neq \varnothing$ and*

$$r_A = \sup_{\lambda \in \sigma(A)} |\lambda|.$$

Proof. If $|\lambda| > r_A$, then

$$r_{A/\lambda} = \lim_{n\to\infty} \left\| \left(\frac{A}{\lambda}\right)^n \right\|^{1/n} = \frac{r_A}{|\lambda|} < 1,$$

then $\mathbb{1} - A/\lambda$ is invertible by Theorem 1.6.5 and so is $\lambda\mathbb{1} - A$. Thus $\lambda \notin \sigma(A)$, then

$$\sup_{\lambda \in \sigma(A)} |\lambda| \le r_A.$$

Suppose $\sup_{\lambda\in\sigma(A)}|\lambda| < r_A$. Then there is $\varepsilon > 0$ s.t. for any $\lambda \in \mathbb{C}$ with $|\lambda| = r_A - \varepsilon$, we have $\lambda \notin \sigma(A)$, i.e. $\lambda\mathbb{1} - A$ is bijective and hence $(\lambda\mathbb{1} - A)^{-1} \leq M$ for some $M > 0$. By the holomorphic functional calculus, for any $n \geq 1$,

$$A^n = \frac{1}{2\pi i}\int_{|\lambda|=r_A-\varepsilon} \lambda^n(\lambda\mathbb{1} - A)^{-1}\,\mathrm{d}\lambda,$$

thus

$$\begin{aligned}\|A^n\| &\leq \frac{1}{2\pi}\int_{|\lambda|=r_A-\varepsilon} |\lambda|^n \left\|(\lambda\mathbb{1} - A)^{-1}\right\|\,\mathrm{d}\lambda \\ &\leq \frac{|r_A-\varepsilon|^n M}{2\pi}\int_{|\lambda|=r_A-\varepsilon} \mathrm{d}\lambda \\ &\leq C|r_A-\varepsilon|^{n+1}\end{aligned},$$

and then

$$\|A^n\|^{1/n} \leq |r_A-\varepsilon|^{1+1/n}C^{1/n},$$

let $n \to \infty$, we have

$$r_A \leq r_A - \varepsilon,$$

i.e. $r_A < r_A$, leading to a contradiction! □

5.2 Spectrum of compact operators

Proposition 5.2.1. *Suppose X is an infinite-dimensional Banach space, $T \in K(X)$, then $0 \in \sigma(T)$.*

Proof. Assume $0 \notin \sigma(T)$, then $T = -(0\mathbb{1} - T)$ is bijective, and by the Inverse operator theorem (2.2.6), $T^{-1} \in \mathcal{L}(X)$, then $\mathbb{1} = T^{-1}\circ T$ is compact (4.2.6), hence

$$\overline{B_1(0)} = \overline{\mathbb{1}(\overline{B_1(0)})}$$

is compact, which contradicts Theorem 1.2.7. □

Definition 5.2.2. Let $A \in \mathcal{L}(X)$, define the eigenspace w.r.t. λ by $E_\lambda(A) := \mathrm{Ker}\,(\lambda I - A)$.

Theorem 5.2.3. *Let X be Banach, $T \in K(X)$, then for any $\varepsilon > 0$,*

$$\bigoplus_{\lambda\in P_\sigma(T):|\lambda|\geq\varepsilon} E_\lambda(T)$$

is finite-dimensional, i.e. there is a finite number of $\lambda_i \in P_\sigma(T)$ s.t. for each i,

$$|\lambda_i| \geq \varepsilon, \quad \dim E_{\lambda_i}(T) < \infty.$$

Proof. Step 1. Suppose not, then there is a sequence of linearly independent vectors $\{x_j\}_{j=1}^{\infty}$ s.t.

$$Tx_j = \lambda_j x_j, \quad \forall j \geq 1,$$

where $\lambda_j \in P_\sigma(T)$ and $|\lambda_j| \geq \varepsilon$. Let $Y_k = \text{Span}(\{x_1, x_2, \cdots, x_k\})$. Choose $y_1 \in Y_1$ s.t. $\|y_1\| = 1$. By Riesz's Lemma (1.5.6), there is $y_2 \in Y_2$ s.t. $\|y_2\| = 1$ and

$$\|y_2 - z\| \geq \frac{1}{2}, \quad \forall z \in Y_1.$$

Iteratively, we can find a sequence $\{y_k\}_{k=1}^{\infty} \subseteq X$ s.t. $y_{k+1} \in Y_{k+1}$, $\|y_{k+1}\| = 1$ and

$$\|y_{k+1} - z\| \geq \frac{1}{2}, \quad \forall z \in Y_k.$$

Step 2. Claim: $\{T(\frac{y_k}{\lambda_k})\}_{k=1}^{\infty}$ does not contain any Cauchy subsequences.
Let $y_k = \sum_{j=1}^{k} a_j x_j$, then

$$T(\frac{y_k}{\lambda_k}) = a_k x_k + \frac{1}{\lambda_k}\sum_{j=1}^{k-1} a_j T x_j = a_k x_k + \frac{1}{\lambda_k}\sum_{j=1}^{k-1} a_j \lambda_j x_j = y_k + z_k,$$

for some $z_k \in Y_{k-1}$. Then for any n with $n < k$, we have

$$\left\| T(\frac{y_k}{\lambda_k}) - T(\frac{y_n}{\lambda_n}) \right\| = \|y_k + z_k - (y_n + z_n)\| = \|y_k + (z_k - y_n - z_n)\| \geq \frac{1}{2},$$

because $z_k - y_n - z_n \in Y_{k-1}$. Thus the claim is proved.
Step 3. Since

$$\left\| \frac{y_k}{\lambda_k} \right\| = \frac{\|y_k\|}{|\lambda_k|} \leq \frac{1}{\varepsilon},$$

$\{\frac{y_k}{\lambda_k}\}_{k=1}^{\infty}$ is bounded, then by the compactness of T, $\{T(\frac{y_k}{\lambda_k})\}_{k=1}^{\infty}$ must have a convergent subsequence, i.e. a Cauchy subsequence, which is a contradiction! □

Remark. This theorem shows that the point spectrum of a compact operator is at most a sequence λ_i converging to 0.

Proposition 5.2.4. *Let X be Banach, $T \in K(X)$, then* $\text{Im}\,(\lambda I - T)$ *is closed for all* $\lambda \neq 0$.

Proof. For any $y \in \text{Im}\,(\lambda \mathbb{1} - T)$, there is $x \in X$ s.t. $y = (\lambda \mathbb{1} - T)x$. Let

$$Z_y := (\lambda \mathbb{1} - T)^{-1}(\{y\}) = \{z \in X : (\lambda \mathbb{1} - T)z = y\},$$

then $Z_y = x + E_\lambda(T)$. Define

$$\alpha(y) = \inf_{z \in Z_y} \|z\|.$$

Claim: There is $C > 0$ s.t.

$$\alpha(y) \leq C \|y\|, \quad \forall y \in \operatorname{Im}(\lambda \mathbb{1} - T).$$

Based on the claim, now we can prove this proposition. Let $\{y_n\}_{n=1}^{\infty} \subseteq \operatorname{Im}(\lambda \mathbb{1} - T)$ s.t. $y_n \to y_\infty \in X$, we want to show $y_\infty \in \operatorname{Im}(\lambda \mathbb{1} - T)$. Since $y_n \to y_\infty$, $\{y_n\}_{n=1}^{\infty}$ is bounded, by the above claim, there is $\{x_n\}_{n=1}^{\infty} \subseteq X$ s.t. $(\lambda \mathbb{1} - T)x_n = y_n$ and

$$\sup_n \|x_n\| \leq C \sup_n \|y_n\| < \infty$$

for some $C > 0$, i.e. it is bounded. Then by the compactness of T, there is a subsequence $\{x_{\sigma(n)}\}_{n=1}^{\infty} \subseteq \{x_n\}_{n=1}^{\infty}$ s.t. $\{Tx_{\sigma(n)}\}_{n=1}^{\infty}$ converges, let $u \in X$ be the limit. Then

$$y_\infty = \lim_{n\to\infty} y_n = \lim_{n\to\infty} (\lambda \mathbb{1} - T)x_n = \lim_{n\to\infty} (\lambda \mathbb{1} - T)x_{\sigma(n)} = \lambda \lim_{n\to\infty} x_{\sigma(n)} - u,$$

so

$$\lim_{n\to\infty} x_{\sigma(n)} = \frac{y_\infty + u}{\lambda} := x_\infty \in X,$$

then by the continuity of $\lambda \mathbb{1} - T$, we have

$$y_\infty = \lim_{n\to\infty} (\lambda \mathbb{1} - T)x_{\sigma(n)} = (\lambda \mathbb{1} - T) \lim_{n\to\infty} x_{\sigma(n)} = (\lambda \mathbb{1} - T)x_\infty,$$

i.e. $y_\infty \in \operatorname{Im}(\lambda \mathbb{1} - T)$. Therefore $\operatorname{Im}(\lambda \mathbb{1} - T)$ is closed. □

Corollary 5.2.5. *Let X be Banach, $T \in K(X)$, then $C_\sigma(T) \subseteq \{0\}$.*

Proof. If $\lambda \in C_\sigma(T)$ and $\lambda \neq 0$, then $\lambda \mathbb{1} - T$ is injective and

$$\operatorname{Im}(\lambda \mathbb{1} - T) = \overline{\operatorname{Im}(\lambda \mathbb{1} - T)} = X,$$

i.e. $\lambda \mathbb{1} - T$ is surjective, thus $\lambda \mathbb{1} - T$ is bijective, a contradiction! □

Definition 5.2.6. Let X, Y be Banach spaces, $A \in \mathcal{L}(X, Y)$ is called a Fredholm operator, denoted by $A \in F(X, Y)$ if both $\operatorname{Ker}(A)$ and $\operatorname{Ker}(A^*)$ are finite-rank and $\operatorname{Im}(A)$ is closed.

Definition 5.2.7. Define the index mapping $\operatorname{Ind} : F(X, Y) \to \mathbb{Z}$ by

$$\operatorname{Ind}(A) := \dim \operatorname{Ker}(A) - \dim \operatorname{Ker}(A^*), \quad \forall A \in F(X, Y).$$

Theorem 5.2.8. Ind *is continuous.*

Corollary 5.2.9. Ind *is locally constant.*

Corollary 5.2.10. *Let X be Banach, $T \in K(X)$ and $\lambda \neq 0$, then $\lambda I - T$ is injective if and only if it is surjective.*

Proof. Define $f : [0,1] \to F(X)$ by $f(s) = \lambda\mathbb{1} - sT$ for all $s \in [0,1]$. Then $\mathrm{Ind} \circ f : [0,1] \to \mathbb{Z}$ is continuous, thus $\mathrm{Ind} \circ f$ is constant. Since

$$\mathrm{Ind}(f(0)) = \dim \mathrm{Ker}\,(\lambda\mathbb{1}) - \dim \mathrm{Ker}\,((\lambda\mathbb{1})^*) = 0 - 0 = 0,$$

we have $\mathrm{Ind}(f(1)) = \mathrm{Ind}(f(0)) = 0$, i.e.

$$\dim \mathrm{Ker}\,(\lambda\mathbb{1} - T) = \dim \mathrm{Ker}\,((\lambda\mathbb{1} - T)^*) = \dim \mathrm{Coker}(\lambda\mathbb{1} - T) = \dim \mathrm{Im}\,(\lambda\mathbb{1} - T). \qquad \square$$

Corollary 5.2.11. *If $T \in K(X)$, then*

$$\sigma(T) = \{0\} \cup P_\sigma(T).$$

5.3 Spectrum of self-adjoint operators

Definition 5.3.1. Suppose H is a Hilbert space, $A \in \mathcal{L}(H)$ is self-adjoint (s.a.) if

$$A^\dagger = A,$$

i.e. $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in H$.

Remark. For any $x \in H$, $\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}$, so $\langle Ax, x\rangle \in \mathbb{R}$.

Proposition 5.3.2. *Suppose H is a Hilbert space, $A \in \mathcal{L}(H)$ is s.a., then*

$$P_\sigma(A) \subseteq \mathbb{R}.$$

Proof. Let $\lambda \in P_\sigma(A)$, $v \in E_\lambda(A) \setminus \{0\}$, then

$$\lambda \|v\|^2 = \langle \lambda v, v\rangle = \langle Av, v\rangle = \langle v, A^\dagger v\rangle = \langle v, Av\rangle = \langle v, \lambda v\rangle = \overline{\lambda} \|v\|^2,$$

so $\lambda = \overline{\lambda}$, i.e. $\lambda \in \mathbb{R}$. $\square$

Actually, we have a stronger result (Propostion 5.3.5).

Lemma 5.3.3. *Suppose H is a Hilbert space, $A \in \mathcal{L}(H)$ is s.a. Then $R_\sigma(A) = \varnothing$.*

Proof. Let $\lambda \in R_\sigma$, then $\lambda\mathbb{1} - A$ does not have a dense image, then by Theorem 4.1.5, $(\lambda\mathbb{1} - A)^\dagger = \overline{\lambda}\mathbb{1} - A^\dagger = \overline{\lambda}\mathbb{1} - A$ is injective, i.e. $\overline{\lambda} \in P_\sigma(A)$. By Proposition 5.3.2, $\lambda = \overline{\lambda} \in P_\sigma(A)$, contradiction!. $\square$

Lemma 5.3.4. *Suppose H is a Hilbert space, $A \in \mathcal{L}(H)$ is s.a. Let $\lambda \in \mathbb{C}$, if there is $c \geq 0$ s.t.*

$$\|(A - \lambda\mathbb{1})x\| \geq c\,\|x\|, \quad \forall x \in H,$$

then $\lambda \notin \sigma(A)$.

Proof. Let $\lambda \in \sigma(A)$, by Lemma 5.3.3, $\lambda \in P_\sigma(A) \sqcup C_\sigma(A)$. The inequality implies $A - \lambda\mathbb{1}$ is injective and has a closed image. Injectivity means $\lambda \notin P_\sigma$, then we must have $\lambda \in C_\sigma(A)$, i.e. $A - \lambda\mathbb{1}$ has a dense image. However, closed image means $A - \lambda\mathbb{1}$ is surjective and hence bijective, then $\lambda \notin \sigma(A)$, contradiction! □

Proposition 5.3.5. *Suppose H is a Hilbert space, $A \in \mathcal{L}(H)$ is s.a., then*

$$\sigma(A) \subseteq \mathbb{R}.$$

Proof. Suppose $\lambda = a + ib \in \sigma(A)$ where $b \neq 0$. Then for any $x \in H \setminus \{0\}$,

$$\langle (A - \lambda\mathbb{1})x, x\rangle = \langle (A - (a+ib))x, x\rangle = \langle (A-a)x, x\rangle + ib\,\|x\|^2,$$

where $\langle (A-a)x, x\rangle = \langle Ax, x\rangle - a\,\|x\|^2 \in \mathbb{R}$. Then

$$|\langle (A - \lambda\mathbb{1})x, x\rangle|^2 = |\langle (A-a)x, x\rangle|^2 + |b|^2\,\|x\|^4,$$

by Cauchy-Schwarz,

$$|b|\,\|x\|^2\,| \leq |\langle (A - \lambda\mathbb{1})x, x\rangle| \leq \|(A - \lambda\mathbb{1})x\|\,\|x\|,$$

i.e.

$$|b|\,\|x\| \leq \|(A - \lambda\mathbb{1})x\|,$$

by Lemma 5.3.4, $\lambda \notin \sigma(A)$, it is a contradiction. □

Proposition 5.3.6. *Suppose H is a Hilbert space, $A \in \mathcal{L}(H)$ is s.a., $\lambda_1, \lambda_2 \in P_\sigma(A)$ s.t. $\lambda_1 \neq \lambda_2$, then*

$$E_{\lambda_1}(A) \perp E_{\lambda_2}(A).$$

Proof. Let $v_1 \in E_{\lambda_1}(A) \setminus \{0\}$, $v_2 \in E_{\lambda_2}(A) \setminus \{0\}$. Then

$$\lambda_1\langle v_1, v_2\rangle = \langle Av_1, v_2\rangle = \langle v_1, Av_2\rangle = \langle v_1, \lambda_2 v_2\rangle = \overline{\lambda_2}\langle v_1, v_2\rangle = \lambda_2\langle v_1, v_2\rangle,$$

therefore $(\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0$, and hence $\langle v_1, v_2\rangle = 0$. □

Proposition 5.3.7. *Suppose H is a separable Hilbert space, $T \in \mathcal{L}(H)$ is compact and s.a. Then*

$$H = \bigoplus_{\lambda \in P_\sigma(T)} E_\lambda(T),$$

and H has a unitary eigenbasis of T, i.e. H has a basis $\{u_j\}_{j=1}^\infty$ s.t. $Tu_j = \lambda_j u_j$, $\|u_j\| = 1$ for all j, and $u_j \perp u_k$ for all $j \neq k$.

Proof. Step 1. By definition of $\|T\|$, i.e. $T = \sup_{\|v\|=1}\|Tv\|$, there is a sequence $\{v_j\}_{j=1}^\infty$ s.t. $\|v_j\| = 1$ and

$$\lim_{j\to\infty}\|Tv_j\| = \|T\|. \tag{$*$}$$

Since $\{v_j\}_{j=1}^{\infty}$ is bounded, by Alaoglu's theorem (3.2.3), there is a subsequence $\{v_\sigma(j)\}_{j=1}^{\infty} \subseteq \{v_j\}_{j=1}^{\infty}$ s.t. $v_\sigma(j) \xrightarrow{\text{w}} v_\infty$ for some $v_\infty \in H$.
Step 2. Claim: $\|v_\infty\| = 1$.
First,

$$\begin{aligned} 1 &= \lim_{j\to\infty} \left\|v_{\sigma(j)}\right\|^2 \\ &= \lim_{j\to\infty} \left\|v_{\sigma(j)} - v_\infty + v_\infty\right\|^2 \\ &= \lim_{j\to\infty} \left(\|v_\infty\|^2 + \left\|v_{\sigma(j)} - v_\infty\right\|^2 + 2\mathrm{Re}\left(\langle v_\infty, v_{\sigma(j)} - v_\infty\rangle\right)\right) \\ &= \|v_\infty\|^2 + \lim_{j\to\infty} \left\|v_{\sigma(j)} - v_\infty\right\|^2, \end{aligned}$$

so $\|v_\infty\| \leq 1$.
Second, by $v_\sigma(j) \xrightarrow{\text{w}} v_\infty$, we have $Tv_\sigma(j) \to Tv_\infty$, then $\|Tv_\sigma(j)\| \to \|Tv_\infty\|$. Also by $(*)$, we have $\|Tv_\sigma(j)\| \to \|T\|$. Thus $\|Tv_\infty\| = \|T\|$. Then

$$\|T\| = \|Tv_\infty\| \leq \|T\| \, \|v_\infty\| \, ,$$

so $\|v_\infty\| \geq 1$. Therefore we have proved the claim. By the above argument, we also find that T achieves its maximum on v_∞
Step 3.

Corollary 5.3.8. *Let H_1, H_2 be Hilbert spaces and H_1 be separable. Let $Fin(H_1, H_2)$ denote the (non-closed) subspace of finite-rank operators in $\mathcal{L}(H_1, H_2)$, then*

$$\overline{Fin(H_1, H_2)} = K(H_1, H_2).$$

5.4 Spectrum of normal operators

Definition 5.4.1. Let H be a Hilbert space. $A \in \mathcal{L}(H)$ is called

(1) normal if $[A, A^\dagger] := AA^\dagger - A^\dagger A = 0$.

(2) unitary if $AA^\dagger = \mathbb{1} = A^\dagger A$.

Remark. 1. If A is unitary, then A preserves the norm because

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle x, A^\dagger Ax\rangle = \|x\|^2 \, .$$

2. A unitary operator is always normal.

Theorem 5.4.2. *Let H be a Hilbert space and $A \in \mathcal{L}(H)$ be normal. Then*

(1) $\|A^n\| = \|A\|^n$

(2) $r_A = \|A\| = \sup_{\lambda\in\sigma(A)} |\lambda|$

(3) *$R_\sigma(A) = R_\sigma(A^\dagger) = \varnothing$ and $P_\sigma(A^\dagger) = \{\overline{\lambda} \in \mathbb{C} : \lambda \in P_\sigma(A)\}$.*

(4) *If A is unitary, then $\sigma(A) \subseteq S^1$.*

Proof. (1) First,

$$\|Ax\|^2 = \langle Ax, Ax\rangle = \langle x, A^\dagger Ax\rangle = \langle x, AA^\dagger x\rangle = \langle AA^\dagger x, x\rangle = \langle A^\dagger x, A^\dagger x\rangle = \left\|A^\dagger x\right\|^2,$$

and

$$\left\|A^\dagger Ax\right\|^2 = \langle A^\dagger Ax, A^\dagger Ax\rangle = \langle AA^\dagger Ax, Ax\rangle = \langle A^\dagger A^2x, Ax\rangle = \langle Ax, A^\dagger A^2x\rangle = \langle A^2x, A^2x\rangle = \left\|A^2x\right\|^2,$$

moreover,

$$\|Ax\|^2 = \langle x, A^\dagger Ax\rangle \le \left\|A^\dagger Ax\right\| \|x\| = \left\|A^2x\right\| \|x\|,$$

so

$$\|A\|^2 = \sup_{\|x\|=1} \|Ax\|^2 \le \sup_{\|x\|=1} \left\|A^2x\right\| \|x\| = \left\|A^2\right\|.$$

It's clear that $\left\|A^2\right\| \le \|A\|^2$, therefore $\left\|A^2\right\| = \|A\|^2$. By induction, we have for any $k \ge 0$,

$$\left\|A^{2^k}\right\| = \|A\|^{2^k}.$$

Next, for any $n \in \mathbb{Z}_+$, we can choose $m \in \mathbb{Z}_+$ s.t. $2^m > n$. Then

$$\|A\|^{2^m-n} \|A\|^n = \|A\|^{2^m} = \left\|A^{2^m}\right\| = \left\|A^{2^m-n}A^n\right\| \le \left\|A^{2^m-n}\right\| \|A^n\|,$$

so we have

$$\|A\|^n \le \|A^n\|.$$

Since we always have $\|A^n\| \le \|A\|^n$, thus $\|A\|^n = \|A^n\|$.
(2)By Theorem 5.1.4,

$$\sup_{\lambda\in\sigma(A)} |\lambda| = r_A = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{n\to\infty} (\|A\|^n)^{1/n} = \|A\|.$$

(3)Let $\lambda \in \mathbb{C}$, then

5.5 Integration

Lemma 5.5.1. *Let X be Banach, and $c : [a, b] \to X$ be continuous. Then there exists a unique $v \in X$ s.t. for all $\varphi \in X^*$,*

$$\varphi(v) = \int_a^b \varphi(c(t))\,\mathrm{d}t.$$

The v is denoted

$$v := \int_a^b c(t)\,\mathrm{d}t.$$

Definition 5.5.2. Let X be Banach, and $c : [a, b] \to X$ be continuous. c is called differentiable at $t \in [a, b]$ if

$$\lim_{h \to 0} \frac{c(t+h) - c(t)}{h}$$

exists. We denote the limit $c'(t)$. c is called differentiable on $[a, b]$ if it is differentiable for any $t \in [a, b]$.

Proposition 5.5.3 (properties of integration)**.**

Definition 5.5.4. Let $\Omega \subseteq \mathbb{C}$ be open and X be a $\mathbb{C}$-Banach space. Let $f : \Omega \to X$ be continuous. f is called holomorphic if

$$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}$$

exists for all $z \in \Omega$.

Definition 5.5.5. Let $\gamma : [a, b] \to \Omega$ be C^1, $f : \Omega \to X$ be continuous, define the integral of f along the curve γ by

$$\int_\gamma f \, \mathrm{d}z = \int_a^b f(\gamma(t))\gamma'(t) \, \mathrm{d}t.$$

Theorem 5.5.6 (Cauchy integral theorem)**.** *Let f be holomorphic in $\Omega \subseteq \mathbb{C}$ and $\overline{B_R(z_0)} \subseteq \Omega$. Then*

$$\frac{1}{2\pi} \int_{|z-z_0|=R} \frac{f(z)}{z-w} \, \mathrm{d}z = \begin{cases} f(w) & \text{if } w \in B_R(z_0) \\ 0 & \text{otherwise} \end{cases}$$

Lemma 5.5.7. *Let $A : \Omega \subseteq \mathbb{C} \to \mathcal{L}(X, Y)$ be weakly continuous, meaning $\varphi \circ A$ is continuous for any $\varphi \in \mathcal{L}(X, Y)^*$. Then TFAE:*

(1) *A is holomorphic*

(2) *$x \mapsto \phi(A(z)x)$ is holomorphic for any $z \in \Omega$, $x \in X$ and $\phi \in Y^*$.*

(3) *for any $\overline{B_r(z_0)} \subseteq \Omega$, $\gamma(t) = z_0 + re^{2\pi i t} : [0, 1] \to \Omega$ and any $w \in B_r(z_0)$, $\phi \in Y^*$, $x \in X$,*

$$\phi(A(w)x) = \frac{1}{2\pi i} \int_\gamma \frac{\phi(A(z)x)}{z-w} \, \mathrm{d}z.$$

Definition 5.5.8. Let X be a complex Banach space, $A \in \mathcal{L}(X)$. Let $U \subseteq \mathbb{C}$ be open s.t. $\sigma(A) \subseteq U$. Let $\gamma = \{\gamma_1, \cdots, \gamma_m\}$ be a collection of simple piecewise smooth closed curves where $\gamma_j : S^1 \to U \setminus \sigma(A)$. For any holomorphic function $f : U \to \mathbb{C}$, define

$$f(A) = \frac{1}{2\pi i} \int_\gamma f(z)(z\mathbb{1} - A)^{-1} \, \mathrm{d}z.$$

Example 5.5.9.

5.6 Functional calculus

Definition 5.6.1. 1. A unital C^* algebra $(\mathcal{A}, *)$ or $\mathcal{A}$ is a complex unital Banach algebra equipped with an antilinear involution $*$ denoted $\mathcal{A} \to \mathcal{A} : a \mapsto a^*$ satisfying for any $a, b \in \mathcal{A}$ and $\lambda \in \mathbb{C}$

   (i) $(ab)^* = b^* a^*$

   (ii) $\mathbb{1}^* = \mathbb{1}$

   (iii) $a^{**} = a$

   (iv) $(\lambda a)^* = \overline{\lambda} a^*$

   (v) $\|a^* a\| = \|a\|^2$

2. $\mathcal{A}$ is called commutative if $ab = ba$ for all $a, b \in \mathcal{A}$.

3. A C^*-homomorphism between 2 unital C^* algebra $\mathcal{A}$ and $\mathcal{B}$ is a map $\varphi \in \mathcal{L}(\mathcal{A}, \mathcal{B})$ s.t.

$$\varphi(\mathbb{1}_{\mathcal{A}}) = \mathbb{1}_{\mathcal{B}}, \quad \varphi(ab) = \varphi(a)\varphi(b), \quad \varphi(a^*) = \varphi(a)^*, \quad \forall a \in \mathcal{A}, b \in \mathcal{B}.$$

Example 5.6.2. Let H be Hilbert, $\mathcal{A} = \mathcal{L}(H)$. For any $a \in \mathcal{A}$, define $a^* := a^\dagger$, then $\mathcal{A}$ is a unital C^* algebra.

Example 5.6.3. Let (M, d) be a metric space, $C_b(M)$ be the set of bounded continuous functions $f : M \to \mathbb{C}$. $C_b(M)$ is a Banach algebra. For any $f \in C_b(M)$, define $f^* = \overline{f}$, then $(C_b(M), *)$ is a commutative C^* algebra.

Lemma 5.6.4. *Let H be a Hilbert space, $A \in \mathcal{L}(H)$. For any polynomial $p : \mathbb{C} \to \mathbb{C}$ defined by*

$$p(z) = \sum_{k=0}^{n} a_k z^k, \quad \forall z \in \mathbb{C}$$

where $a_k \in \mathbb{C}$ and $a_n \neq 0$, we define $p : \mathcal{L}(H) \to \mathcal{L}(H)$ by

$$p(A) = \sum_{k=0}^{n} a_k A^k, \quad \forall A \in \mathcal{L}(H).$$

Then for any polynomials p and q,

(1) $(p + q)(A) = p(A) + q(A)$

(2) $(pq)(A) = p(A)q(A)$

(3) $\sigma(p(A)) = p(\sigma(A))$.

(4) *If A is normal, then $p(A)$ is normal.*

Proof. (1) (2) are obvious. (3) Let $\mu \in \sigma(A)$, then $\mu\mathbb{1} - A$ is not bijective. Since μ is a zero for the polynomial $p(\mu) - p(z)$, thus

$$p(\mu) - p(z) = (\mu - z)q(z),$$

where q is a polynomial of degree $n-1$. Then $p(\mu)\mathbb{1} - p(A) = (\mu\mathbb{1} - A)q(A)$ is not bijective because $\mu\mathbb{1} - A$ is not bijective, and hence $p(u) \in \sigma(p(A))$.
Conversely, if $\tau \in \sigma(p(A))$, by factorization,

$$\tau - p(z) = \alpha \prod_{j=1}^{n} (\mu_j - z),$$

then

$$\tau\mathbb{1} - p(A) = \alpha \prod_{j=1}^{n} (\mu_j\mathbb{1} - A).$$

Since $\tau\mathbb{1} - p(A)$ is not bijective, $\mu_{j_0}\mathbb{1} - A$ is not bijective for some $1 \leq j_0 \leq n$, i.e. $\mu_{j_0} \in \sigma(A)$. Since μ_{j_0} is a zero for $\tau - p(z)$, we have $\tau = p(\mu_{j_0})$, i.e. $\tau \in p(\sigma(A))$.
(4) If A is normal, $AA\dagger = A^\dagger A$. Define $q(A^\dagger) = p(A)^\dagger$, then $AA\dagger = A^\dagger A$ implies

$$p(A)q(A^\dagger) = q(A^\dagger)p(A),$$

i.e. $p(A)p(A)^\dagger = p(A)^\dagger p(A)$ and hence $p(A)$ is normal. □

Theorem 5.6.5. *Let H be a Hilbert space, $A \in \mathcal{L}(H)$ is s.a. Let $\Sigma = \sigma(A) \subseteq \mathbb{R}$ and $C(\Sigma)$ be the set of continuous function $f : \Sigma \to \mathbb{C}$. Then there is a unital C^* algebra homomorphism $\Phi_A \in \mathcal{L}(C(\Sigma), \mathcal{L}(H))$ denoted by*

$$\Phi_A(f) = f(A), \quad \forall f \in C(\Sigma)$$

s.t.

(1) *$1(A) = \mathbb{1}_H$ and $(fg)(A) = f(A)g(A)$ for all $f, g \in C(\Sigma)$*

(2) *$\overline{f}(A) = f(A)^\dagger$*

(3) *$\mathrm{id}_\Sigma(A) = A$*

(4) *If $B \in \mathcal{L}(H)$ s.t. $[A, B] = 0$, then $[f(A), B] = 0$ for all $f \in C(\Sigma)$.*

(5) *$\Phi_A(C(\Sigma))$ is the smallest C^* subalgebra of $\mathcal{L}(H)$ containing A*

(6) *If $\lambda \in \Sigma$ and $x \in E_\lambda(A)$, then $f(A)x = f(\lambda)x$ for all $f \in C(\Sigma)$*

(7) *$f(A)$ is normal and $\sigma(f(A)) = f(\sigma(A))$ for all $f \in C(\Sigma)$*

(8) *If $f \in C(\Sigma, \mathbb{R})$ and $g \in C(f(\Sigma))$, then $(g \circ f)(A) = g(f(A))$*

Proof. Step 1:

Definition 5.6.6. Let H be a Hilbert space, $A \in \mathcal{L}(H)$ is called positive semidefinite (p.s.d.) if

$$\langle Ax, x\rangle \geq 0, \quad \forall x \in H.$$

We write $A \geq 0$ if A is p.s.d.

Corollary 5.6.7. *Let H be a Hilbert space. Let $A \in \mathcal{L}(H)$ be s.a. and p.s.d., then there is a unique $B \in \mathcal{L}(H)$ s.t. B is s.a., p.s.d., and $B^2 = A$.*

Proof. Claim: $\sigma(A) \subseteq [0, \infty)$.
Let $\lambda \in \sigma(A)$. Suppose $\lambda < 0$, then

$$\|(A - \lambda \mathbb{1})x\| \, \|x\| \geq \langle (A - \lambda\mathbb{1})x, x\rangle = \langle Ax, x\rangle - \lambda \|x\|^2 \geq |\lambda| \, \|x\|^2,$$

so for $x \in H \setminus \{0\}$,

$$\|(A - \lambda\mathbb{1})x\| \geq |\lambda| \, \|x\|,$$

by Lemma 5.3.4, $\lambda \notin \sigma(A)$. Thus $\lambda \geq 0$, i.e. $\sigma(A) \subseteq [0, \infty)$.
Define $f : \sigma(A) \to [0, \infty)$ by $f(\lambda) = \sqrt{\lambda}$ for all $\lambda \in \sigma(A)$. Then f is continuous since $\sqrt{\cdot} : [0, \infty) \to [0, \infty)$ is continuous and $\sigma(A)$ is compact.

5.7 Measurable functional calculus

Definition 5.7.1 (Projection valued measure)**.** Let H be Hilbert, $\Sigma \subseteq \mathbb{C}$ be a non-empty closed subset. Let $\mathcal{B}_\Sigma$ denote the collection of Borel subsets of Σ. Then a projection valued Borel measure on Σ is a map $\mathcal{B}_\Sigma \to \mathcal{L}(H) : \Omega \mapsto P_\Omega$ s.t.

(1) For any $\Omega \in \mathcal{B}_\Sigma$, P_Ω is a s.a. unitary projection.

(2) $P_\varnothing = 0$, $P_\Sigma = \mathbb{1}_H$.

(3) For any $\Omega_1, \Omega_2 \in \mathcal{B}_\Sigma$,

$$P_{\Omega_1 \cap \Omega_2} = P_{\Omega_1} P_{\Omega_2} = P_{\Omega_2} P_{\Omega_1}.$$

(4) If $\{\Omega_j\}_{j=1}^\infty \subseteq \mathcal{B}_\Sigma$ are pairwise disjoint and $\Omega = \sqcup_{j=1}^\infty \Omega_j$, then for any $x \in H$,

$$P_\Omega x = \lim_{n\to\infty} \sum_{j=1}^n P_{\Omega_j} x.$$

Definition 5.7.2. Let $\mathcal{B}(\Sigma) := \{f : \Sigma \to \mathbb{C} : \ f \text{ is bounded and Borel measurable}\}$. It's clear $\mathcal{B}(\Sigma)$ is a C^* algebra.

Theorem 5.7.3. *Let H be a complex Hilbert space. Let $\Sigma \subseteq \mathbb{C}$ be closed. Let $\Omega \mapsto P_\Omega$ be a projection valued measure on Σ. For any $x, y \in H$, define the signed Borel measure*

$$\mu_{x,y}(\Omega) = \operatorname{Re}\langle x, P_\Omega y\rangle, \quad \forall \Omega \in \mathcal{B}_\Sigma.$$

Then there is a unique $\Psi \in \mathcal{L}(\mathcal{B}(\Sigma), \mathcal{L}(H))$ s.t. for any $x, y \in H$, $f \in \mathcal{B}(\Sigma)$,

$$\operatorname{Re}\langle x, \Psi(f)y\rangle = \int_\Sigma \operatorname{Re}(f)\,\mathrm{d}\mu_{x,y} + \int_\Sigma \operatorname{Im}(f)\,\mathrm{d}\mu_{x,iy}.$$

Proof. Step 1.

Theorem 5.7.4. *Let H be a complex Hilbert space, $A \in \mathcal{L}(H)$ and $\Sigma = \sigma(A)$. Then there is a unique projection valued measure $\Omega \mapsto P_\Omega$ s.t. for any $x, y \in H$,*

$$\operatorname{Re}\langle x, Ay\rangle = \int_\Sigma \operatorname{Re}(\lambda)\,\mathrm{d}\mu_{x,y}(\lambda) + \int_\Sigma \operatorname{Im}(\lambda)\,\mathrm{d}\mu_{x,iy}(\lambda).$$

5.8 Cyclic vectors

Theorem 5.8.1. *Suppose H is a complex Hilbert space and $A \in \mathcal{L}(H)$ is s.a. Then there is a collection of compact sets*

$$\Sigma_i \subseteq \sigma(A),$$

$i \in I$, Borel measures $\mu_i \in \mathcal{M}(\Sigma_i)$ and an isometric isomorphism

$$U : H \to \bigoplus_{i\in I} L^2(\Sigma_i, \mu_i)$$

s.t. for any $i \in I$, $\psi_i \in L^2(\Sigma_i, \mu_i)$, and $\lambda \in \Sigma_i$,

$$(UAU^{-1}\psi_i)(\lambda) = \lambda\psi_i(\lambda),$$

i.e. UAU^{-1} diagonalizes A. If H is separable, then I is countable.

Definition 5.8.2. Let $A \in \mathcal{L}(H)$ be s.a. $x \in H$ is cyclic for A if

$$H = \overline{\operatorname{Span}(\{A^n x : n \geq 0\})}.$$

Theorem 5.8.3. *Let $A \in \mathcal{L}(H)$ be s.a. and $x \in H$ be cyclic for A. Let $\Sigma = \sigma(A)$. Let μ_x be the measure s.t.*

$$\int_\Sigma f\,\mathrm{d}\mu_x = \langle x, f(A)x\rangle, \quad \forall f \in C(\Sigma).$$

Then

(1) *There is a unique $U \in \mathcal{L}(H, L^2(\Sigma, \mu_x))$ s.t. U is isometric and*

$$U^{-1}\psi = \psi(A)x, \quad \forall \psi \in C(\Sigma).$$

(2) *Let* $f \in \mathcal{B}(\Sigma)$*, then*

$$Uf(A)U^{-1}\psi = f\psi, \quad \forall \psi \in L^2(\Sigma, \mu_x).$$

(3) $(UAU^{-1}\psi)(\lambda) = \lambda\psi(\lambda)$ *for all* $\psi \in L^2(\Sigma, \mu_x)$ *and* $\lambda \in \Sigma$*.*

(4) *If* $\Omega \subseteq \Sigma$ *is relatively open and nonempty, then*

$$\mu_x(\Omega) > 0.$$

Proof. Let $T \in \mathcal{L}(C(\Sigma), H)$ given by

$$T\psi = \psi(A)x, \quad \forall \psi \in C(\Sigma)$$

Claim 1: T is isometric.
For any $\psi \in C(\Sigma)$, we have

$$\begin{aligned}
\|T\psi\|^2 &= \langle \psi(A)x, \psi(A)x \rangle \\
&= \langle x, \psi^\dagger(A)\psi(A)x \rangle \\
&= \langle x, \overline{\psi}(A)\psi(A)x \rangle \\
&= \langle x, |\psi|^2(A)x \rangle \\
&= \int_\Sigma |\psi|^2 \, \mathrm{d}\mu_x \\
&= \|\psi\|_{L^2}^2 ,
\end{aligned}$$

Thus $T \in \mathcal{L}(C(\Sigma, H), H)$ is isometric.

Since any continuous function $\psi \in C(\Sigma)$ can be approximated by simple functions in $L^2(\Sigma, \mu_x)$, we have $C(\Sigma)$ is dense in $L^2(\Sigma, \mu_x)$, then we can extend T uniquely to $\tilde{T} \in \mathcal{L}(L^2(\Sigma, \mu_x), H)$ s.t. $\tilde{T}$ is isometric and

$$\tilde{T}\psi = \psi(A)x, \quad \forall \psi \in L^2(\Sigma, \mu_x).$$

For simplicity, we denote $\tilde{T}$ as T.

Claim 2. $T \in \mathcal{L}(L^2(\Sigma, \mu_x), H)$ is isometric isomorphism.
We only need to show T is surjective. Let $\psi_n(\lambda) = \lambda^n$, then

$$T\psi_n = \psi_n(A)x = A^n x,$$

Therefore

$$H = \overline{\mathrm{Span}(\{A^n x : n \geq 0\})} \subseteq \mathrm{Im}\,(T) \subseteq H,$$

i.e. $\mathrm{Im}\,(T) = H$ and T is surjective.

Then T is bijective and hence invertible, let $U = T^{-1}$.
(1)Obviously true
(2)For any $f \in \mathcal{B}(\Sigma), \psi \in L^2(\Sigma, \mu_x)$, we have $f\psi \in L^2(\Sigma, \mu_x)$, then

$$f(A)U^{-1}\psi = f(A)\psi(A)x = (f\psi)(A)x = U^{-1}(f\psi),$$

so $Uf(A)U^{-1}\psi = f\psi$.
(3) follows from (2).
(4) □

Corollary 5.8.4. *Let H be a complex Hilbert space. Let $x \in H \setminus \{0\}$ and $A \in \mathcal{L}(H)$ be s.a. Define*

$$H_x = \overline{\mathrm{Span}(\{A^n x : n \geq 0\})},$$

then H_x is the smallest closed A-invariant subspace of H containing x. Define $A_x = A|_{H_x}$, let $\Sigma_x = \sigma(A_x)$. Then by Theorem 5.8.3, there is a unique isometric isomorphism $U_x \in \mathcal{L}(H_x, L^2(\Sigma_x, \mu_x))$ s.t.

$$U_x^{-1}\psi = \psi(A_x)x, \quad \forall \psi \in L^2(\Sigma_x, \mu_x).$$

Proof of Theorem 5.8.1. □

Chapter 6

Unbounded operators

6.1 Definition

Definition 6.1.1. Let X, Y be Banach.

(1) An unbounded linear operator from X to Y is a pair $(A, \mathrm{dom}(A))$ where $\mathrm{dom}(A) \subseteq X$ and $A : \mathrm{dom}(A) \to Y$ is linear.

(2) A is densely defined if $\mathrm{dom}(A)$ is dense in X.

(3) A is closed if $\Gamma_A := \{(x, Ax) : x \in \mathrm{dom}(A)\}$ is closed in $X \times Y$.

(4) Let $A : \mathrm{dom}(A) \subseteq X \to Y$ and $B : \mathrm{dom}(B) \subseteq Y \to Z$ are unbounded linear operators, then define $BA : \mathrm{dom}(BA) \subseteq X \to Z$ to be the operator with

$$\mathrm{dom}(BA) = \{x \in \mathrm{dom}(A) : Ax \in \mathrm{dom}(B)\},$$

and $BA(x) = B(Ax)$ for all $x \in \mathrm{dom}(BA)$.

Remark. 1. Recall that $\mathrm{dom}(A)$ is a normed vector space w.r.t. Γ_A norm and A is bounded w.r.t. this norm.
2. If A is closed, then $\mathrm{dom}(A)$ is Banach w.r.t. Γ_A norm.

Example 6.1.2. Let $C([0,1])$ be the set of all continuous functions on $[0,1]$, and $C^1([0,1]) \subseteq C([0,1])$ be the set of all continuously differentiable functions on $[0,1]$. $(C([0,1]), \|\cdot\|_\infty)$ is Banach. $C^1([0,1])$ is dense in $C([0,1])$ by Stone–Weierstrass theorem. Define

$$D : \mathrm{dom}(D) = C^1([0,1]) \to C([0,1])$$

by

$$Df = \frac{\mathrm{d}}{\mathrm{d}x} f, \quad \forall f \in C^1([0,1]).$$

Then D is a densely defined, closed unbounded linear operator.

Definition 6.1.3. Let $(A, \mathrm{dom}(A))$ be a closed unbounded linear operator from X to X.

(1) The spectrum of A is defined as

$$\sigma(A) := \{\lambda \in \mathbb{C} : \lambda \mathbb{1} - A : \mathrm{dom}(A) \to X \text{ is not bijective}\}.$$

(2) And

$$\sigma(A) = P_\sigma(A) \sqcup C_\sigma(A) \sqcup R_\sigma(A)$$

where $P_\sigma(A)$, $C_\sigma(A)$ and $R_\sigma(A)$ are defined analogously to those for bounded linear operators.

(3) Define $\rho(A) = \sigma(A)^c$.

(4) For $\lambda \in \rho(A)$, $R_\lambda(A) := (\lambda \mathbb{1} - A)^{-1} : X \to \mathrm{dom}(A) \subseteq X$ is the resolvent operator of A at λ.

Lemma 6.1.4. *$\sigma(A)$ is closed.*

Proof. We want to show $\rho(A)$ is open. Let $\lambda \in \rho(A)$, then $\lambda \mathbb{1} - A : \mathrm{dom}(A) \to X$ is bijective. Let $\mu \in \mathbb{C}$,

$$\mu \mathbb{1} - A = \lambda \mathbb{1} - A + (\mu - \lambda)\mathbb{1} = (\lambda \mathbb{1} - A)(\mathbb{1} + (\mu - \lambda)(\lambda \mathbb{1} - A)^{-1})$$

is bijective if and only if $\mathbb{1} + (\mu - \lambda)(\lambda \mathbb{1} - A)^{-1}$ is bijective. The latter is bijective if

$$r_{(\lambda-\mu)(\lambda \mathbb{1} - A)^{-1}} < 1,$$

if

$$|\lambda - \mu| \cdot \left\|(\lambda \mathbb{1} - A)^{-1}\right\| < 1,$$

if

$$|\lambda - \mu| < \frac{1}{\|(\lambda \mathbb{1} - A)^{-1}\|}.$$

Therefore, let $\delta = 1/\left\|(\lambda \mathbb{1} - A)^{-1}\right\|$, for any $\mu \in B_\delta(\lambda)$, $\mu \mathbb{1} - A$ is bijective, i.e. $\mu \in \rho(A)$, thus $\rho(A)$ is open. □

Lemma 6.1.5. *$R_\lambda(A) \in \mathcal{L}(X)$.*

Proof. It's clear $R_\lambda(A)$ is linear. Since A is closed, $\lambda \mathbb{1} - A$ is closed, i.e. the graph of $\lambda \mathbb{1} - A$ is closed in $X \times X$. Then since $\lambda \mathbb{1} - A$ is bijective the graph of $R_\lambda(A) = (\lambda \mathbb{1} - A)^{-1} : X \to \mathrm{dom}(A)$ is also closed in $X \times X$. By closed graph theorem, $R_\lambda(A)$ is bounded. □

Lemma 6.1.6. *Let $A : \mathrm{dom}(A) \subsetneq X \to X$ be a densely defined unbounded operator. Let $\mu \in \rho(A)$, then*

$$P_\sigma(R_\mu(A)) = \{\frac{1}{\mu - \lambda} : \lambda \in P_\sigma(A)\},$$

$$C_\sigma(R_\mu(A)) = \{\frac{1}{\mu - \lambda} : \lambda \in C_\sigma(A)\} \cup \{0\},$$

$$R_\sigma(R_\mu(A)) = \{\frac{1}{\mu-\lambda} : \lambda \in R_\sigma(A)\}.$$

(Here $R_\sigma(A)$ means the residual spectrum of A, and $R_\lambda(A)$ means the resolvent operator of A.)

Proof. Since $R_\mu(A) : X \to \mathrm{dom}(A) \subsetneq X$ is bijective, $R_\mu(A) : X \to X$ is not bijective, hence $0 \in \sigma(A)$. Since $-R_\mu(A) = 0\mathbb{1} - R_\mu(A)$ is injective, we have $0 \notin P_\sigma(R_\mu(A))$. Moreover, $\mathrm{Im}\,(0\mathbb{1} - R_\mu(A)) = \mathrm{Im}\,(R_\mu(A)) = \mathrm{dom}(A)$ is dense in X, so $0 \notin R_\sigma(R_\mu(A))$. Therefore $0 \in C_\sigma(R_\mu(A))$.
For $\lambda \neq \mu$ (because $\mu \in \rho(A)$),

$$\begin{aligned}\frac{1}{\mu-\lambda}\mathbb{1} - R_\mu(A) &= \frac{1}{\mu-\lambda}(\mathbb{1} - (\mu-\lambda)(\mu\mathbb{1} - A)^{-1}) \\ &= \left(\frac{1}{\mu-\lambda}(\mu\mathbb{1} - A) - \mathbb{1}\right) R_\mu(A) \\ &= \frac{1}{\mu-\lambda}(\mu\mathbb{1} - A - (\mu-\lambda)\mathbb{1})R_\mu(A) \\ &= \frac{1}{\mu-\lambda}(\lambda\mathbb{1} - A)R_\mu(A).\end{aligned}$$

The left hand side is injective if and only if $\lambda\mathbb{1} - A$ is injective, is surjective if and only if $\lambda\mathbb{1} - A$ is surjective, and has a dense image if and only if $\lambda\mathbb{1} - A$ has a dense image. ☐

Definition 6.1.7. Suppose $A : \mathrm{dom}(A) \subseteq X \to X$ is a closed, densely defined unbounded operator. We say A has compact resolvent if $\rho(A) \neq \varnothing$ and $R_\mu(A)$ is compact for all $\mu \in \rho(A)$.

Corollary 6.1.8. *Suppose A has compact resolvent. Then $\sigma(A) = P_\sigma(A)$ is a discrete subset of $\mathbb{C}$ and $E_\lambda(A)$ has a finite dimension for all $\lambda \in \sigma(A)$.*

6.2 Adjoints of unbounded operators

Definition 6.2.1. Let X, Y be Hilbert spaces. Let $A : \mathrm{dom}(A) \subseteq X \to Y$ be a densely defined unbounded linear operator. The adjoint of A, denoted $A^\dagger : \mathrm{dom}(A^\dagger) \subseteq Y \to X$ is defined as follows:

(i) $\mathrm{dom}(A^\dagger) = \{y \in Y :$ there is $C_y \geq 0$ s.t. $|\langle Ax, y\rangle| \leq C_y \|x\|$ for all $x \in \mathrm{dom}(A)\}$

(ii) For any $y \in \mathrm{dom}(A^\dagger) \subseteq Y$, $\phi : \mathrm{dom}(A) \subseteq X \to \mathbb{C}$ defined by $x \mapsto \langle Ax, y\rangle$ is a bounded linear functional, and can be uniquely extended to all elements in X because $\mathrm{dom}(A)$ is dense in X. By Riesz's representation theorem, there is a unique $v \in X$ s.t. $\phi(x) = \langle x, v\rangle$ for all $x \in X$, and define $A^\dagger y = v$, i.e.

$$\langle Ax, y\rangle = \langle x, A^\dagger y\rangle, \quad \forall x \in \mathrm{dom}(A).$$

A is called self-adjoint (s.a.) if $X = Y$, $\mathrm{dom}(A) = \mathrm{dom}(A^\dagger)$ and

$$A^\dagger x = Ax, \quad \forall x \in \mathrm{dom}(A).$$

Proposition 6.2.2. *$A^\dagger$ is closed.*

Proof. We want to show $\Gamma_{A^\dagger}$ is closed in $Y \times X$. Let $\{y_n\}_{n=1}^\infty \subseteq \mathrm{dom}(A^\dagger)$ s.t. $y_n \to y_\infty$ and $A^\dagger y_n \to x_\infty$. Our goal is to show $y_\infty \in \mathrm{dom}(A^\dagger)$ and $x_\infty = A^\dagger y_\infty$.
For any $x \in \mathrm{dom}(A)$,

$$\langle Ax, y_\infty \rangle = \lim_{n\to\infty} \langle Ax, y_n \rangle = \lim_{n\to\infty} \langle x, A^\dagger y_n \rangle = \langle x, x_\infty \rangle,$$

thus

$$|\langle Ax, y_\infty \rangle| \leq \|x_\infty\| \, \|x\| \,,$$

i.e. $y_\infty \in \mathrm{dom}(A^\dagger)$. And by definition,

$$\langle Ax, y_\infty \rangle = \langle x, A^\dagger y_\infty \rangle, \quad \forall x \in \mathrm{dom}(A)$$

then

$$\langle x, A^\dagger y_\infty - x_\infty \rangle = 0, \quad \forall x \in \mathrm{dom}(A).$$

Since $\mathrm{dom}(A)$ is dense in X, there is $\{x_n\}_{n=1}^\infty \subseteq \mathrm{dom}(A)$ s.t. $x_n \to A^\dagger y_\infty - x_\infty$, so

$$\langle A^\dagger y_\infty - x_\infty, A^\dagger y_\infty - x_\infty \rangle = \langle \lim_{n\to\infty} x_n, A^\dagger y_\infty - x_\infty \rangle = \lim_{n\to\infty} \langle x_n, A^\dagger y_\infty - x_\infty \rangle = 0,$$

which implies $x_\infty = A^\dagger y_\infty$. □

Proposition 6.2.3 (Properties of the adjoint operator)**.** *Let X, Y be Hilbert spaces. Let $A : \mathrm{dom}(A) \subseteq X \to Y$ be a densely defined unbounded linear operator. Then*

(1) *If $P \in \mathcal{L}(X, Y)$, $\lambda \in \mathbb{C}$, then*

$$(A + P)^\dagger = A^\dagger + P^\dagger, \quad (\lambda A)^\dagger = \overline{\lambda} A^\dagger.$$

(2) *A is closeable if and only if $\mathrm{dom}(A^\dagger)$ is dense in Y.*

(3) *A is closed if and only if $A^{\dagger\dagger} = A$*

Lemma 6.2.4. *Let $A : \mathrm{dom}(A) \subseteq X \to Y$ be a densely defined unbounded linear operator. Let $J : Y \times X \to X \times Y$ be $J(y, x) = (-x, y)$, then*

$$\Gamma_A^\perp = J(\Gamma_{A^\dagger}).$$

Proof. If $(x, Ax) \in \Gamma_A$ and $(y, A^\dagger y) \in \Gamma_{A^\dagger}$, then

$$\langle (x, Ax), J(y, A^\dagger y) \rangle_{X\times Y} = \langle (x, Ax), (-A^\dagger y, y) \rangle_{X\times Y} = \langle x, -A^\dagger y \rangle_X + \langle Ax, y \rangle_Y = 0,$$

so $J(\Gamma_{A^\dagger}) \subseteq \Gamma_A^\perp$.
Conversely, suppose $(u,v) \in \Gamma_A^\perp$, we want to show $(z,w) \in J(\Gamma_{A^\dagger})$, i.e. $(w,-z) \in \Gamma_{A^\dagger}$. For any $x \in \mathrm{dom}(A)$, by definition of $\Gamma_A^\perp$,

$$\langle (x, Ax), (z,w)\rangle_{X\times Y} = 0,$$

then

$$\langle x, -z\rangle_X = \langle Ax, w\rangle_Y,$$

thus

$$|\langle Ax, w\rangle_Y| \leq \|z\|_X \cdot \|x\|_X\,,$$

which shows $w \in \mathrm{dom}(A^\dagger)$. Moreover,

$$\langle x, -z\rangle_X = \langle Ax, w\rangle_Y = \langle x, A^\dagger w\rangle_X, \quad \forall x \in \mathrm{dom}(A),$$

thus

$$\langle x, A^\dagger w + z\rangle_X = 0, \forall x \in \mathrm{dom}(A),$$

since $\mathrm{dom}(A)$ is dense in X, we have $A^\dagger w = -z$, therefore $(w,-z) \in \Gamma_{A^\dagger}$. □

Proof of Proposition 6.2.3. (i) Clear.
(ii) Suppose $\mathrm{dom}(A^\dagger)$ is not dense, then $\overline{\mathrm{dom}(A^\dagger)}^\perp \neq \{0\}$. Let $y \in \overline{\mathrm{dom}(A^\dagger)}^\perp \setminus \{0\}$, then for any $w \in \mathrm{dom}(A^\dagger)$,

$$\langle (0,y), (-A^\dagger w, w)\rangle_{X\times Y} = \langle y, w\rangle_Y = 0,$$

thus

$$(0,y) \in (J\Gamma_{A^\dagger})^\perp = (\Gamma_A^\perp)^\perp = \overline{\Gamma_A},$$

which means the projection $\overline{\Gamma_A} \to X$ is not injective, thus by Lemma 2.3.5, A is not closeable. This argument is reversible, so the reverse is also true.
(iii) If $A = A^{\dagger\dagger}$, then by Lemma 6.2.2, A is closed. If A is closed, then

$$\Gamma_A = \overline{\Gamma_A} = (\Gamma_A^\perp)^\perp = (J\Gamma_{A^\dagger})^\perp.$$

Lemma 6.2.5. $\mathbb{1} + A^\dagger A : \mathrm{dom}(A^\dagger A) \to X$ *is bijective.*

Proof. Step 1. Let $u \in \mathrm{dom}(A^\dagger A)$, then $u \in \mathrm{dom}(A)$, and

$$\langle u, A^\dagger A u\rangle = \langle Au, Au\rangle = \|Au\|^2\,,$$

so

$$\|u\| \left\|(\mathbb{1} + A^\dagger A)u\right\| \geq \langle u, (\mathbb{1} + A^\dagger A)u\rangle = \|u\|^2 + \|Au\|^2 \geq \|u\|^2\,.$$

For $u \neq 0$,

$$\left\|(\mathbb{1} + A^\dagger A)u\right\| \geq \|u\|\,,$$

therefore $\mathbb{1} + A^\dagger A$ is injective (and has a closed image).
Step 2. Let $w \in X$, define $\Lambda_w \in (\mathrm{dom}(A)_{\Gamma_A})^*$ by

$$\Lambda_w(x) = \langle x, w\rangle_X, \quad \forall x \in \mathrm{dom}(A)_{\Gamma_A}.$$

Since A is closed, $\mathrm{dom}(A)_{\Gamma_A}$ is closed, hence Hilbert, then there is $v \in \mathrm{dom}(A)_{\Gamma_A}$ s.t.

$$\Lambda_w(x) = \langle x, v\rangle_{\Gamma_A} = \langle x, v\rangle_X + \langle Ax, Av\rangle_Y, \quad \forall x \in \mathrm{dom}(A)_{\Gamma_A}.$$

Therefore

$$\langle x, w\rangle_X = \Lambda_w(x) = \langle x, v\rangle_X + \langle Ax, Av\rangle_Y, \quad \forall x \in \mathrm{dom}(A)_{\Gamma_A},$$

then

$$|\langle Ax, Av\rangle_Y| = |\langle x, w - v\rangle_X| \le \|x\| \, \|w - v\|,$$

so $Av \in \mathrm{dom}(A^\dagger)$. Then

$$\langle x, v\rangle_{\Gamma_A} = \langle x, v\rangle_X + \langle Ax, Av\rangle_Y = \langle x, v\rangle_X + \langle x, A^\dagger Av\rangle_X = \langle x, (\mathbb{1} + A^\dagger A)v\rangle_X,$$

and then

$$\langle x, (\mathbb{1} + A^\dagger A)v\rangle_X = \langle x, w\rangle_X, \quad \forall x \in \mathrm{dom}(A)_{\Gamma_A}.$$

Since $\mathrm{dom}(A)$ is dense in X, we have $(\mathbb{1} + A^\dagger A)v = w$, i.e. $\mathbb{1} + A^\dagger A$ is surjective. □

Lemma 6.2.6. $\mathrm{dom}(A^\dagger A)$ *is dense in* X*.*

Proof. Consider the inclusion map $i : \mathrm{dom}(A)_{\Gamma_A} \to (X, \langle\cdot,\cdot\rangle_X)$. For the adjoint operator $i^\dagger$, since

$$|\langle i(x), y\rangle_X| = |\langle x, y\rangle_X| \le \|y\| \cdot \|y\|, \quad \forall x \in \mathrm{dom}(A), y \in X,$$

we have $\mathrm{dom}(i^\dagger) = X$, and for any $x \in \mathrm{dom}(A)_{\Gamma_A}$ and $y \in \mathrm{dom}(i^\dagger) = X$,

$$\langle i(x), y\rangle_X = \langle x, i^\dagger(y)\rangle_{\Gamma_A}.$$

Define $\Lambda = (\mathbb{1} + A^\dagger A)^{-1} : X \to \mathrm{dom}(A^\dagger A)$, which is bijective by Lemma 6.2.5. Then for any $w \in X, x \in \mathrm{dom}(A)$,

$$\begin{aligned}
\langle x, i^\dagger(w)\rangle_{\Gamma_A} &= \langle i(x), w\rangle_X \\
&= \langle x, (\mathbb{1} + A^\dagger A)\Lambda w\rangle_X \\
&= \langle x, \Lambda w\rangle_X + \langle x, A^\dagger A\Lambda w\rangle_X \\
&= \langle x, \Lambda w\rangle_X + \langle Ax, A\Lambda w\rangle_X \\
&= \langle x, \Lambda w\rangle_{\Gamma_A},
\end{aligned}$$

since $\mathrm{dom}(A)$ is dense in X, we have $i^\dagger(w) = \Lambda w$. Then

$$\mathrm{Im}\,(i^\dagger) = \mathrm{Im}\,(\Lambda) = \mathrm{dom}(A^\dagger A).$$

By Theorem 4.1.5, $\mathrm{Im}\,(i^\dagger)$ is dense if and only if $(i^\dagger)^\dagger$ is injective.
Claim: $(i^\dagger)^\dagger = i$, hence it is injective.
First, for any $y \in \mathrm{dom}(A) = \mathrm{dom}(i)$ and $x \in \mathrm{dom}(i^\dagger) = X$, we have

$$|\langle i^\dagger(x), y\rangle_{\Gamma_A}| = |\langle x, i(y)\rangle_X| = |\langle x, y\rangle_X| \le \|y\| \cdot \|x\|,$$

so $\mathrm{dom}(i^{\dagger\dagger}) = \mathrm{dom}(A)$.
Second, for any $w \in \mathrm{dom}(i^{\dagger\dagger}) = \mathrm{dom}(A)_{\Gamma_A}$ and $x \in \mathrm{dom}(i^\dagger) = X$,

$$\langle i^{\dagger\dagger}(w), x\rangle_X = \langle w, i^\dagger(x)\rangle_{\Gamma_A} = \langle i(w), x\rangle_X,$$

thus $i^{\dagger\dagger}(w) = i(w)$ for all $w \in \mathrm{dom}(i^{\dagger\dagger}) = X$. □

Theorem 6.2.7. *Let X, Y be complex Hilbert spaces. Let $A : \mathrm{dom}(A) \subseteq X \to Y$ be a closed, densely defined unbounded linear operator. Then $A^\dagger A$ is s.a., well-defined.*

Proof. We want show (1) $\mathrm{dom}(A^\dagger A) = \mathrm{dom}((A^\dagger A)^\dagger)$ and (2) for any $x \in \mathrm{dom}(A^\dagger A)$,

$$(A^\dagger A)^\dagger x = A^\dagger A x.$$

Step 1: $\mathrm{dom}(A^\dagger A) \subseteq \mathrm{dom}((A^\dagger A)^\dagger)$.
For any $w \in \mathrm{dom}(A^\dagger A)$, $Aw \in \mathrm{dom}(A^\dagger)$, thus for any $x \in \mathrm{dom}(A^\dagger A)$,

$$\langle A^\dagger A w, x\rangle = \langle Aw, Ax\rangle = \langle w, A^\dagger A x\rangle,$$

so

$$|\langle A^\dagger A w, x\rangle| \le \left\|A^\dagger A x\right\| \cdot \|w\|,$$

i.e. $w \in \mathrm{dom}((A^\dagger A)^\dagger)$.
Step 2: $\mathrm{dom}((A^\dagger A)^\dagger) \subseteq \mathrm{dom}(A^\dagger A)$.
Let $v \in \mathrm{dom}((A^\dagger A)^\dagger)$, then for any $u \in \mathrm{dom}(A^\dagger A)$,

$$\langle v, A^\dagger A u\rangle = \langle (A^\dagger A)^\dagger v, u\rangle,$$

so

$$|\langle v, A^\dagger A u\rangle_X| \le \|u\|_X \cdot \left\|(A^\dagger A)^\dagger v\right\|_X.$$

Define $\Lambda \in (\mathrm{dom}(A)_{\Gamma_A})^*$ by

$$\Lambda u = \langle A^\dagger A u, v\rangle_X + \langle u, v\rangle_X, \quad \forall u \in \mathrm{dom}(A^\dagger A),$$

and since $\mathrm{dom}(A^\dagger A)$ is dense, we can extend the domain continuously to $\mathrm{dom}(A)_{\Gamma_A}$! Then by Riesz's representation theorem, there is $w \in \mathrm{dom}(A)_{\Gamma_A}$ s.t.

$$\Lambda u = \langle u, w\rangle_{\Gamma_A} = \langle u, w\rangle_X + \langle Au, Aw\rangle_X.$$

Then for $u \in \mathrm{dom}(A^\dagger A)$,

$$\langle u, v\rangle_X + \langle A^\dagger A u, v\rangle_X = \langle u, w\rangle_X + \langle Au, Aw\rangle_X,$$

since $Au \in \mathrm{dom}(A^\dagger)$,

$$\langle u, v\rangle_X + \langle A^\dagger A u, v\rangle_X = \langle u, w\rangle_X + \langle A^\dagger A u, v\rangle_X,$$

i.e.

$$\langle (\mathbb{1} + A^\dagger A)u, v - w\rangle_X = 0, \quad \forall u \in \mathrm{dom}(A^\dagger A).$$

Since $\mathbb{1} + A^\dagger A$ is bijection and $\mathrm{dom}(A^\dagger A)$ is dense, we have $v = w$, thus $v \in \mathrm{dom}(A)$. Then for any $u \in \mathrm{dom}(A^\dagger A)$,

$$\langle v, A^\dagger A u\rangle_X = \langle Av, Au\rangle_X,$$

and

$$|\langle Av, Au\rangle_X| = |\langle v, A^\dagger A u\rangle_X| \le \|u\|_X \cdot \left\|(A^\dagger A)^\dagger v\right\|_X,$$

which implies $Av \in \mathrm{dom}(A^\dagger)$, therefore $v \in \mathrm{dom}(A^\dagger A)$.
Step 3. Now we have shown $\mathrm{dom}(A^\dagger A) = \mathrm{dom}((A^\dagger A)^\dagger)$. For any $x, y \in \mathrm{dom}(A^\dagger A)$,

$$\langle (A^\dagger A)^\dagger x, y\rangle = \langle x, A^\dagger A y\rangle = \langle Ax, Ay\rangle = \langle A^\dagger A x, y\rangle,$$

since $\mathrm{dom}(A^\dagger A)$ is dense in X, we have $(A^\dagger A)^\dagger x = A^\dagger A x$. □

6.3 Functional calculus

Theorem 6.3.1. *Let H be a complex Hilbert space.*

(1) *Let $A : \mathrm{dom}(A) \subseteq H \to H$ be s.a. By the same argument for bounded s.a. operator, $\sigma(A) \subseteq \mathbb{R}$. So we can define an invertible operator $U : H \to H$ by*

$$U = (A - i\mathbb{1})(A - i\mathbb{1})^{-1}.$$

Then U is unitary, $\mathbb{1} - U$ is injective, $\mathrm{dom}(A) = \mathrm{Im}\,(\mathbb{1} - U)$ where $A = i(\mathbb{1} + U)(\mathbb{1} - U)^{-1}$. U is called the Cayley transform of A.

(2) *Let $U \in \mathcal{L}(H)$ be unitary s.t. $\mathbb{1} - U$ is injective. Then*

$$A = i(\mathbb{1} + U)(\mathbb{1} - U)^{-1} : \mathrm{dom}(A) \to H,$$

where $\mathrm{dom}(A) = \mathrm{Im}\,(\mathbb{1} - U)$, is s.a. and U is the Cayley transform of A.

Appendix A

Some theorems

A.1 Zorn's lemma

Definition A.1.1. A partially ordered set $(S, \preceq)$ is a pair consisting a set S and a relation $\preceq$ on S called partial order s.t. for any $x, y, z \in S$

(1) $x \preceq x$;

(2) If $x \preceq y$ and $y \preceq x$, then $x = y$;

(3) If $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

Definition A.1.2. Suppose $(S, \preceq)$ is a partial ordered set. $T \subseteq S$ is a subset (thus also a partial ordered set).

1. For any $x, y \in S$, denote $x \prec y$ if $x \preceq y$ and $x \neq y$.
2. $x, y \in S$ are called comparable if either $x \preceq y$ or $y \preceq x$.
3. T is called a chain (or totally ordered set) if any $x, y \in T$ are comparable.
4. $s \in S$ is called an upper bounded of T if for any $x \in T$, s and x are comparable, and moreover $x \preceq s$.
5. $m \in S$ is called a maximal element of S if there is no such $x \in S$ s.t. $m \prec x$.

Theorem A.1.3 (Zorn's lemma)**.** *Suppose $(S, \preceq)$ is a partially ordered set. If every chain in S has an upper bound in S, then S has at least one maximal element.*

The following is a simple application of Zorn's lemma.

Definition A.1.4. Suppose V is a vector space and $B \subseteq V$ is a subset.

1. B is called linearly independent, if for any finite subset $\{v_1, \cdots, v_n\} \subseteq B$, if there is scalars $c_i \in \mathbb{R}$ $(1 \leq i \leq n)$ s.t.

$$c_1 v_1 + c_2 v_2 + \cdots + c_n v_n = 0,$$

then $c_i = 0$ for all $1 \leq i \leq n$.

2. B spans V, denoted by $V = \text{Span}(B)$, if for any $v \in V$, there is a finite subset $\{v_1, \cdots, v_n\} \subseteq B$ and scalars $c_i \in \mathbb{R}$ $(1 \leq i \leq n)$ s.t.

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n.$$

3. B is called a Hamel basis of V if

   (a) B is linearly independent;

   (b) $\text{Span}(B) = V$.

Lemma A.1.5. *Suppose $B \subseteq V$ is a linearly independent subset, if $\text{Span}(B) \subsetneq V$, then for any $v \in V \setminus \text{Span}(B)$, $B \cup \{v\}$ is linearly independent.*

Proof. Take $v \in V \setminus \text{Span}(B)$, it is clear $v \neq 0$. Assume $B \cup \{v\}$ is dependent, then there is a finite subset $\{v_1, \cdots, v_n\} \subseteq B$ and not all 0 scalars $c_i \in \mathbb{R}$ $(1 \leq i \leq n+1)$ s.t.

$$c_1 v_1 + \cdots + c_n v_n + c_{n+1} v = 0.$$

Moreover, $c_{n+1} \neq 0$, otherwise $c_i = 0$ for all $1 \leq i \leq n+1$, then $B \cup \{v\}$ is linearly independent. Therefore

$$v = \frac{c_1}{c_{n+1}} v_1 + \cdots + \frac{c_n}{c_{n+1}} v_n,$$

i.e. $v \in \text{Span}(B)$. □

Theorem A.1.6. *Every vector space has a Hamel basis.*

Proof. Let V be a vector space, define

$$\mathcal{P} := \{B \subseteq V : \ B \text{ is linearly independent}\},$$

then $(\mathcal{P}, \subseteq)$ is a partially ordered set. For any chain $\mathcal{C} \subseteq \mathcal{P}$, define

$$B_{\mathcal{C}} = \bigcup_{B \in \mathcal{C}} B,$$

then $B_{\mathcal{C}}$ is linearly independent, thus $B_{\mathcal{C}} \in \mathcal{P}$ and then it is an upper bound of $\mathcal{C}$. By Zorn's lemma (A.1.3), there is a maximal element $B_m \in \mathcal{P}$. It's clear $\text{Span}(B_m) = V$, otherwise, we can find $v \in V \setminus \text{Span}(B_m)$ s.t. $B_m \cup \{v\} \in \mathcal{P}$, which means B_m is not a maximal element. □

A.2 Compact sets

Theorem A.2.1. *Suppose (X, d) is a metric space and $A \subseteq X$. TFAE*

(1) *A is pre-compact.*

(2) *Every sequence in A has a convergent subsequence.*

(3) *A is totally bounded and every Cauchy sequence in A converges in X.*

Definition A.2.2. $\mathcal{F} \subseteq C(X)$ is equi-continuous if for every $\varepsilon > 0$, there is $\delta > 0$ s.t. for all $x, y \in X$ satisfying

$$d(x, y) < \delta,$$

we have

$$|f(x) - f(y)| < \varepsilon, \quad \forall f \in \mathcal{F}.$$

Theorem A.2.3. *Suppose (X, d) is a compact metric space and $\mathcal{F} \subseteq C(X)$. Then*

(1) *$\mathcal{F}$ is pre-compact if and only if it is bounded and equi-continuous.*

(2) *$\mathcal{F}$ is compact if and only if it is closed, bounded, and equi-continuous.*