# Functional Analysis

Shitao Liu
(Notes by Shuai Wei)

September 17, 2023

# Contents

# Chapter 1

# Metric Spaces

In this chapter we introduce metric spaces which can be seen as a generalization of the real numbers studied in advanced calculus. A metric space is a set $X$ with a metric defined on it. More precisely, the metric associates with any pair of elements of $X$ a distance. The metric is defined axiomatically, with the axioms being motivated from the corresponding properties of real numbers. A metric on a set also induces topological properties such as open and closed sets, which can lead to the study of more abstract topological spaces. Other important properties of metric spaces, such as separability and completeness, will also be investigated in this chapter. To demonstrate some applications of completeness, we introduce the Banach fixed point theorem and the Baire category theorem in the context of metric spaces at the end of this chapter.

Let $\overline{\mathbb{R}} = \mathbb{R} \sqcup \{\infty\}$.

## 1.1 Definition and Examples

**Definition 1.1.** Let $X$ be a set and $d : X \times X \to \mathbb{R}$ be a function on $X \times X := \{(x, y) \mid x, y \in X\}$ that satisfies for all $x, y, z \in X$,

(0) $d(x, y) \geqslant 0$ "nonnegativity".

(1) $d(x, y) = 0$ if and only if $x = y$ "definiteness".

(2) $d(x, y) = d(y, x)$ "symmetry".

(3) $d(x, y) \leqslant d(x, z) + d(z, y)$ "triangle-inequality".

Then the pair $(X, d)$ is called a *metric space*, where $X$ is called the *underlying set*, $d$ called the *metric* or *distance*.

**Remark.** Condition (0) is redundant. To see this, let $y = x$ in (3), then $0 = d(x, x) \leqslant d(x, z) + d(z, x) = 2d(x, z)$ for any $x, z \in X$.

**Example 1.2.** (a) Let $p \in \mathbb{R}$ and $|\cdot|^p : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined as $(x, y) \mapsto |x - y|^p$. Then $(\mathbb{R}, |\cdot|^p)$ is a metric space if and only if $0 < p \leqslant 1$.

(b) Let $p \in \overline{\mathbb{R}}$ and $n \in \mathbb{N}$. Define for $p \in \mathbb{R}^{\geqslant 1}$,

$$\begin{aligned} d_p : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \left( \sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}}, \end{aligned}$$

and

$$\begin{aligned} d_\infty : \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \max_{i=1,\dots,n} |x_i - y_i|. \end{aligned}$$

(1) If $n = 1$, then $d_p = |\cdot|$ for $p \in \overline{\mathbb{R}} \setminus \{0\}$.

(2) Let $n \in \mathbb{Z}^{\geqslant 2}$. Then $(\mathbb{R}^n, d_p)$ is a metric space if and only if $p \in \overline{\mathbb{R}}^{\geqslant 1}$. The metric $d_2$ is called the *Euclidean metric* on $\mathbb{R}^n$.

**Definition 1.3.** (a) Let

$$\begin{aligned} l^\infty &:= \{x : \mathbb{N} \to \mathbb{R} \mid x = \{x_i\}_{i \geqslant 1} \text{ is a bounded sequence}\} \\ &= \left\{ x : \mathbb{N} \to \mathbb{R} \mid x = \sup_{i \in \mathbb{N}} |x_i| < \infty \right\}. \end{aligned}$$

(b) Let

$$l^p = \left\{ x : \mathbb{N} \to \mathbb{R} \mid x = \{x_i\}_{i \geqslant 1} \text{ such that } \sum_{i=1}^{\infty} |x_i|^p < \infty \right\}.$$

**Example 1.4.** $(l^p, d_p)$ is a metric space if and only if $p \in \overline{\mathbb{R}}^{\geqslant 1}$, where for $p \in \mathbb{R}^{\geqslant 1}$,

$$\begin{aligned} d_p : l^p \times l^p &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \left( \sum_{i=1}^{\infty} |x_i - y_i|^p \right)^{\frac{1}{p}}, \end{aligned}$$

and

$$\begin{aligned} d_\infty : l^\infty \times l^\infty &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \sup_{i \in \mathbb{N}} |x_i - y_i|. \end{aligned}$$

**Remark.** $d_\infty$ is well-defined.

**Fact 1.5.** Let $p \in \overline{\mathbb{R}}$. Then $x \in (l^p, d_p)$ if and only if $d_p(x, 0) < \infty$.

**Theorem 1.6.** *$l^p \subsetneq l^q$ if and only if $1 \leqslant p < q \leqslant \infty$.*

**Example 1.7.** $(S, d_S)$ is a metri space, where $S = \{x : \mathbb{N} \to \mathbb{R} \mid x = \{x_i\}_{i \in \mathbb{N}}$ is a (real) sequence$\}$ and

$$\begin{aligned} d_S : S \times S &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|x_i - y_i|}{1 + |x_i - y_i|}. \end{aligned}$$

*Proof.* $d_S$ is well-defined since $\sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|x_i - y_i|}{1+|x_i - y_i|} \leqslant \sum_{i=1}^{\infty} \frac{1}{2^i} = 1$. □

**Remark.** (a) $\frac{1}{2^i}$ can be replaced by $a_i > 0$ such that $\sum_{i=1}^{\infty} a_i < \infty$ (converges).

(b) $\frac{t}{1+t}$ can be replaced by any function $f$ which is bounded, increasing, concave and $f(0) = 0$.

**Example 1.8.** Let

$$C[a,b] = \{x : [a,b] \to \mathbb{R} \mid x = x(t) \text{ is a continuous function}\}.$$

$(C[a,b], d_p)$ is a metric space if and only if $p \in \overline{\mathbb{R}}^{\geqslant 1}$, where for $p \in \mathbb{R}^{\geqslant 1}$,

$$\begin{aligned} d_p : C[a,b] \times C[a,b] &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto \left( \int_a^b |x(t) - y(t)|^p \right)^{\frac{1}{p}}, \end{aligned}$$

and

$$\begin{aligned} d_\infty : C[a,b] \times C[a,b] &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto \max_{t \in [a,b]} |x(t) - y(t)|. \end{aligned}$$

Check both are well-defined.

**Definition 1.9.** Let

$$B[a,b] = \{x : [a,b] \to \mathbb{R} \mid x = x(t) \text{ is a bounded function}\}.$$

**Example 1.10.** $(B[a,b], d_\infty)$ and $(B(a,b), d_\infty)$ are metric spaces, where

$$\begin{aligned} d_\infty : B(a,b) \times B(a,b) &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto \sup_{t \in (a,b)} |x(t) - y(t)|. \end{aligned}$$

**Example 1.11.** $(X, d_{\text{disc}})$ is a *discrete metric space*, where

$$\begin{aligned} d_{\text{disc}} : X \times X &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto \begin{cases} 0, & x = y, \\ 1, & x \neq y. \end{cases} \end{aligned}$$

**Assumption 1.12.** Let $X$ be a metrix space.

Next we look at some basic ways to create new metric spaces from the ones that we already have.

**Definition 1.13.** Let $(X,d)$ be a metric space and $Y \subseteq X$, then $(Y, d_{Y \times Y})$ is a *metric subspace* of $(X,d)$, and $d_{Y \times Y}$ is called the metric *induced* by $d$.

**Example 1.14.** Let $Y = \{x \in l^\infty \mid x = \{x_n\},\ x_n = 0 \text{ or } 1, \forall n \in \mathbb{N}\} \subseteq l^\infty$. We have $(Y, d_\infty|_{Y \times Y})$ is a metric subspace of $l^\infty$ and $d_\infty|_{Y \times Y} = d_{\text{disc}}$.

**Theorem 1.15** (Finite product)**.** *Let $(X_i, d_i)$ be a metric space for $i = 1, \ldots, n$. Then $(\prod_{i=1}^n X_i, d_p)$ is a metric space where*

$$d_p : \prod_{i=1}^n X_i \times \prod_{i=1}^n X_i \longrightarrow \mathbb{R}$$

$$(x, y) \longmapsto \begin{cases} \left(\sum_{i=1}^n d_i(x_i, y_i)^p\right)^{\frac{1}{p}}, & \text{when } p \in \mathbb{R}^{\geqslant 1}, \\ \max\limits_{1 \leqslant i \leqslant n} d_i(x_i, y_i), & \text{when } p = \infty. \end{cases}$$

*Proof.* Let $p \in \mathbb{R}^{\geqslant 1}$. Let $x = \{x_i\}_{i=1}^n, y = \{y_i\}_{i=1}^n, z = \{z_i\}_{i=1}^n \in \prod_{i=1}^n X_i$. Let

$$f = (d_1(x_1, z_1), \ldots, d_n(x_n, z_n)) \in \mathbb{R}^n \text{ and } g = (d_1(z_1, y_1), \ldots, d_n(z_n, y_n)) \in \mathbb{R}^n.$$

Then by Minkowski inequality, $d_p(x, y) = (\sum_{i=1}^n d_i(x_i, y_i)^p)^{\frac{1}{p}} \leqslant (\sum_{i=1}^n (d_i(x_i, z_i) + d_i(z_i, y_i))^p)^{\frac{1}{p}} = \|f + g\|_p \leqslant \|f\|_p + \|g\|_p = d_p(x, z) + d_p(y, z)$. □

**Corollary 1.16.** Let $p \in \overline{\mathbb{R}}$. Then $(\mathbb{R}^n, d_p)$ is a metric space if and only if $n = 1$ or $n \in \mathbb{Z}^{\geqslant 2}$ and $p \in \overline{\mathbb{R}}^{\geqslant 1}$.

*Proof.* Take $(X_i, d_i) = (\mathbb{R}, |\cdot|)$ for $i = 1, \ldots, n$. □

**Theorem 1.17** (Countable space)**.** *Let $(X_i, d_i)$ be a metric space for each $i \in \mathbb{N}$. Then $\left(\prod_{i=1}^\infty X_i, d_i\right)$ is a metric space, where*

$$d : \prod_{i=1}^\infty X_i \times \prod_{i=1}^\infty X_i \longrightarrow R$$

$$(x, y) \longmapsto \sum_{i=1}^\infty \frac{1}{2^i} \frac{d_i(x_i, y_i)}{1 + d_i(x_i, y_i)}.$$

**Corollary 1.18.** $(S, d_S)$ is a metric space.

*Proof.* Take $(X_i, d_i) = (\mathbb{R}, |\cdot|)$ for each $i \in \mathbb{N}$. □

## 1.2 Topology of Metric Spaces

In this section we introduce some basic topological concepts that are of fundamental importance in studying metric spaces. Most of these concepts are motivated from the geometry of Euclidean spaces Rn and should become quite familiar when looking at them that way.

**Assumption 1.19.** Let $(X, d)$ be a metric space and $A \subseteq X$.

**Definition 1.20.** Define

(a)

$$d(x, A) = \inf\{d(x, y) \mid y \in A\}.$$

(b)

$$d(A, B) = \inf\{d(x, y) \mid x \in A \text{ and } y \in B\}.$$

(c)
$$\operatorname{diam}(A) = \sup\{d(x,y) \mid x,y \in A\}.$$

(d) $A$ is bounded if $\operatorname{diam}(A) < \infty$.

**Definition 1.21.** Let $x_0 \in X$ and $\epsilon > 0$. Then we define

(a) "*Open $\epsilon$-ball* around $x_0$":
$$B_\epsilon^d(x_0) = \{x \in X \mid d(x,x_0) < \epsilon\}.$$

(b) "*Closed $\epsilon$-ball* around $x_0$":
$$\overline{B_\epsilon^d}(x_0) = \{x \in X \mid d(x,x_0) \leqslant \epsilon\}.$$

(c) "*$\epsilon$-sphere* around $x_0$":
$$\partial B_\epsilon^d(x_0) = \{x \in X \mid d(x,x_0) = \epsilon\} = \overline{B_\epsilon^d}(x_0) \smallsetminus B_\epsilon^d(x_0).$$

**Remark.** (a) $B_\epsilon^d(x_0) \neq \emptyset$ and $\overline{B_\epsilon^d}(x_0) \neq \emptyset$ for any metric $d$ and $\epsilon > 0$.

(b) $\partial B_\epsilon^d(x_0) = \emptyset$ for $d = \text{disc}$ and any $\epsilon \neq 1$.

**Definition 1.22.** (a) $x_0 \in X$ is an *interior point* of $A$ if there exists $\epsilon > 0$ such that $B_\epsilon^d(x_0) \subseteq A$.

(b) The *interior* of $A$:
$$\operatorname{Int}(A) := \{x \in X \mid x \text{ is an interior point of } A\}.$$

(c) $x_0 \in X$ is an *accumulation point* (*limiting point*) of $A$ if for $\epsilon > 0$, there exists $x_0 \neq x \in A$ such that $x \in B_\epsilon^d(x_0)$.

(d) The *derived set* of $A$:
$$A' = \{x \in X \mid x \text{ is an accumulation point of } A\}.$$

(e) The *closure* of $A$:
$$\overline{A} := A \cup A'.$$

**Example 1.23.** Let $X = (0,1)$ and $A = X$. Then $\overline{A} = (0,1)$ by definition.

**Lemma 1.24.** (a) $\operatorname{Int}(A) \subseteq A \subseteq \overline{A}$.

(b) $x_0 \in \overline{A}$ if and only if for $\epsilon > 0$, there exists $x \in A$ such that $x \in B_\epsilon^d(x_0)$, i.e., $B_\epsilon^d(x_0) \cap A \neq \emptyset$.

(c) If $A \subseteq B$, then $\operatorname{Int}(A) \subseteq \operatorname{Int}(B)$ and $\overline{A} \subseteq \overline{B}$.

(d) $\operatorname{Int}(A) = \left(\overline{A^c}\right)^c$.

(e) $\overline{A} = (\operatorname{Int}(A^c))^c$.

**Definition 1.25.** (a) $A$ is *open* if $A = \operatorname{Int}(A)$ or $A \subseteq \operatorname{Int}(A)$.

(b) $A$ is *closed* if $A = \overline{A}$ or $\overline{A} \subseteq A$.

**Lemma 1.26.** $A$ is open if and only if $A^c$ is closed and $A$ is closed if and only if $A^c$ is open.

**Theorem 1.27.** *$B_\epsilon(x_0)$ is open and $\overline{B_\epsilon}(x_0)$ is closed for all $x_0 \in X$ and $\epsilon > 0$.*

*Proof.* Let $x \in B_\epsilon(x_0)$, then $d(x, x_0) < \epsilon$. Let $\epsilon' := \epsilon - d(x, x_0) > 0$. Note for $y \in B_{\epsilon'}(x)$, we have $d(y, x_0) \leqslant d(y, x) + d(x, x_0) < \epsilon' + d(x, x_0) = \epsilon$. So $B_{\epsilon'}(x) \subseteq B_\epsilon(x_0)$. Then $x \in \operatorname{Int}(B_\epsilon(x_0))$.

Let $x \in \overline{\overline{B_\epsilon}(x_0)}$, then we have $B_{\frac{1}{n}}(x) \cap \overline{B_\epsilon}(x_0) \neq \emptyset$ for $n \in \mathbb{N}$. So there exists $\{y_n\}$ such that $d(y_n, x) < \frac{1}{n}$ and $d(y_n, x_0) \leqslant \epsilon$ for $n \in \mathbb{N}$. Then $d(x, x_0) \leqslant d(x, y_n) + d(y_n, x_0) < \frac{1}{n} + \epsilon$ for $n \in \mathbb{N}$. So $d(x, x_0) \leqslant \epsilon$. Thus, $x \in \overline{B_\epsilon}(x_0)$. □

**Remark.** (a) A metric space is a topology space with the topology being the collection of all open sets.

(b) In metric space, certain sets could be open and closed. For example, in $(X, d_{\text{disc}})$, any set $A \subseteq X$ is both open and closed.

**Proposition 1.28.** (a) $\emptyset$ and $X$ are open and closed.

(b) If $\{A_\alpha\}_{\alpha \in I}$ is open, then $\bigcup_{\alpha \in I} A_\alpha$ is open.

(c) If $\{A_i\}_{i=1}^n$ is open, then $\bigcap_{i=1}^n A_i$ is open.

(d) If $\{A_\alpha\}_{\alpha \in I}$ is closed, then $\bigcap_{\alpha \in I} A_\alpha$ is closed.

(e) If $\{A_i\}_{i=1}^n$ is closed, then $\bigcup_{i=1}^n A_i$ is closed.

*Proof.* (a) Since $\emptyset \subseteq \operatorname{Int}(\emptyset)$, $\emptyset$ is open. Since for $x \in X$, $B_\epsilon(x) \subseteq X$ for $\epsilon > 0$, $X$ is open.

(b) Let $x \in \bigcup_{\alpha \in I} A_\alpha$. Then there exists $\alpha_0 \in I$ such that $x \in A_{\alpha_0} = \operatorname{Int}(A_{\alpha_0}) \subseteq \operatorname{Int}\left(\bigcup_{\alpha \in I} A_\alpha\right)$.

(c) Let $x \in \bigcap_{i=1}^n A_i = \operatorname{Int}(A_i)$. Then $x \in A_i$ for $i = 1, \ldots, n$. So there exists $\epsilon_i > 0$ such that $B_{\epsilon_i}(x) \subseteq A_i$ for each $i = 1, \cdot, n$. Let $\epsilon = \min_{1 \leqslant i \leqslant n} \{\epsilon_i\} > 0$. Then $B_\epsilon(x) \subseteq A_i$ for each $i = 1, \ldots, n$. So $B_\epsilon(x) \subseteq \bigcap_{i=1}^n A_i$. Thus, $x \in \operatorname{Int}(\bigcap_{i=1}^n A_i)$. □

**Theorem 1.29.** *(a)* $\operatorname{Int}(A)$ *is the largest open set contained in $A$.*

*(b) $\overline{A}$ is the smallest closed set containing $A$.*

*Proof.* (a) Let $x \in \operatorname{Int}(A)$. Then there exists $\epsilon_x > 0$ such that $B_{\epsilon_x}(x) \subseteq A$. Since $B_{\epsilon_x}(x) = \operatorname{Int}(B_{\epsilon_x}(x)) \subseteq \operatorname{Int}(A)$, we have $\bigcup_{x \in \operatorname{Int}(A)} B_{\epsilon_x}(x) \subseteq \operatorname{Int}(A) = \bigcup_{x \in \operatorname{Int}(A)} \{x\} \subseteq \bigcup_{x \in \operatorname{Int}(A)} B_{\epsilon_x}(x)$. So $\operatorname{Int}(A) = \bigcup_{x \in \operatorname{Int}(A)} B_{\epsilon_x}(x)$, which implies $\operatorname{Int}(A)$ is open. Let $B \subseteq A$ such that $B$ is open. $B = \operatorname{Int}(B) \subseteq \operatorname{Int}(A)$, which implies $\operatorname{Int}(A)$ is the largest such set.

(b) Since $\overline{A} = (\operatorname{Int}(A^c))^c$, we have $\overline{A}$ is closed. Let $C \supseteq A$ such that $C$ is closed. Then $\overline{C} \supseteq \overline{A}$, which implies $\overline{A}$ is the smallest such set. □

**Corollary 1.30.** (a) $\operatorname{Int}(\operatorname{Int}(A)) = \operatorname{Int}(A)$.

(b) $\overline{\overline{A}} = \overline{A}$.

(c) If $A$ is open, then $A$ is a union of open balls.

**Proposition 1.31.** (a) $\overline{A \cup B} = \overline{A} \cup \overline{B}$ and $\operatorname{Int}(A \cap B) = \operatorname{Int}(A) \cap \operatorname{Int}(B)$.

(b) $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$ and $\operatorname{Int}(A \cup B) \supseteq \operatorname{Int}(A) \cup \operatorname{Int}(B)$.

## 1.3 Separable Spaces

In this section we introduce the concept of separability for a metric space. This is a topo- logical property that may also give a limitation on the size of metric spaces. More precisely, separable metric spaces can not have a size that is larger than the size of real numbers. In order to define separability, we need to introduce the notion of denseness. Roughly speaking, a subset $A$ is dense in a metric space $(X, d)$ if for every point in $X$, it is either in $A$ or arbitrarily close to a member of $A$.

**Assumption 1.32.** Let $(X, d)$ be a metric space and $A \subseteq X$.

**Definition 1.33.** $A$ is said to be *dense* in $(X, d)$ if $\overline{A} = X$.

**Remark.** $A$ is dense in $X$ if and only if $x \in \overline{A}$ for $x \in X$ if and only if for $x \in X$ and any $\epsilon > 0$, we have $B_\epsilon(x) \cap A \neq \emptyset$ if and only if for $x \in X$ and any $\epsilon > 0$, there exists $\alpha \in A$ such that $d(x, \alpha) < \epsilon$.

**Definition 1.34.** $A$ is said to be *separable* if $X$ contains a countable dense subset.

**Remark.** Separability depends on the metric.

**Example 1.35.** $(X, d_{\text{disc}})$ is separable if and only if $X$ is countable.

*Proof.* $\Longleftarrow$ Clearly.

$\Longrightarrow$ Let $A \subsetneq X$. Then there exists $x \in X \setminus A$. So $B_{1/2}(x) = \{x\} \not\subseteq A$. Thus, $\overline{A} \neq X$ and so the dense subset of $X$ is $X$. Since $X$ is countable, $X$ is the countable dense subset of $X$. □

**Example 1.36.** (a) $(\mathbb{R}, |\cdot|)$ is separable since $\mathbb{Q}$ is a countable dense subset of $\mathbb{R}$.

(b) $(\mathbb{R}^n, d_p)$ is separable since $\{(q_1, \ldots, q_n) \mid q_i \in \mathbb{Q}, \forall i = 1, \ldots, n\}$ is a countable dense subset of $\mathbb{R}^n$.

(c) $(\mathbb{R}, d_{\text{disc}})$ is not separable.

**Example 1.37.** (a) $(l^p, d_p)$ is a separable space for $p \in \mathbb{R}^{\geqslant 1}$.

(b) $(l^\infty, d_\infty)$ is not separable.

*Proof.* (a) Let $A = \{\{x_n\} \mid x_n \in \mathbb{Q}, x_n \neq 0 \text{ for finitely many } n \in \mathbb{N}\}$. Since $A \cong \bigcup_{n\in\mathbb{N}} \mathbb{Q}^n$, we have $A \subseteq l^p$ is countable. Let $x = \{x_i\} \in l^p$ and $\epsilon > 0$. Then $\sum_{i=1}^{\infty} |x_i|^p < \infty$. So there exists $N \in \mathbb{N}$ such that $\sum_{i=N}^{\infty} |x_i|^p < \epsilon/2$. Pick $a = \{q_1, \ldots, q_n, 0, 0, \cdots\} \in A$ such that $|x_i - q_i| < \frac{\epsilon}{(2N)^p}$. Then $d_p(x, a) = (\sum_{i=1}^{\infty} |x_i - q_i|^p)^{1/p} < \epsilon$. So $A$ is dense in $(l_p, d_p)$.

(b) Consider $Y = \{\{y_n\} = y \in l^\infty \mid y_n = 0 \text{ or } 1, \forall n \in \mathbb{N}\}$. There is a 1-1 correspondence between $x \in [0, 1]$ and a sequence coming from its binary representaion $\sum_{n=1}^{\infty} \frac{y_n}{2^n}$.

Since $d_\infty|_Y$ is a discrete metric, we have $\{B_{1/2}(y) \mid y \in Y\}$ is a collection of uncountable disjoint open balls. So any dense subset of $(l^\infty, d_\infty)$ is not countable. □

**Example 1.38.** $(C[0, 1], d_p)$ is separable for $p \in \overline{\mathbb{R}}^{\geqslant 1}$.

*Proof.* (a) Let $p=\infty$. It is enough to consider $Q[0,1]=\left\{\sum_{i=0}^{\text{finite}} q_i t^i \;\middle|\; q_i\in\mathbb{Q}, \forall i, t\in[0,1]\right\}$, which is countable since $Q[0,1]\cong\bigcup_{n=1}^{\infty}\mathbb{Q}^n$. Let $x\in C[0,1]$ and $\epsilon>0$. By the Weierstrass Approximation theorem, $P[0,1]$ is dense in $C[0,1]$, so there exists $p=\sum_{i=1}^{n} p_i t^i\in P[0,1]$ such that $d_\infty(x,p)<\frac{\epsilon}{2}$. Since $Q$ is dense in $\mathbb{R}$, we can pick $q(t)=\sum_{i=0}^{n} q_i t^i\in Q[0,1]$ such that $|q_i-p_i|<\frac{\epsilon}{2(n+1)}$ for $i=1,\dots,n$. Then $d_\infty(x,q)\leqslant d_\infty(x,p)+d_\infty(p,q)\leqslant\frac{\epsilon}{2}+\max_{t\in[0,1]}\left|\sum_{i=0}^{n}(p_i-q_i)t^i\right|\leqslant\frac{\epsilon}{2}+\max_{t\in[0,1]}\sum_{i=0}^{n}|p_i-q_i|t^i\leqslant\frac{\epsilon}{2}+\sum_{i=0}^{n}|p_i-q_i|=\epsilon$.

(b) Let $p\in\mathbb{R}^{\geqslant 1}$. Observe that $d_p(x,y)=\left(\int_0^1|x(t)-y(t)|^p dt\right)^{1/p}\leqslant\left(\int_0^1|d_\infty(x,y)|^p dt\right)^{1/p}=d_\infty(x,y)$ for $x,y\in C[0,1]$. Then, clearly, we have the result. □

**Example 1.39.** $(B[0,1],d_\infty)$ is not separable.

*Proof.* Let $Y=\left\{y_s\in B[0,1]\;\middle|\; y_s(t)=\begin{cases}1 & \text{if } s\neq t\\ 0 & \text{if } s=t\end{cases}\right\}$. Then $|Y|=|[0,1]|$, which implies $Y$ is uncountable. Also, note $d_\infty(y_{s_1},y_{s_2})=\sup_{t\in[0,1]}|y_{s_1}(t)-y_{s_2}(t)|=\begin{cases}1 & \text{if } s_1\neq s_2\\ 0 & \text{if } s_1=s_2\end{cases}$ for $s_1,s_2\in[0,1]$. So $d_\infty|_Y$ is a discrete metric. Then $\{B^d_{1/2}(y_s)\mid y_s\in Y\}$ is a collection of uncountable disjoint open balls. Thus, any dense subset of $(B[0,1],d_\infty)$ is not countable. □

## 1.4 Continuous Mapping and Sequences

Let $(X,d_X)$ and $(Y,d_Y)$ be metric spaces and $A\subseteq X$.

**Definition 1.40.** Let $f:(X,d_X)\to(Y,d_Y)$. We say $f$ is *continuous* at $x_0$ if for $\epsilon>0$, there exists $\delta(x_0,\epsilon)>0$ such that $f(x)\in B^{d_Y}_\epsilon(f(x_0))$ whenever $x\in B^{d_X}_\delta(x_0)$, i.e., $f\left(B^{d_X}_\delta(x_0)\right)\subseteq B^{d_Y}_\epsilon(f(x_0))$, i.e., $B^{d_X}_\delta(x_0)\subseteq f^{-1}\left(B^{d_Y}_\epsilon(f(x_0))\right)$.

**Theorem 1.41.** *Let $f:(X,d_X)\to(Y,d_Y)$. Then $f$ is continuous if and only if for any $U\subseteq Y$ open, $f^{-1}(U)\subseteq X$ open.*

*Proof.* $\Longleftarrow$ Let $f$ be continuous and $U\subseteq Y$ be open. If $f^{-1}(U)=\emptyset$, then it is open. Assume now $f^{-1}(U)\neq\emptyset$. Let $x_0\in f^{-1}(U)$. Then $f(x_0)\in f\left(f^{-1}(U)\right)\subseteq U$. Since $U$ is open, there exists $\epsilon>0$ such that $B_\epsilon(f(x_0))\subseteq U$. Since $f$ is continuous at $x_0$, there exists $\delta>0$ such that $B_\delta(x_0)\subseteq f^{-1}(B_\epsilon(f(x_0)))\subseteq f^{-1}(U)$. So $f^{-1}(U)$ is open.

$\Longleftarrow$ Let $x_0\in X$ and $\epsilon>0$. Then $B_\epsilon(f(x_0))\subseteq Y$ open. So by assumption, $f^{-1}(B_\epsilon(f(x_0)))\subseteq X$ is open. Also, since $x_0\in f^{-1}f(x_0)\subseteq f^{-1}(B_\epsilon(f(x_0)))$, there exists $\delta>0$ such that $B_\delta(x_0)\subseteq f^{-1}(B_\epsilon(f(x_0)))$. So $f$ is continuous at $x_0$. □

**Definition 1.42.** $x$ is a *sequence* in $X$ denoted as $x=\{x_n\}$ if $x$ is a mapping from $\mathbb{N}$ to $X$, where $x_n=x(n)$ for $n\in\mathbb{N}$.

(a) $\{x_n\}$ *converges* to $x_0\in X$ if for $\epsilon>0$, there exists $N\in\mathbb{N}$ such that $d(x_n,x_0)<\epsilon$ for $n\geqslant N$, i.e., $\{x_n\}_{n\geqslant N}\subseteq B^d_\epsilon(x_0)$. Denote it as

$$\lim_{n\to\infty}x_n=x_0 \text{ and } x_n\xrightarrow{d}x_0 \text{ as } n\to\infty.$$

(b) $\{x_n\}$ is *Cauchy* if for $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $x_n \in B^d_\epsilon(x_m)$ whenever $m, n > N$, i.e., $d(x_n, x_m) \xrightarrow{|\cdot|} 0$ as $m, n \to \infty$.

(c) $\{x_n\}$ is *bounded* if $\{x_n\} \subseteq B_r(x_0)$ for some $x_0 \in X$ and $r > 0$.

**Remark.** $\{x_n\} \subseteq X$ is bounded if and only if $\operatorname{diam}(\{x_n\}) = \sup_{m,n\in\mathbb{N}} \{d(x_m, x_n)\} < \infty$.

*Proof.* $\Longrightarrow$ Since $\{x_n\} \subseteq X$ is bounded, there exists $x_0 \in X$ and $r > 0$ such that $d(x_n, x_0) < r$ for each $n \in \mathbb{N}$. Then $d(x_m, d_n) \leqslant d(x_m, x_0) + d(x_n, x_0) \leqslant 2r < \infty$ for $m, n \in \mathbb{N}$.

$\Longleftarrow$ We can choose any $x_n$ as a compared point. □

**Theorem 1.43.** *(a) Convergent sequence has a unique limit.*

*(b) "Convergent sequence" $\subseteq$ "Cauchy sequence" $\subseteq$ "bounded sequence".*

*Proof.* (a) Assume $x_n \to x$ and $x_n \to y$ as $n \to \infty$. Then $d(x, y) \leqslant d(x, x_n) + d(x_n, y)$ for $n \in \mathbb{N}$. So $d(x, y) \leqslant \lim_{n\to\infty}(d(x, x_n) + d(x_n, y)) = \lim_{n\to\infty} d(x, x_n) + \lim_{n\to\infty}(x_n, y) = 0 + 0 = 0$.

(b) Let $x_n \to x$ as $n \to \infty$. Then $d(x_n, x_m) \leqslant d(x_n, x) + d(x, x_m)$. So $0 \leqslant \lim_{m,n\to\infty} d(x_n, x_m) \leqslant \lim_{n\to\infty} d(x_n, x) + \lim_{m\to\infty} d(x, x_m) = 0 + 0 = 0$. So $\{x_n\}$ is Cauchy.

Let $\{x_n\}$ be Cauchy. Then for $\epsilon = 1$, there exists $N \in \mathbb{N}$ such that $d(x_n, x_m) < 1$ whenever $n, m \geqslant N$. So $d(x_n, x_N) < 1$ for $n \geqslant N$. Let $r = \max\{1, d(x_1, x_N), \ldots, d(X_{N-1}, X_N)\}$. Then $\{x_n\} \subseteq B^d_{2r}(x_N)$. □

**Theorem 1.44.** *(a) $x \in A'$ if and only if there exists $\{x_n\} \subseteq A$ with $x_n \neq x$ for $n \in \mathbb{N}$ such that $x_n \xrightarrow{d} x \in X$ as $n \to \infty$.*

*(b) $x \in \overline{A}$ if and only if there exists $\{x_n\} \subseteq A$ such that $x_n \xrightarrow{d} x \in X$ as $n \to \infty$.*

*(c) $A$ is closed if and only if if $\{x_n\} \subseteq A$ and $x_n \xrightarrow{d} x$ as $n \to \infty$, then $x \in A$.*

*Proof.* (a) $x \in A'$ if and only if for $\epsilon > 0$, there exists $x \neq y \in A$ such that $y \in B^d_\epsilon(x)$ if and only if there exists $x \neq x_n \in A$ such that $x_n \in B^d_{1/n}(x)$ for $n \in \mathbb{N}$ if and only if there exists $\{x_n\} \subseteq A$ with $x_n \neq x$ for $n \in \mathbb{N}$ such that $x_n \xrightarrow{d} x$ as $n \to \infty$.

(b) It is similar to (1).

(c) $\Longrightarrow$ By (2), $x_n \xrightarrow{d} x \in \overline{A} = A$ as $n \to \infty$.

$\Longleftarrow$ Let $x \in \overline{A}$. Then by (2), there exists $\{x_n\} \subseteq A$ such that $x_n \xrightarrow{d} x$ as $n \to \infty$. So By assumption, $x \in A$. □

**Theorem 1.45.** *Let $f : (X, d_X) \to (Y, d_Y)$. Then $f$ is continuous at $x_0 \in X$ if and only if $(x_n) \xrightarrow{d_Y} f(x_0)$ whenever $x_n \xrightarrow{d_X} x_0$ as $n \to \infty$.*

*Proof.* $\Longrightarrow$ Let $f$ be continuous at $x_0 \in X$ and $\epsilon > 0$. Then there exists $\delta(x_0, \epsilon) > 0$ such that $B^{d_X}_\delta(x_0) \subseteq f^{-1}\left(B^{d_Y}_\epsilon(f(x_0))\right)$. Also, since $x_n \xrightarrow{d} x$ as $n \to \infty$, there exists $N \in \mathbb{N}$ such that $\{x_n\}_{n\geqslant N} \subseteq B^{d_X}_\delta(x_0) \subseteq f^{-1}\left(B^{d_Y}_\epsilon(f(x_0))\right)$. So $f\left(\{x_n\}_{n\geqslant N}\right) \subseteq ff^{-1}\left(B^{d_Y}_\epsilon(f(x_0))\right) \subseteq B^{d_Y}_\epsilon(f(x_0))$, i.e., $\{f(x_n)\}_{n\geqslant N} \subseteq B^{d_Y}_\epsilon(f(x_0))$. Thus, $f(x_n) \xrightarrow{d} f(x_0)$ as $n \to \infty$.

$\Longleftarrow$ Suppose $f$ is not continuous at $x_0$. Then there exists $\epsilon > 0$ such that for $\delta > 0$, there exists $x_\delta$ such that $x_\delta \in B_\delta^{d_X}(x_0)$ but $f(x_\delta) \notin B_\epsilon^{d_Y}(f(x_0))$. So there exists $\{x_n\} \subseteq X$ such that $x_n \in B_{1/n}^{d_X}(x_0)$ but $f(x_n) \notin B_\epsilon^{d_Y}(f(x_0))$ for $n \in \mathbb{N}$. Thus, we have $x_n \xrightarrow{d_X} x_0$ but $f(x_n) \overset{d_Y}{\nrightarrow} f(x_0)$ as $n \to \infty$, a contradiction. $\square$

## 1.5 Completeness

In real analysis we have encountered the notion of completeness when we prove that every Cauchy sequence in R converges. In this section we generalize this concept to metric spaces. We will see that unlike R, certain metric spaces are not complete. Roughly speaking, a space is complete if there are no elements "missing" from it. A simple example is that the rational numbers $\mathbb{Q}$ is not complete, because for instance, e is "missing" from it, even though we can construct a Cauchy sequence of rational numbers that converges to e (e.g. $\{x_n = (1+\frac{1}{n})^n\}$). Remarkably, it turns out that it is always possible to"fill all the holes" in an incomplete metric space, which leads to the completion of a given metric space.

**Assumption 1.46.** Let $(X, d)$ be a metric space.

**Definition 1.47.** $(X, d)$ is said to be *complete* if every Cauchy sequence in $(X, d)$ converges, i.e., there exists $x \in X$ such that $x_n \xrightarrow{d} x$ for $\{x_n\}$ Cauchy.

**Example 1.48.** $(\mathbb{R}, |\cdot|)$ is complete.

**Example 1.49.** $(\mathbb{R}^n, d_2)$ is complete for $n \in \mathbb{N}$.

*Proof.* Let $\{x^{(m)}\}$ be Cauchy in $(\mathbb{R}^n, d_2)$, then $d_2(x^{(m)}, x^{(k)}) \to 0$ as $m, k \to \infty$. Since $\left|x_i^{(m)} - x_i^{(k)}\right| \leqslant \left(\sum_{i=1}^n \left|x_i^{(m)} - x_i^{(k)}\right|^2\right)^{\frac{1}{2}}$, we have $\left|x_i^{(m)} - x_i^{(k)}\right| \to 0$ as $m, k \to \infty$ for each $i = 1, \ldots, n$, i.e., $\{x_i^{(m)}\}$ is Cauchy in $(\mathbb{R}, |\cdot|)$ for each $i = 1, \ldots, n$. Since $\mathbb{R}$ is complete, there exists $x_i \in \mathbb{R}$ such that $x_i^{(m)} \to x_i$ as $m \to \infty$ for each $i = 1, \ldots, n$. Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Then $d_2(x^{(m)}, x) = \left(\sum_{i=1}^n \left|x_i^{(m)} - x_i\right|^2\right)^{\frac{1}{2}} \to 0$ as $m \to \infty$, i.e., $\{x^{(m)}\}$ converges in $(\mathbb{R}^n, d_2)$. $\square$

**Example 1.50.** $(X, d_{\text{disc}})$ is complete.

*Proof.* Let $\{x_n\}$ be Cauchy in $(X, d_{\text{disc}})$. Then there exists $N \in \mathbb{N}$ such that $d_{\text{disc}}(x_n, x_m) < \frac{1}{2}$ for all $m, n \geqslant N$. So $x_m = x_n$ for $m, n \geqslant N$. Thus, $\{x_n\}$ converges in $X$. $\square$

**Example 1.51.** $(l^p, d_p)$ with $p \in \overline{\mathbb{R}}^{\geqslant 1}$ is complete.

*Proof.* (a) Assume $p = \infty$. Let $\{x^n\}$ be a Cauchy sequence in $(l^\infty, d_\infty)$ and $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $\left|x_i^{(n)} - x_i^{(m)}\right| \leqslant \sup_{i\in\mathbb{N}} \left|x_i^{(m)} - x_i^{(m)}\right| = d_\infty(x^{(n)}, x^{(m)}) < \epsilon$ for all $m, n \geqslant N$. Since $\epsilon > 0$ is arbitrary, $\{x_i^{(n)}\}$ is Cauchy in $(\mathbb{R}, |\cdot|)$ for each $i \in \mathbb{N}$. Since $(\mathbb{R}, |\cdot|)$ is complete, there exists $x_i \in \mathbb{R}$ such that $x_i^{(n)} \to x_i$ for $i \in \mathbb{N}$. Let $x = (x_1, x_2, \ldots,)$. Observe $\left|x_i^{(n)} - x_i\right| = \lim_{m\to\infty} \left|x_i^{(n)} - x_i^{(m)}\right| \leqslant \epsilon$ for $i \in \mathbb{N}$ and any $n \geqslant N$. So $d_\infty(x^{(n)}, x) = \sup_{i\in\mathbb{N}} \left|x_i^{(n)} - x_i\right| \leqslant \epsilon$ for

$n \geqslant N$. Since $\epsilon > 0$ is arbitrary, $d_\infty(x^{(n)}, x) \to 0$ as $n \to \infty$. Notice $|x_i| \leqslant \left|x_i - x_i^{(N)}\right| + \left|x_i^{(N)}\right| \leqslant \epsilon + \left|x_i^{(N)}\right|$ for $i \in \mathbb{N}$. So $x \in l^\infty$. Thus, $\{x^{(n)}\}$ converges in $(l^\infty, d_\infty)$.

(b) Assume $p \in \mathbb{R}^{\geqslant 1}$. Let $\{x^{(n)}\}$ be Cauchy in $(l^p, d_p)$ and $\epsilon > 0$. Then for each $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $\left|x_i^{(n)} - x_i^{(m)}\right| \leqslant \left(\sum_{i=1}^\infty \left|x_i^{(n)} - x_i^{(m)}\right|^p\right)^{\frac{1}{p}} = d_p(x^{(n)}, x^{(m)}) < \epsilon$ for all $m, n \geqslant N$ and each $i \in \mathbb{N}$. So $\{x_i^{(n)}\}$ is a Cauchy sequence for each $i \in \mathbb{N}$. Since $(\mathbb{R}, |\cdot|)$ is complete, there exists $x_i \in \mathbb{R}$ such that $x_i^{(n)} \to x_i$ for $i \in \mathbb{N}$. Let $x = (x_1, x_2, \cdots)$. Observe $\sum_{i=1}^k \left|x_i^{(n)} - x_i^{(m)}\right|^p < \epsilon^p$ for all $m, n \geqslant N$ each $k \in \mathbb{N}$. So $\sum_{i=1}^k \left|x_i^{(n)} - x_i\right|^p = \lim_{m\to\infty} \sum_{i=1}^k \left|x_i^{(n)} - x_i^{(m)}\right|^p \leqslant \epsilon^p$ for each $n \geqslant N$ and each $k \in \mathbb{N}$. Hence $\sum_{i=1}^\infty \left|x_i^{(n)} - x_i\right|^p \leqslant \epsilon^p$ for each $n \geqslant N$. It implies

- $\left(\sum_{i=1}^\infty \left|x_i^{(n)} - x_i\right|^p\right)^{\frac{1}{p}} \leqslant \epsilon$ for each $n \geqslant N$, i.e., $x^{(n)} \xrightarrow{d_p} x$ as $n \to \infty$ and
- $x^{(n)} - x \in l^p$. By the Minkowski inequality, $\|x\| = \left\|x^{(n)} + (x - x^{(n)})\right\|_p \leqslant \left\|x^{(n)}\right\|_p + \left\|x - x^{(n)}\right\|_p < \infty$, i.e., $x \in l^p$.

Thus, $\{x^{(n)}\}$ converges in $(l^p, d_p)$. □

**Example 1.52.** $(B[0,1], d_\infty)$ is complete.

*Proof.* Let $\{x_n\}$ be Cauchy in $(B[0,1], d_\infty)$ and $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $|x_n(t) - x_m(t)| \leqslant \sup_{t\in[0,1]} |x_n(t) - x_m(t)| = d_\infty(x_n, x_m) < \epsilon$ for all $m, n \geqslant N$ and each $t \in [0,1]$. So $\{x_n(t)\}$ is Cauchy in $(\mathbb{R}, |\cdot|)$ for each $t \in [0,1]$. Since $(\mathbb{R}, |\cdot|)$ is complete, there exists $x_t \in \mathbb{R}$ such that $x_n(t) \to x_t$ for each $t \in [0,1]$. Let $x : [0,1] \to \mathbb{R}$ given by $x(t) = \lim_{n\to\infty} x_n(t)$. Observe $|x_n(t) - x(t)| = \lim_{m\to\infty} |x_n(t) - x_m(t)| \leqslant \epsilon$ for $n \geqslant N$ and any $t \in [0,1]$. So $d_\infty(x_n, x) = \sup_{t\in[0,1]} |x_n(t) - x(t)| \leqslant \epsilon$ for $n \geqslant N$. So $x_n \xrightarrow{d_\infty} x$ as $n \to \infty$. Note $|x(t)| \leqslant |x(t) - x_N(t)| + |x_N(t)| \leqslant \epsilon + M_t$, for some $M_t \in \mathbb{R}$ given $x_N \in B[0,1]$, for $t \in [0,1]$. So $x \in B([0,1])$. Thus, $\{x^{(n)}\}$ converges in $(B[0,1], d_\infty)$. □

**Theorem 1.53.** *Let $A \subseteq X$.*

*(a) If $(A, d)$ is complete, then $A = \overline{A}$.*

*(b) If $A = \overline{A}$ and $(X, d)$ is complete, then $(A, d)$ is complete,*

*Proof.* (a) Let $x \in \overline{A} \subseteq X$. Then there exists $\{x_n\} \subseteq A \subseteq X$ such that $x_n \to x$ as $n \to \infty$. So $\{x_n\}$ converges in $(X, d)$ and then $\{x_n\}$ is Cauchy in $(X, d)$ and hence Cauchy in $(A, d)$. Since $(A, d)$ is complete, there exists $x_0 \in A$ such that $x_n \to x_0$ as $n \to \infty$. By the uniqueness of limit, $x = x_0 \in A$. So $\overline{A} \subseteq A$.

(b) Let $\{x_n\}$ be Cauchy in $(A, d)$. Then $\{x_n\}$ is Cauchy in $(X, d)$. Since $(X, d)$ is complete, there exists $x \in X$ such that $x_n \to x \in \overline{A}$ as $n \to \infty$. Since $A$ is closed, $x \in A$. Thus, $\{x_n\}$ converges in $A$. □

**Example 1.54.** $(l^p, d_q)$ with $1 \leqslant p < q \leqslant \infty$ is not complete.

*Proof.* Let $A = \{\{x_n\} \mid x_n \in \mathbb{Q}, x_n \neq 0 \text{ for finitely many } n \in \mathbb{N}\}$.

(a) Assume $q < \infty$. Recall $A$ is dense in $l^q$. Note $A \subsetneq l^p \subsetneq l^q$. So $l^q = \overline{A} \subseteq \overline{l^p} \subseteq \overline{l^q} = l^q$. Hence $\overline{l^p} = l^q \supsetneq l^p$ in $(l^\infty, d_\infty)$. So $(l^p, d_p)$ is not complete.

(b) Assume $q = \infty$. We know $A$ is dense in $c_0$. Note $A \subsetneq l^p \subsetneq c_0 \subsetneq l^\infty$. So $c_0 = \overline{A} \subseteq \overline{l^p} \subseteq \overline{c_0} = c_0$. Hence $\overline{l^p} = c_0 \subsetneq l^\infty$. So $l^p \subsetneq \overline{l^p}$ in $(l^\infty, d_\infty)$. Thus, $(l^p, d_\infty)$ is not complete. □

**Example 1.55.** (a) $C([0,1], d_\infty)$ is complete.

(b) $(C[0,1], d_p)$ with $p \in \mathbb{R}^{\geqslant 1}$ is not complete.

*Proof.* (a) Method 1. Since $C[0,1] \subseteq B[0,1]$ and $(B[0,1], d_\infty)$ is complete, it is enough to show $C[0,1]$ is closed in $(B[0,1], d_\infty)$. Let $\{x_n\} \subset C[0,1]$ and $x_n \xrightarrow{d_\infty} x$ as $n \to \infty$. Let $\{t_n\} \subseteq \mathbb{R}$ and $t_n \xrightarrow{|\cdot|} t$ as $n \to \infty$. Note $|x(t_n) - x(t)| \leqslant |x(t_n) - x_n(t_n)| + |x_n(t_n) - x_n(t)| + |x_n(t) - x(t)| \to 0$ as $n \to \infty$. So $x \in C[0,1]$. Thus, $C[0,1]$ is closed in $(B[0,1], d_\infty)$.

Method 2. Let $\{x_n\}$ be Cauchy in $(C[0,1], d_\infty)$ and $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $|x_n(t) - x_m(t)| \leqslant \max_{t\in[0,1]}, |x_n(t) - x_m(t)| = d_\infty(x_n, x_m) < \epsilon$ for all $m, n \geqslant N$ and each $t \in [0,1]$. So $\{x_n(t)\}$ is Cauchy in $(\mathbb{R}, |\cdot|)$ for each $t \in [0,1]$. Since $(\mathbb{R}, |\cdot|)$ is complete, there exists $x_t \in \mathbb{R}$ such that $x_n(t) \to x_t$ for each $t \in [0,1]$. Let $x : [0,1] \to \mathbb{R}$ given by $x(t) = \lim_{n\to\infty} x_n(t)$. Observe $|x_n(t) - x(t)| = \lim_{m\to\infty} |x_n(t) - x_m(t)| \leqslant \epsilon$ for $n \geqslant N$ and any $t \in [0,1]$. So $d_\infty(x_n, x) = \max_{t\in[0,1]} |x_n(t) - x(t)| \leqslant \epsilon$ for $n \geqslant N$. So $x_n$ converges to $x$ uniformly and hence $x \in C[0,1]$. Thus, $\{x^{(n)}\}$ converges in $(C[0,1], d_\infty)$.

(b) Let $x_n(t) = \begin{cases} 0 & \text{if } 0 \leqslant t \leqslant \frac{1}{2} \\ (n+1)(t - \frac{1}{2}) & \text{if } \frac{1}{2} < t < \frac{1}{2} + \frac{1}{n+1} \\ 1 & \text{if } \frac{1}{2} + \frac{1}{n+1} \leqslant t \leqslant 1 \end{cases}$ for $n \in \mathbb{N}$. Then $\{x_n\} \subseteq C[0,1]$. Let $m, n \in \mathbb{N}$ with $n \geqslant m$. Then $|x_n(t) - x_m(t)| = x_n(t) - x_m(t) \begin{cases} \leqslant 1 & \text{if } \frac{1}{2} \leqslant t \leqslant \frac{1}{2} + \frac{1}{m+1} \\ = 0 & \text{otherwise} \end{cases}$. So $d_p(x_n, x_m) = \left(\int_0^1 |x_n(t) - x_m(t)|^p\right)^{\frac{1}{p}} \leqslant \left(\frac{1}{m+1}\right)^{\frac{1}{p}} \to 0$ as $m \to 0$. Hence $\{x_n\}$ is Cauchy in $C[0,1]$. Suppose there is $x \in C[0,1]$ such that $d_p(x_n, x) \to 0$ as $n \to \infty$. Then

$$\left(\int_0^{\frac{1}{2}} |x_n(t) - x(t)|^p dt + \int_{\frac{1}{2}+\frac{1}{m+1}}^1 |x_n(t) - x(t)|^p dt\right)^{\frac{1}{p}} \leqslant \left(\int_0^1 |x_n(t) - x(t)|^p dt\right)^{\frac{1}{p}} = d_p(x_n, x).$$

Let $n \geqslant m \to 0$, then $\lim_{n\to\infty} \left(\int_0^{\frac{1}{2}} |x_n(t) - x(t)|^p dt\right)^{\frac{1}{p}} \leqslant 0 = \lim_{n\to\infty} d_p(x_n, x)$. So $x(t) = \begin{cases} 0 & \text{if } 0 \leqslant t \leqslant \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} < t \leqslant 1 \end{cases}$. Thus, $x \notin C[0,1]$, a contradiction. □

**Example 1.56.** $(P[0,1], d_\infty)$ is not complete.

*Proof.* By Weierstrass Approximation theorem, $P[0,1] \subseteq C[0,1]$ is dense. Then $\overline{P[0,1]} = C[0,1] \supsetneq P[0,1]$. Also, since $(C[0,1], d_\infty)$ is complete, we have $(P[0,1], d_\infty)$ is not complete. □

### 1.5.1 Completion of a Metric Space

**Definition 1.57.** Let $f : (X, d_X) \to (Y, d_Y)$.

(a) $f$ is said to be *isometric* or an *isometry* if $d_Y(f(x), f(y)) = d_X(x, y)$ for all $x, y \in X$, i.e., $f$ preserve distance.

(b) If $f$ is a bijective isometry, then $(X, d_X)$ and $(Y, dY)$ are called *isometric spaces.*

**Remark.** If $f$ is an isometry, then $f$ is 1-1 and continuous.

**Definition 1.58.** Let $(X, d)$ be a metric space. A complete metric space $(\widetilde{X}, \widetilde{d})$ is called a completion of $(X, d)$ if there exists an isometry $f : (X, d) \to (\widetilde{X}, \widetilde{d})$ such that $f(X)$ is dense in $(\widetilde{X}, \widetilde{d})$.

**Example 1.59.** Let $I$ be the identity mapping $I : \mathbb{Q} \hookrightarrow \mathbb{R}$. Since $\overline{I(\mathbb{Q})} = \overline{\mathbb{Q}} = \mathbb{R}$, it is a completion.

**Lemma 1.60.** (a) If $x_n \to x$, $y_n \to y$, then $d(x_n, y_n) \to d(x, y)$.

(b) If $\{x_n\}$ and $\{y_n\}$ are Cauchy, then there exists $r \in (\mathbb{R}, |\cdot|)$ such that $d(x_n, y_n) \to r$.

*Proof.* (a) $|d(x_n, y_n) - d(x, y)| \leqslant d(x_n, x) + d(y_n, y)$.

(b) $|d(x_n, y_n) - d(x_m, y_m)| \leqslant d(x_n, x_m) + d(y_n, y_m) \to 0$. □

**Theorem 1.61.** *Every metric space has a completion and all the completion are isometric spaces.*

*Proof.*

- Step 1. Construct $\widetilde{X}$. Let $C = \{\{x_n\} \mid \{x_n\} \text{ is Cauchy in } X\}$. Define "$\sim$" on $C$ by $\{x_n\} \sim \{x'_n\}$ if $\lim_{n\to\infty} d(x_n, x'_n) = 0$. Easy to check "$\sim$" is an equivalent relation. Let $\widetilde{X} = \{\,[x] \mid [x] \text{ is an equivalent class w.r.t "}\sim\text{"}\}$.

- Step 2. Define
$$\widetilde{d} : \widetilde{X} \times \widetilde{X} \to \mathbb{R}$$
$$([x], [y]) \mapsto \lim_{n\to\infty} d(x_n, y_n),$$
where $\{x_n\} \in [x]$ and $\{y_n\} \in [y]$. First, $\lim_{n\to\infty} d(x_n, y_n)$ exists by the above lemma. Let $\{x_n\}\ \{x'_n\} \subseteq [x]$, and $\{y_n\},\ \{y'_n\} \subseteq [y]$, then $|d(x_n, y_n) - d(x'_n, y'_n)| \leqslant d(x_n, x'_n) + d(y_n, y'_n) \to 0 + 0 = 0$. So $\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(x'_n, y'_n)$. Thus, $\widetilde{d}$ is well-defined. Easy to check $\widetilde{d}$ is a metric.

- Step 3. Define
$$f : X \to \widetilde{X}$$
$$x \mapsto [(x)],$$
where $[(x)]$ is the equivalent class that contains the constant (cauchy) sequence $(x) = \{x, x, \cdots\}$. Then for $x, y \in X$, $\widetilde{d}(f(x), f(y)) = \widetilde{d}([(x)], [(y)]) = \widetilde{d}((x), (y)) = \lim_{n\to\infty} d(x, y) = d(x, y)$. So $f$ is an isometry.

- Step 4. Let $\{x_n\} \in [x] \in \widetilde{X}$. For $\epsilon > 0$, since $\{x_n\}$ is Cauchy in $(X, d)$, there exists $N \in \mathbb{N}$ such that as $n \geqslant N$, $d(x_n, X_N) < \frac{\epsilon}{2}$. Then $\widetilde{d}([x], f(X_N)) = \widetilde{d}([x], [(x_N)]) = \lim_{n\to\infty} d(x_n, x_N) \leqslant \frac{\epsilon}{2} < \epsilon$. So $f(X)$ is dense in $(\widetilde{X}, d)$.

- Step 5. Let $\{[x]_n\}$ be Cauchy in $(\widetilde{X}, \widetilde{d})$. Since $f(X)$ is dense in $(\widetilde{X}, \widetilde{d})$, there exists $\{z_n\} \subseteq X$ such that $\widetilde{d}\,([x]_n, f(z_n)) = \widetilde{d}\,([x]_n, [(z_n)]) < \frac{1}{n}$ for $n \in \mathbb{N}$. Since $f$ is an isometry,

$$\begin{aligned} d(z_n, z_m) = \widetilde{d}\,(f(z_n), f(z_m)) &\leqslant \widetilde{d}\,(f(z_n), [x]_n) + \widetilde{d}\,([x]_n, [x]_m) + \widetilde{d}\,([x]_m, f(z_m)) \\ &< \frac{1}{n} + \widetilde{d}\,([x]_n, [x]_m) + \frac{1}{m} \to 0, \text{ as } n, m \to \infty. \end{aligned}$$

So $\{z_n\}$ is Cauchy in $(X, d)$. Let $[x]$ be the equivalent class that contains $\{z_n\}$. Note

$$\begin{aligned} \widetilde{d}\,([x]_n, [x]) &\leqslant \widetilde{d}\,([x]_n, [(z_n)]) + \widetilde{d}\,([(z_n)], [x]) \\ &\leqslant \frac{1}{n} + \lim_{m\to\infty} d(z_n, z_m), \forall n \in \mathbb{N}. \end{aligned}$$

So $[x]_n \xrightarrow{\widetilde{d}} [x]$. Thus, $(\widetilde{X}, \widetilde{d})$ is complete.

- Let $(\hat{X}, \hat{d})$ be another completion of $(X, d)$, then there exists an isometry $g : X \to \hat{X}$ such that $g(X)$ is dense in $(\hat{X}, \hat{d})$.

$$\begin{array}{ccc} X & \xrightarrow{g} & \hat{X} \\ & \searrow^{f} & \downarrow{\scriptstyle f\circ g^{-1}} \\ & & \widetilde{X} \end{array}$$

For $\hat{x},\ \hat{y} \in \hat{X}$, since $g(X)$ is dense in $(\hat{X}, \hat{d})$, there exist $\{x_n\},\ \{y_n\} \subseteq X$ such that $g(x_n) \to \hat{x}$ and $g(y_n) \to \hat{y}$. Since $g$ is an isometry (then 1-1), $g^{-1} : (g(X), \hat{d}) \to (X, d)$ is also an isometry, so $g^{-1}$ is continuous. Thus, we get

$$\begin{aligned} &\widetilde{d}\left((f\circ g^{-1})(\hat{x}), (f\circ g^{-1})(\hat{y})\right) = \widetilde{d}\left((f\circ g^{-1})\left(\lim_{n\to\infty} g(x_n)\right), (f\circ g^{-1})\left(\lim_{n\to\infty} g(y_n)\right)\right) \\ &= \widetilde{d}\left(f\left(\lim_{n\to\infty} x_n\right), f\left(\lim_{n\to\infty} y_n\right)\right) = \widetilde{d}\left(\lim_{n\to\infty} f(x_n), \lim_{n\to\infty} f(y_n)\right) = \lim_{n\to\infty} \widetilde{d}(f(x_n), f(y_n)) \\ &= \lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} \hat{d}(g(x_n), g(y_n)) = \hat{d}\left(\lim_{n\to\infty} g(x_n), \lim_{n\to\infty} g(y_n)\right) = \hat{d}(\hat{x}, \hat{y}). \end{aligned}$$

So $f\circ g^{-1}$ is an isometry. Let $[x] \in \widetilde{X}$. Since $f(X)$ is dense in $(\widetilde{X}, \widetilde{d})$, there exists $\{x_n\} \subseteq X$ such that $\lim_{n\to\infty} f(x_n) = [x]$, w.r.t. $\widetilde{d}$. So $f(x_n)$ is Cauchy in $(\widetilde{X}, \widetilde{d})$. Since $f$ is an isometry, $\{x_n\}$ is Cauchy in $(X, d)$. Since $g$ is an isometry, $\{g(x_n)\}$ is Cauchy in $(\hat{X}, \hat{d})$. Since $(\hat{X}, \hat{d})$ is complete, $\lim_{n\to\infty} g(x_n)$ exists w.r.t. $\hat{d}$. So $(f\circ g^{-1})\left(\lim_{n\to\infty} g(x_n)\right) = \lim_{n\to\infty}(f\circ g^{-1})g(x_n) = \lim_{n\to\infty} f(x_n) = [x]$. Hence $f\circ g^{-1}$ is onto. Thus, $f\circ g^{-1} : (\widetilde{X}, \widetilde{d}) \to (\hat{X}, \hat{d})$ is a bijective isometry, as desired. □

## 1.6 Application of completeness

Let $(X, d)$ be a metric space.

**Definition 1.62.** Let $f : X \to X$.

(a) $x \in X$ is called a *fixed* point of $f$ if $f(x) = x$.

(b) $f$ is called a *contradiction* on $X$ if there exists $\alpha \in (0,1)$ such that $d(f(x), f(y)) \leqslant \alpha d(x,y)$ for $x, y \in X$.

**Remark.** If $f$ is a contration, then so is $f^n$ for $n \in \mathbb{N}$. The converse is not true. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = \begin{cases} -x & \text{if } x \geqslant 0 \\ \frac{x}{2} & \text{if } x < 0 \end{cases}$. Since $|f(x) - f(y)| = |x - y|$, $f$ is not a contraction. On the other hand, $f^2(x) = f(f(x)) = \begin{cases} -\frac{x}{2} & \text{if } x \geqslant 0 \\ \frac{x}{4} & \text{if } x < 0 \end{cases}$. So for $x, y \in \mathbb{R}$, we have $\left|f^x(x) - f^2(y)\right| \leqslant \frac{1}{2}|x - y|$, i.e., $f^2$ is a contraction on $\mathbb{R}$.

**Theorem 1.63** (Banach Fixed Point Theorem)**.** *Let* $(X, d)$ *be nonempty and complete, and* $f$ *be a contraction on* $X$*. Then* $f$ *has a unique fixed point.*

*Proof.* Choose $x_0 \in X$. Define $x_n = f(x_{n-1}) = f^2(x_{n-2}) = f^{n-1}(x_1) = f^n(x_0)$ for $n \in \mathbb{N}$. Then $\{x_n\} \subseteq X$. Since $f$ is a contraction, there exists $0 < \alpha < 1$ such that $d(f(x), f(y)) \leqslant d(x,y)$ for $x, y \in X$. Note $d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leqslant \alpha d(x_n, x_{n-1}) \leqslant \cdots \leqslant \alpha^n d(x_1, x_0)$. Then for $n > m$, $d(x_n, x_m) \leqslant d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \cdots + d(x_{n-1}, x_n) \leqslant (\alpha^m + \alpha^{m+1} + \cdots + \alpha^{n-1}) d(x_1, x_0) \leqslant \alpha^m \frac{1-\alpha^{n-m}}{1-\alpha} d(x, x_0) \leqslant \frac{\alpha^m}{1-\alpha} d(x, x_0) \to 0$ as $m \to \infty$. So $\{x_n\}$ is Cauchy in $(X, d)$. Since $(X, d)$ is complete, there exists $x \in X$ such that $x_n \xrightarrow{d} x$ as $n \to \infty$. Let $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \frac{\epsilon}{2}$ for $n \geqslant N$. So $d(x, f(x)) \leqslant d(x, x_{N+1}) + d(x_{N+1}, f(x)) < \frac{\epsilon}{2} + d(f(X_N), f(x)) \leqslant \frac{\epsilon}{2} + \alpha d(x_N, x) \leqslant \frac{\epsilon}{2} + \alpha \frac{\epsilon}{2} < \epsilon$. Since $\epsilon > 0$ is arbitrary, $d(x, f(x)) = 0$. So $f(x) = x$. Suppose there exist two $x_1, x_2 \in X$ such that $f(x_1) = x_1$ and $f(x_2) = x_2$. Then $d(x_1, x_2) = d(f(x_1), f(x_2)) \leqslant \alpha d(x_1, x_2)$. So $d(x_1, x_2) = 0$ and thus $x_1 = x_2$. ☐

**Remark.** In the above theorem, if there is $n \in \mathbb{N}$ such that $f^n$ is a contraction, then $f$ has a unique fixed point. Proof. Since $f^n$ is a contraction, it has a unique fixed point $x \in X$. Then $f^n(x) = x$ and so $f^n(f(x)) = f^{n+1}(x) = f(f^n(x)) = f(x)$, i.e., $f(x)$ is a fixed point of $f^n$. By the uniqueness, $f(x) = x$. Suppose there are two $x_1, x_2 \in X$ such that $f(x_1) = x_1$ and $f(x_2) = x_2$. Then $f^n(x_1) = f^{n-1}(f(x_1)) = f^{n-1}(x_1) = \cdots = f(x_1) = x_1$ and similarly, $f^n(x_2) = x_2$, i.e., $x_1$ and $x_2$ are both a fixed point of $f^n$. By the uniqueness, $x_1 = x_2$.

**Example 1.64.** Consider the integral equation $x(t) - \lambda \int_0^1 e^{t-s} x(s) ds = y(t)$, where $y \in C[0,1]$ and $|\lambda| < 1$. Then the equation has a unique solution $x \in [0,1]$. Rewrite the equation as $e^{-t}x(t) - \lambda \int_0^1 e^{-s} x(s) ds = e^{-t} y(t)$. Let $z(t) = e^{-t} x(t)$ and $w(t) = e^{-t} y(t)$. Then $z(t) - \lambda \int_0^1 z(s) ds = w(t)$. Define $f : C[0,1] \to C[0,1]$ by $f(z) = w + \lambda \int_0^1 z(s) ds$. Let $z_1, z_2 \in [0,1]$. Note

$$\begin{aligned} d_\infty(f(z_1), f(z_2)) &= \max_{t \in [0,1]} \left| w(t) + \lambda \int_0^1 z_1(s) ds - w(t) - \lambda \int_0^1 z_2(s) ds \right| \\ &= \max_{t \in [0,1]} \left| \lambda \int_0^1 z_1(s) ds - \lambda \int_0^1 z_2(s) ds \right| = |\lambda| \left| \int_0^1 (z_1(s) - z_2(s)) ds \right| \\ &\leqslant |\lambda| \int_0^1 |z_1(s) - z_2(s)| ds \leqslant |\lambda| \int_0^1 d_\infty(z_1, z_2) ds = |\lambda| d_\infty(z_1, z_2). \end{aligned}$$

**Definition 1.65.** Let $A \subseteq X$. $A$ is called *nowhere dense* in $X$ if $\operatorname{Int}(\overline{A}) = \emptyset$.

**Example 1.66.** A single point $\{x\} \subseteq \mathbb{R}$ is nowhere dense in $(\mathbb{R}, |\cdot|)$.

**Example 1.67.**

$$A = \left\{ f \in C[0,1] \mid \exists\ x_0 \in [0,1] \text{ and } M > 0 \text{ s.t. } \left| \frac{f(x) - f(x_0)}{x - x_0} \right| \leqslant M, \forall x \in [0,1] \right\}$$

is nowhere dense in $C[0,1]$.

**Example 1.68.** The cantor set is nowhere dense in $(\mathbb{R}, |\cdot|)$.

**Theorem 1.69** (Baire Catogory Theorem)**.** *Let $(X,d)$ be complete. Then $X$ cannot be written as a countable union of nowhere dense sets.*

*Proof.* Suppose not, then $x = \bigcup_{n=1}^{\infty} A_n$ with $A_n$ nowhere dense in $X$. Since $A_1$ is nowhere dense, $\overline{A}_1 \neq \emptyset$ and then $\overline{A}_1^c \neq \emptyset$ and open. Pick $x_1$ and $B_1 := B_{\epsilon_1}^d(x_1) \subseteq \overline{A}_1^c$ with $\epsilon_1 < \frac{1}{2}$. Since $A_2$ is nowhere dense, $\overline{A}_2 \not\supseteq B_{\frac{\epsilon_1}{2}}^d(x_1)$ and then $\overline{A}_2^c \cap B_{\frac{\epsilon_1}{2}}^d(x_1) \neq \emptyset$ and open. Pick $x_2$ and $B_2 := B_{\epsilon_2}^d(x_2) \subseteq \overline{A}_2^c \cap B_{\frac{\epsilon_1}{2}}^d(x_1)$ open with $\epsilon_2 < \frac{\epsilon_1}{2} < \frac{1}{2^2}$. Since $A_3$ is nowhere dense, $\overline{A}_3 \not\supseteq B_{\frac{\epsilon_2}{2}}^d(x_2)$ and then $\overline{A}_3^c \cap B_{\frac{\epsilon_2}{2}}^d(x_2) \neq \emptyset$ and open. Repeat this, we get a sequence of open balls $B_n$'s that satisfies $B_n \supseteq B_{\frac{\epsilon_n}{2}}^d(x_n) \supseteq B_{n+1}$ with $\epsilon_n < \frac{1}{2^n}$ and $B_n \cap A_n = \emptyset$. Let $n, m \in \mathbb{N}$ with $n > m$. Note $d(x_n, x_m) \leqslant d(x_m, x_{m+1}) + \cdots + d(x_{n-1}, x_n) < \frac{\epsilon_m}{2} + \frac{\epsilon_{m+1}}{2} + \cdots + \frac{\epsilon_{n-1}}{2} < \frac{1}{2}\left(\frac{1}{2^m} + \frac{1}{2^{m+1}} + \cdots + \frac{1}{2^{n-1}}\right) = \frac{1}{2}\frac{1}{2^m}\frac{1-\left(\frac{1}{2}\right)^{n-m}}{1-\frac{1}{2}} < \frac{1}{2^m} \to 0$ as $m \to 0$. So $\{x_n\}$ is Cauchy. Since $(X,d)$ is complete, there exists $x \in X$ such that $x_n \to x$ as $n \to \infty$. Also, since $B_n \subseteq B_m$, $d(x_n, x_m)$. Letting $n$ be such that $d(x_n, x) < \frac{\epsilon_m}{2}$, we have $d(x, x_m) \leqslant d(x, x_n) + d(x_n, x_m) < \frac{\epsilon_m}{2} + \frac{\epsilon_m}{2} < \epsilon_m$. So $x \in B_m$ for $m \in \mathbb{N}$. Then $x \notin A_m$ for $m \in \mathbb{N}$ and so $x \notin \bigcup_{n=1}^{\infty} A_n = X$, a contradiction. □

**Example 1.70.** $[0,1]$ is not countable. Suppose $[0,1]$ is countable, then $[0,1] = \bigcup_{x \in [0,1]} \{x\}$ which is a countable union of nowhere dense sets w.r.t. $|\cdot|$, contradicted with Baire Category theorem.

**Example 1.71.** There exists $f \in C[0,1]$ such that $f$ is not differentiable at every $x \in [0,1]$.

*Proof.* Suppose not. Let $f \in C[0,1]$. Then there exists a point $x_0 \in [0,1]$ such that $f$ is differentiable at $x_0$. So there is $n \in \mathbb{N}$ such that $|f'(x_0)| \leqslant n - 1$, i.e., $\lim_{x \to x_0} \left| \frac{f(x) - f(x_0)}{x - x_0} \right| \leqslant n - 1$. Then there exists $\delta > 0$ such that $\left| \frac{f(x) - f(x_0)}{x - x_0} \right| \leqslant n$ whenever $0 < |x - x_0| < \delta$. Choose $m \in \mathbb{N}$ such that $\frac{1}{m} < \delta$. So $\left| \frac{f(x) - f(x_0)}{x - x_0} \right| \leqslant n$ whenever $0 < |x - x_0| < \frac{1}{m}$. Then $f \in A_{n \times m} := \{f \in C[0,1], \frac{|f(x) - f(x_0)|}{|x - x_0|} \leqslant n, 0 < |x - x_0| < \frac{1}{m}, n, m \in \mathbb{N}\}$. So $C[0,1] = \bigcup_{n,m \in \mathbb{N}} A_{n \times m}$, where $A_{n \times m}$ is nowhere dense in $C[0,1]$ for $n, m \in \mathbb{N}$, contradicted with the Baire Category theorem. □

**Example 1.72.** If $(V, \|\cdot\|)$ is Banach over any field $k$ and $V_i \subsetneq V$ is (topological) closed for $i \in \mathbb{N}$, then $\bigcup_{i=1}^{\infty} V_i \subsetneq V$. Suppose not. Since the proper subspaces $\{V_1, V_2, \ldots,\}$ of the vector space $V$ all have an empty interior and are closed, by Baire category theorem, $V$ is not complete, a contradiction. For instance, let $k$ be a field, note $k[x] = \bigcup_{i=1}^{\infty} V_i$, where $V_i = \{\deg(f) = i \mid f \in k[x]\}$ with $\dim_k(V_i) = i$. So $V_i$ is finite dimensional, $V_i$ is closed for each $i \in \mathbb{N}$. Thus, $k[x]$ can not be equipped with a complete norm.

# Chapter 2

# Normed Linear Spaces

## 2.1 Definitions and Examples

**Definition 2.1.** $X$ is called a *vector space* (*linear space*) $(X, +, \cdot)$ over a scalar field $\mathbb{K}$ ($\mathbb{R}$, $\mathbb{C}$, etc) if there exist two algebraic operations

$$\begin{aligned} \text{“}+\text{”} : X \times X &\to X \\ (x, y) &\mapsto x + y \end{aligned} \qquad\qquad \begin{aligned} \text{“}\cdot\text{”} : \mathbb{K} \times X &\to X \\ (a, x) &\mapsto ax \end{aligned}$$

that satisfies for all $a, b \in \mathbb{K}$ and any $x, y \in X$,

(a) $(X, +)$ is an abelian group with identity $\mathbf{0}$,

(b) $a \cdot (b \cdot x) = (ab) \cdot x$,

(c) $a \cdot (x + y) = a \cdot x + a \cdot y$,

(d) $(a + b) \cdot x = a \cdot x + b \cdot x$,

(e) $1_{\mathbb{K}} \cdot x = x$, where $1_{\mathbb{K}}$ is the multiplicative identity of $\mathbb{K}$.

**Remark.** Notice that there is no definition of the “product” of two elements in $X$. We typically choose $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$.

**Assumption 2.2.** Let $(X, +, \cdot)$ be a $\mathbb{K}$-vector space and $A \subseteq X$.

**Definition 2.3.** (a) $Y \subseteq X$ is called a *subspace*, denoted by $Y \leqslant X$ if for $a_1, a_2 \in \mathbb{K}$ and any $y_1, y_2 \in Y$, $a_1 y_1 + a_2 y_2 \in Y$.

(b) The *span* of $A$, denoted by $\operatorname{span}\{A\}$ or $\langle A \rangle$, is

$$\langle A \rangle = \left\{ \sum_{i=1}^{\text{finite}} a_i x_i \;\middle|\; a_i \in \mathbb{K}, x_i \in A \right\}.$$

(c) $A$ is *linearly independent*, if $\sum_{i=1}^{n} a_i x_i = 0$ with $n \in \mathbb{N}$, $a_i \in \mathbb{K}$ and $x_i \in A$ for $i = 1, \ldots, n$, then $a_i = 0$ for $i = 1, \ldots, n$.

(d) $A \subseteq X$ is a *Hamel basis* of $X$ if $A$ is linearly independent and $\langle A \rangle = X$.

(e) The *dimension* of $X$ is defined as $\dim X = |A|$, where $A$ is a Hamel basis of $X$.

A couple of natural questions one may ask are that whether a vector space always has a Hamel basis, and whether all Hamel bases of the same vector space have the same number of elements (so that the dimension of $X$ is well-defined). The answers to both questions are affirmative. However, their proofs require more tools (e.g., Zorn's Lemma) from the set theory so we omit them and only state the results here.

**Theorem 2.4.** *(a) Every nonempty vector space has a Hamel basis.*

*(b) If $A_1$ and $A_2$ are Hamel bases of $X$, then $|A_1| = |A_2|$.*

**Definition 2.5.** (a) $(X, d, +, \cdot)$ is called a *metric linear space* if

(1) $(X, d)$ is a metric space,

(2) $(X, +, \cdot)$ is a vector space,

(3) $+$ and $\cdot$ is continuous.

(b) $(X, d, +, \cdot)$ is a *translation and scaling invariant metric linear space* if

(1) $(X, d)$ is a metric space,

(2) $(X, +, \cdot)$ vector space,

(3) $d(x+z, y+z) = d(x, y)$ and $d(ax, ay) = |a|d(x, y)$ for all $a \in \mathbb{R}$ and $x, y \in X$.

(c) $(X, \|\cdot\|, +, \cdot)$ is a *normed linear space* (NLS)

(1) $(X, +, \cdot)$ is a vector space,

(2) the "norm" $\|\cdot\| : X \to \mathbb{K}$ satisfies for all $a \in \mathbb{K}$ and $x, y \in X$,

i. $\|x\| \geqslant 0$,

ii. $\|x\| = 0$ if and only if $x = \mathbf{0}$,

iii. $\|ax\| = |a|\|x\|$,

iv. $\|x + y\| \leqslant \|x\| + \|y\|$.

**Remark.** Condition i is redundant. By iii and iv, $0 = \|\mathbf{0}\| = \|x - x\| \leqslant \|x\| + \|-x\| = 2\|x\|$, i.e., $\|x\| \geqslant 0$.

**Theorem 2.6.** *If $X$ is a translation and scaling invariant metric linear space, then $X$ is a metric linear space.*

*Proof.* It is enough to show $+$ and $\cdot$ are continuous. Note $+ : X \times X \to X$ and $d_{X\times X} : (X \times X) \times (X \times X) \to \mathbb{K}$ given by $d_{X\times X}((x, y), (z, w)) = d(x, z) + d(y, w)$. Since $d(+(x, y), +(z, w)) = d(x + y, z + w) \leqslant d(x + y, y + z) + d(y + z, z + w) = d(x, z) + d(y, w) = d_{X\times X}((x, y), (z, w))$, we have $+$ is Lipschitz continuous. Note $\cdot : \mathbb{K} \times X \to X$ and $d_{\mathbb{K}\times X} : (\mathbb{K} \times X) \times (\mathbb{K} \times X) \to \mathbb{K}$ give by $d_{\mathbb{K}\times X}((a, x), (b, y)) = |a - b| + d(x, y)$. Let $(a_n, x_n) \to (a, x)$ as $n \to \infty$ in $\mathbb{K} \times X$. Then $|a_n - a| + d(x_n, x) \to 0$ as $n \to \infty$. So $a_n \to a$ and $d(x_n, x) \to 0$ as $n \to \infty$. Hence $d(\cdot(a_n, x_n), \cdot(a, x)) = d(a_n{\cdot}x_n, a{\cdot}x) \leqslant d(a_n{\cdot}x_n, a_n{\cdot}x) + d(a_n{\cdot}x, a{\cdot}x) = |a_n|d(x_n, x) + d(a_n{\cdot}x - ax, \mathbf{0}) = |a_n|d(x_n, x) + |a_n - a|d(x, \mathbf{0}) \to 0$ as $n \to \infty$. Thus, $\cdot$ is continuous. $\square$

**Theorem 2.7.** *$X$ is a translation and scaling invariant metric space if and only if $X$ is a NLS.*

*Proof.* $\Longrightarrow$ Let $(X, d, +, \cdot)$ be a translation and scaling invariant metric linear space. Define $\|\cdot\| : X \to \mathbb{K}$ by $\|x\| = d(x, \mathbf{0})$. Then for all $a \in \mathbb{K}$ and $x, y \in X$,

(a) $\|x\| = 0$ if and only if $d(x, \mathbf{0}) = 0$ if and only if $x = \mathbf{0}$,

(b) $\|a \cdot x\| = d(a \cdot x, \mathbf{0}) = d(a \cdot x, a \cdot \mathbf{0}) = |a| d(x, \mathbf{0}) = |a|\|x\|$,

(c) $\|x + y\| = d(x + y, \mathbf{0}) \leqslant d(x + y, y) + d(y, \mathbf{0}) = d(x, \mathbf{0}) + d(y, \mathbf{0}) = \|x\| + \|y\|$.

So $(X, \|\cdot\|, +, \cdot)$ is a NLS.

$\Longleftarrow$ Let $(X, \|\cdot\|, +, \cdot)$ be a NLS. Define $d : X \times X \to \mathbb{K}$ by $d(x, y) = \|x - y\|$. Then

(a) $d(x, y) = 0$ if and only if $\|x - y\| = 0$ if and only if $x - y = \mathbf{0}$ if and only if $x = y$,

(b) $d(x, y) = \|x - y\| = \|-(y - x)\| = \|y - x\| = d(y, x)$,

(c) $d(x, y) = \|x - y\| = \|(x - z) + (z - y)\| \leqslant \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$.

Also, note $d(x + z, y + z) = \|(x + z) - (y + z)\| = \|x - y\| = d(x, y)$ and $d(ax, ay) = \|ax - ay\| = |a|\|x - y\|$. So $(X, d, +, \cdot)$ is a translation and scaling invariant metric linear space. □

**Remark.** Let $(X, \|\cdot\|, +, \cdot)$ be a NLS. Then $\|\cdot\| : X \to \mathbb{R}$ is Lipschitz continuous.

*Proof.* Note $|\|x\| - \|y\|| = |d(x, \mathbf{0}) - d(y, \mathbf{0})| \leqslant d(x, y) + d(\mathbf{0}, \mathbf{0}) = \|x - y\|$. □

**Example 2.8.** $(X, d_{\text{disc}}, +, \cdot)$ is not a metric linear space.

*Proof.* Note $(X, d_{\text{disc}})$ is a metric space and an $\mathbb{K}$-vector space. Note $+ : X \times X \to X$. Let $(x_n, y_n) \to (x, y)$ as $n \to \infty$ in $(X \times X, d_{X\times X})$. Then $x_n \xrightarrow{d_{\text{disc}}} x$ and $y_n \xrightarrow{d_{\text{disc}}} y$ as $n \to \infty$. So there exists $n \in \mathbb{N}$ such that $x_n = x$ and $y_n = y$ for $n \geqslant N$. Then $x_n + y_n = x + y$ for $n \geqslant N$. Hence $+(x_n, y_n) = x_n + y_n \xrightarrow{\text{disc}} x + y = +(x, y)$ as $n \to \infty$. So $+$ is continuous. Note $\cdot : \mathbb{K} \times X \to X$. Let $(\frac{1}{n}, x_n) \to (0, x)$ as $n \to \infty$ in $(\mathbb{K} \times X, d_{\mathbb{K}\times X})$. But $\cdot(\frac{1}{n}, x_n) = \frac{1}{n} \cdot x_n \overset{d_{\text{disc}}}{\nrightarrow} \mathbf{0} = 0 \cdot x = \cdot(0, x)$. □

**Example 2.9.** $(S, d_S, +, \cdot)$ is a MLS but not a NLS.

*Proof.* Note $(S, d_S)$ is a metric space and clearly an $\mathbb{K}$-vector space. Note $+ : S \times S \to S$. Let $(x^{(n)}, y^{(n)}) \to (x, y)$ as $n \to \infty$ in $(S \times S, d_{S\times S})$. Then $x^{(n)} \xrightarrow{d_S} x$ and $y^{(n)} \xrightarrow{d_S} y$ as $n \to \infty$. So $x_i^{(n)} \xrightarrow{|\cdot|} x_i$ and $y_i^{(n)} \xrightarrow{|\cdot|} y_i$ as $n \to \infty$ for $i \in \mathbb{N}$. Hence $x_i^{(n)} + y_i^{(n)} \xrightarrow{|\cdot|} x_i + y_i$ as $n \to \infty$ for $i \in \mathbb{N}$. So $+(x^{(n)}, y^{(n)}) = x^{(n)} + y^{(n)} \xrightarrow{d_S} x + y = +(x, y)$ as $n \to \infty$. Thus, $+$ is contniuous. Note $\cdot : \mathbb{K} \times S \to S$. Let $(a_n, x^{(n)}) \to (a, x)$ as $n \to \infty$ in $(\mathbb{K} \times S, d_{\mathbb{K}\times S})$. Then $a_n \xrightarrow{|\cdot|} a$ and $x^{(n)} \xrightarrow{d_S} x$ as $n \to \infty$. So $x_i^{(n)} \xrightarrow{\cdot} x_i$ as $n \to \infty$ for $i \in \mathbb{N}$. Hence $a_n \cdot x_i^{(n)} \xrightarrow{|\cdot|} a \cdot x_i$ as $n \to \infty$ for $i \in \mathbb{N}$. So $\cdot(a_n, x^{(n)}) = a_n \cdot x^{(n)} \xrightarrow{d_S} a \cdot x = \cdot(a, x)$ as $n \to \infty$. Thus, $\cdot$ is continuous. Easy to see $(S, d_S, +, \cdot)$ is translation invariant, but it is not scaling invariant. □

**Example 2.10.** $(l^p, d_p, +, \cdot)$ for $p \in \overline{\mathbb{R}}^{\geqslant 1}$, $(C[0,1], d_p, +, \cdot)$ for $p \in \overline{\mathbb{R}}^{\geqslant 1}$ and $(B[0,1], d_\infty)$ are translation and scaling invariant metric linear spaces and hence NLSs, with

$$\|x\|_p = \begin{cases} \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{\frac{1}{p}}, & \text{if } x \in l^p, \\ \sup_{n \in \mathbb{N}} |x_n| & \text{if } x \in l^\infty, \\ \left(\int_0^1 |x(t)|^p dt\right)^{\frac{1}{p}} & \text{if } x \in C[0,1], \\ \sup_{t \in [0,1]} x(t) & \text{if } x \in C[0,1] \text{ or } B[0,1]. \end{cases}$$

**Definition 2.11.** A *Banach space* is a complete NLS, i.e.,

(a) $(X, \|\cdot\|)$ is a NLS.

(b) $X$ is complete w.r.t. $d : X \times X \to \mathbb{R}$ given by $d(x,y) = \|x - y\|$.

**Example 2.12.** $(l^p, d_p, +, \cdot)$ for $p \in \overline{\mathbb{R}}^{\geqslant 1}$, $(C[0,1], d_\infty, +, \cdot)$ and $(B[0,1], d_\infty, +, \cdot)$ are Banach spaces.

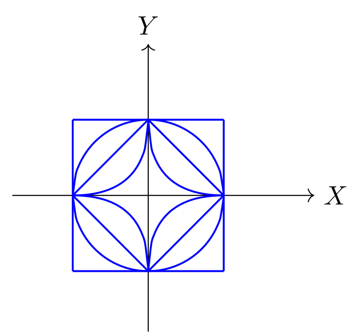
**Definition 2.13.** Let $X$ be a vector space and $A \subseteq X$. $A$ is said to be *convex* if for $x, y \in A$ and any $\lambda[0,1]$, $\lambda x + (1-\lambda)y \in A$.

**Theorem 2.14.** *Let $(X, \|\cdot\|)$ be a NLS. Then $B_\epsilon(\mathbf{0})$ and $\overline{B_\epsilon}(\mathbf{0})$ are convex for $\epsilon > 0$.*

*Proof.* Let $x, y \in B_\epsilon(\mathbf{0}) = \{z \in X \mid \|z\| < \epsilon\}$ and $\lambda \in [0,1]$. Then $\|x\|, \|y\| < \epsilon$ and so $\|\lambda x + (1-\lambda)y\| \leqslant \|\lambda x\| + \|(1-\lambda)y\| \leqslant |\lambda|\|x\| + |1-\lambda|\|y\| < \lambda\epsilon + (1-\lambda)\epsilon = \epsilon$. □

**Example 2.15** (Unit desks in $(\mathbb{R}^2, \|\cdot\|_p)$)**.** Note

$$\overline{B_1}(\mathbf{0}) = \{(x,y) \mid \|(x,y)\|_p \leqslant 1\} = \left\{(x,y) \;\middle|\; \begin{cases} x^p + y^p \leqslant 1 & \text{if } 1 \leqslant p < \infty, \\ \max\{|x|, |y|\} \leqslant 1 & \text{if } p = \infty. \end{cases}\right\}$$




(a) When $p = \infty$, it is a square,

(b) When $p = 2$, it is a circle,

(c) When $p = 1$, it is a tiltable square

(d) When $0 < p < 1$, the unit desk is concave, so $(\mathbb{R}^2, \|\cdot\|_p)$ is not a NLS, which we have showed it is not a metric space.

## 2.2 Sequence series and Schander basis

**Assumption 2.16.** Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be NLSs.

**Definition 2.17.** Let $\{x_n\} \subseteq X$.

(a) $\{x_n\}$ is convergent if there exists $x \in X$ such that $\|x_n - x\| \to 0$ as $n \to \infty$. Denote is as $\lim_{n\to\infty} x_n = x$ or $x_n \xrightarrow{\|\cdot\|} x$ as $n \to \infty$.

(b) $\{x_n\}$ is Cauchy if for $\epsilon > 0$, thre exists $N \in \mathbb{N}$ such that $\|x_n - x_m\| < \epsilon$ whenever $n \geqslant N$, i.e., $\|x_n - x_m\| \to 0$ as $n, m \to \infty$.

(c) The series $\sum_{n=1}^{\infty} x_n$ is *convergent* if there exists $x \in X$ such that $\left\|\sum_{i=1}^{N} x_i - x\right\| \to 0,$ as $N \to \infty$.

(d) The series $\sum_{n=1}^{\infty}$ is absolutely convergent if $\sum_{n=1}^{\infty}\|x_n\| < \infty$.

**Theorem 2.18.** $(X, \|\cdot\|)$ *is Banach if and only if every absolutely convergent series in* $(X, \|\cdot\|)$ *converges.*

*Proof.* $\Longrightarrow$ Let $\{x_n\} \subseteq X$ such that $\sum_{n=1}^{\infty}\|x_n\| < \infty$. Let $S_N = \sum_{n=1}^{N} x_n$ for $N \in \mathbb{N}$. Let $N, M \in \mathbb{N}$ with $N > M$. Since $\sum_{n=1}^{\infty}\|x_n\| < \infty$, $\|S_N - S_M\| = \left\|\sum_{n=M+1}^{N} x_n\right\| \leqslant \sum_{n=M+1}^{N}\|x_n\| \leqslant \sum_{n=M+1}^{\infty}\|x_n\| \to 0$ as $M \to \infty$. So $\{S_N\}$ is Cauchy in $(X, \|\cdot\|)$. Since $(X, \|\cdot\|)$ is Banach, there exists $x \in X$ such that $\sum_{n=1}^{\infty} x_n = \lim_{N\to\infty} \sum_{n=1}^{N} x_n = \lim_{N\to\infty} S_N = x$, i.e., $\sum_{n=1}^{\infty}$ converges.

$\Longleftarrow$ Let $\{x_n\}$ be Cauchy in $(X, \|\cdot\|)$. Then there exists $n_1 \in \mathbb{N}$ such that $\|x_n - x_{n_1}\| < \frac{1}{2}$ for $n \geqslant n_1$. For $i \in \mathbb{Z}^{\geqslant 2}$, $n_i \in \mathbb{N}$ such that $\|x_n - x_{n_i}\| < \frac{1}{2^i}$ for $n \geqslant n_i \geqslant n_{i-1}$. In particular, $\left\|x_{n_{i+1}} - x_{n_i}\right\| < \frac{1}{2^i}$ for $i \in \mathbb{N}$. So $\sum_{k=1}^{\infty}\left\|x_{n_{k+1}} - x_{n_k}\right\| \leqslant \sum_{k=1}^{\infty} \frac{1}{2^k} = 1 < \infty$. Hence $\sum_{k=1}^{\infty} x_{n_{k+1}} - x_{n_k}$ is absolutely convergent. By assumption, there exists $x \in X$ such that $\sum_{k=1}^{\infty} x_{n_{k+1}} - x_{n_k} = x$, i.e., $\lim_{k\to\infty}(x_{n_{k+1}} - x_{n_1}) = x$, i.e., $\lim_{k\to\infty} x_{n_k} = x + x_{n_1}$, i.e., $\{x_{n_k}\}$ converges in $(X, \|\cdot\|)$. Also, since $\{x_n\}$ is Cauchy in $(X, \|\cdot\|)$, we have $\{x_n\}$ converges. Thus, $(X, \|\cdot\|)$ is Banach. $\square$

**Definition 2.19.** A sequence $\{e_n\} \subseteq X$ is a *Schauder basis* of $(X, \|\cdot\|)$ if for $x \in X$, there exists a unique sequence of coefficients $\{a_n\} \subseteq \mathbb{R}$ such that $x = \sum_{n=1}^{\infty} a_n e_n$.

**Theorem 2.20.** *If* $X$ *has a Schauder basis, then* $X$ *is separable.*

**Example 2.21.** $(l^{\infty}, \|\ldots\|)$ and $(B[0,1], \|\ldots,\|_{\infty})$ do not have Schauder bases.

## 2.3 Finite Dimension NLS

In this section we consider normed linear spaces that are of finite dimensions. These spaces are important since they often appear in many considerations such as in linear algebra and approximation theory. Intuitively such spaces should be simpler than infinite dimensional normed linear spaces. In this and the subsequent section, we will see for certain aspects finite dimensional normed linear spaces do become nicer than infinite dimensional spaces. A main reason of that is in a finite dimensional space we always have a Hamel basis with finitely many linearly independent vectors to work with. We start by looking at an important property of such vectors that will be used throughout this section.

**Assumption 2.22.** Let $X$ be a vector space.

**Theorem 2.23** (Linear combination theorem)**.** *Let* $(X, \|\cdots\|)$ *be a normed linear space and* $\{x_i\}_{i=1}^n \subseteq X$ *be linearly independent. Then there exists* $c > 0$ *such that* $\|\sum_{i=1}^n a_i x_i\| \geqslant c \sum_{i=1}^n |a_i|$.

**Theorem 2.24.** *Every finite dimensional NLS is complete.*

**Definition 2.25.** Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms defined on $X$.

(a) $\|\cdot\|_1$ is said to be *stronger* than $\|\cdot\|_2$ if there exists $M > 0$ such that $\|x\|_2 \leqslant M x_1$ for all $x \in X$.

(b) $\|\cdot\|_2$ are said to be *equivalent* if there exist $m, M > 0$ such that $m\|x\|_2 \leqslant \|x\|_1 \leqslant M|x|_2$ for all $x \in X$. Namely, $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$ and $\|\cdot\|_2$ is also stronger than $\|\cdot\|_1$.

**Lemma 2.26.** Let $\|\cdot\|_1$ be stronger than $\|\cdot\|_2$.

(a) The identity map $i : (X, \|\cdot\|) \to \|\cdot\|_2$ is Lipschitz continuous.

(b) If $\{x_n\}$ converges in $(X, \|\cdot\|_1)$, then $\{x_n\}$ converges in $\|\cdot\|_2$.

(c) If $\{x_n\}$ is Cauchy in $(X, \|\cdot\|_1)$, then $\{x_n\}$ is Cauchy in $\|\cdot\|_2$.

(d) $A \subseteq (X, \|\cdot\|_1)$ is dense, then $A \subseteq \|\cdot\|_2$ is dense.

(e) $A \subseteq \|\cdot\|_2$ is open (closed), then $A \subseteq (X, \|\cdot\|_1)$ is open (closed).

**Theorem 2.27.** *All norms on a finite dimensional vector space are equivalent.*

**Remark.** It implies that convergence or divergence of a sequence in a finite dimension vector space does not depend on the particular choice of a norm on that space.

## 2.4 Compactness

**Theorem 2.28.** *Let* $(X, \|\cdot\|)$ *be finite dimensional and* $K \subseteq X$. *Then* $K$ *is compact if and only if* $K$ *is closed and bounded.*

**Theorem 2.29.** *Let* $(X, d_X)$ *and* $(Y, d_Y)$ *be metric spaces and* $f : X \to Y$ *is continuous. If* $K \subseteq X$ *is compact, then* $f(K) \subseteq Y$ *is compact.*

**Lemma 2.30** (Riesz's Lemma)**.** Let $(X, \|\cdot\|)$ be a normed linear space and $Y \lneqq X$ closed. Then for $\theta \in (0, 1)$, there exists $x \in X$ and $\|x\| = 1$ such that $d(x, Y) \geqslant \theta$.

**Corollary 2.31.** Let $(X, \|\cdot\|)$ be a normed linear space, $Y \lneqq X$ and $\dim Y < \infty$. Then there exists $x \in X$ with $\|x\| = 1$ such that $d(x, Y) = 1$.

**Theorem 2.32.** *Let* $(X, \|\cdot\|)$ *be nonzero. Then* $\overline{B_1}(0)$ *is compact if and only if* $\dim X < \infty$.

Let $X$ be a metric space and $A \subseteq X$.

**Definition 2.33.** $A$ is *totally bounded* if and only if for $\epsilon > 0$, there exists $\{x_i\}_{i=1}^n = \{x_i(\epsilon)\}_{i=1}^n \subseteq A$ such that $A \subseteq \bigcup_{i=1}^n B_\epsilon(x_i)$, i.e., $d(a, x_i) < \epsilon$ for some $i \in \{1, \ldots, n\}$.

**Theorem 2.34.** *$A$ is compact if and only if $A$ is totally bounded and complete.*

**Theorem 2.35.** *(a) If* $\overline{A}$ *is compact, then* $A$ *is totally bounded.*

*(b) If $A$ is totally bounded and $X$ is complete, then $A$ is compact. "precompact"*

*(c) If $A$ is totally bounded, then $A$ is separable.*

## 2.5 Bounded linear operator on NLS

**Definition 2.36.** Let $X$ and $Y$ be two vector spaces and $T : X \to Y$ is a *linear operator* if $\mathcal{D}(T) \leqslant X$ and $T(a_1x_1 + a_2x_2) = a_1T(x_1) + a_2T(x_2)$ for $a_1, a_2 \in \mathbb{R}$ and any $x_1, x_2 \in X$.

**Remark.** If both $X$ and $Y$ are of finite dimension, then $T$ can be represented as a matrix.

**Theorem 2.37.** *Let $X$ and $Y$ be two vector spaces and $T : X \to Y$ a linear operator.*

*(a)* $T(0) = 0$,

*(b)* $\mathrm{Im}(T) \leqslant Y$,

*(c)* $\mathrm{Ker}(T) \leqslant \mathcal{D}(T) \leqslant X$,

*(d)* $\dim \mathcal{D}(T) = \dim \mathrm{Ker}(T) + \dim \mathrm{Im}(T)$.

*(e)* $\mathrm{Ker}(T) = \{0\}$ *if and only if $T$ is 1-1 if and only if there exists* $T^{-1} : \mathrm{Im}(T) \to \mathcal{D}(T)$. *In this case, $T^{-1}$ is also linear and* $\dim(\mathcal{D}(T)) = \dim(\mathrm{Im}(T))$.

**Assumption 2.38.** Throughout this section, we always assume $\mathcal{D}(T) = X$ unless otherwise indicated.

**Definition 2.39.** Let
$$\mathcal{L}(X, Y) = \{T : X \to Y \mid T \text{ is linear}\}.$$

**Theorem 2.40.** *$\mathcal{L}(X, Y)$ is an $\mathbb{K}$-vector space.*

**Assumption 2.41.** Let $(X, \|\cdot\|)$ and $(Y, \|\cdot\|)$ be two normed linear spaces.

**Definition 2.42.** $T : X \to Y$ is a *bounded operator* if there exists $M > 0$ such that $\|T(x)\|_Y \leqslant M\|x\|_X$ for $x \in \mathcal{D}(T)$.

**Remark.** $T : X \to Y$ is a bounded operator if and only if if $A \subseteq \mathcal{D}(T)$ bounded, then $T(A) \subseteq Y$ bounded.

**Theorem 2.43.** *Let $T : X \to Y$ be linear. Then $T$ is bounded if and only if $T$ is Lipschitz continuous if and only if $T$ is continuous if and only if $T$ is continuous at $x_0 \in \mathcal{D}(T)$.*

**Corollary 2.44.** Let $T : X \to Y$ be linear and bounded, then $\mathrm{Ker}(T)$ is closed in $X$.

**Definition 2.45.**

$$\begin{aligned}\mathcal{B}(X, Y) &:= \{T : X \to Y \mid T \text{ is linear and bounded}\}\\ &= \{T : X \to Y \mid T \text{ is linear and continous}\}\\ &=: \mathcal{C}(X, Y)\\ &\leqslant \mathcal{L}(X, Y).\end{aligned}$$

**Theorem 2.46.** *$(\mathcal{B}(X, Y), \|\cdot\|)$ is a NLS, with*

$$\|T\| := \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|T(x)\|_Y = \sup_{\|x\|_X \leqslant 1} \|T(x)\|_Y.$$

**Theorem 2.47.** *If $\dim X < \infty$, then $\mathcal{B}(X, Y) = \mathcal{L}(X, Y)$.*

**Theorem 2.48.** *$\mathcal{B}(X, Y)$ is a Banach if $Y$ is a Banach.*

## 2.6 Bounded linear functional and Dual space

In last section, we considered (bounded and linear) operators which map between normed linear spaces. When the range space lies in the scalar field $\mathbb{R}$ (or $\mathbb{C}$), such operators are even more important and so frequently used that a special name, functional, is designated to them. Collection of all bounded linear functionals of a given normed linear space also plays an important role, and is called the dual space. One of the core branches in mathematics, functional analysis, was initiated from the analysis of functionals.

**Definition 2.49.** Let $X$ be a (real) vector space.

(a) The *algebraic dual* of $X$ is $X^\vee = \mathcal{L}(X, \mathbb{R})$.

(b) The *topological dual* of $X$ is $X' = \mathcal{B}(X, \mathbb{R}) \leqslant \mathcal{L}(X, \mathbb{R})$.

**Remark.** (a) Since the topological dual is more commonly used, usually we simply call $X'$ the dual of $X$.

(b) $X'$ is a Banach space. If $\dim X < \infty$, then $X^\vee = X'$.

**Definition 2.50.** $\{x_n\}$ converges *weakly* to $x$, denoted by $x_n \xrightarrow{w} x$ or $x_n \rightharpoonup x$ as $n \to \infty$, if $|T(x_n) - T(x)| \to 0$ for $T \in X'$.

Weak convergence has various applications in analysis, for instance, in the theory of partial diferential equations. The concept exactly illustrates a basic principle of functional analysis, namely, the investigation of spaces is often related to that of their dual spaces. The following result indicates some of the basic relationships between strong and weak convergences.

**Theorem 2.51.** *If $x_n \to x$ as $n \to \infty$, then $x_n \rightharpoonup x$ as $n \to \infty$.*

**Theorem 2.52.** *If $\dim(X) < \infty$, then $x_n \to x$ as $n \to \infty$ if and only if $x_n \rightharpoonup x$ as $n \to \infty$.*

# Chapter 3

# Inner Product Spaces

## 3.1 Definition and Examples

In normed linear spaces we can add vectors and multiply a vector by scalars, just as in the usual vector algebra in $\mathbb{R}^n$. In addition, the norm of a vector generalizes the basic concept of the length of a vector in $\mathbb{R}^n$. However, one important aspect from vector algebra that is missing in normed linear spaces is an analogue of the "dot" product, and many geometric properties (e.g. orthogonality) that may be described by the dot product. Inner product spaces and Hilbert spaces (complete inner product space) are the vectors spaces in which such generalizations can be done. As we will see in this chapter, such spaces are special normed linear spaces, but their theory is richer and retains many features of Euclidean spaces, with a central concept being orthogonality.

## 3.2 Definition and Examples

**Definition 3.1.** Let $X$ be a vector space over $\mathbb{K}\,(=\mathbb{R}$ or $\mathbb{C})$. We call $(X, \langle\cdot,\cdot\rangle)$ an *inner product space* if we can define the *inner product* $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{K}$ that satisfies for all $x, y, z \in X$ and $a \in \mathbb{K}$,

(a) $\langle x, x\rangle \geqslant 0$ and $\langle x, x\rangle = 0$ if and only if $x = \mathbf{0}$, "positive definiteness".

(b) $\langle x, y\rangle = \overline{\langle y, x\rangle}$, "conjugacy symmetry".

(c) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ and $\langle ax, z\rangle = a\langle x, z\rangle$, "linearity in first argument".

**Remark.** (a) Note the definition of an inner product depends on the scalar field we use, and using real number $\mathbb{R}$ is a special case of using complex number $\mathbb{C}$. Therefore throughout this chapter we will choose complex number $\mathbb{C}$ as the scalar field. This is different from the norm defined in Chapter 2 where choosing real number $\mathbb{R}$ as the scalar field in general does not lead to much difference from choosing complex number $\mathbb{C}$.

(b) We have $\langle \mathbf{0}, x\rangle = \langle x, \mathbf{0}\rangle = 0$ for any $x \in X$ by conjugate symmetry since $\langle \mathbf{0}, x\rangle = \langle 0 \cdot \mathbf{0}, x\rangle = 0\langle \mathbf{0}, x\rangle = 0$.

(c) We have

$$\langle x, ay+bz\rangle = \overline{\langle ay+bz, x\rangle} = \bar{a}\overline{\langle y, x\rangle} + \bar{b}\overline{\langle z, x\rangle} = \bar{a}\langle x, y\rangle + \bar{b}\langle x, z\rangle, \forall x, y, z \in X \text{ and } a, b \in \mathbb{K}.$$

So if $\mathbb{K} = \mathbb{R}$, $\langle\cdot\rangle$ is bilinear. If $\mathbb{K} = \mathbb{C}$, $\langle\cdot\rangle$ is sesquilinear.

**Lemma 3.2** (Cauchy-Schwartz inequality)**.** Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space. Then

$$|\langle x, y\rangle|^2 \leqslant \langle x, x\rangle\langle y, y\rangle, \forall x, y \in X.$$

**Theorem 3.3.** $(X, \|\cdot\|)$ *is a real inner product space if and only if* $(X, \|\cdot\|)$ *is a NLS with the norm satisfying the* parallelogram identity

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2), \forall x, y \in X.$$

**Remark.** Thus whenever there is an inner product, it automatically generates a norm by the formula $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$, and as a result we can consider the length of a vector, distance between vectors, and convergence, etc. In particular, the Cauchy-Schwartz inequality may be stated as $|\langle x, y\rangle| \leqslant \|x\|\|y\|$. On the other hand, a norm can only generate an inner product (by the polarization identity) when it satisfies the parallelogram identity.

**Definition 3.4.** A complete inner product space is called a *Hilbert space*.

**Example 3.5.** $\mathbb{R}^n$ is a Hibert space with the inner product defined by

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i, \forall x = (x_1, \ldots, x_n) \text{ and } y = (y_1, \ldots, y_n) \in \mathbb{R}^n.$$

**Example 3.6.** $\mathbb{C}^n$ is a Hibert space with the inner product defined by

$$\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}, \forall x = (x_1, \ldots, x_n) \text{ and } y = (y_1, \ldots, y_n) \in \mathbb{C}^n.$$

**Example 3.7.** $(l^p, \|\cdot\|_p)$ with $p \in \overline{\mathbb{R}}^{\geqslant 1}$ can not be an inner product space unless $p = 2$, in which case it is also a Hilbert space.

**Example 3.8.** (a) $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ is not an inner product space.

(b) $(\mathcal{C}[0,1], \|\cdot\|_2)$ is an inner product space, but not a Hilbert space.

**Lemma 3.9** (Continuity of inner product)**.** Let $X$ be an inner product space. Let $x_n \to x$ and $y_n \to y$ in $X$, as $n \to \infty$, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$ in $\mathbb{C}$, as $n \to \infty$.

## 3.3 Orthogonal Complement and Direct Sum

One distinguished feature of inner product spaces or Hilbert spaces is that they may be decomposed as the (direct) sum of appropriate smaller subspaces which are orthogonal to each other. Such decomposition is motivated from the Euclidean geometry where we can “project” a vector onto a plane or an axis by drawing perpendicular lines. We start with some basic notations.

**Assumption 3.10.** Let $X$ be an inner product space unless otherwise indicated..

**Definition 3.11.** Let $A, B \subseteq X$ and $x, y \in X$.

(a) The *angle* between $X$ and $y$ is defined by

$$\angle(x, y) = \arccos\left(\frac{|\langle x, y\rangle|}{\|x\|\|y\|}\right) \in [0, \pi].$$

(b) $x$ and $y$ are *orthogonal*, denoted by $x \perp y$, if $\angle(x, y) = \frac{\pi}{2}$ or $\langle x, y\rangle = 0$.

(c) $x$ is *orthogonal* to $A$, denoted by $x \perp A$, if and only if $\langle x, y\rangle = 0$ for $y \in A$.

(d) $A$ is *orthogonal* to $B$, denoted by $A \perp B$, if $\langle x, y\rangle = 0$ for $x \in A$ and any $y \in B$.

(e) The *orthogonal complement* of $A$, denoted by $A^\perp$, is defined by

$$A^\perp = \{x \in X \mid \langle x, y\rangle = 0, \forall y \in A\}.$$

**Theorem 3.12.** *Let $Y \subseteq X$ be convex and complete. Then for $x \in X$, there exists a unique $y_x \in Y$ such that $\|x - y_x\| = d(x, Y)$.*

**Remark.** The convexity and completeness assumptions may be replaced by stronger conditions. The commonly used ones are "$Y$ is a complete subspace of $X$", and "$X$ is Hilbert, $Y$ is a closed subspace of X" (note that linearity (subspace) implies convexity), as we will see in the rest of this chapter.

**Theorem 3.13.** *Let $Y \leqslant X$ be complete. Then for any $x \in X$, there exists $y_x \in Y$ such that $\|x - y_x\| = d(x, Y)$ if and only if $(x - y_x) \perp Y$.*

**Definition 3.14.** Let $X$ be a vector space and $Y, Z \leqslant X$. Then we say $X$ is a *direct sum* of $Y$ and $Z$, denoted by $X = Y \oplus Z$, if $X = Y + Z$ and $Y \cap Z = \{0\}$.

**Lemma 3.15.** Let $X$ be a vector space and $Y, Z \leqslant X$. Then $X = Y \oplus Z$ if and only if every $x \in X$ has a unique expression $x = y_x + z_x$ for some $y_x \in Y$ and $z_x \in Z$.

**Theorem 3.16.** *Let $Y \leqslant X$ be complete, then $X = Y \oplus Y^\perp$.*

**Theorem 3.17.** *Let $Y \leqslant X$ complete. Then there exists $P_Y : X \to Y$ such that*

*(a) $P_Y \in \mathcal{B}(X, Y)$ and $\|P_Y\| = 1$.*

*(b) $P_Y^2 = P_Y$.*

*(c) $\mathrm{Im}(P_Y) = Y$ and $\mathrm{Ker}(P_Y) = Y^\perp$.*

**Lemma 3.18.** Let $\emptyset \neq A, B \subseteq X$, then $A \perp B$ if and only if $\overline{\langle A\rangle} \perp \overline{\langle B\rangle}$.

**Corollary 3.19.** Let $A \subseteq X$, then $A^\perp$ is a closed subspace of $X$.

## 3.4 Orthonormal Sets and Sequences

**Assumption 3.20.** Let $X$ be an inner product space.

**Definition 3.21.** (a) $A \subseteq X$ is an *orthonormal set* if $\langle x, y\rangle = 0$ for any $x, y \in A$ such that $x \neq y$, and $\langle x, x\rangle = 1$.

(b) $A \subseteq X$ is an *orthonormal basis* of $X$ if $A$ is an orthonormal set and $\overline{\langle A\rangle} = X$.

**Remark.** If an orthonormal set $A$ is countable, e.g., $A = \{x_n\}_{n=1}^{\infty}$, then $A$ is also called an *orthonormal sequence*. In this case, the orthonormality condition may be written in terms of the *Kronecker delta* notation:

$$\langle x_i, x_j\rangle = \delta_{i,j} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = 1 \end{cases}, \forall i, j \in \mathbb{N}.$$

**Theorem 3.22.** *Let* $\dim X = n \in \mathbb{N}$. *Then there exists an orthonormal basis* $\{e_i\}_{i=1}^{n}$ *of* $X$. *In fact,* $\{e_i\}_{i=1}^{n}$ *is also a Hamel basis. Then for* $x \in X$, $x = \sum_{i=1}^{n}\langle x, e_i\rangle e_i$ *and* $\|x\|^2 = \sum_{i=1}^{n}|\langle x, e_i\rangle|^2$.

**Theorem 3.23** (Finite Dimension Approximation)**.** *Let* $\{e_i\}_{i=1}^{n} \subseteq X$ *be an orthonormal set. Then for* $x \in X$, *there exists a unique* $y_x \in Y := \langle\{e_i\}_{i=1}^{n}\rangle$ *such that* $\|x - y_x\| = d(x, Y)$ *and* $x - y_x \perp Y$. *In fact,* $y_x = \sum_{i=1}^{n}\langle x, e_i\rangle e_i$, *so*

$$x = \sum_{i=1}^{n}\langle x, e_i\rangle e_i + z_x \in Y \oplus Y^{\perp}.$$

*Proof.* $Y$ is complete since $Y$ is a finite dimensional NLS. □

**Theorem 3.24** (Bessel's inequality)**.** *Let* $\{e_i\}_{i=1}^{\infty}$ *be an orthonormal sequence. Then* $\sum_{i=1}^{\infty}|\langle x, e_i\rangle|^2 \leqslant \|x\|^2$ *for all* $x \in X$.

**Lemma 3.25.** Let $\{e_i\}_{i=1}^{\infty} \subseteq X$ be an orthonormal sequence. Then $\sum_{i=1}^{\infty}|\langle x, e_i\rangle|^2 = \|x\|^2$ for all $x \in \overline{\langle\{e_i\}_{i=1}^{\infty}\rangle}$.

**Theorem 3.26.** *Let* $\{e_i\}_{i=1}^{\infty} \subseteq$ *be an orthonormal sequence. The followings are equivalent.*

*(a)* $X = \overline{span\{e_i\}_{i=1}^{\infty}}$, *"*$\{e_i\}_{i=1}^{\infty}$ *is an orthonormal basis".*

*(b)* $\|x\|^2 = \sum_{i=1}^{\infty}|\langle x, e_i\rangle|^2$ *for* $x \in X$, *"Parseval's indetity".*

*(c)* $x = \sum_{i=1}^{\infty}\langle x, e_i\rangle e_i$ *for* $x \in X$, *"Fouries series".*

*(d)* $\langle x, y\rangle = \sum_{i=1}^{\infty}\langle x, e_i\rangle\overline{\langle y, e_i\rangle}$ *for* $x, y \in X$, *"Plancherel' identity".*

## 3.5 Dual space of Hilbert spaces and Adjoint operator

In Chapter 2 we defined the dual space of a normed linear space $X$ as the space of all bounded linear functionals on $X$. As inner product spaces are special normed linear spaces, we may also consider their dual spaces, namely, if $X$ is an inner product space, we have $X' = \mathcal{B}(X, \mathbb{C})$, with the complex numbers $\mathbb{C}$ being the scalar field. In particular, in this section we will see that if $X$ is a Hilbert space, then its dual space can be identified as itself, i.e., $X' \cong X$. The proof of this isomorphism is based on the celebrated Riesz Representation Theorem, which says every bounded linear functional defined on a Hilbert space has a unique representation in terms of taking inner product with a fixed element that depends on the functional. The Riesz Representation Theorem also leads to a general representation of sesquilinear forms on Hilbert spaces and enables us to define an important class of operators called the (Hilbert) adjoint operator.

**Assumption 3.27.** Let $X$ be an inner product space unless otherwise indicated.

**Lemma 3.28.** For $x \in X$,

$$\begin{aligned} T_x : X &\to \mathbb{C} \\ y &\mapsto \langle y, x \rangle \end{aligned}$$

is a bounded linear functional on $X$ and $\|T_x\| = \|x\|$.

**Theorem 3.29** (Riesz Representation Theorem)**.** *Let $X$ be a Hilbert space. Then for any $T \in X'$, there exists a unique $x_T \in X$ such that $T = T_{x_T}$. i.e., $T(y) = \langle y, x_T \rangle$ for all $y \in X$. In addtion, $\|T\| = \|x_T\|$.*

**Corollary 3.30.** Let $X$ be a Hilbert space. Then $X \cong X'$.

**Definition 3.31.** Let $X$ be $Y$ be a NLSs. Then $B : X \times Y \to \mathbb{C}$ is a *bounded sesquilinear form* if

(a) $B$ is sequilinear: $B(ax_1 + bx_2, y_1) = aB(x_1, y_1) + bB(x_2, y_1)$ and $B(x_1, ay_1 + by_2) = \bar{a}B(x_1, y_1) + \bar{b}B(x_1, y_2)$.

(b) $B$ is bounded: ther exists $M > 0$ such that $|B(x, y)| \leqslant M\|x\|\|y\|$ for $x \in X$ and $y \in Y$. In this case, we define the *norm* of $B$ as

$$\|B\| := \sup_{x \neq \mathbf{0}, y \neq \mathbf{0}} \frac{|B(x, y)|}{\|x\|\|y\|} = \sup_{\|x\|=1,\ \|y\|=1} |B(x, y)|.$$

**Theorem 3.32** (RRT for bounded sesqulinear form)**.** *Let $X, Y$ be Hilbert spaces and $B : X \times Y \to \mathbb{C}$ be a bounded sesquilinear form. Then there exists a unique $S \in \mathcal{B}(X, Y)$ such that $B(x, y) = \langle Sx, y \rangle_Y$ for $x \in X$ and $y \in Y$. In addition, $\|B\| = \|S\|$.*

**Definition 3.33.** Let $X, Y$ be Hilbert spaces and $T \in \mathcal{B}(X, Y)$. We call $T^* \in \mathcal{B}(Y, X)$ the *(Hilbert) adjoint operator* of $T$ if

$$\langle Tx, y \rangle_Y = \langle x, T^* y \rangle_X.$$

**Theorem 3.34.** *Let $X, Y$ be Hilbert spaces and $T \in \mathcal{B}(X, Y)$. Then there exists a unique $T^* \in B(Y, X)$. Moreover, $\|T^*\| = \|T\|$.*

**Theorem 3.35.** *Let $X, Y, Z$ be Hilbert spaces and $T, S \in \mathcal{B}(X, Y)$ and $U \in \mathcal{B}(Y, Z)$.*

*(a)* $(T + S)^* = T^* + S^*$.

*(b)* $(aT)^* = \bar{a}T^*$ *for any* $a \in \mathbb{C}$.

*(c)* $(T^*)^* = T$.

*(d)* $(UT)^* = T^*U^*$.

**Theorem 3.36.** *Let $X, Y$ be Hilbert spaces and $T \in \mathcal{B}(X, Y)$.*

*(a)* $\operatorname{Ker}(T) = \operatorname{Im}(T^*)^\perp$ *and* $\operatorname{Ker}(T^*) = \operatorname{Im}(T)^\perp$.

*(b)* $\operatorname{Ker}(T)^\perp = \overline{\operatorname{Im}(T^*)}$ *and* $\operatorname{Ker}(T^*)^\perp = \overline{\operatorname{Im}(T)}$.

*(c)* $\operatorname{Ker}(T^*T) = \operatorname{Ker}(T)$ *and* $\operatorname{Ker}(TT^*) = \operatorname{Ker}(T^*)$.

**Corollary 3.37.** Let $X, Y$ be Hilbert spaces and $T \in \mathcal{B}(X, Y)$. Then

$$X = \operatorname{Ker}(T) \oplus \operatorname{Ker}(T)^\perp = \operatorname{Im}(T^*)^\perp \oplus \overline{\operatorname{Im}(T^*)},$$

$$Y = \operatorname{Ker}(T^*) \oplus \operatorname{Ker}(T^*)^\perp = \operatorname{Im}(T)^\perp \oplus \overline{\operatorname{Im}(T)}.$$

# Chapter 4

# Fundamental Theorems for Normed and Banach Spaces

Let $X$ be a set.

## 4.1 Zorn's Lemma

**Definition 4.1.** A *partial order* on $X$ is a binary relation on $X \times X$ denoted "$\leqslant$" that satisfies: for any $x, y, z \in X$,

(a) $x \leqslant x$, "reflexivity";

(b) if $x \leqslant y$ and $y \leqslant x$, then $x = y$, "anti-symmetry";

(c) if $x \leqslant y$ and $y \leqslant z$, then $x \leqslant z$, "transitivity".

**Remark.** A partial ordered set is also called a poset.

Let $(X, \leqslant)$ be a poset.

**Definition 4.2.** $X$ is *totally ordered* if for any $x, y \in X$, we must have $x \leqslant y$ or $y \leqslant x$.

**Definition 4.3.** $Y \subseteq X$ is called a *chain* if $Y$ is a totally ordered.

**Example 4.4.** (a) $\mathbb{R}$ is totally ordered w.r.t. the usual "$\leqslant$".

(b) $\mathscr{P}(X)$ is partially ordered not totally ordered w.r.t. the usual "$\subseteq$".

(c) Let $X = \mathbb{R} \times \mathbb{R}$. Define "$\leqslant$" as following: $(x_1, x_2) \leqslant (y_1, y_2)$ if $x_1 \leqslant y_1$ and $x_2 \leqslant y_2$. Then $X$ is partially ordered not totally ordered.

(d) Let $X = \mathbb{R} \times \mathbb{R}$. Define "$\leqslant$" as following: $(x_1, x_2) \leqslant (y_1, y_2)$ if $x_1 < y_1$ or $x_1 = y_1$ and $x_2 \leqslant y_2$. Then $X$ is totally ordered.

**Definition 4.5.** Let $Y \subseteq X$.

(a) $u \in X$ is called an *upper bound* of $Y$ if $y \leqslant u$ for any $y \in Y$.

(b) $m \in X$ is called a *maximal element* of $X$ if it satisfies that if $m \leqslant x$ for some $x \in X$, then $x = m$.

**Remark.** $\mathbb{R}$ has neither upper bound nor maximal element.

**Lemma 4.6** (Zorn's lemma)**.** Let $X$ be a partially ordered set. If every chain of $X$ has an upper bound, then $X$ has a maximal element.

**Remark.** Zorn's lemma is equivalent to the Axiom of Choice and the Well-Ordering Principal.

**Definition 4.7.** Let $X$ be an $\mathbb{R}$-vector space. Let $p : X \to \mathbb{R}$ such that for any $x, y \in X$,

(a) $p(x + y) \leqslant p(x) + p(y)$, "subadditivity".

(b) $p(ax) = ap(x)$ for any $a > 0$, "positive homogeneity".

Then $p$ is called a *sublinear functional* on $X$.

**Example 4.8.** If $X$ is a normed linear space, then $\|\cdot\|$ is a sublinear functional on $X$.

## 4.2 The Hahn-Banach Theorems

**Theorem 4.9** (The Hahn-Banach Theorem)**.** *Let $X$ be an $\mathbb{R}$-vector space and $p$ be a sublinear functional on $X$. If $Y \leqslant X$ and $\varphi : Y \to \mathbb{R}$ is a linear functional that satisfies $\varphi(y) \leqslant p(y)$ for any $y \in Y$, then there exists a linear extension $\widetilde{\varphi} : X \to \mathbb{R}$, i.e., $\widetilde{\varphi}$ is linear and $\widetilde{\varphi}|_Y = \varphi$ such that $\widetilde{\varphi}(x) \leqslant p(x)$ for any $x \in X = \mathscr{D}(\widetilde{\varphi})$.*

*Proof.* Let $\mathcal{B} = \{\psi : \mathscr{D}(\psi) \to \mathbb{R} \mid Y \subseteq \mathscr{D}(\psi), \psi \text{ is linear}, \ \psi|_Y = \varphi, \ \psi(x) \leqslant p(x), \forall x \in \mathscr{D}(\psi)\}$.

Define "$\leqslant$" on $\mathcal{B}$ as follows: $\psi_1 \leqslant \psi_2$ if and only if $\psi_2$ is a linear extension of $\psi_1$, i.e., $\mathscr{D}(\psi_1) \subseteq \mathscr{D}(\psi_2)$ and $\psi_2|_{\mathscr{D}(\psi_1)} = \psi_1$. Let $\psi_1, \psi_2, \psi_3 \in \mathcal{B}$.

Since $\mathscr{D}(\psi_1) \subseteq \mathscr{D}(\psi_1)$ and $\psi_1|_{\mathscr{D}(\psi_1)} = \psi_1$, we have $\psi_1 \leqslant \psi_1$.

Let $\psi_1 \leqslant \psi_2$ and $\psi_2 \leqslant \psi_1$. Then $\mathscr{D}(\psi_1) \subseteq \mathscr{D}(\psi_2)$ and $\psi_2|_{\mathscr{D}(\psi_1)} = \psi_1$ and $\mathscr{D}(\psi_2) \subseteq \mathscr{D}(\psi_1)$ and $\psi_1|_{\mathscr{D}(\psi_2)} = \psi_2$. So $\psi_1 = \psi_1|_{\mathscr{D}(\psi_1)} = \psi_1|_{\mathscr{D}(\psi_2)} = \psi_2$.

Let $\psi_1 \leqslant \psi_2 \leqslant \psi_3$. Then $\mathscr{D}(\psi_1) \subseteq \mathscr{D}(\psi_2)$ and $\psi_2|_{\mathscr{D}(\psi_1)} = \psi_1$ and $\mathscr{D}(\psi_2) \subseteq \mathscr{D}(\psi_3)$ and $\psi_3|_{\mathscr{D}(\psi_2)} = \psi_2$. So $\mathscr{D}(\psi_1) \subseteq \mathscr{D}(\psi_2) \subseteq \mathscr{D}(\psi_3)$ and $\psi_3|_{\mathscr{D}(\psi_1)} = \psi_3|_{\mathscr{D}(\psi_2|_{\mathscr{D}(\psi_1)})} = \psi_2|_{\mathscr{D}(\psi_1)} = \psi_1$. Hence $\psi_1 \leqslant \psi_3$.

So "$\leqslant$" is a partial order on $\mathcal{B}$.

Let $\mathcal{C} \subseteq \mathcal{B}$ be a chain. Define $\widetilde{\varphi} : \bigcup_{\psi \in \mathcal{C}} \mathscr{D}(\psi) \to \mathbb{R}$ given by $\widetilde{\varphi}(x) = \psi(x)$ if $x \in \mathscr{D}(\psi)$. Let $x, y \in \bigcup_{\psi \in \mathcal{C}} \mathscr{D}(\psi)$. Since $\mathcal{C}$ is a chain, there exists $\psi \in \mathcal{C}$ such that $x, y \in \mathscr{D}(\psi)$. Since $\mathscr{D}(\psi)$ is an $\mathbb{R}$-vector space, $ax + by \in \mathscr{D}(\psi) \subseteq \bigcup_{\psi \in \mathcal{C}} \mathscr{D}(\psi)$ for any $a, b \in \mathbb{R}$. So $\mathscr{D}(\widetilde{\varphi})$ is an $\mathbb{R}$-vector space.

Let $x \in \mathscr{D}(\psi_1) \cap \mathscr{D}(\psi_2)$, where $\psi_1, \psi_2 \in \mathcal{C}$. Since $\mathcal{C}$ is a chain, either $\psi_1 \leqslant \psi_2$ or $\psi_2 \leqslant \psi_1$, say $\psi_1 \leqslant \psi_2$. Then $\mathscr{D}(\psi_1) \cap \mathscr{D}(\psi_2) = \mathscr{D}(\psi_1)$ and $\psi_2|_{\mathscr{D}(\psi_1)} = \psi_1$. So $\psi_1(x) = \psi_2(x)$ for any $x \in \mathscr{D}(\psi_1) = \mathscr{D}(\psi_1) \cap \mathscr{D}(\psi_2)$. Hence $\widetilde{\varphi}$ is well-defined.

Similar argument gives $\widetilde{\varphi}(ax + by) = a\widetilde{\varphi}(x) + b\widetilde{\varphi}(y)$. So $\widetilde{\varphi}$ is linear.

In addition, $\widetilde{\varphi}$ is clearly an upper bound of $\mathcal{C}$. So by Zorn's lemma, $\mathcal{B}$ has a maximal element $\widetilde{\varphi}$.

Claim $\widetilde{\varphi}$ is what we need. It suffices to show $\mathscr{D}(\widetilde{\varphi}) = X$. Suppose not, there exists $0 \neq x_0 \in X \setminus \mathscr{D}(\widetilde{\varphi})$. Consider $\mathcal{Z} = \operatorname{span}\{\mathscr{D}(\widetilde{\varphi}), x_0\}$. Since $\mathscr{D}(\widetilde{\varphi})$ is an $\mathbb{R}$-vector space, every $z \in \mathcal{Z}$ has a representation $z = y + cx_0$ with $y \in \mathscr{D}(\widetilde{\varphi})$ and $c \in \mathbb{R}$. Let $z = y_1 + c_1x_0 = y_2 + c_2x_0$ with $y_1, y_2 \in \mathscr{D}(\widetilde{\varphi})$ and $c_1, c_2 \in \mathbb{R}$, then $y_1 - y_2 = (c_2 - c_1)x_0 \in \mathscr{D}(\widetilde{\varphi}) \cap \operatorname{span}\{x_0\} = \{0\}$, so $y_1 = y_2$ and since $x_0 \neq 0$, $c_1 = c_2$. Hence this representation is unique. Define $\varphi_0 : \mathcal{Z} \to \mathbb{R}$ given by $\varphi_0(y + cx_0) = \widetilde{\varphi}(y) + ca$ for some $a \in \mathbb{R}$. Clearly, $\varphi_0|_{\mathscr{D}(\widetilde{\varphi})} = \widetilde{\varphi}$. Note $\varphi_0$ is linear since for any $y_1, y_2 \in \mathscr{D}(\widetilde{\varphi})$ and any $b_1, b_2, c_1, c_2 \in \mathbb{R}$, we have $\varphi_0(b_1(y_1 + c_1x_0) + b_2(y_2 + c_2x_0)) = \varphi_0(b_1y_1 + b_2y_2 + (b_1c_1 + b_2c_2)x_0) = \widetilde{\varphi}(b_1y_1 + b_2y_2) + (b_1c_1 + b_2c_2)a = b_1\widetilde{\varphi}(y_1) + b_2\widetilde{\varphi}(y_2) + b_1c_1a + b_2c_2a = b_1(\widetilde{\varphi}(y_1) + c_1a) + b_2(\widetilde{\varphi}(y_2) + c_2a) = b_1\varphi_0(y_1 + c_1x_0) + b_2\varphi_0(y_2 + c_2x_0)$. Claim $\varphi_0(z) \leqslant p(z)$ for any $z \in \mathcal{Z}$. To show it, it is equivalent to show $\varphi_0(y + cx_0) \leqslant p(y + cx_0)$ for any $y \in \mathscr{D}(\widetilde{\varphi})$ and any $c \in \mathbb{R}$. If $c = 0$, clearly, $\varphi_0(y) = \widetilde{\varphi}(y) = \varphi(y) \leqslant p(y)$ for any $y \in Y$. Assume $c \neq 0$ now. It suffices to show the case $c = \pm 1$ because then for any $c > 0$, we have $\varphi_0(y + cx_0) = c\varphi_0(y/c + x_0) \leqslant cp(y/c + x_0) = p(y + cx_0)$ since $\varphi_0$ is linear and similarly for $c < 0$. Thus, we only need to show $\varphi_0(y + x_0) \leqslant p(y + x_0)$ and $\varphi_0(y' - x_0) \leqslant p(y' - x_0)$ for any $y, y' \in \mathscr{D}(\widetilde{\varphi})$, i.e., $\widetilde{\varphi}(y') - p(y' - x_0) \leqslant c \leqslant p(y + x_0) - \widetilde{\varphi}(y)$ for any $y, y' \in \mathscr{D}(\widetilde{\varphi})$, i.e., $c$ can be selected if and only if $\widetilde{\varphi}(y') - p(y' - x_0) \leqslant p(y + x_0) - \widetilde{\varphi}(y)$ for any $y, y' \in \mathscr{D}(\widetilde{\varphi})$ if and only if $\widetilde{\varphi}(y) + \widetilde{\varphi}(y') \leqslant p(y + x_0) + p(y' - x_0)$ for any $y, y' \in \mathscr{D}(\widetilde{\varphi})$ which is always true since $\widetilde{\varphi}(y) + \widetilde{\varphi}(y') = \widetilde{\varphi}(y + y') \leqslant p(y + y') \leqslant p(y + x_0) + p(y' - x_0)$.

Therefore, we have $\varphi_0 \in \mathcal{B}$ and $\widetilde{\varphi} \leqslant \varphi_0$. By the maximality of $\widetilde{\varphi}$, we have $\varphi = \varphi_0$, a contradiction since $x_0 \notin D(\widetilde{\varphi})$ but $x_0 \in \mathcal{Z} = \mathscr{D}(\varphi_0)$. □

Let $X$ be a NLS.

**Corollary 4.10** (Hahn-Banach for NLS)**.** Let $Y \leqslant X$ and $\varphi \in Y'$. Then there exists $\widetilde{\varphi} \in X'$ such that $\widetilde{\varphi}|_Y = \varphi$ and $\|\widetilde{\varphi}\| = \|\varphi\|$.

*Proof.* Note $|\varphi(y)| \leqslant \|\varphi\|\|y\|$ for any $y \in Y$. Define $p(x) = \|\varphi\|\|x\|$ for any $x \in X$. Since $\varphi \in Y'$, $p$ is well-defined. Then $\varphi(y) \leqslant p(y)$ for any $y \in Y$. Also, note $p$ is sublinear on $Y$ since for any $x, y \in X$ and $a > 0$, we have

(a) $p(x + y) = \|\varphi\|\|x + y\| \leqslant \|\varphi\|(\|x\| + \|y\|) = \|\varphi\|\|x\| + \|\varphi\|\|y\| = p(x) + p(y)$;

(b) $p(ax) = \|\varphi\|\|ax\| = a\|\varphi\|\|x\| = ap(x)$.

Since $\varphi \in Y'$, $\varphi : Y \to \mathbb{R}$ is a linear functional. Then by Hahn-Banach theorem, there exists $\widetilde{\varphi} : X \to \mathbb{R}$ such that $\widetilde{\varphi}$ is a linear extension of $\varphi$ and $\widetilde{\varphi}(x) \leqslant p(x) = \|\varphi\|\|x\|$ for any $x \in X$. Then $\|\widetilde{\varphi}\| = \sup_{x \in X, \|x\|=1} |\widetilde{\varphi}(x)| \leqslant \sup_{x \in X, \|x\|=1} \|\varphi\|\|x\| = \|\varphi\|$. On the other hand, $\|\widetilde{\varphi}\| \geqslant \|\varphi\|$ is obvious. So $\|\widetilde{\varphi}\| = \|\varphi\|$. □

**Corollary 4.11.** Let $x \in X$. Then there exists $\widetilde{\varphi} \in X'$ such that $\|\widetilde{\varphi}\| = 1$ and $\widetilde{\varphi}(x) = \|x\|$.

*Proof.* Let $Y = \operatorname{span}\{x\} \leqslant X$ and define $\varphi : Y \to \mathbb{R}$ given by $\varphi(ax) = a\|x\|$. Easy to verify $\varphi$ is linear. Note $\varphi$ is bounded since $|\varphi(ax)| = |a|\|x\| = \|ax\|$ for any $a \in \mathbb{R}$ and $x \in X$. So $\|\varphi\| \leqslant 1$. Actually, $\|\varphi\| = 1$ since $\|\varphi\| = \sup_{ax \neq 0} \frac{|\varphi(ax)|}{\|ax\|} = \sup_{a \neq 0} \frac{|a\|x\||}{\|ax\|} = \sup_{a \neq 0} \frac{|a|\|x\|}{|a|\|x\|} = \sup_{a \neq 0} \|x_0\|^2 = 1$. Thus, by previous corollary, there exists $\widetilde{\varphi} \in X'$ such that $\widetilde{\varphi}|_Y = \varphi$ and $\|\widetilde{\varphi}\| = \|\varphi\| = 1$. Since $x \in Y$, $\widetilde{\varphi}(x) = \varphi(x) = \|x\|$. □

**Corollary 4.12.** Let $x \in X$. Then $\|x\| = \sup_{\varphi \in X', \|\varphi\|=1} |\varphi(x)|$. Conclude that if there is $x \in X$ such that $\varphi(x) = 0$ for any $\varphi \in X'$, then $x = 0$.

*Proof.* "$\geqslant$". Since $|\varphi(x)| \leqslant \|\varphi\|\|x\|$ for any $\varphi \in X'$, $\sup_{\varphi \in X', \|\varphi\|=1} |\varphi(x)| \leqslant \sup_{\varphi \in X', \|\varphi\|=1} \|\varphi\|\|x\| = \sup_{\varphi \in X', \|\varphi\|=1} \|x\| = \|x\|$.

"$\leqslant$". For any $x \in X$, by previous corollary, there exists $\widetilde{\varphi} \in X'$ such that $\|\widetilde{\varphi}\| = 1$ and $\widetilde{\varphi}(x) = \|x\|$. So $\sup_{\varphi \in X', \|\varphi\|=1} |\varphi(x)| \geqslant |\widetilde{\varphi}(x)| = \|x\|$. □

**Remark.** Equivalently, $\|x\| = \sup_{\varphi\in X', x\neq 0} \frac{|\varphi(x)|}{\|\varphi\|} = \sup_{\varphi\in X', \|\varphi\|\leqslant 1}|\varphi(x)|$, which all compute the norm of an element the other way around.

**Theorem 4.13.** *Let $Y \leqslant X$. Then $d(x, Y) := \inf_{y\in Y}\|x - y\| = \sup_{\varphi\in X', \|\varphi\|\leqslant 1, \varphi|_Y=0}|\varphi(x)|$ for any $x \in X$. In particular, we recover the previous corollary if $Y = \{0\}$.*

*Proof.* "$\geqslant$". Let $\varphi \in X'$ such that $\|\varphi\| \leqslant 1$ and $\varphi|_Y = 0$. Since $\varphi$ is linear, $|\varphi(x)| = |\varphi(x) - \varphi(y)| = |\varphi(x - y)| \leqslant \|\varphi\|\|x - y\| \leqslant \|x - y\|$ for any $y \in Y$. So $|\varphi(x)| \leqslant \inf_{y\in Y}\|x - y\| = d(x, Y)$. Hence $\sup_{\varphi\in X', \|\varphi\|\leqslant 1, \varphi|_Y=0}|\varphi(x)| \leqslant d(x, Y)$.

"$\leqslant$". If $x \in Y$, then $d(x, Y) = 0$ and clearly it holds. Assume $x \notin Y$ now. Consider $Y_0 = \operatorname{span}\{Y, x\} = \{y + ax \mid y \in Y, a \in \mathbb{R}\}$. Similar to the proof of Hahn-Banach theorem, every $y_0 \in Y_0$ has a representation $y_0 = y + ax$ with $y \in Y$ and $a \in \mathbb{R}$. Define $\varphi_0 : Y_0 \to \mathbb{R}$ by $\varphi_0(y + ax) = ad(x, Y)$. Clearly, $\varphi_0$ is linear and $\varphi_0|_Y = 0$. Let $y \in Y$ and $a \in \mathbb{R}$. Since $-\frac{y}{a} \in Y$, we have $|\varphi_0(y + ax)| = |a|d(x, Y) \leqslant |a|\|x - (-\frac{y}{a})\| = \|y + ax\|$. So $\varphi_0$ is bounded and hence $\varphi_0 \in Y_0'$. By previous corollary, there exists an extension $\widetilde{\varphi} \in X'$ such that $\widetilde{\varphi}|_{Y_0} = \varphi_0|_{Y_0} = 0$ and $\|\widetilde{\varphi}\| = \|\varphi_0\| \leqslant 1$. Since $x \in Y_0$, we have $\sup_{\varphi\in X', \|\varphi\|\leqslant 1, \varphi|_Y=0}|\varphi(x)| \geqslant |\widetilde{\varphi}(x)| = |\varphi_0(x)| = |\varphi_0(0 + x)| = d(x, Y)$. □

**Corollary 4.14.** Let $Y \leqslant X$ and $x \in X$. Then $d(x, Y) \leqslant \|x\|$.

*Proof.* Follow from two previous result or follow from $0 \in Y$ and $\|x\| = d(x, 0)$. □

**Remark.** This is a kind of dual variational problem.

## 4.3 Geometric Hahn-Banach Theorem

Let $X$ be a real **vector space**.

**Definition 4.15.** Let $Y \subseteq X$. $x_0 \in Y$ is an *internal* point of $Y$ if for any $x \in X$, there exists $\epsilon(x) > 0$ such that $x_0 + tx \in Y$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$.

**Remark.** If $x_0 \in Y \subseteq X$ is an interior point, then it is an internal point. But the converse may be false.

Since $x_0 \in Y$ is an interior point, there exists $r > 0$ such that $B_r(x_0) \subseteq Y$. For any $0 \neq x \in X$, choose $\epsilon(x) = \frac{r}{\|x\|}$ and then $x_0 + tx \in B_r(x_0) \subseteq Y$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$.

Consider the set $A \subseteq \mathbb{R}^2$ consisting of the union of

- the region delimited by the graphs of $\sqrt{x}$ and $-\sqrt{x}$ over $[0, \infty)$,
- the region delimited by the graphs of $\sqrt{-x}$ and $-\sqrt{-x}$ over $(-\infty, 0]$,
- the $y$-axis.

Then $0 \in A$ is internal but not interior.

**Theorem 4.16.** *Let $Y \subseteq X$ be convex and $x_0 \in Y$. If $X$ is finite dimensional, then $x_0$ is an interior point of $Y$ if and only if $x_0$ is an internal point of $Y$.*

**Lemma 4.17.** Let $K \subseteq X$ be convex and contain 0 as an internal point. Define $p_K : X \to \mathbb{R}$ given by $p_K(x) = \inf\{m > 0 \mid \frac{x}{m} \in K\}$. Then

(a) $p_K$ is well-defined.

(b) $p_K$ is a sublinear functional on $X$.

(c) If $x \in K$, then $p_K(x) \leqslant 1$.

(d) $x$ is an internal point of $K$ if and only if $p_K(x) < 1$.

*Proof.* (a) Let $x \in X$. It suffices to show $p_K(x) < \infty$. Since 0 is an internal point of $K$, there exists $\epsilon(x) > 0$ such that $0 + tx = tx \in K$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$. Pick $m = \frac{1}{\epsilon} + 1$. Then $\frac{1}{m} = \frac{\epsilon}{1+\epsilon} < \epsilon$. So $\frac{x}{m} = \frac{\epsilon}{1+\epsilon} x \in K$. Thus, $p_K(x) \leqslant m = \frac{1}{\epsilon} + 1$.

(b) $p_K(ax) = ap_K(x)$ for any $a > 0$ and any $x \in X$ follows from the definition of $p_K(x)$.
Let $x, y \in X$, $\alpha = p_K(x)$ and $\beta = p_K(y)$. Let $\epsilon > 0$. Then by the definition of the infimum, there exist $m_x \in \{m > 0 \mid \frac{x}{m} \in K\}$ with $m_x < \alpha + \epsilon$ and $m_y \in \{m > 0 \mid \frac{x}{m} \in K\}$ with $m_y < \alpha + \epsilon$, i.e., there exist $0 < \epsilon_x, \epsilon_y < \epsilon$ such that $\frac{x}{\alpha+\epsilon_x} \in K$ and $\frac{y}{\beta+\epsilon_y} \in K$. Since $K$ is convex, $\frac{x+y}{\alpha+\beta+\epsilon_x+\epsilon_y} = \frac{\alpha+\epsilon_x}{\alpha+\beta+\epsilon_x+\epsilon_y}\frac{x}{\alpha+\epsilon_x} + \frac{\beta+\epsilon_y}{\alpha+\beta+\epsilon_x+\epsilon_y}\frac{y}{\beta+\epsilon_y} \in K$. So $p_K(x+y) \leqslant \alpha + \beta + \epsilon_x + \epsilon_y \leqslant \alpha + \beta + 2\epsilon$. Since $\epsilon > 0$ is arbitrary, $p_K(x+y) \leqslant \alpha + \beta = p_K(x) + p_K(y)$.

(c) Let $x \in K$. Then $x = \frac{x}{1} \in K$. So $p_K(x) \leqslant 1$. Converse is not true. For example, $X = \mathbb{R}$, $K = (-1, 1)$ and $x = 1$ with $p_K(1) = 1$.

(d) $\Longrightarrow$ Let $x$ be an internal point of $K$. Then there exists $\epsilon(x) > 0$ such that $x + tx \in K$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$. Pick $\alpha$ such that $\frac{1}{1+\epsilon} < \alpha < 1$, i.e., $1/\alpha - 1 < \epsilon$. Then $x/\alpha = x + (1/\alpha - 1)x \in K$. Thus, $p_K(x) \leqslant \alpha < 1$.

$\Longleftarrow$ Let $p_K(x) < 1$. Then there exists $\alpha \in (0, 1)$ such that $\frac{x}{\alpha} \in K$. Let $y \in X$. Since $X$ is a real vector space, $\frac{y}{1-\alpha} \in X$. Since 0 is an internal point $K$, there exists $\epsilon(y, \alpha) > 0$ such that $0 + t\frac{y}{1-\alpha} \in K$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$. Since $K$ is convex, $x + ty = \alpha\frac{x}{\alpha} + (1-\alpha)\frac{ty}{1-\alpha} \in K$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$. Since $y \in Y$ is arbitrary, we have $x$ is an internal point of $K$. □

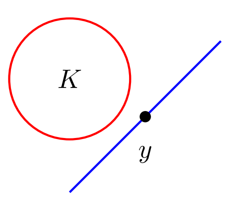
**Definition 4.18.** Let $\varphi$ be a linear functional on $X$ and $c \in \mathbb{R}$.

(a) The set $\{x \in X \mid \varphi(x) = c\}$ is called a *hyperplane* of $X$ w.r.t. $\varphi$.

(b) The sets $\{x \in X \mid \varphi(x) > c\}$ and $\{x \in X \mid \varphi(x) < c\}$ are called a *half spaces* of $X$ w.r.t. $\varphi$.

**Remark.** The set of solutions of $\{\varphi(x) = c \mid x \in X\}$ forms a hyperplane w.r.t. $\varphi$.

**Theorem 4.19** (Hahn-Banach Geometric Version). *Let $K \subseteq X$ be convex and contains 0 as an internal point. Then for any $y \notin K$, there exists $\varphi : X \to \mathbb{R}$ linear such that $\varphi(y) = 1$ and if every point in $K$ is an internal point, then $\varphi(x) < 1$ for any $x \in K$. Namely, $y$ can be separated from $K$ by the hyperplane $\{x \in X \mid \varphi(x) = 1\}$.*



*Proof.* Since $K \subseteq X$ is convex and contains 0 as an internal point, previous lemma is applicable. Consider $Y = \text{span}\{y\}$ and define $\varphi_0 : Y \to \mathbb{R}$ by $\varphi_0(ay) = a$. Easy to check $\varphi_0$ is linear. If $a \leqslant 0$, $\varphi_0(ay) = a \leqslant 0 \leqslant p_K(ay)$. If $a > 0$, since $y \notin K$, we have $p_K(y) \geqslant 1$ and then $\varphi_0(ay) = a \leqslant ap_K(y) = p_K(ay)$, i.e., $\varphi_0(y) \leqslant p_K(y)$ for any $y \in Y$. Hence $\varphi_0(ay) \leqslant p_K(ay)$ for any $a \in \mathbb{R}$. Also, $p_K$ is sublinear on $X$, by Hahn-Banach theorem (analytic version), there exists $\varphi : X \to \mathbb{R}$ linear and $\varphi|_Y = \varphi_0$ such that $\varphi(x) \leqslant p_K(x)$ for any $x \in X$. In particular, since $y \in Y$, $\varphi(y) = \varphi_0(y) = 1$ and since every point $x \in K$ is an internal point, $\varphi(x) \leqslant p_K(x) < 1$ for any $x \in K$. □

**Corollary 4.20.** Let $K \subseteq X$ be convex and contains at least one internal point. Then for any $y \notin K$, there exists $\varphi : X \to \mathbb{R}$ linear such that $\varphi(x) \leqslant \varphi(y)$ for any $x \in K$.

*Proof.* Let $x_0 \in K$ be an internal point and consider $\widetilde{K} = K - \{x_0\} = \{x - x_0 \mid x \in K\} \subseteq X$. For any $x_1 - x_0, x_2 - x_0 \in \widetilde{K}$ with $x_1, x_2 \in K$, since $K$ is convex, we have $\lambda(x_1 - x_0) + (1-\lambda)(x_2 - x_0) = \underbrace{\lambda x_1 + (1-\lambda)x_2}_{\in K} - x_0 \in \widetilde{K}$. So $\widetilde{K} \subseteq X$ is convex. Since $x_0 \in K$, $0 = x_0 - x_0 \in \widetilde{K}$. Claim 0 is an internal point of $\widetilde{K}$. Need to show for any $x - x_0 \in \widetilde{K}$, there exists $\epsilon = \epsilon(x, x_0) > 0$ such that $tx + (1-t)x_0 - x_0 = 0 + t(x - x_0) \in \widetilde{K}$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$, i.e., $tx + (1-t)x_0 \in K$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$. Since $x_0$ is an interior point, there exists $0 < \epsilon_0 = \epsilon_0(x) < 1$ such that $x_0 + tx \in K$ for any $t \in \mathbb{R}$ with $|t| < \epsilon_0 < 1$. Also, since $K$ is convex and $x_0 \in K$, for any $x \in X$, $tx + (1-t)x_0 \in K$ for any $t \in \mathbb{R}$ with $|t| < \epsilon_0 < 1$. Hence 0 is an internal point of $\widetilde{K}$. Now, let $y \notin K$, then $y - x_0 \notin \widetilde{K}$. So by a similar proof as the above theorem, there exists $\varphi : X \to \mathbb{R}$ linear and $\varphi(y - x_0) = 1$ such that $\varphi(x) \leqslant p_{\widetilde{K}}(x)$ for any $x \in X$. Then for any $x \in K$, we have $x - x_0 \in \widetilde{K} \subseteq X$ and then $\varphi(x) - \varphi(x_0) = \varphi(x - x_0) \leqslant p_{\widetilde{K}}(x - x_0) \leqslant 1 = \varphi(y - x_0) = \varphi(y) - \varphi(x_0)$, i.e., $\varphi(x) \leqslant \varphi(y)$. □

**Corollary 4.21.** Let $A, B \subseteq X$ be nonempty and convex with $A \cap B = \emptyset$ and at least one has an internal point. Then there exists $c \in \mathbb{R}$ and $\varphi : X \to \mathbb{R}$ linear such that $\varphi(a) \leqslant c \leqslant \varphi(b)$ for any $a \in A$ and $b \in B$. Namely, $A$ and $B$ can be separated by the hyperplane $\{x \in X \mid \varphi(x) = c\}$.

*Proof.* Let $K = A - B = \{a - b \mid a \in A, b \in B\}$. Let $a_1 - b_1, a_2 - b_2 \in K$ and $\lambda \in (0,1)$. Since $\lambda(a_1 - b_1) + (1-\lambda)(a_2 - b_2) = \lambda a_1 + (1-\lambda)a_2 - (\lambda b_1 + (1-\lambda)b_2) \in K$, we have $K$ is convex. Wlog, assume $A$ has an internal point $x_0$. Then for any $x \in X$, there exists $\epsilon(a)$ such that $x_0 + tx \in A$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$. Since $B \neq \emptyset$, there exists $b \in B$ and we have for any $x \in X$, $x_0 - b + tx = (x_0 + tx) - b \in A - B = K$ for any $t \in \mathbb{R}$ with $|t| < \epsilon$. So $x_0 - b$ is an internal point of $K$. Since $A \cap B \neq \emptyset$, $0 \notin K$. By previous corollary, there exists $\varphi : X \to \mathbb{R}$ linear such that for any $a \in A$ and $b \in B$, $\varphi(a) - \varphi(b) = \varphi(a - b) \leqslant \varphi(0) = 0$, i.e., $\varphi(a) \leqslant \varphi(b)$. Pick $c \in [\sup_{a \in A} \varphi(a), \inf_{b \in B} \varphi(b)]$. □

**Remark.** The above corollary is related to OR.

## 4.4 The adjoint operator

Let $X$ and $Y$ be NLS's, $T \in B(X, Y)$.

**Recall 4.22.** Let $X, Y$ be Hilbert space and $T \in \mathcal{B}(X, Y)$. Then there exists a unique (Hilbert) adjoint $T^* \in \mathcal{B}(Y, X)$ such that $\langle Tx, y\rangle_Y = \langle x, T^*y\rangle_X$ for any $x \in X$ and $y \in Y$. In addition, $\|T^*\| = \|T\|$.

**Remark.** Let $\psi \in Y'$. Define $\varphi : X \to \mathbb{R}$ by $\varphi(x) = \psi(Tx)$. Then

(a) $\varphi$ is linear since both $\psi$ and $T$ are linear.

(b) $\varphi$ is bounded since $|\varphi(x)| = |\psi(Tx)| \leqslant \|\psi\| \cdot \|Tx\|_Y \leqslant \|\psi\| \cdot \|T\| \cdot \|x\|_X$ for any $x \in X$.

**Definition 4.23.** The (NLS) *adjoint* of $T$, denoted by $T^*$, is defined as $T^* : Y' \to X'$ by $T^*(\psi) = \psi \circ T$, i.e., $(T^*\psi)(x) = \psi(Tx)$ for any $\psi \in Y'$ and $x \in X$.

**Theorem 4.24.** *$T^* \in \mathcal{B}(Y', X')$ and $\|T^*\| = \|T\|$.*

*Proof.* Easy to check $T^*$ is linear. Since $\|T^*\psi\| = \sup_{\|x\|=1}|(T^*\psi)(x)| = \sup_{\|x\|=1}|\psi(Tx)| \leqslant \sup_{\|x\|=1}\|\psi\|\|Tx\|_Y \leqslant \sup_{\|x\|=1}\|\psi\|\|T\|\|x\| = \|\psi\|\|T\|$ for any $\psi \in Y'$, we have $T^*$ is bounded and $\|T^*\| \leqslant \|T\|$.

On the other hand, let $x \in X$ and consider $Tx \in Y$. Then by previous corollary, there exists $\psi \in Y'$ with $\|\psi\| = 1$ such that $\psi(T(x)) = \|Tx\|$. So $\|Tx\| = \psi(T(x)) = (T^*\psi)(x) \leqslant |T^*\psi(x)| \leqslant \|T^*\psi\|\|x\| \leqslant \|T^*\|\|\psi\|\|x\| = \|T^*\|\|x\|$. Since $x \in X$ is arbitrary, $\|T\| \leqslant \|T^*\|$. □

**Proposition 4.25.** (a) $(T + S)^* = T^* + S^*$ for any $T, S \in \mathcal{B}(X, Y)$.

(b) $(aT)^* = aT^*$ for any $a \in \mathbb{R}$ and $T \in \mathcal{B}(X, Y)$.

(c) If $T \in \mathcal{B}(X, Y)$ and $T^{-1}$ exists **with** $T^{-1} \in B(Y, X)$, then $(T^*)^{-1}$ also exists with $(T^*)^{-1} \in \mathcal{B}(X', Y')$ and $(T^*)^{-1} = (T^{-1})^*$.

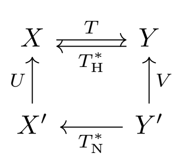
(d) $(T \circ S)^* = S^* \circ T^*$ for any $S \in \mathcal{B}(X, Y)$ and $T \in \mathcal{B}(Y, Z)$

*Proof.* (3) Let $\psi_1, \psi_2 \in Y'$. Since $T^{-1}$ exists, $T^*(\psi_1) = T^*(\psi_2)$ if and only if $(T^*\psi_1)(x) = (T^*\psi_2)(x)$ for any $x \in X$ if and only if $\psi_1(Tx) = \psi_2(Tx)$ for any $x \in X$ if and only if $\psi_1 \circ T = \psi_2 \circ T$ if and only if $\psi_1 = \psi_2$. So $T^*$ is 1-1.

Let $\varphi \in X'$. Then $\varphi \circ T^{-1} \in Y'$. Since $T^*(\varphi \circ T^{-1}) = \varphi \circ T^{-1} \circ T = \varphi$, we have $T^*$ is onto.

Hence $(T^*)^{-1}$ exists. Claim. $(T^*)^{-1} = (T^{-1})^*$. First, $(T^*)(T^{-1})^* = \mathrm{id}_{X'}$ if and only if $T^*(T^{-1})^*\varphi = \varphi$ for any $\varphi \in X'$ which is true since $T^*(T^{-1})^*\varphi = (\varphi \circ T^{-1}) \circ T = \varphi$. Similarly, we can prove $(T^{-1})^*(T^*) = \mathrm{id}_{Y'}$. □

**Remark.** Let $X, Y$ be Hilbert spaces and $T \in \mathcal{B}(X, Y)$. Define $T^*_{\mathrm{H}}$ by $\langle Tx, y\rangle = \langle x, T^*_{\mathrm{H}}y\rangle$, then $T^*_{\mathrm{H}} \in \mathcal{B}(Y, X)$ and define $T^*_{\mathrm{N}}$ by $T^*_{\mathrm{N}}(\psi) = \psi \circ T$, then $T^*_{\mathrm{N}} \in \mathcal{B}(Y', X')$.
Relation between $T^*_{\mathrm{H}}$ and $T^*_{\mathrm{N}}$.



Let $\psi \in Y'$. Then $\varphi := T_{\mathrm{N}}^* \psi \in X'$. By R.R.T., there exists a unique $y_\psi \in Y$ and $x_\varphi \in X$ such that $\psi(y) = \langle y, y_\psi \rangle$ and $\varphi(x) = \langle x, x_\varphi \rangle$ for any $y \in Y$ and $x \in X$, and $\|y_\psi\| = \|\psi\|$ and $\|x_\varphi\| = \|\varphi\|$. Define $U : X' \to X$ by $U(\varphi) = x_\varphi$ and $V : Y' \to Y$ by $V(\psi) = y_\psi$. Then $U, V$ are bijective, linear and isometry. So we can define $T^* : Y \to X$ by $T^*(y_\psi) = x_\varphi$. Then $T^* = U \circ T_{\mathrm{N}}^* \circ V^{-1}$ and $T^* \in \mathcal{B}(Y, X)$. Let $x \in X$, then $Tx \in Y$ and so $\langle Tx, y_\psi \rangle_Y = \psi(Tx) = (T_{\mathrm{N}}^* \psi)(x) = \varphi(x) = \langle x, x_\varphi \rangle_X = \langle x, T^* y_\psi \rangle_X$. Since the Hilbert adjoint is unique, $T_{\mathrm{H}}^* = T^* = U \circ T_{\mathrm{N}}^* \circ V^{-1}$.

**Example 4.26.** Let $X = l^p$ and $Y = l^q$, where $p, q \in \mathbb{R}^{>1}$ and $\frac{1}{p} + \frac{1}{q} = 1$. Recall $X' \cong Y$ and $Y' \cong X$.

Let $y = \{y_n\} \in l^q$ and $\varphi_y : l^p \to \mathbb{R}$ be defined by $\varphi_y(x) = \sum_{n=1}^\infty x_n y_n$.

Let $x = \{x_n\} \in l^p$ and $\varphi_x : l^q \to \mathbb{R}$ be defined by $\varphi_x(y) = \sum_{n=1}^\infty x_n y_n$.

If $T \in \mathcal{B}(X, Y)$, let $T : l^p \to l^q$ be given by $T(x) = y$, and $T^* \in \mathcal{B}(Y', X')$ with $T^* : (l^q)' \to (l^p)'$ by $T^*(\varphi_x) = \varphi_y$.

**Example 4.27.** Let $X = (\mathcal{C}[0,1], \|\cdot\|_\infty)$, $g \in X$ and define $T_g : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ by $T_g(f) = f \circ g$. Then $T_g \in \mathcal{B}(\mathcal{C}[0,1], \mathcal{C}[0,1])$. Let $x \in [0,1]$ and define $\varphi_x : \mathcal{C}[0,1] \to \mathbb{R}$ by $\varphi_x(f) = f(x)$. Then $\varphi_x \in (\mathcal{C}[0,1])'$. Since $T_g^*(\varphi_x)(f) = \varphi_x(T_g f) = \varphi_x(f \circ g) = (f \circ g)(x) = \varphi_{g(x)}(f)$ for any $f \in \mathcal{C}[0,1]$, we have $T_g^*(\varphi_x) = \varphi_{g(x)}$.

## 4.5 Reflexive spaces

Let $X$ be a NLS.

**Definition 4.28.** Given $x \in X$, we may define $l_x : X' \to \mathbb{R}$ by sending $\varphi$ to $\varphi(x)$.

**Lemma 4.29.** $l_x \in X''$ and $\|l_x\| = \|x\|$.

*Proof.* Easy to see $l_x$ is linear. Since $|l_x(\varphi)| = |\varphi(x)| \leqslant \|x\|\|\varphi\|$, we have $\|l_x\| \leqslant \|x\|$. In fact, $\|l_x\| = \sup_{\varphi \in X', \|\varphi\|=1} |l_x(\varphi)| = \sup_{\varphi \in X', \|\varphi\|=1} |\varphi(x)| = \|x\|$. □

**Definition 4.30.** The map $C : X \to X''$ defined by $C(x) = l_x$ is called the *canonical mapping* from $X$ to $X''$.

**Lemma 4.31.** The canonical mapping $C$ is an isometrical isomorphism ("$\cong$") between $X$ and $\mathrm{Im}(C)$, i.e., bijective linear and isometric.

*Proof.* Easy to check $C$ is linear. Since $\|C(x)\| = \|l_x\| = \|x\|$, it is an isometry and then 1-1. Clearly, it is onto. □

**Definition 4.32.** $X$ is called *reflexive* if $\mathrm{Im}(C) = X''$.

**Remark.** (a) If $X$ is reflexive, then $X \cong X''$. But if $X \cong X'$, $X$ may be not reflexive. Counterexample, R. (James (1951)).

(b) If $X$ is reflexive, then $X$ is complete.

**Example 4.33.** $\mathbb{R}^n$, $l^p$, $L^p[0,1]$ with $p \in \mathbb{R}^{>1}$, finite dimensional NLS's, and Hilbert spaces are reflexive.

**Example 4.34.** $l^1$, $l^\infty$. $L^1[0,1]$, $L^\infty[0,1]$ and $\mathcal{C}[0,1]$ are not reflexive.

**Theorem 4.35.** *Every Hilbert space is reflexive.*

*Proof.* Let $H$ be a real Hilberl space. Want to show for any $l \in H''$, there exists $x \in H$ such that $C(x) = l_x = l$.
Recall for $\varphi \in H'$, by R.R.T., there exists a unique $x_\varphi \in H$ such that $\varphi(x) = \langle x, x_\varphi \rangle_H$ for any $x \in X$ and $\|\varphi\| = \|x_\varphi\|$. We can define $U : H' \to H$ by $U(\varphi) = x_\varphi$, then $U$ is linear, bijective and isometry.
Claim $H' = \mathcal{B}(H, \mathbb{R})$ is a Hilbert space w.r.t. $\langle \varphi_1, \varphi_2 \rangle_{H'} = \langle U\varphi_1, U\varphi_2 \rangle_H$. First, since $\mathbb{R}$ is complete, $H'$ is complete. Secondly, for any $\varphi_1, \varphi_2 \in H'$ and $a \in \mathbb{R}$, since $H$ is Hilbert and $U$ is linear,

(a) $\langle \varphi_1, \varphi_1 \rangle_{H'} = \langle U\varphi_1, U\varphi_1 \rangle_H \geqslant 0$.

(b) $\langle \varphi_1, \varphi_2 \rangle_{H'} = \langle U\varphi_1, U\varphi_2 \rangle_H = \langle U\varphi_2, U\varphi_1 \rangle_H = \langle \varphi_2, \varphi_1 \rangle_{H'}$.

(c) the linearity in both arguments.

So $H'$ is Hilbert. Then by R.R.T., for any $l \in (H')' = H''$, there exists a unique $\varphi_l \in H'$ such that $l(\varphi) = \langle \varphi, \varphi_l \rangle_{H'} = \langle U\varphi, U\varphi_l \rangle_H = \langle x_\varphi, x_{\varphi_l} \rangle_H = \langle x_{\varphi_l}, x_\varphi \rangle_H = \varphi(x_{\varphi_l}) = l_{x_{\varphi_l}}(\varphi)$ for any $\varphi \in H'$. Hence $l = l_{x_{\varphi_l}}$ and so $C$ is onto. □

**Recall 4.36.** $(X, d)$ is separable if $X$ has a countable dense subset. For example, $l^\infty$ is not separable.

**Theorem 4.37.** *If $X'$ is separable, so is $X$.*

*Proof.* Since $X'$ is separable, $\partial B_1(\mathbf{0}) = \{\varphi \in X' \mid \|\varphi\| = 1\} \subseteq X'$ has a countable dense subset, say $\{\varphi_n\} \subseteq \partial B_1(\mathbf{0})$. For each $n \in \mathbb{N}$, since $\sup_{x \in X, \|x\|=1} |\varphi_n(x)| = \|\varphi_n\| = 1$, there exists $x_n \in X$ with $\|x_n\| = 1$ such that $|\varphi_n(x_n)| \geqslant \frac{1}{2}$. Let $Y = \overline{\operatorname{span}\{x_n\}_{n=1}^\infty}$. Claim $Y = X$ and hence $X$ is separable since it has a countable dense subset $\{\sum_{n=1}^{\text{finite}} a_n x_n \mid a_n \in \mathbb{Q}, \forall n\}$.
Suppose $Y \neq X$, then $Y \leqslant X$ is closed, by HW#1, there exists $\varphi \in X'$ and $\|\varphi\| = 1$ such that $\varphi|_Y = 0$. Note $\varphi \in \partial B_1(\mathbf{0})$ and $\frac{1}{2} \leqslant |\varphi_n(x_n)| = |\varphi_n(x_n) - \varphi(x_n)| = |(\varphi_n - \varphi)(x_n)| \leqslant \|\varphi_n - \varphi\| \|x_n\| = \|\varphi_n - \varphi\|$ for any $n \in \mathbb{N}$, contradicting the fact that $\{\varphi_n\}$ is dense in $\partial B_1(\mathbf{0})$. Thus, $Y = X$. □

**Remark.** The converse is not true, e.g. $l^1$ is separable, but $l^\infty$ is not.

**Corollary 4.38.** If $X$ is separable but $X'$ is not, then $X$ is not reflexive.

*Proof.* Suppose $X$ is reflexive, then $X \cong X''$. Also, since $X$ is separable, $X''$ is separable. So $X'$ is separable, a contradiction. □

**Example 4.39.** $(\mathcal{C}[0,1], \|\cdot\|_\infty)$ is not reflexive. We know $\mathcal{C}^\infty[0,1]$ is separable.
Claim $(\mathcal{C}[0,1])'$ is not separable. It suffices to construct uncountably many disjoint open balls in $(\mathcal{C}[0,1])'$. Let $x \in [0,1]$ and define $\varphi_x : \mathcal{C}[0,1] \to \mathbb{R}$ by $\varphi_x(f) = f(x)$. Then $\varphi_x \in (\mathcal{C}[0,1])'$ with $\|\varphi_x\| = 1$. In addition, for $x \neq y \in [0,1]$, claim $\|\varphi_x - \varphi_y\| = 2$. Note $\|(\varphi_x - \varphi_y)f\| = |\varphi_x(f) - \varphi_y(f)| \leqslant |f(x) - f(y)| \leqslant 2\|f\|_\infty$ and we can find a set of spline functions to make its norm $\geqslant 1$.

## 4.6 Uniform Boundedness Principal

Let $X, Y$ be NLS's.

**Definition 4.40.** Let $(X, d)$ be a metric space and $A \subseteq X$. $A$ is of *first category* in $X$ if $A$ can be written as a countable union of nowhere dense sets. Otherwise, $A$ is of *second category* in $X$.

**Remark.** Finite union of nowhere dense sets is still nowhere dense. But not necessarily for countable union, e.g. $\mathbb{Q} = \bigcup_{n=1}^{\infty} F_n$, where $F_n = \{\frac{m}{n} \mid m \in \mathbb{Z}\}$ for any $n \in \mathbb{N}$.

**Theorem 4.41** (Baire Category Theorem). *Let $(X, d)$ be a complete metric space. Then $X$ is of second category in itself.*

**Theorem 4.42** (Uniform Boundedness Principal). *Let $X$ be Banach and $\{T_\alpha\}_{\alpha \in I} \subseteq \mathcal{B}(X, Y)$. If for any $x \in X$, $\{T_\alpha x\}_{\alpha \in I}$ is bounded, then $\{\|T_\alpha\|\}_{\alpha \in I}$ is also bounded.*

*Proof.* Let $A_N = \{x \in X \mid \|T_\alpha x\|_Y \leqslant N, \forall \alpha \in I\}$ for $N \in \mathbb{N}$. Fix $N \in \mathbb{N}$ and let $\{x_n\} \subseteq A_N$ such that $x_n \to x$ as $n \to \infty$. Since $\{T_\alpha\}_{\alpha \in I}$ and $\|\cdot\|$ are continuous, $\|T_\alpha x\|_Y = \|T_\alpha(\lim_{n\to\infty} x_n)\|_Y = \|\lim_{n\to\infty} T_\alpha x_n\|_Y = \lim_{n\to\infty} \|T_\alpha x_n\|_Y \leqslant N$ for $\alpha \in I$ since $\{x_n\} \subseteq A_N$. So $x \in A_N$ and then $A_N \subseteq X$ is closed.

Let $x \in X$. Since $\{T_\alpha x\}_{\alpha \in I}$ is bounded, there exists $N = N_x \in \mathbb{N}$ such that $\|T_\alpha x\|_Y \leqslant N$ for $\alpha \in I$. So $x \in A_N$ and hence $X = \bigcup_{N=1}^{\infty} A_N$. Since $X$ is complete, by Baire Category theorem, there exists $N_0 \in \mathbb{N}$ such that $\operatorname{Int}(A_{N_0}) = \operatorname{Int}(\overline{A}_{N_0}) \neq \emptyset$ since $A_{N_0}$ is closed. So there exists $x_0 \in A_{N_0}$ and $r_0 > 0$ such that $B_{r_0}(x_0) \subseteq A_{N_0}$. Let $z_x = x_0 + \frac{r_0}{2} \frac{x}{\|x_0\|_X}$. Then $z_x \in B_{r_0}(x_0) \subseteq A_{N_0}$. So $\|T_\alpha x\|_Y = \left\|T_\alpha(\frac{2\|x_0\|_X}{r_0}(z_x - x_0))\right\|_Y = \frac{2\|x_0\|_X}{r_0}\|T_\alpha z_x - T_\alpha x_0\|_Y \leqslant \frac{2\|x_0\|_X}{r_0} \cdot 2N_0 = \frac{4\|x_0\|_X N_0}{r_0} =: M$ for $\alpha \in I$. Thus, $\|T_\alpha\| \leqslant M$ for $\alpha \in I$. □

**Corollary 4.43.** If $\sup_{\alpha \in I}\|T_\alpha\| = \infty$, then there exists $x_0 \in X$ such that $\sup_{\alpha \in I}\|T_\alpha(x_0)\|_Y = \infty$.

**Corollary 4.44** (Banach-Steinhaus Theorem). Let $X$ be Banach, $\{T_n\}_{n \in \mathbb{N}} \subseteq \mathcal{B}(X, Y)$, $\{T_n x\}$ be convergent for any $x \in X$ and $T : X \to Y$ given by $Tx = \lim_{n\to\infty} T_n(x)$. Then $T \in \mathcal{B}(X, Y)$.

*Proof.* Let $a_1, a_2 \in \mathbb{R}$ and $x_1, x_2 \in X$. Since $\{T_n\}_{n \in \mathbb{N}} \subseteq \mathcal{B}(X, Y)$ and $\{T_n x\}$ is convergent for any $x \in X$, $T(a_1 x_1 + a_2 x_2) = \lim_{n\to\infty} T_n(a_1 x_1 + a_2 x_2) = \lim_{n\to\infty}(a_1 T_n(x_1) + a_2 T_n(x_2)) = a_1 \lim_{n\to\infty} T_n(x_1) + a_2 \lim_{n\to\infty} T_n(x_2) = a_1 T(x_1) + a_2 T(x_2)$. So $T$ is linear.

Let $x \in X$. Since $\{T_n x\}$ converges, $\{\|T_n x\|\}$ is bounded. Then by Uniform Boundedness Principal, $\{\|T_n\|\}$ is bounded. So there exists $M > 0$ such that $\|T_n\| \leqslant M$ for $n \in \mathbb{N}$. Since $\{T_n\}$ and $\|\cdot\|$ are continuous, $\|Tx\| = \|\lim_{n\to\infty} T_n x\| = \lim_{n\to\infty}\|T_n x\| \leqslant \lim_{n\to\infty}\|T_n\|\|x\| \leqslant M\|x\|$. Thus, $\|T\| \leqslant M$. □

**Remark.** We may not have $\|T_n\| \to \|T\|$ as $n \to \infty$.

**Example 4.45.** Consider $c_{00} = \{\{x_n\} \subseteq l^\infty \mid \{x_n\} \text{ has finitely many nonzero terms}\}$. Then $(c_{00}, \|\cdot\|_\infty)$ is not complete.

For $n \in \mathbb{N}$, define $\varphi_n : c_{00} \to \mathbb{R}$ by $\varphi_n(\{x_m\}) = n x_n$. Then $\{\varphi_n\} \subseteq \mathcal{B}(c_{00}, \mathbb{R}) = c_{00}'$.

Let $x = \{x_m\} \in c_{00}$. Then there exists $N_x \in \mathbb{N}$ such that $x_m = 0$ for any $m > N_x$. So $|\varphi_n(\{x_m\})| = n|x_n| \leqslant N_x \sup_{1 \leqslant m \leqslant N_x}|x_m| =: M_x$ for $n \in \mathbb{N}$. On the other hand, since $|\varphi_n(\{x_m\})| = n|x_n| \leqslant n\|\{x_m\}\|_\infty$ and $|\varphi_n(\{0, 0, \ldots, 0, \underset{\uparrow n^{\text{th}}}{1}, 0, \cdots\})| = n$ with $\|\{0, 0, \ldots, 0, \underset{\uparrow n^{\text{th}}}{1}, 0, \cdots\}\|_\infty = 1$, we have $\|\varphi_n\| = n$ for $n \in \mathbb{N}$. So by Uniform Boundedness Principal, $(c_{00}, \|\cdot\|_\infty)$ cannot be complete.

**Theorem 4.46.** *Let $Y \subseteq X$. If $\{\varphi(y) \mid y \in Y\}$ is bounded in $\mathbb{R}$ for any $\varphi \in X'$, then $Y$ is bounded.*

*Proof.* By assumption, $\{l_y(\varphi)\}_{y\in Y} = \{\varphi(y) \mid y \in Y\}$ is bounded with $\{l_y\}_{y\in Y} \subseteq \mathcal{B}(X', \mathbb{R})$. Since $X'$ is Banach, by Uniform Boundedness Principal, we have $\{\|y\|\}_{y\in Y} = \{\|l_y\|\}_{y\in Y}$ is bounded. $\square$

## 4.7 Weak Convergence

Let $X$ be a NLS.

**Definition 4.47.** Let $\{x_n\} \subseteq X$. We say $\{x_n\}$ *converges weakly* to $x \in X$, denoted $x_n \rightharpoonup x$ if $|\varphi(x_n) - \varphi(x)| \to 0$ as $n \to \infty$ for $\varphi \in X'$.

**Theorem 4.48.** *Let $\{x_{n_k}\} \subseteq \{x_n\} \subseteq X$ and $x_n \rightharpoonup x \in X$ as $n \to \infty$. Then*

*(a) $x$ is unique;*

*(b) $x_{n_k} \to x$ as $k \to \infty$;*

*(c) $\{x_n\}$ is bounded.*

*Proof.* (a) Suppose $x_n \rightharpoonup y \in X$ as $n \to \infty$. Let $\varphi \in X'$. Then $\varphi(x_n) \to \varphi(x)$ and $\varphi(x_n) \to \varphi(y)$ as $n \to \infty$. Since convergent sequence has a unique limit, $\varphi(x) = \varphi(y)$, i.e., $\varphi(x - y) = 0$. Since $\varphi \in X'$ is arbitrary, $x = y$.

(b) Since $x_n \rightharpoonup x$, $\varphi(x_n) \to \varphi(x)$ as $n \to \infty$ for $\varphi \in X'$. So $\varphi(x_{n_k}) \to \varphi(x)$ as $k \to \infty$ for $\varphi \in X'$. Hence $x_{n_k} \rightharpoonup x$ as $k \to \infty$.

(c) Since $x_n \rightharpoonup$, $\{\varphi(x_n)\}$ converges in $\mathbb{R}$ for $\varphi \in X'$. So $\{\varphi(x_n)\}$ is bounded in $\mathbb{R}$ for $\varphi \in X'$. Hence by a corollary of Uniform Boundedness Principal, we have $\{x_n\}$ is bounded. $\square$

**Theorem 4.49.** *Let $\{x_n\} \subseteq X$.*

*(a) If $x_n \to x \in X$, then $x_n \rightharpoonup x$ as $n \to \infty$.*

*(b) The converse of (a) is not true in general.*

*(c) If $\dim_{\mathbb{R}} X < \infty$, $x_n \to x \in X$ if and only if $x_n \rightharpoonup x \in X$ as $n \to \infty$.*

*Proof.* (a) Since $x_n \to x \in X$, $\|x_n - x\| \to 0$ as $n \to \infty$. So for any $\varphi \in X'$, $|\varphi(x_n) - \varphi(x)| = |\varphi(x_n - x)| \leqslant \|\varphi\|\|x_n - x\| \to 0$ as $n \to \infty$.

(b) Counter example. Let $X$ be a Hilbert space with an orthonormal basis $\{e_n\}$. Claim. $e_n \rightharpoonup \mathbf{0}$ but $e_n \nrightarrow \mathbf{0}$ as $n \to \infty$. Since $\|e_m - e_n\| = \sqrt{2}$ for all $1 \leqslant m < n < \infty$, we have $e_n \nrightarrow \mathbf{0}$ as $n \to \infty$. Let $\varphi \in X'$. By R.R.T., there exists a unique $z \in X$ such that $\varphi(x) = \langle x, z\rangle$ for any $x \in X$. by Bessel's inequality, $\sum_{n=1}^{\infty}\left|\langle e_n, z\rangle^2\right| \leqslant \|z\|^2 < \infty$, i.e., $\sum_{n=1}^{\infty}|\langle e_n, z\rangle|^2$ converges. So $\langle e_n, z\rangle \to 0$ as $n \to \infty$. Hence $\varphi(e_n) = \langle e_n, z\rangle \to 0 = \varphi(\mathbf{0})$ as $n \to \infty$. Thus, $e_n \rightharpoonup \mathbf{0}$ as $n \to \infty$.

(c) It suffices to show if $x_n \rightharpoonup x \in X$, then $x_n \to x$. Let $\dim_{\mathbb{R}} X = m$ and $\{e_i\}_{i=1}^m \subseteq X$ be a basis. Then $x = \sum_{i=1}^m a_i e_i$ for some $a_1, \ldots, a_m \in \mathbb{R}$ and for $n \in \mathbb{N}$, $x_n = \sum_{i=1}^m a_i^{(n)}$ for some $a_1^{(n)}, \ldots, a_m^{(n)} \in \mathbb{R}$. By Exercise#6 in Homework 1, there exists $\varphi_i \in X'$ for $i = 1, \ldots, m$ such that $\varphi_i(e_j) = \delta_{ij}$ for $1 \leqslant j \leqslant m$. Then $\varphi_i(x) = \varphi_i(\sum_{j=1}^m a_j e_j) = a_i$ and $\varphi_i(x_n) = \varphi_i(\sum_{j=1}^m a_j^{(n)} e_j) = a_i^{(n)}$ for $i = 1, \ldots, m$. Since $x_n \rightharpoonup x$, $a_i^{(n)} = \varphi_i(x_n) \to \varphi_i(x) = a_i$ as $n \to \infty$ for $i = 1, \ldots, m$. So $\|x_n - x\| = \left\|\sum_{i=1}^m (a_i^{(n)} - a_i)e_i\right\| \leqslant \sum_{i=1}^m \left|a_i^{(n)} - a_i\right|\|e_i\| \to 0$ as $n \to \infty$. $\square$

**Theorem 4.50.** *Let $A \subseteq X$ be closed and convex. Let $\{x_n\} \subseteq A$ with $x_n \rightharpoonup x$. Then $x \in A$.*

*Proof.* Suppose $x \notin A$. Since $A$ is closed, $x \in A^c$ open. So there exists $\epsilon > 0$ such that $B_\epsilon(x) \subseteq A^c$, i.e., $B_\epsilon(x) \cap A = \emptyset$. Also, since $B_\epsilon(x)$ is open, by Geometric Hahn-Banach theorem, there exists $0 \neq \varphi \in X'$ and $c \in \mathbb{R}$ such that $\varphi(a) \leqslant c \leqslant \varphi(b)$ for $a \in A$ and $b \in B_\epsilon(x)$. So $\varphi(a) \leqslant \varphi(x + \epsilon y) = \varphi(x) + \epsilon\varphi(y)$ for $a \in A$ and $y \in B_1(\mathbf{0})$. Since $\{x_n\} \subseteq A$, $\varphi(x_n) - \varphi(x) \leqslant \epsilon\varphi(y)$ for $n \in \mathbb{N}$ and $y \in B_1(\mathbf{0})$.

Since $0 \neq \varphi \in X'$, there exists $0 \neq z \in X$ such that $\varphi(z) \neq 0$. Also, since $\varphi$ is linear, $\varphi(x_n) - \varphi(x) \leqslant \pm\epsilon\varphi(\frac{z}{\|z\|}) = \pm\frac{\epsilon}{\|z\|}|\varphi(z)|$. Let $n \to \infty$, since $x_n \rightharpoonup x$, we have $0 \leqslant -\frac{\epsilon}{\|z\|}|\varphi(z)|$, a contradicton. □

**Corollary 4.51.** Let $f : X \to \mathbb{R}$ be continuous and convex. If $\{x_n\} \subseteq X$ such that $x_n \rightharpoonup x$ as $n \to \infty$, then $f(x) \leqslant \liminf_{n\to\infty} f(x_n)$.

*Proof.* Let $M = \liminf_{n\to\infty} f(x_n)$. Let $\epsilon > 0$. Set $A_\epsilon = \{x \in X \mid f(x) \leqslant M + \epsilon\}$, then $A_\epsilon$ is closed since $f$ is continuous, and is convex since $f$ is convex. Since $\liminf_{n\to\infty} f(x_n) = M$, there exists $\{x_{n_k}\} \subseteq \{x_n\}$ such that $f(x_{n_k}) \leqslant M + \epsilon$ for $k \in \mathbb{N}$. So $\{x_{n_k}\} \subseteq A_\epsilon$. Also, since $x_n \rightharpoonup x$ as $n \to \infty$, $x_{n_k} \rightharpoonup x$ as $k \to \infty$. Hence, by the previous theorem, $x \in A_\epsilon$. So $f(x) \leqslant M + \epsilon$. Since $\epsilon > 0$ is arbitrary, $f(x) \leqslant M$. □

**Definition 4.52.** $Y \subseteq X$ is called *weakly sequentially compact* (W.S.C.) if every sequence in $Y$ has a subsequence that converges weakly in $Y$.

**Lemma 4.53.** If $Y$ is weakly sequentially compact, then $Y$ is bounded.

*Proof.* Suppose not, then there is $\{y_n\} \subseteq Y$ such that $\|y_n\| \geqslant n$ for each $n \in \mathbb{N}$. Since $Y$ is weakly sequentially compact, there is $\{y_{n_k}\} \subseteq y_n$ such that $\{y_{n_k}\}$ converges weakly. So $\{y_{n_k}\}$ is bounded, contradicted with $\|y_{n_k}\| \geqslant n_k$. □

**Theorem 4.54** (Banach-Alaoglu)**.** *Let $X$ be reflexive. Then $\overline{B}_1(\mathbf{0}) \subseteq X$ is weakly sequentially compact.*

*Proof.* Let $\{y_n\} \subseteq \overline{B}_1(\mathbf{0})$ and $Y = \overline{\operatorname{span}\{y_n\}}$. Then $Y \leqslant X$ is closed and separable with a countable dense subset $\{\sum_{i=1}^{\text{finite}} a_i y_i \mid a_i \in \mathbb{Q}, \forall i\} \cong \mathbb{Q}^{\mathbb{N}}$. Since $X$ is reflexive and $Y \leqslant X$ is closed, by Exercise#2 in Homework 3, $Y$ is reflexive, i.e., $Y \cong Y''$. Also, since $Y$ is separable, we have $Y''$ is separable.

Since $Y'$ is isomorphic to a subspace of $Y''$, $Y'$ is separable. So there exists $A = \{\varphi_m \in Y' \mid m \in \mathbb{N}\}$ countable dense in $Y'$. For a fixed $m \in \mathbb{N}$, $\{\varphi_m(y_n)\}_{n=1}^\infty$ is bounded in $\mathbb{R}$ since $|\varphi_m(y_n)| \leqslant \|\varphi_m\|\|y_n\| \leqslant \|\varphi_m\| < \infty$ for $n \in \mathbb{N}$. So there exists $\{y_{n_k}\} \subseteq \{y_n\}$ such that $\{\varphi_1(y_{n_k})\}_{k=1}^\infty$ converges. Then there exists $\{y_{n_{k_l}}\} \subseteq \{y_{n_k}\}$ such that $\{\varphi_2(y_{n_{k_l}})\}_{l=1}^\infty$ converges and clearly $\{\varphi_1(y_{n_{k_l}})\}_{l=1}^\infty$ converges. By diagonal argument, there exists $\{y_{n_k}\} \subseteq \{y_n\}$ such that $\{\varphi_m(y_{n_k})\}_{k=1}^\infty$ converges for all $m \in \mathbb{N}$.

Define $l : A \to \mathbb{R}$ by $l(\varphi_m) = \lim_{k\to\infty} \varphi_m(y_{n_k})$. Easy to check $l$ is linear. Since $|l(\varphi_m)| = |\lim_{k\to\infty} \varphi_m(y_{n_k})| = \lim_{k\to\infty}|\varphi_m(y_{n_k})| \leqslant \lim_{k\to\infty}\|\varphi_m\|\|y_{n_k}\| \leqslant \|\varphi_m\|$, $l$ is bounded. Hence $l \in Y''$. Since $Y$ is reflexive, there exists $y \in Y$ such that $l(\varphi) = C(y)(\varphi) = l_y(\varphi) = \varphi(y)$ for any $\varphi \in Y'$. So $\lim_{k\to\infty} \varphi_m(y_{n_k}) = l(\varphi_m) = \varphi_m(y)$ for $m \in \mathbb{N}$, i.e., $\varphi_m(y_{n_k}) \to \varphi_m(y)$ as $k \to \infty$. Since $A$ is dense in $Y'$, $\varphi(y_{n_k}) \to \varphi(y)$ for all $\varphi \in Y'$. Thus, $y_{n_k} \rightharpoonup y \in Y$ as $k \to \infty$. □

**Theorem 4.55.** *If*

*(a)* $\{\|x_n\|\}$ *bounded and,*

*(b) for any* $\varphi \in A$, $\varphi(x_n) \to \varphi(x)$ *as* $n \to \infty$,

*then* $x_n \rightharpoonup x$ *as* $n \to \infty$.

*Proof.* By Problem#1 in Homework 4. □

**Remark.** (a) If $X$ is Banach and $\overline{B}_1(\mathbf{0}) \subseteq X$ is weakly sequentially compact, then $X$ is reflexive.

(b) If $K \subseteq X$ is closed, bounded and convex, then $K$ is weakly sequentially compact.

(c) If $X$ is reflexive, then any bounded sequence in $X$ must have a weakly convergent subsequence.

**Corollary 4.56.** Let $X$ be reflexive, $\emptyset \neq K \subseteq X$ closed and convex. If $y \notin K$, then there exists $x_0 \in X$ such that $\|y - x_0\| = d(y, K)$.

*Proof.* □

**Definition 4.57.** A sequence $\{T_n\} \subseteq \mathcal{B}(X, Y)$ is called

(a) *uniformly convergent* if $\{T_n\}$ converges in $\mathcal{B}(X, Y)$, i.e., there exists $T \in \mathcal{B}(X, Y)$ such that $\|T_n - T\| \to 0$ as $n \to 0$;

(b) *strongly convergent* if $\{T_n x\}$ converges strongly in $Y$ for $x \in X$, i.e., there exists $T : X \to Y$ such that $\|T_n x - T_n x\| \to 0$ as $n \to \infty$ for $x \in X$;

(c) *weakly convergent* if $\{T_n x\}$ converges weakly in $Y$ for $x \in X$, i.e., there exists $T : X \to Y$ such that $|\varphi(T_n x) - \varphi(Tx)| \to 0$ as $n \to 0$ for $x \in X$ and $\varphi \in Y'$.

**Remark.** Uniform convergence implies strong convergence and strong convergence implies weakly convergence.

**Example 4.58.** (a) For $n \in \mathbb{N}$, define $T_n : l^2 \to l^2$ by $x = \{x_n\} \mapsto \{\underbrace{0, \ldots, 0}_{n \text{ times}}, x_{n+1}, x_{n+2}, \cdots\}$. Easy to show $\{T_n\} \subseteq \mathcal{B}(l^2, l^2)$ and $\|T_n\| = 1$ for $n \in \mathbb{N}$. So $\{T_n\}$ does not converge to $\mathbf{0}$ uniformly.

For any $x = \{x_n\} \in l^2$, $\|T_n x - \mathbf{0}x\|_2 = \|T_n x\|_2 = \left(\sum_{i=n+1}^{\infty} x_i^2\right)^{1/2} \to 0$ as $n \to \infty$. So $T_n \to \mathbf{0}$ strongly.

(b) For $n \in \mathbb{N}$, define $T_n : l^2 \to l^2$ by $x = \{x_n\} \mapsto \{\underbrace{0, \ldots, 0}_{n \text{ times}}, x_1, x_2, x_3, \cdots\}$. Easy to check $\{T_n\} \subseteq \mathcal{B}(l^2, l^2)$ and $\|T_n\| = 1$.

Let $\varphi \in (l^2)'$. Since $l^2$ is Hilbert, by R.R.T., there exists $y = \{y_n\} \in l^2$ such that $\varphi(x) = \langle x, y\rangle = \sum_{i=1}^{\infty} x_i y_i$ for $x \in l^2$. So by Cauchy Schwarz inequality, $|\varphi(T_n x)| = |\varphi(\{\underbrace{0, \ldots, 0}_{n \text{ times}}, x_1, x_2, x_3, \cdots\})| = |\sum_{i=1}^{\infty} x_i y_{n+i}| \leqslant (\sum_{i=1}^{\infty} |x_i|^2)^{\frac{1}{2}} (\sum_{i=n+1}^{\infty} |y_i|^2)^{\frac{1}{2}} = \|x\|_2 (\sum_{i=n+1}^{\infty} |y_i|^2)^{\frac{1}{2}} \to 0$ as $n \to \infty$ for $x \in l^2$. So $T_n \to T$ weakly.

In addition, note for $x = \{1, 0, 0, \cdots\} \in l^2$, $\|T_n x - T_m x\| = \sqrt{2}$ for $m, n \in \mathbb{N}$ with $m \neq n$. So $T_n$ does not converges to $\mathbf{0}$ strongly.

**Lemma 4.59.** If $X$ is Banach and $T$ is a strong convergence limit of $\{T_n\} \subseteq \mathcal{B}(X,Y)$, then $T \in \mathcal{B}(X,Y)$.

*Proof.* It direcly follows from Banach-Steinhaus Theorem. $\square$

**Remark.** Note $\|T_n\|$ may not converge to $\|T\|$.

**Example 4.60.** For $n \in \mathbb{N}$, let $T_n : c_{00} \to l^2$ be given by $x = \{x_n\} \mapsto \{x_1, 2x_2, 3x_3, \ldots, nx_n, x_{n+1}, x_{n+2}, \cdots\}$. Claim. $T_n \to T$ strongly, where $T : c_{00} \to l^2$ is defined by $x = \{x_n\} \mapsto \{nx_n\}$.

Note $\|T_n x - Tx\|_2 = ||\{\underbrace{0, \ldots, 0}_{n \text{ times}}, -nx_{n+1}, -(n+1)x_{n+2}, \cdots\}||_2 \to 0$ for $x = \{x_n\} \in c_{00}$. So $T_n \to T$ strongly. But $T \notin \mathcal{B}(X,Y)$ since it is unbounded.

**Definition 4.61.** A sequence $\{\varphi_n\} \subseteq X'$ is

(a) *strongly convergent* to $\varphi \in X'$ if $\|\varphi_n - \varphi\|_{X'} \to 0$ as $n \to \infty$.

(b) *weak* convergent* to $\varphi \in X'$, denoted $\varphi_n \xrightarrow{w^*} \varphi$ if $|\varphi_n(x) - \varphi(x)| \to 0$ as $n \to \infty$ for any $x \in X$.

**Proposition 4.62.** Let $\{\varphi_n\} \subseteq X'$.

(a) If $\varphi_n \rightharpoonup \varphi$, then $\varphi_n \xrightarrow{w^*} \varphi$.

(b) If $\varphi_n \xrightarrow{w^*} \varphi$ and $X$ is reflexive, then $\varphi_n \rightharpoonup \varphi$.

*Proof.* (a) Note for $x \in X$, $|\varphi_n(x) - \varphi(x)| = |l_x(\varphi_n) - l_x(\varphi)| \to 0$ since $\varphi_n \rightharpoonup \varphi$.

(b) Let $l \in X''$. Since $X$ is reflexive, there exists $x \in X$ such that $l = C(x) = l_x$. So $|l(\varphi_n) - l(\varphi)| = |l_x(\varphi_n) - l_x(\varphi)| = |\varphi_n(x) - \varphi(x)| \to 0$ since $\varphi_n \xrightarrow{w^*} \varphi$. $\square$

**Definition 4.63.** Let $Y \subseteq X'$ be *weak* sequentially compact* if for any $\{\varphi_n\} \subseteq Y$, there exists $\{\varphi_{n_k}\} \subseteq \{\varphi_n\}$ such that $\varphi_{n_k} \xrightarrow{w^*} \varphi$.

**Theorem 4.64.** *Let $X$ be separable. Then $\overline{B}_1^{\|\cdot\|_{X'}}(\mathbf{0}) \subseteq X'$ is weak* sequentially compact.*

*Proof.* Let $\{\varphi_n\} \subseteq \overline{B}_1^{\|\cdot\|_{X'}}(\mathbf{0})$. Let $\{x_m\} \subseteq X$ be countable and dense. For a fixed $m \in \mathbb{N}$, $\{\varphi_n(m)\}_{n=1}^{\infty}$ is bounded in $\mathbb{R}$ since $|\varphi_n(x_m)| \leqslant \|\varphi_n\|\|x_m\| \leqslant \|x_m\|$ for $n \in \mathbb{N}$. So there exists $\{\varphi_{n_k}\} \subseteq \{\varphi_n\}$ such that $\{\varphi_{n_k}(x_1)\}$ converges. Then there exists $\{\varphi_{n_{k_l}}\} \subseteq \{\varphi_{n_k}\}$ such that $\{\varphi_{n_{k_l}}\} \subseteq \{\varphi_{n_k}\}$ such that $\varphi_{n_{k_l}}(x_2)$ converges. By "diagonal argument", we can find $\{\varphi_{n_k}\} \subseteq \{\varphi_n\}$ such that $\{\varphi_{n_k}(x_m)\}$ converges for all $m \geqslant 1$.

Define $\varphi : \{x_m\} \to \mathbb{R}$ by $\varphi(x_m) = \lim_{k\to\infty} \varphi_{n_k}(x_m)$. By a similar result to problem#1 in Homework 4, we can extend from $\{x_m\}$ to $X$ by defining $\varphi : X \to \mathbb{R}$ by $x \mapsto \lim_{k\to\infty} \varphi_{n_k}(x)$. So $\varphi_{n_k} \xrightarrow{w^*} \varphi$.

In addition, since $|\varphi(x)| = |\lim_{k\to\infty} \varphi_{n_k}(x)| \leqslant \lim_{k\to\infty} \|\varphi_{n_k}\|\|x\| \leqslant \|x\|$ for $x \in X$. So $\|\varphi\| \leqslant 1$ and thus $\varphi \in \overline{B}_1^{\|\cdot\|_{X'}}(\mathbf{0})$. $\square$

## 4.8 Open mapping and closed graph theorem

Let $X$ and $Y$ be NLS's.

**Definition 4.65.** Let $X$ and $T$ metric spaces. $T: X \to Y$ is called an *opening mapping* if for any $A \subseteq X$ open, $T(A) \subseteq Y$ open.

**Remark.** A continuous mapping may not be an open mapping. For example, consider $f: (0,1) \to \mathbb{R}$ given by $f(x) = 1$, and $\sin(\cdot) : (0, 2\pi) \to \mathbb{R}$. Then $f(0,1)$ and $\sin(0, 2\pi)$ are not open in $\mathbb{R}$, respectively.

**Lemma 4.66.** Let $X$ and $Y$ be Banach and $T \in \mathcal{B}(X,Y)$. If $T$ is onto, then $T(B_1(\mathbf{0}))$ contains an open ball that includes $\mathbf{0} \in Y$.

*Proof.* First note for $x \in X$, take $k \in \mathbb{N}$ with $k > 2\|x\|$, i.e., $\left\|\frac{x}{k}\right\| < \frac{1}{2}$, i.e., $x \in kB_{\frac{1}{2}}(\mathbf{0})$. So there exists $\epsilon_0 > 0$ and $y_0 \in \overline{T(B_{\frac{1}{2}}(\mathbf{0}))}$ such that $B_{\epsilon_0}(y_0) \subseteq \overline{T(B_{\frac{1}{2}}(\mathbf{0}))}$, i.e., $B_{\epsilon_0}(\mathbf{0}) = B_{\epsilon_0}(y_0) - \{y_0\} \subseteq \overline{T(B_{\frac{1}{2}}(\mathbf{0}))} - \{y_0\}$. Claim. $\overline{T(B_{\frac{1}{2}}(\mathbf{0}))} - \{y_0\} \subseteq \overline{T(B_1(\mathbf{0}))}$.

Proof of the Claim. Let $y \in \overline{T(B_{\frac{1}{2}}(\mathbf{0}))} - \{y_0\}$. Then $y + y_0 \in \overline{T(B_{\frac{1}{2}}(\mathbf{0}))}$. So there exists $\{T(w_n)\} \subseteq T(B_{\frac{1}{2}}(\mathbf{0}))$ such that $T(w_n) \to y + y_0$ as $n \to \infty$. Also, since $y_0 \in \overline{B_{\frac{1}{2}}(\mathbf{0})}$, there exists $\{T(z_n)\} \subseteq T(B_{\frac{1}{2}}(\mathbf{0}))$ such that $T(z_n) \to y$ as $n \to \infty$. So $T(w_n - z_n) = Tw_n - Tz_n \to (y + y_0) - y_0 = y$ with $\|w_n - z_n\| \leqslant \|w_n\| + \|z_n\| < \frac{1}{2} + \frac{1}{2} = 1$. Then $\{w_n - z_n\} \subseteq B_1(\mathbf{0})$ and so $y \in \overline{T(B_1(\mathbf{0}))}$.

Hence we get $B_{\epsilon_0}(\mathbf{0}) \subseteq \overline{T(B_1(\mathbf{0}))}$ and so $B_{\frac{\epsilon_0}{2^n}}(\mathbf{0}) \subseteq T(B_{\frac{1}{2^n}}(\mathbf{0}))$ for $n \in \mathbb{N}$ since $T$ is linear.

Claim. $B_{\frac{\epsilon_0}{2}}(\mathbf{0}) \subseteq T(B_1(\mathbf{0}))$. Let $y \in B_{\frac{\epsilon_0}{2}}(\mathbf{0}) \subseteq \overline{T(B_{\frac{1}{2}}(\mathbf{0}))}$. Then there exists $T(x_1) \in T(B_{\frac{1}{2}}(\mathbf{0}))$ such that $\|y - Tx_1\| < \frac{\epsilon_0}{2^2}$, i.e., $y - Tx_1 \in B_{\frac{\epsilon_0}{2^2}}(\mathbf{0}) \subseteq \overline{T(B_{\frac{1}{2^2}}(\mathbf{0}))}$. Then there exists $T(x_2) \in T(B_{\frac{1}{2^2}}(\mathbf{0}))$ such that $\|y - Tx_1 - Tx_2\| \leqslant \frac{\epsilon_0}{2^3}$, i.e., $y - Tx_1 - Tx_2 \in B_{\frac{\epsilon_0}{2^3}}(\mathbf{0}) \subseteq \overline{T(B_{\frac{1}{2^3}}(\mathbf{0}))}$. Repeat this, for $n \in \mathbb{N}$, we can find $\{x_n\} \subseteq X$ with $\|x_n\| < \frac{1}{2^n}$ such that $\|y - Tx_1 - Tx_2 - \cdots - Tx_n\| \leqslant \frac{\epsilon_0}{2^{n+1}}$. Let $z_n = \sum_{i=1}^n x_i \in X$. Then for $n, m \in \mathbb{N}$ with $n > m$, $\|z_n - z_m\| = \|\sum -i = m + 1^n X_i\| \leqslant \sum_{i=m+1}^n \frac{1}{2^i} \leqslant \frac{1}{2^{m+1}} \to 0$. So $\{z_n\} \subseteq X$ is Cauchy. Also, since $X$ is complete, there exists $x \in X$ such that $\sum_{i=1}^n x_i = z_n \to x$ as $n \to \infty$. Note $\|y - T(\sum_{i=1}^n x_i)\| \to 0$ as $n \to \infty$. Since $T$ is continuous, we have $y = \lim_{n\to\infty} T(\sum_{i=1}^n x_i) = T(\lim_{n\to\infty} \sum_{i=1}^n x_i) = T(x)$. Since $\|x\| = \|\sum_{i=1}^\infty x_i\| \leqslant \sum_{i=1}^\infty \|x_i\| \leqslant \sum_{i=1}^\infty \frac{1}{2^i} = 1$, we have $x \in B_1(\mathbf{0})$. So $y = Tx \in T(B_1(\mathbf{0}))$. ☐

**Theorem 4.67** (Open Mapping Theorem)**.** *Let $X$ and $Y$ be Banach and $T \in \mathcal{B}(X,Y)$. If $T$ is onto, then $T$ is an open mapping.*

*Proof.* Let $A \subseteq X$ be open. Let $x \in A$. Then there exists $r > 0$ such that $B_r(x) \subseteq A$. So $A - \{x\} \supseteq B_r(x) - \{x\} = B_r(\mathbf{0})$, i.e., $\frac{1}{r}(A - \{x\}) \supseteq B_1(\mathbf{0})$. Also, since $T$ is linear, $\frac{1}{r}(T(A) - Tx) = T(\frac{1}{r}(A - \{x\})) \supseteq T(\mathcal{B}_1(\mathbf{0})) \supseteq \mathcal{B}_\epsilon^{\|\cdot\|_Y}(\mathbf{0})$ for some $\epsilon > 0$ by previous lemma, i.e., $T(A) - T(x) \supseteq B_{r\epsilon}(\mathbf{0})$, i.e., $T(A) \supseteq B_{r\epsilon}(0) + T(x) = B_{r\epsilon}(T(x))$. ☐

**Corollary 4.68** (Inverse Mapping Theorem)**.** Let $X$ and $Y$ be Banach and $T \in \mathcal{B}(X,Y)$. If $T$ is bijective, then $T^{-1} \in \mathcal{B}(Y, X)$.

*Proof.* Since $T$ is bijective and $T$ is linear and then $T^{-1} : Y \to X$ is also linear. Since $T$ is onto, by Open Mapping Theorem, $T$ is an open mapping. So $T^{-1}$ is continuous and thus $T^{-1}$ is bounded. ☐

**Definition 4.69.** A linear operator $T : \mathcal{D}(T) \to Y$ with domain $\mathcal{D}(T) \subseteq X$, is called a *closed operator* if the *graph of the operator*

$$g(T) := \{(x, y) \in X \times Y \mid x \in \mathcal{D}(T), y = Tx\}$$

is closed in $X \times Y$.

**Remark.** (a) Recall $X \times Y$ is a NLS with the norm $\|(x, y)\|_{X\times Y} := \|x\|_X + \|y\|_Y$. In particular, if $X, Y$ are both Banach, then so is $X \times Y$.

(b) Most of the (linear) operators in practical purpose are closed.

**Lemma 4.70.** *$T$ is closed if and only if if $\{x_n\} \subseteq \mathcal{D}(T)$, $x_n \xrightarrow{\|\cdot\|_X} x \in X$ and $Tx_n \xrightarrow{\|\cdot\|_Y} y \in Y$, then $x \in \mathcal{D}(T)$ and $y = Tx$.*

*Proof.* $\Longrightarrow$ Since $x_n \to x$ and $Tx_n \to y$ as $n \to \infty$, we have $\|(x_n, Tx_n) - (x, y)\|_{X\times Y} = \|(x_n - x, Tx_n - y)\|_{X\times Y} = \|x_n - x\|_X + \|Tx_n - y\|_Y \to \infty$ as $n \to \infty$, i.e., $\{(x_n, T(x_n))\} \subseteq g(T)$ converges to $(x, y)$, then $(x, y) \in g(T)$ since $g(T)$ is closed. So $x \in \mathcal{D}(T)$ and $y = Tx$.

$\Longleftarrow$ Let $(x, y) \in \overline{g(T)}$. Then there exists $\{(x_n, Tx_n)\} \subseteq g(T)$ such that $(x_n, T(x_n)) \xrightarrow{\|\cdot\|_{X\times Y}} (x, y)$. Similarly, $x_n \to x$ and $Tx_n \to y$ as $n \to \infty$. By assumption, $x \in \mathcal{D}(T)$ and $y = Tx$. So $(x, y) \in g(T)$. ☐

**Example 4.71.** Let $T : \mathcal{C}^1[0, 1] \to \mathcal{C}[0, 1]$ be given by $f \mapsto f'$. Then $T$ is unbounded but closed.

*Proof.* Let $n \in \mathbb{N}$ and $f_n(x) = x^n \in \mathcal{C}^1[0, 1]$. Then $\|f_n\|_\infty = 1$ and $\|Tf_n\|_\infty = \left\|nx^{n-1}\right\|_\infty = n$. So $\|T\| \geqslant n$. Thus, $T$ is unbounded.

Let $\{f_n\} \subseteq \mathcal{C}^1[0, 1]$, $f_n \xrightarrow{\|\cdot\|_\infty} f$ and $T(f_n) = f_n' \xrightarrow{\|\cdot\|_\infty} g$. Let $x \in [0, 1]$. Then $\int_0^x g(t)dt = \int_0^x \lim_{n\to\infty} f_n'(t)dt = \lim_{n\to\infty} \int_0^x f_n'(t)dt = \lim_{n\to\infty}(f_n(x) - f_n(0)) = f(x) - f(0)$ since $f_n' \xrightarrow{\|\cdot\|_\infty} g$ and length$([0, x]) = x < \infty$, i.e., $f(x) = f(0) + \int_0^x g(t)dt$. By Fundamental Theorem of Calculus, $f \in \mathcal{C}^1[0, 1]$ and $T(f) = f' = g$. Thus, by previous lemma, $T$ is closed. ☐

**Example 4.72.** Let id : $\mathcal{D}(T) \to \mathcal{D}(T) \subseteq X$ and $\mathcal{D}(T) \subseteq X$ is dense. If $\{x_n\} \subseteq \mathcal{D}(T)$ and $x_n \to x \in X \setminus \mathcal{D}(T)$, then $\mathrm{id}(x_n) = x_n \to x$. But since $x \notin \mathcal{D}(T)$, $T$ is not closed.

**Theorem 4.73** (Closed Graph Theorem)**.** *Let $X, Y$ be Banach and $T : X \supseteq \mathcal{D}(T) \to Y$ be a closed linear operator. If $\mathcal{D}(T)$ is closed in $X$, then $T$ is bounded.*

*Proof.* Since $X$ is Banach and $\mathcal{D}(T) \subseteq X$ is closed, $\mathcal{D}(T)$ is Banach. Since $X, Y$ are Banach, $X \times Y$ is Banach. Also, since $T$ is closed, $g(T)$ is closed in $X \times Y$. So $g(T)$ is Banach.

Now consider $P : g(T) \to \mathcal{D}(T)$ given by $(x, Tx) \mapsto x$. Easy to check $P \in \mathcal{B}(g(T), \mathcal{D}(T))$ and $P$ is bijective. Then by Inverse Mapping Theorem, $P^{-1} : \mathcal{D}(T) \to g(T)$ given by $x \mapsto (x, Tx)$ is bounded. So there exists $M > 0$ such that $\left\|P^{-1}(x)\right\|_{X\times Y} = \|(x, Tx)\|_{X\times Y} \leqslant M\|x\|_X$ for $x \in \mathcal{D}(T)$. Hence $\|Tx\|_Y \leqslant M\|x\|_X$ for $x \in \mathcal{D}(T)$. ☐

**Fact 4.74.** Let $T : \mathcal{D}(T) \subseteq X \to Y$ be linear and bounded. If $\mathcal{D}(T)$ is closed in $X$, then $T$ is closed.

*Proof.* If $\{x_n\} \subseteq \mathcal{D}(T)$, $x_n \to x$ and $Tx_n \to y$, then $x \in \mathcal{D}(T)$ since $\mathcal{D}(T) \subseteq X$ is closed, and $Tx_n \to Tx$ since $T$ is bounded. So by the uniqueness of limit, we have $Tx = y$. ☐

**Theorem 4.75** (Two-Norm theorem)**.** *If $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ are both Banach, and one norm is stronger than the other, then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

*Proof.* Wlog., assume $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$. Then there exists $M > 0$ such that $\|x\|_1 \leqslant M\|x\|_2$ for all $x \in X$. Consider $\mathrm{id} : (X, \|\cdot\|1) \to (X, \|\cdot\|_2)$. Obviously, id is linear. To show id is bounded, it suffices to show id is closed by Closed Graph Theorem since $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ are Banach. If $\{x_n\} \subseteq X$, $x_n \xrightarrow{\|\cdot\|_1} x$ and $\mathrm{id}\, x_n = x_n \xrightarrow{\|\cdot\|_2} y$, then obviously $x \in X$ and $\mathrm{id}\, x = y$ since $\|\mathrm{id}(x) - y\|_1 = \|x - y\|_1 \leqslant \|x - x_n\|_1 + \|x_n - y\|_1 \leqslant \|x - x_n\|_1 + M\|x_n - y\|_2 \to 0$ as $n \to \infty$. □

**Example 4.76.** $(\mathcal{C}[0, 1], \|\cdot\|_1)$ is not complete, where $\|f\|_1 = \int_0^1 |f(t)| dt$.

*Proof.* Suppose not. Since $\|f\|_1 = \int_0^1 |f(t)| dt \leqslant \|f\|_\infty$ for all $f \in \mathcal{C}^1[0, 1]$, we have $\|\cdot\|_\infty$ is stronger than $\|\cdot\|_1$. Also, since $(\mathcal{C}[0, 1], \|\cdot\|_\infty)$ is complete, by Two-Norm Theorem, $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are equivalent. So there is a $M > 0$ such that $\|f\|_\infty \leqslant M\|f\|$ for all $f \in \mathcal{C}[0, 1]$, which is impossible. □

# Chapter 5

# Spectral Theory of Linear Operators

Let $X, Y$ be complex NLS's.

## 5.1 Basic definitions and examples

**Definition 5.1.** Let $T : X \supseteq \mathcal{D}(T) \to X$ be a linear operator.

(a) For $\lambda \in \mathbb{C}$, if $T_\lambda = T - \lambda I : \mathrm{Im}(T_\lambda) \to \mathcal{D}(T_\lambda)$ has an inverse, then we call the inverse $\mathcal{R}_\lambda(T) = T_\lambda^{-1} = (T - \lambda I)^{-1}$ the *resolvent* of $T$.

(b) $\lambda \in \mathbb{C}$ is a *regular value* of $T$ if the following conditions hold.

(1) $\mathcal{R}_\lambda(T)$ exists.

(2) $\mathcal{R}_\lambda(T)$ is bounded.

(3) $\mathcal{R}_\lambda(T)$ is densely defined, i.e., $\mathcal{D}(\mathcal{R}_\lambda(T)) \subseteq X$ is dense.

The collection of all regular values is called the *resolvent* set of $T$, denoted as $\rho(T)$.

(c) The complement of $\rho(T)$, denoted as $\sigma(T) = \mathbb{C} \setminus \rho(T)$, is called the *spectrum* of $T$ and can be divided into the following 3 cases:

(1) *Point (discrete) spectrum* of $T$:

$$\sigma_\rho(T) = \{\lambda \in \sigma(T) \mid \mathcal{R}_\lambda(T) \text{ does not exist}\}.$$

(2) *Continuous spectrum of* $T$:

$$\sigma_c(T) = \{\lambda \in \sigma(T) \mid \mathcal{R}_\lambda(T) \text{ exists, densely defined, but unbounded}\}.$$

(3) *Residual spectrum* of $T$:

$$\sigma_r(T) = \{\lambda \in \sigma(T) \mid \mathcal{R}_\lambda(T) \text{ exists, but not densely defined}\}.$$

$\sigma_c(T) \cup \sigma_r(T)$ are called the *essential spectrum*.

**Remark.** (a) Name "resolvent" comes from solving $T_\lambda x = y$.

(b) $\mathcal{R}_\lambda(T)$ is linear.

(c) $\mathbb{C} = \rho(T) \cup \sigma(T) = \rho(T) \cup \sigma_\rho(T) \cup \sigma_c(T) \cup \sigma_r(T)$.

(d) If $\lambda \in \sigma_\rho(T)$, $\lambda$ is also called an *eigenvalue* of $T$. Note $\mathcal{R}_\lambda(T) = (T-\lambda I)^{-1} : \mathcal{D}(\mathcal{R}_\lambda(T)) \to \mathcal{D}(T_\lambda)$ exists if and only if $T_\lambda = T - \lambda I$ is 1-1 if and only if $\text{Ker}(T_\lambda) = \{\mathbf{0}\} = \{x \mid (T - \lambda I)x = \mathbf{0}\}$. Thus if $\mathbf{0} \neq x$ satisfying $(T - \lambda I)x = \mathbf{0}$, then $\lambda \in \sigma_\rho(T)$ and $x$ is called an *eigenvector* corresponding to $\lambda$.

**Lemma 5.2.** Let $x_1, \ldots, x_n$ be eigenvectors corresponding to different eigenvalues $\lambda_1, \ldots, \lambda_n$, then $x_1, \ldots, x_n$ are linearly independent.

*Proof.* Suppose not, reorder them and let $x_m$ with $m \in \{1, \ldots, n\}$ be the first vector that can be written as a linear combination of its previous vectors. Namely, $x_m = \sum_{i=1}^{m-1} a_i x_i$ with $a_i \in \mathbb{C}$ for $i = 1, \ldots, m-1$. Then $0 = (T - \lambda_m I)x_m = \sum_{i=1}^{m-1} a_i(Tx_i - \lambda_m x_i) = \sum_{i=1}^{m-1} a_i(\lambda_i - \lambda_m)x_i$. Also, by the minimality of $m$, $x_1, \ldots, x_{m-1}$ are independent, so $a_i(\lambda_i - \lambda_m) = 0$ for $i = 1, \ldots, m-1$. Also, since $\lambda_1, \ldots, \lambda_m$ are distinct, $a_i = 0$ for $i = 1, \ldots, m-1$. So $x_m = 0$, a contradiction. □

**Example 5.3.** Let $\dim_{\mathbb{C}} X = n \geqslant 1$ with a standard basis $\{e_1, \ldots, e_n\}$. Let $T : X \to X$ be given by $T(e_i) = \sum_{j=1}^{n} a_{ij} e_j$ with $a_{i1}, \ldots, a_{in} \in \mathbb{C}$ for $i = 1, \ldots, n$. Then $T$ is linear and can be represented as a matrix $A = [a_{ij}]$. Let $\lambda \in \sigma_\rho(T)$. Then $(T - \lambda I)x = 0$ for some $x \neq \mathbf{0}$. So $(A - \lambda I_n)x = 0$. Hence $\det(T - \lambda I_n) = 0$ and it has a complex root since $\det(A - \lambda I_n) \in \mathbb{C}[\lambda]$ has degree $n$.

If $\lambda \notin \sigma_\rho(T)$ , then $\mathcal{R}_\lambda(T)$ exists and is defined on the entire $X$ and bounded. So $\lambda \in \rho(T)$, i.e., $\sigma_c(T) = \sigma_r(T) = \emptyset$.

**Example 5.4.** Consider the right-shift operator $T : l^2 \to l^2$. Recall $T \in \mathcal{B}(l^2, l^2)$ and $\|T\| = 1$. Claim. $0 \in \sigma_r(T) \subseteq \mathbb{C}$. Let $T^{-1} : \text{Im}(T) \to$ be the left inverse (left-shift) of $T$. Then $T^{-1} = (T - 0I)^{-1} = \mathcal{R}_0(T)$. So $\mathcal{D}(\mathcal{R}_0(T)) = \mathcal{D}(T^{-1}) = \text{Im}(T) = \{x \in l^2 \mid x = \{0, x_1, x_2, \cdots\}\}$ is not dense in $l^2$. Hence $0 \in \sigma_r(T)$.

**Theorem 5.5.** *Let $X$ be Banach and $T \in \mathcal{B}(X, X)$.*

*(a) If $\mathcal{R}_\lambda(T)$ exists and is defined on $X$, then $\mathcal{R}_\lambda(T)$ is bounded.*

*(b) If $\lambda \in \rho(T)$, then $\mathcal{R}_\lambda(T)$ is defined on the entire $X$ (and also bounded by (a)).*

*Proof.* (a) Since $T \in \mathcal{B}(X, X)$, $T - \lambda I \in \mathcal{B}(X, X)$. Since $\mathcal{R}_\lambda(T) = (T - \lambda I)^{-1}$ exists and is defined on $X$, and $X$ is Banach, by Inverse Mapping Theorem, $\mathcal{R}_\lambda(T) \in \mathcal{B}(X, X)$.

(b) Since $T \in \mathcal{B}(X, X)$ and $\mathcal{D}(T) = X$ is closed, we have $T$ is closed by previous Fact. Claim. $T - \lambda I$ is also closed. Let $\{x_n\} \subseteq X$ with $x_n \to x \in X$, and $(T - \lambda I)x_n \to y$ as $n \to \infty$. Then $Tx_n = (T - \lambda I)x_n + x_n \to y + x = Tx$ as $n \to \infty$ since $T$ is closed. So $(T - \lambda I)x = Tx - \lambda x = \lim_{n\to\infty} Tx_n - \lambda \lim_{n\to\infty} x_n = \lim_{n\to\infty}(T - \lambda I)x_n = y$.

Let $\{(T - \lambda I)x_n\} \subseteq \text{Im}(T - \lambda I) = \mathcal{D}(\mathcal{R}_\lambda(T))$ such that $(T - \lambda I)x_n \to y$, where $\{x_n\} \subseteq X$. Since $X$ is complete, $y \in X$. Since $\lambda \in \sigma_\rho(T)$, $\mathcal{R}_\lambda(T)$ is bounded. So $\mathcal{R}_\lambda(T)((T - \lambda I)x_n) \to \mathcal{R}_\lambda(T)y \in X$. Then $(T - \lambda I)^{-1}y = x$, i.e., $y = (T - \lambda I)x$ for some $x \in X$. So $y \in \text{Im}(T - \lambda I) = \mathcal{D}(\mathcal{R}_\lambda(T))$. Hence $\mathcal{D}(\mathcal{R}_\lambda(T))$ is closed. Also, since $\lambda \in \rho(T)$, $\mathcal{D}(\mathcal{R}_\lambda(T)) = \overline{\mathcal{D}(\mathcal{R}_\lambda(T))} = X$. □

## 5.2 Spectral Properties of Bounded Linear Operators

Let $X$ be Banach and $T \in \mathcal{B}(X, X)$.

**Recall 5.6.** $(X, \|\cdot\|)$ is Banach if and only if every absolutely convergent series converges.

**Lemma 5.7.** If $\|T\| < 1$, then $(I - T)^{-1}$ exists, $(I - T)^{-1} \in \mathcal{B}(X, X)$ and $(I - T)^{-1} = \sum_{n=0}^{\infty} T^n$.

*Proof.* Since $X$ is Banach, $\mathcal{B}(X, X)$ is Banach. To show $\sum_{n=0}^{\infty} T^n$ converges, it suffices to show $\sum_{n=0}^{\infty} \|T^n\| < \infty$, which is true since $\sum_{n=0}^{\infty} \|T^n\| \leqslant \sum_{n=0}^{\infty} \|T\|^n = \frac{1}{1-\|T\|} < \infty$. So $\sum_{n=0}^{\infty} T^n \in \mathcal{B}(X, X)$. Only need to show $(I - T)^{-1} = \sum_{n=0}^{\infty} T^n$. This can be seen from $(I - T)(\sum_{n=0}^{\infty} T^n) = \sum_{n=0}^{\infty}((I - T)T^n) = \lim_{N\to\infty}(I - T)(\sum_{n=0}^{N} T^n) = \lim_{N\to\infty}(I - T^{N+1}) = I - \mathbf{0} = I$ since $\left\|T^{N+1} - \mathbf{0}\right\| \leqslant \|T\|^{N+1} \to 0$ as $N \to \infty$. □

**Theorem 5.8.** $\sigma(T) \subseteq \overline{B}_{\|T\|}(0) = \{\lambda \in \mathbb{C} \mid |\lambda| \leqslant \|T\|\}$. *So* $\rho(T) \neq \emptyset$.

*Proof.* Prove by contrapositive. Since $\left\|\frac{T}{\lambda}\right\| < 1$, by previous lemma, we have $\mathcal{R}_\lambda(T) = (T - \lambda I)^{-1} = (-\lambda(I - \frac{1}{\lambda}T))^{-1} = -\frac{1}{\lambda}(I - \frac{T}{\lambda})^{-1} = -\frac{1}{\lambda}\sum_{n=0}^{\infty}\frac{T^n}{\lambda^n} = -\sum_{n=0}^{\infty}\frac{T^n}{\lambda^{n+1}} \in \mathcal{B}(X, X)$. So $\lambda \in \rho(T)$. □

**Theorem 5.9.** *$\sigma(T)$ is closed in $\mathbb{C}$. So $\sigma(T)$ is compact, combining with the previous theorem.*

*Proof.* It suffices to show $\rho(T)$ is open. Let $\lambda_0 \in \rho(T)$. Let $\lambda \in \mathbb{C}$. Then $T - \lambda I = T - \lambda_0 I + (\lambda_0 - \lambda)I = (T - \lambda_0 I)(I + (\lambda_0 - \lambda)(T - \lambda_0 I)^{-1})$. So $T_\lambda = T_{\lambda_0}(I - (\lambda - \lambda_0)\mathcal{R}_{\lambda_0}(T))$. Hence for $\|(\lambda - \lambda_0)\mathcal{R}_{\lambda_0}(T)\| = |\lambda - \lambda_0|\|\mathcal{R}_{\lambda_0}(T)\| < 1$, i.e., $|\lambda - \lambda_0| < \frac{1}{\left\|\mathcal{R}_{\lambda_0}(T)\right\|}$, we have $\mathcal{R}_\lambda(T) = T_\lambda^{-1} = (I - (\lambda - \lambda_0)\mathcal{R}_{\lambda_0}(T))^{-1}T_{\lambda_0}^{-1} = \sum_{n=0}^{\infty}(\lambda - \lambda_0)^n\mathcal{R}_{\lambda_0}(T)^n\mathcal{R}_{\lambda_0}(T) = \sum_{n=0}^{\infty}(\lambda - \lambda_0)^n\mathcal{R}_{\lambda_0}^{n+1}(T)$ and so $\|\mathcal{R}_\lambda(T)\| \leqslant \|\mathcal{R}_{\lambda_0}(T)\|\sum_{n=0}^{\infty}\|(\lambda - \lambda_0)\mathcal{R}_{\lambda_0}(T)\|^n = \frac{\left\|\mathcal{R}_{\lambda_0}(T)\right\|}{1-\left\|(\lambda-\lambda_0)\mathcal{R}_{\lambda_0}(T)\right\|}$, hence $\lambda \in \rho(T)$. So $B_{\frac{1}{\left\|\mathcal{R}_{\lambda_0}(T)\right\|}}(\lambda_0) \subseteq \rho(T)$. Thus, $\rho$ is open. □

**Theorem 5.10.** *Let $\lambda, \mu \in \rho(T)$. Then*

*(a)* $\mathcal{R}_\lambda(T) - \mathcal{R}_\mu(T) = (\lambda - \mu)\mathcal{R}_\lambda(T)\mathcal{R}_\mu(T)$. *"resolvent equation"*

*(b)* $\mathcal{R}_\lambda(T)S = S\mathcal{R}_\lambda(T)$ *if* $S \in \mathcal{B}(X, X)$ *and* $ST = TS$.

*(c)* $\mathcal{R}_\lambda(T)\mathcal{R}_\mu(T) = \mathcal{R}_\mu(T)\mathcal{R}_\lambda(T)$.

*Proof.* (a) $\mathcal{R}_\lambda(T) - R_\mu(T) = \mathcal{R}_\lambda(T)T_\mu\mathcal{R}_\mu(T) - \mathcal{R}_\lambda(T)T_\lambda\mathcal{R}_\mu(T) = \mathcal{R}_\lambda(T)(T_\mu - T_\lambda)\mathcal{R}_\mu(T) = (\lambda - \mu)\mathcal{R}_\lambda(T)\mathcal{R}_\mu(T)$.

(b) If $ST = TS$, then $ST_\lambda = S(T - \lambda I) = ST - \lambda S = TS - \lambda S = (T - \lambda I)S = T_\lambda S$. So $\mathcal{R}_\lambda(T)S = \mathcal{R}_\lambda(T)ST_\lambda\mathcal{R}_\lambda(T) = \mathcal{R}_\lambda(T)T_\lambda S\mathcal{R}_\lambda(T) = S\mathcal{R}_\lambda(T)$.

(c) Since $\mu \in \rho(T)$, $\mathcal{R}_\mu(T) \in \mathcal{B}(X, X)$ by previous theorem. Since $TT = TT$, by (b), we have $\mathcal{R}_\mu(T)T = T\mathcal{R}_\mu(T)$. Again, by (b), $\mathcal{R}_\lambda(T)\mathcal{R}_\mu(T) = \mathcal{R}_\mu(T)\mathcal{R}_\lambda(T)$. □

### 5.2.1 Recall

Let $X, Y, X_1, X_2, \{X_n\}$ be Banach and $T : X \to Y$ be linear.

**Definition 5.11.** (a) A *domain* $\mathcal{D} \subseteq \mathbb{C}$ is an open connected set.

(b) A complex-valued function of a complex variable $f(\lambda)$ is *holomorphic* (*analytic*) *on* $\mathcal{D}$ if $f$ is differentiable at every $\lambda \in \mathcal{D}$, i.e., $f'(\lambda) = \lim_{\Delta\lambda \to 0} \frac{f(\lambda+\Delta\lambda)-f(\lambda)}{\Delta\lambda}$ exists for $\lambda \in \mathcal{D}$.

(c) $f$ is *holomorphic at* $\lambda_0 \in \mathbb{C}$ if $f$ is holomorphic on a neighborhood around $\lambda_0$.

(d) $f$ is *entire* if $f$ is holomorphic on $\mathbb{C}$.

**Theorem 5.12.** *(a) $f$ is holomorphic on $\mathcal{D}$ if and only if $f(\lambda) = \sum_{n=0}^{\infty} c_n(\lambda - \lambda_0)^n$ with $c_n \in \mathbb{C}$ for $i \geqslant 0$ for any $\lambda \in B_r^{|\cdot|}(\lambda_0)$, where $r$ is* radius of convergence.

*(b) (Liouville) Any bounded entire function is a constant function.*

**Definition 5.13.** Let $E \subseteq \mathbb{C}$ be open. Then an operator function $S : E \to \mathcal{B}(X, X)$ is called *locally holomorphic* if for $x \in X$ and $\varphi \in X'$, the complex function $f(\lambda) := \varphi(S(\lambda)x)$ is holomorphic on $E$.

In fact, $S$ is holomorphic if $E$ is a domain.

**Theorem 5.14.** *(a) $\mathcal{R} : \rho(T) \to \mathcal{B}(X, X)$ given by $S(\lambda) = \mathcal{R}_\lambda(T)$ is locally holomorphic.*

*(b) For $\lambda \in \rho(T)$, $\|\mathcal{R}_\lambda(T)\| \geqslant \frac{1}{d(\lambda)}$, where $d(\lambda) = \inf_{\mu \in \sigma(T)}|\lambda - \mu|$.*

*Proof.* (a) Want to show for $x \in X$ and $\varphi \in X'$, $f(\lambda) := \varphi(\mathcal{R}_\lambda(T)x)$ is holomorphic on $\rho(T)$. Let $\lambda_0 \in \rho(T)$. Let $\lambda \in \mathbb{C}$. Similar to the proof of previous theorem, for $\lambda \in B_{\frac{1}{\left\|\mathcal{R}_{\lambda_0}(T)\right\|}}(\lambda_0)$, $f(\lambda) = \varphi(\mathcal{R}_\lambda(T)) = \varphi(\sum_{n=0}^{\infty}(\lambda - \lambda_0)^n R_{\lambda_0}^{n+1}(T)x) = \sum_{n=0}^{\infty} \varphi(\mathcal{R}_{\lambda_0}^{n+1}(T)x)(\lambda - \lambda_0)^n$ with $\varphi(\mathcal{R}_{\lambda_0}^{n+1}(T)x) \in \mathbb{C}$ for $n \geqslant 0$. So $f$ is holomorphic on $\rho(T)$.

(b) If $|\mu - \lambda| < \frac{1}{\|\mathcal{R}_\lambda(T)\|}$, then $\mu \in \rho(T)$. So if $\mu \in \sigma(T)$, then $|\mu - \lambda| \geqslant \frac{1}{\|\mathcal{R}_\lambda(T)\|}$. Hence $d(\lambda) = \inf_{\mu \in \sigma(T)}|\mu - \lambda| \geqslant \frac{1}{\|\mathcal{R}_\lambda(T)\|}$. □

**Lemma 5.15.** (a) If $T$ is compact, then $T$ is bounded, i.e., $K(X, Y) \subseteq \mathcal{B}(X, Y)$.

(b) If $\dim(X) = \infty$, then id is not complete.

(c) $\overline{K} \subseteq X$ is compact if and only if any sequence in $K$ has a convergent subsequence which converges to a point in $\overline{K}$.

**Lemma 5.16.** For $\{y_n\} \subseteq Y$, $\{y_n\}$ has a convergent subsequence which converges to $\overline{Y}$, then $\overline{Y}$ is compact.

*Proof.* Let $\{x_n\} \subseteq \overline{Y}$. Then for $n \in \mathbb{N}$, there exists $\{y_{n,m}\}_{m=1}^{\infty} \subseteq Y$ such that $y_{n,m} \to x_n$ as $m \to \infty$. For $n \in \mathbb{N}$, there exists $M_n \in \mathbb{N}$ such that $\|y_{n,M_n} - x_n\| < \frac{1}{n}$. Consider $\{y_{n,M_n}\}_{n=1}^{\infty} \subseteq Y$, by assumption, there is $\{y_{n_k,M_{n_k}}\}_{k=1}^{\infty} \subseteq \{y_{n,M_n}\}_{n=1}^{\infty}$ such that $y_{n_k,M_{n_k}} \to y \in \overline{Y}$ as $k \to \infty$. Then $\|x_{n_k} - y\| \leqslant \left\|x_{n_k} - y_{n_k,M_{n_k}}\right\| + \left\|y_{n_k,M_{n_k}} - y\right\| \leqslant \frac{1}{n_k} + \left\|y_{n_k,M_{n_k}} - y\right\| \to 0$ as $k \to \infty$. So $x_{n_k} \to y \in \overline{Y}$ as $k \to \infty$. Thus, $\overline{Y}$ is compact. □

**Theorem 5.17.** *$T$ is compact if and only if for any $\{x_n\} \subseteq X$ bounded, $\{T(x_n)\}$ has a convergent subsequence.*

*Proof.* $\Longrightarrow$ Since $T$ is compact and $\{x_n\}$ is bounded, $\overline{\{T(x_n)\}}$ is compact. So $\{T(x_n)\} \subseteq \overline{\{T(x_n)\}}$ has a convergent subsequence.

$\Longleftarrow$ For any $B \subseteq X$ bounded, consider $\{y_n\} \subseteq T(B)$, where $y_n = T(x_n)$ and $\{x_n\} \subseteq B$ is bounded. By assumption, $\{y_n\}$ has a convergent subsequence. So $\overline{T(B)}$ is compact by previous lemma. Thus, $T$ is compact. $\square$

**Remark.** If $T_1, T_2$ are compact linear, then $T_1 + T_2$ and $aT_1$ for $a \in \mathbb{C}$ are also compact linear. Thus, $K(X, Y)$ is a vector space.

**Theorem 5.18.** *(a) If $T$ is bounded and $\dim(T(X)) < \infty$ ("operators with finite rank"), then $T$ is compact.*

*(b) If $\dim(X) < \infty$, then $T$ is compact.*

*Proof.* (a) Let $\{x_n\} \subseteq X$ be bounded. Since $T$ is bounded, $\{T(x_n)\}$ is also bounded since $\|Tx_n\| \leqslant \|x_n\|$ and $\{x_n\}$ is uniformly bounded. So $\overline{\{T(x_n)\}}$ is closed and bounded in $T(X)$. So $\overline{\{T(x_n)\}}$ is compact since $\dim(X) < \infty$. Thus, $\{T(x_n)\} \subseteq \overline{\{T(x_n)\}}$ has a convergent subsequence and hence $T$ is compact.

(b) Since $\dim(X) < \infty$ and $T$ is linear, $T$ is bounded from MATH8210. On the other hand, since $\dim(T(X)) \leqslant \dim(X) < \infty$, $T$ is compact by (a). $\square$

**Remark.** $K(X, Y) \subseteq \mathcal{B}(X, Y)$ is closed since $Y$ is Banach.

**Theorem 5.19.** *If $\{T_n\}$ are compact linear and $\|T_n - T\| \to 0$ as $n \to 0$. Then $T$ is compact linear.*

*Proof.* Linearity of $T$ is obvious. Let $\{x_n\} \subseteq X$ be bounded. Then there exists $M > 0$ such that $\|x_n\| < M$. Since $\{T_n\}$ is compact, by "diagonal argument", there exists $\{x_{m_k}\} \subseteq \{x_m\}$ such that $\{T_n(x_{m_k})\}_{k=1}^{\infty}$ converges for all $n \in \mathbb{N}$. Let $\epsilon > 0$. Since $\|T_n - T\| \to 0$ as $n \to \infty$, there exists $N \in \mathbb{N}$ such that $\|T_N - T\| < \frac{\epsilon}{3M}$. Since $\{T_N(x_{m_k})\}_{k=1}^{\infty}$ is Cauchy, there exists $N'$ such that $|T_N(x_{m_k}) - T_N(x_{m_l})| < \frac{\epsilon}{3}$ for $k, l \geqslant N'$. So $\|T(x_{m_k}) - T(x_{m_l})\| \leqslant \|T(x_{m_k}) - T_N(x_{m_k})\| + \|T_N(x_{m_k}) - T_N(x_{m_l})\| + \|T_N(x_{m_l}) - T(x_{m_l})\| \leqslant \frac{\epsilon}{3M}M + \frac{\epsilon}{3} + \frac{\epsilon}{3M}M = \epsilon$ for $k, l \geqslant N'$. Hence $\{T(x_{m_k})\}$ is Cauchy in $Y$. Also, since $Y$ is Banach, $\{T(x_{m_k})\}$ converges. Thus, $T$ is compact. $\square$

**Example 5.20.** If $\{T_n\}$ is compact linear and $T_n \to T$ strongly as $n \to \infty$, then $T$ may not be compact.

*Proof.* For $n \in \mathbb{N}$, let $T_n : l^2 \to l^2$ be given by $x = \{x_m\} \mapsto \{x_1, \ldots, x_n, 0, 0, \cdots\}$. Then $\{T_n\}$ are bounded linear with finite rank for $n \in \mathbb{N}$. For $x \in l^2$, $\|T_n x - \mathrm{id}(x)\|_2 = (\sum_{i=n+1}^{\infty} |x_i|^2)^{\frac{1}{2}} \to 0$ as $n \to \infty$ since $x \in l^2$. So $T_n \to \mathrm{id}$ strongly as $n \to \infty$, but id is not compact since $\dim(l^2) = \infty$. $\square$

**Example 5.21.** Define $T : l^2 \to l^2$ by $x = \{x_n\} \mapsto \{x_1, \frac{x_2}{2}, \frac{x_3}{3}, \ldots, \frac{x_n}{n}, \cdots\}$. Then $T$ is compact linear.

*Proof.* For $n \in \mathbb{N}$, let $T_n : l^2 \to l^2$ be given by $x = \{x_m\} \mapsto \{x_1, \frac{x_2}{2}, \frac{x_3}{3}, \ldots, \frac{x_n}{n}, 0, 0, \cdots\}$. Then for $n \in \mathbb{N}$, $T_n$ are compact linear since $T_n$ is bounded and $\dim(T_n(l^2)) = n < \infty$. For $n \in \mathbb{N}$ and $x \in l^2$, $\|(T_n - T)x\|^2 = \|T_n x - Tx\|_2^2 = \sum_{i=n+1}^{\infty} \frac{x_i^2}{i^2} \leqslant \frac{1}{(n+1)^2} \sum_{i=n+1}^{\infty} x_i^2 \leqslant \frac{1}{(n+1)^2} \|x\|_2^2$. Then $\|T_n - T\| \leqslant \frac{1}{n+1}$. So $T_n \to T$ uniformly. Thus, $T$ is compact linear. $\square$

**Example 5.22.** Define $T : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ by $Tx = \int_0^1 K(t,s)x(s)ds$, where $K : [0,1] \times [0,1] \to \mathbb{R}$ is continuous. Then $T$ is compact linear.

*Proof.* Linearity and boundedness of $T$ is obvious. Let $\{x_n\} \subseteq \mathcal{C}[0,1]$ be bounded. Then there exists $M > 0$ such that $\|x_n\| \leqslant M$ for $n \in \mathbb{N}$. So $\|Tx_n\|_\infty \leqslant \|T\|\|x_n\|_\infty < \|T\|M$, i.e., $\{T(x_n)\}$ is uniformly bounded. Let $\epsilon > 0$. Since $K \in \mathcal{C}([0,1] \times [0,1])$ and $[0,1] \times [0,1]$ is compact, $K$ is uniformly continuous. So there exists $\delta = \delta(\epsilon) > 0$ such that $|K(t_1,s) - K(t_2,s)| < \frac{\epsilon}{M}$ whenever $t_1, t_2 \in [0,1]$ with $|t_1 - t_2| < \delta$ for $s \in [0,1]$. Hence for $n \in \mathbb{N}$, $|T(x_n)(t_1) - T(x_n)(t_2)| = \left|\int_0^1 K(t_1,s)x_n(s)ds - \int_0^1 K(t_2,s)x_n(s)ds\right| \leqslant \int_0^1 |K(t_1,s) - K(t_2,s)|\|x_n\|ds < \frac{\epsilon}{M} \cdot M = \epsilon$ whenever $t_1, t_2 \in [0,1]$ with $|t_1 - t_2| < \delta$. So $\{T(x_n)\}$ is equicontinuous. Thus, by Arzela-Ascoli theorem, $\{T(x_n)\}$ has a convergent subsequence. □

**Theorem 5.23** (Arzela-Ascoli). *Let* $\{f_n\} \subseteq (\mathcal{C}[0,1], \|\cdot\|)$. *If*

*(a)* $\|f_n\|_\infty \leqslant M$ *for* $n \in \mathbb{N}$*;*

*(b)* $\{f_n\}$ *is equicontinuous, i.e., for* $\epsilon > 0$*, there exists* $\delta = \delta(\epsilon) > 0$ *such that* $|f_n(t_1) - f_n(t_2)| < \epsilon$ *whenever* $t_1, t_2 \in [0,1]$ *with* $|t_1 - t_2| < \delta$ *for* $n \in \mathbb{N}$*;*

*then* $\{f_n\}$ *has a convergent subsequence.*

**Lemma 5.24.** If $\{x_n\} \subseteq X$ satisfies for any subsequence $\{x_{n_k}\} \subseteq \{x_n\}$, there exsits $\{x_{n_{k_l}}\} \subseteq \{x_{n_k}\}$ such that $x_{n_{k_l}} \to x_0$ as $l \to \infty$, then $x_n \to x_0$ as $n \to \infty$.

*Proof.* Suppose not, then there exists $\epsilon > 0$ such that for $k \in \mathbb{N}$, we have $\|x_{n_k} - x_0\| \geqslant \epsilon$ for some $n_k \geqslant k$ and $n_k \geqslant n_{k-1}$, where $n_0 = 0$. So $\{x_{n_k}\}$ does not have any subsequence convergent to 0, a contradiction. □

**Theorem 5.25.** *Let* $T : X \to Y$ *be compact linear. If* $x_n \rightharpoonup x$*, then* $T(x_n) \to T(x)$ *as* $n \to \infty$*.*

*Proof.* Let $\varphi \in Y'$. Then $\varphi \circ T$ is bounded linear since both $\varphi$ and $T$ are bounded linear, i.e., $\varphi \circ T \in X'$. So $\varphi(T(x_n)) \to \varphi(T(x))$. Hence $Tx_n \rightharpoonup Tx$ as $n \to \infty$. Let $\{Tx_{n_k}\} \subseteq \{Tx_n\}$ with $\{x_{n_k}\} \subseteq \{x_n\}$. Since $x_n \rightharpoonup x$, $\{x_{n_k}\}$ is bounded. Also, since $T$ is compact, $\{Tx_{n_k}\}$ has a convergent subsequence $\{Tx_{n_{k_l}}\}$. Since $Tx_n \rightharpoonup Tx$ as $n \to \infty$, $Tx_{n_{k_l}} \rightharpoonup Tx$ as $l \to \infty$. Thus, by previous lemma, $Tx_n \to Tx$ as $n \to \infty$. □

**Theorem 5.26.** *Let* $T : X \to Y$ *be compact linear. Then* $\mathrm{Im}(T) = T(X)$ *is separable.*

*Proof.* Note $X = \bigcup_{n=0}^\infty \overline{B}_n(\mathbf{0})$. Then $T(X) = T(\bigcup_{n=0}^\infty \overline{B}_n(\mathbf{0})) = \bigcup_{n=0}^\infty T(\overline{B}_n(\mathbf{0}))$. Let $n \in N$. Then $\overline{B}_n \subseteq X$ is bounded. So $\overline{T(\overline{B_n}(\mathbf{0})}$ is compact since $T$ is compact. Then $T(\overline{B}_n(\mathbf{0}))$ is totally bounded. So $T(\overline{B}_n(\mathbf{0}))$ is separable. Then there exists countable dense subset $D_n \subseteq \overline{T(\overline{B}_n(\mathbf{0}))}$. So $\bigcup_{n=1}^\infty D_n$ is countable dense in $\bigcup_{n=1}^\infty T(\overline{B}_n(\mathbf{0})) = T(X)$ since $\bigcup_{n=1}^\infty U_n$ is countable and $\overline{\bigcup_{n=1}^\infty D_n} \supseteq \bigcup_{n=1}^\infty \overline{D}_n = \bigcup_{n=1}^\infty T(\overline{B}_n(\mathbf{0})) = T(X)$. □

**Theorem 5.27.** *Let* $T : X \to Y$ *be compact linear. Then* $T^* : Y' \to X'$ *is also compact linear.*

*Proof.* Since $X'$ is complete, it suffices to show for any $B \subseteq Y'$, say $\|\varphi\| \leqslant M$ for $\varphi \in B$, $T^*(B)$ is totally bounded. Let $\epsilon > 0$. Want to show there exist $\varphi_1, \ldots, \varphi_n \in B$ such that for $\varphi \in B$, $\sup_{\|x\| \leqslant 1} |\varphi(Tx) - \varphi_k(Tx)| = \|T^*(\varphi) - T^*(\varphi_k)\|_X \leqslant \epsilon$ for some $k \in \{1, \ldots, n\}$. Since $\overline{B}_1(\mathbf{0}) \subseteq X$ is bounded, $T(\overline{B}_1(\mathbf{0}))$ is totally bounded in $Y$. Hence there exist $x_1, \ldots, x_m \in \overline{B}_1(\mathbf{0})$ such that for $x \in \overline{B}_1(\mathbf{0})$, $\|Tx - Tx_j\|_Y \leqslant \frac{\epsilon}{3M}$ for some $j \in \{1, \ldots, m\}$.

Define $l : Y' \to \mathbb{R}^m$ by $l(\varphi) = (\varphi(Tx_1), \ldots, \varphi(Tx_m))$. Since $T$ is compact, it is bounded. Also, since $\varphi$ is bounded, $l$ is bounded. Also, since $\dim(l(Y')) \leqslant m$, $l$ is compact. Also, since $B \subseteq Y'$ is bounded, $l(B)$ is totally bounded in $\mathbb{R}^m$. Hence there exist $\varphi_1, \ldots, \varphi_n \in B$ such that for $\varphi \in B$, $\|l(\varphi) - l(\varphi_k)\|_{\mathbb{R}^m} \leqslant \frac{\epsilon}{3}$ for some $k \in \{1, \ldots, n\}$.

Thus, for $x \in \overline{B}_1(\mathbf{0})$ and $\varphi \in B$, $|\varphi(Tx) - \varphi_k(Tx)| \leqslant |\varphi(Tx) - \varphi(Tx_j)| + |\varphi(Tx_j) - \varphi_k(Tx_j)| + |\varphi_k(Tx_j) - \varphi_k(Tx)| \leqslant \|\varphi\|\|Tx - Tx_j\| + \|l(\varphi) - l(\varphi_k)\| + \|\varphi_k\|\|Tx_j - Tx| \leqslant M \cdot \frac{\epsilon}{3M} + \frac{\epsilon}{3} + M \cdot \frac{\epsilon}{3M} = \epsilon$. Therefore, $\sup_{\|x\| \leqslant 1} |\varphi(Tx) - \varphi_k(Tx)| \leqslant \epsilon$. □