

# FUNCTIONAL ANALYSIS

CHI-WAI LEUNG

ABSTRACT. This is a note on the course MATH 4010: Functional analysis in 2024-25, 1st term.

## 1. NORMED SPACES

**Definition 1.1.** Let  $X$  be a vector space over a field  $\mathbb{K}$ , where  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$ . A function  $\|\cdot\| : X \rightarrow \mathbb{R}$  is called a norm on  $X$  if it satisfies the following conditions.

- (i)  $\|x\| \geq 0$  for all  $x \in X$ .
- (ii)  $\|x\| = 0$  if and only if  $x = 0$ .
- (iii)  $\|\alpha x\| = |\alpha| \|x\|$  for  $x \in X$  and  $\alpha \in \mathbb{K}$ .
- (iii) (Triangle inequality)  $\|x - y\| \leq \|x - z\| + \|z - y\|$  for all  $x, y, z \in X$ .

In this case, the pair  $(X, \|\cdot\|)$  is called a normed space.

**Example 1.2.** The following are important examples of finite dimensional normed spaces.

- (i) Let  $\ell_\infty^{(n)} = \{(x_1, \dots, x_n) : x_i \in \mathbb{K}, i = 1, 2, \dots, n\}$ . Put  $\|(x_1, \dots, x_n)\|_\infty = \max\{|x_i| : i = 1, \dots, n\}$ .
- (ii) Let  $\ell_p^{(n)} = \{(x_1, \dots, x_n) : x_i \in \mathbb{K}, i = 1, 2, \dots, n\}$ . Put  $\|(x_1, \dots, x_n)\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$  for  $1 \leq p < \infty$ .

**Proposition 1.3.** If  $X$  is a normed space, then the addition  $(x, y) \in X \times X \mapsto x + y \in X$  and the scalar multiplication  $(\alpha, x) \in \mathbb{K} \times X \mapsto \alpha x \in X$  both are continuous maps.

**Notation 1.4.** From now on,  $(X, \|\cdot\|)$  always denotes a normed space over a field  $\mathbb{K}$ .

For  $r > 0$  and  $x \in X$ , let

- (i)  $B(x, r) := \{y \in X : \|x - y\| < r\}$  (called an open ball with the center at  $x$  of radius  $r$ ) and  $B^*(x, r) := \{y \in X : 0 < \|x - y\| < r\}$
- (ii)  $B(x, r) := \{y \in X : \|x - y\| \leq r\}$  (called a closed ball with the center at  $x$  of radius  $r$ ).

Put  $B_X := \{x \in X : \|x\| \leq 1\}$  and  $S_X := \{x \in X : \|x\| = 1\}$  the closed unit ball and the unit sphere of  $X$  respectively.

**Definition 1.5.** We say that a sequence  $(x_n)$  in  $X$  converges to an element  $a \in X$  if  $\lim \|x_n - a\| = 0$ , i.e., for any  $\varepsilon > 0$ , there is  $N \in \mathbb{N}$  such that  $\|x_n - a\| < \varepsilon$  for all  $n \geq N$ .

In this case,  $(x_n)$  is said to be convergent and  $a$  is called a limit of the sequence  $(x_n)$ .

**Definition 1.6.** Let  $A$  be a subset of  $X$ .

- (i) A point  $z \in X$  is called a limit point of  $A$  if for any  $\varepsilon > 0$ , there is an element  $a \in A$  such that  $0 < \|z - a\| < \varepsilon$ , that is,  $B^*(z, \varepsilon) \cap A \neq \emptyset$  for all  $\varepsilon > 0$ .  
Furthermore, if  $A$  contains the set of all its limit points, then  $A$  is said to be closed in  $X$ .
- (ii) The closure of  $A$ , denoted by  $\overline{A}$ , is defined by

$$\overline{A} := A \cup \{z \in X : z \text{ is a limit point of } A\}.$$

**Remark 1.7.** Using the notations as above, a point  $z \in \bar{A}$  if and only if  $B(z, r) \cap A \neq \emptyset$  for all  $r > 0$ . This is equivalent to saying that there is a sequence  $(x_n)$  in  $A$  such that  $x_n \rightarrow z$ . In fact, this can be shown by considering  $r = \frac{1}{n}$  for  $n = 1, 2, \dots$

**Proposition 1.8.** Using the notations as before, we have the following assertions.

- (i)  $A$  is closed in  $X$  if and only if its complement  $X \setminus A$  is open in  $X$ .
- (ii) The closure  $\bar{A}$  is the smallest closed subset of  $X$  containing  $A$ . The "smallest" in here means that if  $F$  is a closed subset containing  $A$ , then  $\bar{A} \subseteq F$ .  
Consequently,  $A$  is closed if and only if  $\bar{A} = A$ .

*Proof.* If  $A$  is empty, then the assertions (i) and (ii) both are obvious. Now assume that  $A \neq \emptyset$ . For part (i), let  $C = X \setminus A$  and  $b \in C$ . Suppose that  $A$  is closed in  $X$ . If there exists an element  $b \in C \setminus \text{int}(C)$ , then  $B(b, r) \not\subseteq C$  for all  $r > 0$ . This implies that  $B(b, r) \cap A \neq \emptyset$  for all  $r > 0$  and hence,  $b$  is a limit point of  $A$  since  $b \notin A$ . It contradicts to the closedness of  $A$ . Thus,  $C = \text{int}(C)$  and thus,  $C$  is open.

For the converse of (i), assume that  $C$  is open in  $X$ . Assume that  $A$  has a limit point  $z$  but  $z \notin A$ . Since  $z \notin A$ ,  $z \in C = \text{int}(C)$  because  $C$  is open. Hence, we can find  $r > 0$  such that  $B(z, r) \subseteq C$ . This gives  $B(z, r) \cap A = \emptyset$ . This contradicts to the assumption of  $z$  being a limit point of  $A$ . Thus,  $A$  must contain all of its limit points and hence, it is closed.

For part (ii), we first claim that  $\bar{A}$  is closed. Let  $z$  be a limit point of  $\bar{A}$ . Let  $r > 0$ . Then there is  $w \in B^*(z, r) \cap \bar{A}$ . Choose  $0 < r_1 < r$  small enough such that  $B(w, r_1) \subseteq B^*(z, r)$ . Since  $w$  is a limit point of  $A$ , we have  $\emptyset \neq B^*(w, r_1) \cap A \subseteq B^*(z, r) \cap A$ . Hence,  $z$  is a limit point of  $A$ . Thus,  $z \in \bar{A}$  as required. This implies that  $\bar{A}$  is closed.

It is clear that  $\bar{A}$  is the smallest closed set containing  $A$ .

The last assertion follows from the minimality of the closed sets containing  $A$  immediately.

The proof is complete.  $\square$

A sequence  $(x_n)$  in  $X$  is called a **Cauchy sequence** if for any  $\varepsilon > 0$ , there is  $N \in \mathbb{N}$  such that  $\|x_m - x_n\| < \varepsilon$  for all  $m, n \geq N$ . We have the following simple observation.

**Lemma 1.9.** Every convergent sequence in  $X$  is a Cauchy sequence.

The following notation plays an important role in mathematics.

**Definition 1.10.** A normed space  $X$  is called a **Banach space** if it is a complete normed space, i.e., every Cauchy sequence in  $X$  is convergent.

**Proposition 1.11.** Let  $X$  be a normed space. Then the following assertions are equivalent.

- (i)  $X$  is a Banach space.
- (ii) If a series  $\sum_{n=1}^{\infty} x_n$  is absolutely convergent in  $X$ , i.e.,  $\sum_{n=1}^{\infty} \|x_n\| < \infty$ , implies that the series  $\sum_{n=1}^{\infty} x_n$  converges in the norm.

*Proof.* (i)  $\Rightarrow$  (ii) is obvious.

Now suppose that Part (ii) holds. Let  $(y_n)$  be a Cauchy sequence in  $X$ . It suffices to show that  $(y_n)$  has a convergent subsequence. In fact, by the definition of a Cauchy sequence, there is a subsequence  $(y_{n_k})$  such that  $\|y_{n_{k+1}} - y_{n_k}\| < \frac{1}{2^k}$  for all  $k = 1, 2, \dots$ . By the assumption, the series  $\sum_{k=1}^{\infty} (y_{n_{k+1}} - y_{n_k})$  converges in the norm, and hence the sequence  $(y_{n_k})$  is convergent in  $X$ . The proof is complete.  $\square$

Throughout the note, we write a sequence of numbers as a function  $x : \{1, 2, \dots\} \rightarrow \mathbb{K}$ . The following examples are important classes in the study of functional analysis.

**Example 1.12.** Put

$$c_0 := \{(x(i)) : x(i) \in \mathbb{K}, \lim_{i \rightarrow \infty} |x(i)| = 0\} \text{ (the null sequence space);}$$

$$\ell_\infty := \{(x(i)) : x(i) \in \mathbb{K}, \sup_i |x(i)| < \infty\} \text{ (the bounded sequence space);}$$

and

$$c_{00} := \{(x(i)) : \text{there are only finitely many } x(i)\text{'s are non-zero}\} \text{ (the finite sequence space).}$$

The sup-norm  $\|\cdot\|_\infty$  on  $\ell_\infty$  is defined by  $\|x\|_\infty := \sup_i |x(i)|$  for  $x \in \ell_\infty$ . Then  $\ell_\infty$  is a Banach space.

Now if  $c_{00}$  is endowed with the sup-norm defined above, then  $c_{00}$  is dense in  $c_0$ , i.e.,  $\overline{c_{00}} = c_0$ . Consequently,  $c_0$  is a closed subspace of  $\ell_\infty$ . In particular,  $c_0$  is Banach space too.

*Proof.* We first claim that  $\overline{c_{00}} \subseteq c_0$ . Let  $z \in \ell_\infty$ . It suffices to show that if  $z \in \overline{c_{00}}$ , then  $z \in c_0$ , i.e.,  $\lim_{i \rightarrow \infty} z(i) = 0$ . Let  $\varepsilon > 0$ . Then there is  $x \in B(z, \varepsilon) \cap c_{00}$  and hence, we have  $|x(i) - z(i)| < \varepsilon$  for all  $i = 1, 2, \dots$ . Since  $x \in c_{00}$ , there is  $i_0 \in \mathbb{N}$  such that  $x(i) = 0$  for all  $i \geq i_0$ . Therefore, we have  $|z(i)| = |z(i) - x(i)| < \varepsilon$  for all  $i \geq i_0$ . Therefore,  $z \in c_0$  is as desired.

For the reverse inclusion, let  $w \in c_0$ . We need to show that  $B(w, r) \cap c_{00} \neq \emptyset$  for all  $r > 0$ . Let  $r > 0$ . Since  $w \in c_0$ , there is  $i_0$  such that  $|w(i)| < r$  for all  $i \geq i_0$ . If we let  $x(i) = w(i)$  for  $1 \leq i < i_0$  and  $x(i) = 0$  for  $i \geq i_0$ , then  $x \in c_{00}$  and  $\|x - w\|_\infty := \sup_{i=1,2,\dots} |x(i) - w(i)| < r$  is as required.  $\square$

**Example 1.13.** For  $1 \leq p < \infty$ . Put

$$\ell_p := \{(x(i)) : x(i) \in \mathbb{K}, \sum_{i=1}^{\infty} |x(i)|^p < \infty\}.$$

In addition,  $\ell_p$  is equipped with the norm  $\|x\|_p := \left(\sum_{i=1}^{\infty} |x(i)|^p\right)^{\frac{1}{p}}$  for  $x \in \ell_p$ . Then  $\ell_p$  becomes a Banach space under the norm  $\|\cdot\|_p$ .

**Example 1.14.** Let  $X$  be a locally compact Hausdorff space, for example,  $\mathbb{K}$ . Let  $C_0(X)$  be the space of all continuous  $\mathbb{K}$ -valued functions  $f$  on  $X$  which are vanish at infinity, i.e., for every  $\varepsilon > 0$ , there is a compact subset  $D$  of  $X$  such that  $|f(x)| < \varepsilon$  for all  $x \in X \setminus D$ . Now  $C_0(X)$  is endowed with the sup-norm, i.e.,

$$\|f\|_\infty = \sup_{x \in X} |f(x)|$$

for every  $f \in C_0(X)$ . Then  $C_0(X)$  is a Banach space. (**Try to prove this fact for the case  $X = \mathbb{R}$ . Just use the knowledge from MATH 2060 !!!**)

**Proposition 1.15.** Let  $(X, \|\cdot\|)$  be a normed space. Then there is a normed space  $(X_0, \|\cdot\|_0)$ , together with a linear map  $i : X \rightarrow X_0$ , satisfies the following conditions.

- (i)  $X_0$  is a Banach space.
- (ii) The map  $i$  is an isometry, that is,  $\|i(x)\|_0 = \|x\|$  for all  $x \in X$ .
- (iii) the image  $i(X)$  is dense in  $X_0$ , that is,  $\overline{i(X)} = X_0$ .

Moreover, such pair  $(X_0, i)$  is unique up to isometric isomorphism in the following sense.

If  $(W, \|\cdot\|_1)$  is a Banach space and an isometry  $j : X \rightarrow W$  is an isometry such that  $\overline{j(X)} = W$ , then there is an isometric isomorphism  $\psi$  from  $X_0$  onto  $W$  such that

$$j = \psi \circ i : X \rightarrow X_0 \rightarrow W.$$

In this case, the pair  $(X_0, i)$  is called the completion of  $X$ .

**Example 1.16.** Proposition 1.15 cannot give an explicit form of the completion of a given normed space. The following examples are basically due to the uniqueness of the completion.

- (i) If  $X$  is a Banach space, then the completion of  $X$  is itself.
- (ii) The completion of the finite sequence space  $c_{00}$  is the null sequence space  $c_0$ .
- (iii) The completion of  $C_c(\mathbb{R})$  is  $C_0(\mathbb{R})$ .

## 2. FINITE DIMENSIONAL NORMED SPACES

Throughout this section, let  $(X, \|\cdot\|)$  is a normed space. Put  $S_X$  the unit sphere of  $X$ , i.e.,  $S_X = \{x \in X : \|x\| = 1\}$ .

**Definition 2.1.** Two norms  $\|\cdot\|$  and  $\|\cdot\|'$  on a vector space  $X$  are equivalent, denoted by  $\|\cdot\| \sim \|\cdot\|'$ , if there are positive numbers  $c_1$  and  $c_2$  such that  $c_1\|\cdot\| \leq \|\cdot\|' \leq c_2\|\cdot\|$  on  $X$ .

**Example 2.2.** Consider the norms  $\|\cdot\|_1$  and  $\|\cdot\|_\infty$  on  $\ell^1$ . We want to show that  $\|\cdot\|_1$  and  $\|\cdot\|_\infty$  are not equivalent. In fact, if we put  $x_n(i) := (1, 1/2, \dots, 1/n, 0, 0, \dots)$  for  $n, i = 1, 2, \dots$ . Then  $x_n \in \ell^1$  for all  $n$ . Note that  $(x_n)$  is a Cauchy sequence with respect to the norm  $\|\cdot\|_\infty$  but it is not a Cauchy sequence with respect to the norm  $\|\cdot\|_1$ . Hence  $\|\cdot\|_1 \not\sim \|\cdot\|_\infty$  on  $\ell^1$ .

**Proposition 2.3.** All norms on a finite dimensional vector space are equivalent.

*Proof.* Let  $X$  be a finite dimensional vector space and let  $\{e_1, \dots, e_n\}$  be a vector basis of  $X$ . For each  $x = \sum_{i=1}^n \alpha_i e_i$  for  $\alpha_i \in \mathbb{K}$ , define  $\|x\|_0 = \max_{i=1}^n |\alpha_i|$ . Then  $\|\cdot\|_0$  is a norm on  $X$ . The result is obtained by showing that all norms  $\|\cdot\|$  on  $X$  are equivalent to  $\|\cdot\|_0$ .

Note that for each  $x = \sum_{i=1}^n \alpha_i e_i \in X$ , we have  $\|x\| \leq (\sum_{1 \leq i \leq n} \|e_i\|) \|x\|_0$ . It remains to find  $c > 0$

such that  $c\|\cdot\|_0 \leq \|\cdot\|$ . In fact, let  $S_X := \{x \in X : \|x\|_0 = 1\}$  be the unit sphere of  $X$  with respect to the norm  $\|\cdot\|_0$ . Note that by using the Weierstrass Theorem on  $\mathbb{K}$ , we see that  $S_X$  is compact with respect to the norm  $\|\cdot\|_0$ .

Define a real-valued function  $f$  on the unit sphere  $S_X$  of  $X$  by

$$f : x \in S_X \mapsto \|x\|.$$

Note that  $f > 0$  and  $f$  is continuous with respect to the norm  $\|\cdot\|_0$ . Hence, there is  $c > 0$  such that  $f(x) \geq c > 0$  for all  $x \in S_X$ . This gives  $\|x\| \geq c\|x\|_0$  for all  $x \in X$  as desired. The proof is complete.  $\square$

**Corollary 2.4.** We have the following assertions.

- (i) All finite dimensional normed spaces are Banach spaces. Consequently, any finite dimensional subspace of a normed space is closed.
- (ii) The closed unit ball of any finite dimensional normed space is compact.

*Proof.* Let  $(X, \|\cdot\|)$  be a finite dimensional normed space. Using the notations as in the proof of Proposition 2.3 above, we see that  $\|\cdot\|$  must be equivalent to the norm  $\|\cdot\|_0$ .  $X$  is clearly complete with respect to the norm  $\|\cdot\|_0$  and so is complete in the original norm  $\|\cdot\|$ . The Part (i) follows. For Part (ii), it is clear that the compactness of the closed unit ball of  $X$  is equivalent to saying that any closed and bounded subset is compact. Therefore, Part (ii) follows from the simple observation that any closed and bounded subset of  $X$  with respect to the norm  $\|\cdot\|_0$  is compact. The proof is complete.  $\square$

In the remainder of this section, we want to show that the converse of Corollary 2.4(ii) holds. Before this result, we need the following useful result.

**Lemma 2.5. Riesz's Lemma:** *Let  $Y$  be a closed proper subspace of a normed space  $X$ . Then for each  $\theta \in (0, 1)$ , there is an element  $x_0 \in S_X$  such that  $d(x_0, Y) := \inf\{\|x_0 - y\| : y \in Y\} \geq \theta$ .*

*Proof.* Let  $u \in X - Y$  and  $d := \inf\{\|u - y\| : y \in Y\}$ . Note that since  $Y$  is closed,  $d > 0$  and hence we have  $0 < d < \frac{d}{\theta}$  because  $0 < \theta < 1$ . This implies that there is  $y_0 \in Y$  such that  $0 < d \leq \|u - y_0\| < \frac{d}{\theta}$ . Now put  $x_0 := \frac{u - y_0}{\|u - y_0\|} \in S_X$ . We are going to show that  $x_0$  is as desired. Indeed, let  $y \in Y$ . Since  $y_0 + \|u - y_0\|y \in Y$ , we have

$$\|x_0 - y\| = \frac{1}{\|u - y_0\|} \|u - (y_0 + \|u - y_0\|y)\| \geq d/\|u - y_0\| > \theta.$$

Thus,  $d(x_0, Y) \geq \theta$ . □

**Remark 2.6.** The Riesz's lemma does not hold when  $\theta = 1$ . The following example can be found in the Diestel's interesting book without proof (see [6, Chapter 1 Ex.3(i)]).

Let  $X = \{x \in C([0, 1], \mathbb{R}) : x(0) = 0\}$  and  $Y = \{y \in X : \int_0^1 y(t)dt = 0\}$ . Both  $X$  and  $Y$  are endowed with the sup-norm. Note that  $Y$  is a closed proper subspace of  $X$ . We are going to show that for any  $x \in S_X$ , there is  $y \in Y$  such that  $\|x - y\|_\infty < 1$ . Thus, the Riesz's Lemma does not hold as  $\theta = 1$  in this case.

In fact, let  $x \in S_X$ . Since  $x(0) = 0$  with  $\|x\|_\infty = 1$ , we can find  $0 < a < 1/4$  such that  $|x(t)| \leq 1/4$  for all  $t \in [0, a]$ .

We fix  $0 < \varepsilon < 1/4$  first. Since  $x$  is uniform continuous on  $[a, 1]$ , we can find a partitions  $a = t_0 < \dots < t_n = 1$  on  $[a, 1]$  such that  $\sup\{|x(t) - x(t')| : t, t' \in [t_{k-1}, t_k]\} < \varepsilon/4$ . Now for each  $(t_{k-1}, t_k)$ , if  $\sup\{x(t) : t \in [t_{k-1}, t_k]\} > \varepsilon$ , then we set  $\phi(t) = \varepsilon$ . In addition, if  $\inf\{x(t) : t \in [t_{k-1}, t_k]\} < -\varepsilon$ , then we set  $\phi(t) = -\varepsilon$ . From this, one can construct a continuous function  $\phi$  on  $[a, 1]$  such that  $\|\phi - x\|_{[a, 1]} < 1$  and  $|\phi(x)| < 2\varepsilon$  for all  $x \in [a, 1]$ . Hence, we have  $|\int_a^1 \phi(t)dt| \leq 2\varepsilon(1 - a)$ .

As  $|x(t)| < 1/4$  on  $[0, a]$ , so if we choose  $\varepsilon$  small enough such that  $(1 - a)(2\varepsilon) < a/4$ , then we can find a continuous function  $y_1$  on  $[0, a]$  such that  $|y_1(t)| < 1/4$  on  $[0, a]$  with  $y_1(0) = 0$ ;  $y_1(a) = x(a)$  and  $\int_0^a y_1(t)dt = -\int_a^1 \phi(t)dt$ . Now we define  $y = y_1$  on  $[0, a]$  and  $y = \phi$  on  $[a, 1]$ . Then  $\|y - x\|_\infty < 1$  and  $y \in Y$  is as desired.

**Theorem 2.7.**  *$X$  is a finite dimensional normed space if and only if the closed unit ball  $B_X$  of  $X$  is compact.*

*Proof.* The necessary condition has been shown by Proposition 2.4(ii).

Now assume that  $X$  is of infinite dimension. Fix an element  $x_1 \in S_X$ . Let  $Y_1 = \mathbb{K}x_1$ . Then  $Y_1$  is a proper closed subspace of  $X$ . The Riesz's lemma gives an element  $x_2 \in S_X$  such that  $\|x_1 - x_2\| \geq 1/2$ . Now consider  $Y_2 = \text{span}\{x_1, x_2\}$ . Then  $Y_2$  is a proper closed subspace of  $X$  since  $\dim X = \infty$ . To apply the Riesz's Lemma again, there is  $x_3 \in S_X$  such that  $\|x_3 - x_k\| \geq 1/2$  for  $k = 1, 2$ . To repeat the same step, there is a sequence  $(x_n) \in S_X$  such that  $\|x_m - x_n\| \geq 1/2$  for all  $n \neq m$ . Thus,  $(x_n)$  is a bounded sequence without any convergence subsequence. Hence,  $B_X$  is not compact. The proof is complete. □

Recall that a metric space  $Z$  is said to be *locally compact* if for any point  $z \in Z$ , there is a compact neighborhood of  $z$ . Theorem 2.7 implies the following corollary immediately.

**Corollary 2.8.** *Let  $X$  be a normed space. Then  $X$  is locally compact if and only if  $\dim X < \infty$ .*

In view of the proof of Theorem 2.7, we call a subset  $A$  of a normed space  $X$  a *r-discrete set* if  $\|x - y\| > r$  whenever  $x, y \in A$  with  $x \neq y$ .

**Proposition 2.9.** *Let  $X$  be an infinite dimensional normed space and  $r > 0$ . If  $A := \{x_i\}_{i \in I}$  is an infinite bounded  $r$ -discrete subset of  $X$ , then  $|I| \leq \dim X$ . Consequently, if  $\{B_i : i \in I\}$  is a collection of pairwise disjoint  $r/2$ -balls in a bounded subset of  $X$ , then  $|I| \leq \dim X$ .*

*Proof.* Let  $\mathcal{C}$  be the collection of linearly independent subsets of  $A$ . Then the Zorn's Lemma gives a maximal element  $\{x_i : i \in J\}$  in  $\mathcal{C}$  for a subset  $J$  of  $I$ . We are going to show that  $|J| = |I|$ . Suppose not. We assume that  $|J| < |I|$ . Now for each finite subset  $F$  of  $J$ , put  $V_F$  the linear span of the set  $F$ . Since  $V_F$  is of finite dimension, the set  $V_F \cap A$  has finite volume with respect to the Lebesgue measure in  $V_F$ . Then  $V_F \cap A$  is a finite set because it is a  $r$ -discrete set. Therefore, if we let  $\mathcal{F}(J) := \{F \subseteq J : |F| < \infty\}$ , then we have

$$\left| \bigcup_{F \in \mathcal{F}(J)} V_F \cap A \right| = |\mathcal{F}(J)| = |J|.$$

Hence, if  $|J| < |I|$ , then  $\bigcup_{F \in \mathcal{F}(J)} V_F \cap A \subsetneq A$ . Hence, there is an element  $x_{i_0} \in A$  so that  $x_{i_0} \notin V_F \cap A$

for all  $F \in \mathcal{F}(J)$ . This implies that  $x_{i_0}$  is linearly independent to the set  $(x_i)_{i \in J}$ . This contradicts to the maximality of  $(x_i)_{i \in J}$ . Hence,  $|I| = |J| \leq \dim X$  by the definition of dimension.  $\square$

### 3. BOUNDED LINEAR OPERATORS

**Proposition 3.1.** *Let  $T$  be a linear operator from a normed space  $X$  into a normed space  $Y$ . Then the following statements are equivalent.*

- (i)  $T$  is continuous on  $X$ .
- (ii)  $T$  is continuous at  $0 \in X$ .
- (iii)  $\sup\{\|Tx\| : x \in B_X\} < \infty$ .

In this case, let  $\|T\| = \sup\{\|Tx\| : x \in B_X\}$  and  $T$  is said to be bounded.

*Proof.* (i)  $\Rightarrow$  (ii) is obvious.

For (ii)  $\Rightarrow$  (i), suppose that  $T$  is continuous at 0. Let  $x_0 \in X$ . Let  $\varepsilon > 0$ . Then there is  $\delta > 0$  such that  $\|Tw\| < \varepsilon$  for all  $w \in X$  with  $\|w\| < \delta$ . Therefore, we have  $\|Tx - Tx_0\| = \|T(x - x_0)\| < \varepsilon$  for any  $x \in X$  with  $\|x - x_0\| < \delta$ . Part (i) follows.

For (ii)  $\Rightarrow$  (iii), since  $T$  is continuous at 0, there is  $\delta > 0$  such that  $\|Tx\| < 1$  for any  $x \in X$  with  $\|x\| < \delta$ . Now for any  $x \in B_X$  with  $x \neq 0$ , we have  $\|\frac{\delta}{2}x\| < \delta$ . Therefore, we see have  $\|T(\frac{\delta}{2}x)\| < 1$  and hence, we have  $\|Tx\| < 2/\delta$ . Part (iii) follows.

Finally, we need to show (iii)  $\Rightarrow$  (ii). Note that by the assumption of (iii), there is  $M > 0$  such that  $\|Tx\| \leq M$  for all  $x \in B_X$ . Thus, for each  $x \in X$ , we have  $\|Tx\| \leq M\|x\|$ . This implies that  $T$  is continuous at 0. The proof is complete.  $\square$

**Corollary 3.2.** *Let  $T : X \rightarrow Y$  be a bounded linear map. Then we have*

$$\sup\{\|Tx\| : x \in B_X\} = \sup\{\|Tx\| : x \in S_X\} = \inf\{M > 0 : \|Tx\| \leq M\|x\|, \forall x \in X\}.$$

*Proof.* Let  $a = \sup\{\|Tx\| : x \in B_X\}$ ,  $b = \sup\{\|Tx\| : x \in S_X\}$  and  $c = \inf\{M > 0 : \|Tx\| \leq M\|x\|, \forall x \in X\}$ .

Clearly, we have  $b \leq a$ . Now for each  $x \in B_X$  with  $x \neq 0$ , then we have  $b \geq \|T(x/\|x\|)\| = (1/\|x\|)\|Tx\| \geq \|Tx\|$ . Thus, we have  $b \geq a$  and thus,  $a = b$ .

Now if  $M > 0$  satisfies  $\|Tx\| \leq M\|x\|$ ,  $\forall x \in X$ , then we have  $\|Tw\| \leq M$  for all  $w \in S_X$ . Hence, we have  $b \leq M$  for all such  $M$ , and so we have  $b \leq c$ . Finally, it remains to show  $c \leq b$ . Note that by the definition of  $b$ , we have  $\|Tx\| \leq b\|x\|$  for all  $x \in X$ . Thus,  $c \leq b$ .  $\square$

**Proposition 3.3.** *Let  $X$  and  $Y$  be normed spaces. Let  $B(X, Y)$  be the set of all bounded linear maps from  $X$  into  $Y$ . For each element  $T \in B(X, Y)$ , let*

$$\|T\| = \sup\{\|Tx\| : x \in B_X\}.$$

*be defined as in Proposition 3.1.*

*Then  $(B(X, Y), \|\cdot\|)$  becomes a normed space.*

*Furthermore, if  $Y$  is a Banach space, then so is  $B(X, Y)$ .*

*In particular, if  $Y = \mathbb{K}$ , then  $B(X, \mathbb{K})$  is a Banach space. In this case, put  $X^* := B(X, \mathbb{K})$  and call it the **dual space** of  $X$ .*

*Proof.* We can directly check that  $B(X, Y)$  is a normed space (**Do It By Yourself!**).

We want to show that  $B(X, Y)$  is complete if  $Y$  is a Banach space. Let  $(T_n)$  be a Cauchy sequence in  $B(X, Y)$ . Then for each  $x \in X$ , it is easy to see that  $(T_n x)$  is a Cauchy sequence in  $Y$ . Thus,  $\lim T_n x$  exists in  $Y$  for each  $x \in X$  because  $Y$  is complete. Hence, we can define a map  $Tx := \lim T_n x \in Y$  for each  $x \in X$ . Clearly,  $T$  is a linear map from  $X$  into  $Y$ .

We need show that  $T \in B(X, Y)$  and  $\|T - T_n\| \rightarrow 0$  as  $n \rightarrow \infty$ . Let  $\varepsilon > 0$ . Since  $(T_n)$  is a Cauchy sequence in  $B(X, Y)$ , there is a positive integer  $N$  such that  $\|T_m - T_n\| < \varepsilon$  for all  $m, n \geq N$ . Hence, we have  $\|(T_m - T_n)(x)\| < \varepsilon$  for all  $x \in B_X$  and  $m, n \geq N$ . Taking  $m \rightarrow \infty$ , we have  $\|Tx - T_n x\| \leq \varepsilon$  for all  $n \geq N$  and  $x \in B_X$ . Therefore, we have  $\|T - T_n\| \leq \varepsilon$  for all  $n \geq N$ . From this, we see that  $T - T_N \in B(X, Y)$  and thus,  $T = T_N + (T - T_N) \in B(X, Y)$  and  $\|T - T_n\| \rightarrow 0$  as  $n \rightarrow \infty$ . Therefore,  $\lim_n T_n = T$  exists in  $B(X, Y)$ .  $\square$

**Remark 3.4.** *By using Proposition 3.1, we can show that if  $f : X \rightarrow \mathbb{K}$  is any linear functional defined on a vector space  $X$ , then  $X$  can be endowed with a norm so that  $f$  is bounded.*

*In fact, if we fix a vector basis  $(e_i)_{i \in I}$  for  $X$  and put  $\|x\|_\infty := \max_{i \in I} |a_i|$  as  $x = \sum_{i \in I} a_i e_i \in X$ , (note that it is a finite sum), where  $a_i \in \mathbb{K}$ , then the function  $\|\cdot\|_\infty$  is a norm on  $X$ . Now for each  $x \in X$ , set*

$$\|x\|_1 := |f(x)| + \|x\|_\infty.$$

*Clearly, the function  $\|\cdot\|_1$  is a norm on  $X$ . In addition, we have  $|f(x)| \leq \|x\|_1$  for all  $x \in X$ . Hence,  $f$  is bounded on  $X$  with respect to the norm  $\|\cdot\|_1$  as required.*

**Proposition 3.5.** *Let  $X$  and  $Y$  be normed spaces. Suppose that  $X$  is of finite dimension  $n$ . Then we have the following assertions.*

(i) *Any linear operator from  $X$  into  $Y$  must be bounded.*

(ii) *If  $T_k : X \rightarrow Y$  is a sequence of linear operators such that  $T_k x \rightarrow 0$  for all  $x \in X$ , then  $\|T_k\| \rightarrow 0$ .*

*Proof.* Using Proposition 2.3 and the notations as in the proof, then there is  $c > 0$  such that

$$\sum_{i=1}^n |\alpha_i| \leq c \left\| \sum_{i=1}^n \alpha_i e_i \right\|$$

for all scalars  $\alpha_1, \dots, \alpha_n$ . Therefore, for any linear map  $T$  from  $X$  to  $Y$ , we have

$$\|Tx\| \leq \left( \max_{1 \leq i \leq n} \|Te_i\| \right) c \|x\|$$

for all  $x \in X$ . This gives the assertions (i) and (ii) immediately.  $\square$

**Remark 3.6.** *The assumption of  $X$  of finite dimension in Proposition 3.5 cannot be removed. For example, if for each positive integer  $k$ , we define  $f_k : c_0 \rightarrow \mathbb{R}$  by  $f_k(x) := x(k)$ , then  $f_k$  is bounded for each  $k$  and*

$$\lim_{k \rightarrow \infty} f_k(x) = \lim_{k \rightarrow \infty} x(k) = 0$$

*for all  $x \in c_0$ . However  $f_k \not\rightarrow 0$  because  $\|f_k\| \equiv 1$  for every  $k$ .*

**Proposition 3.7.** *Let  $Y$  be a closed subspace of  $X$  and  $X/Y$  be the quotient space. For each element  $x \in X$ , put  $\bar{x} := x + Y \in X/Y$  the corresponding element in  $X/Y$ . Define*

$$(3.1) \quad \|\bar{x}\| = \inf\{\|x + y\| : y \in Y\}.$$

*If we let  $\pi : X \rightarrow X/Y$  be the natural projection, i.e.,  $\pi(x) = \bar{x}$  for all  $x \in X$ , then  $(X/Y, \|\cdot\|)$  is a normed space and  $\pi$  is bounded with  $\|\pi\| \leq 1$ . In particular,  $\|\pi\| = 1$  as  $Y$  is a proper closed subspace.*

*Furthermore, if  $X$  is a Banach space, then so is  $X/Y$ .*

*In this case, we call  $\|\cdot\|$  in (3.1) the quotient norm on  $X/Y$ .*

*Proof.* Note that since  $Y$  is closed, we can directly check that  $\|\bar{x}\| = 0$  if and only if  $x \in Y$ , i.e.,  $\bar{x} = \bar{0} \in X/Y$ . It is easy to check the other conditions of the definition of a norm. Thus,  $X/Y$  is a normed space. Moreover,  $\pi$  is clearly bounded with  $\|\pi\| \leq 1$  by the definition of the quotient norm on  $X/Y$ .

Furthermore, if  $Y \subsetneq X$ , then by using the Riesz's Lemma 2.5, we see that  $\|\pi\| = 1$ .

We show the last assertion. Suppose that  $X$  is a Banach space. Let  $(\bar{x}_n)$  be a Cauchy sequence in  $X/Y$ . It suffices to show that  $(\bar{x}_n)$  has a convergent subsequence in  $X/Y$ .

Indeed, since  $(\bar{x}_n)$  is a Cauchy sequence, we can find a subsequence  $(\bar{x}_{n_k})$  of  $(\bar{x}_n)$  such that

$$\|\bar{x}_{n_{k+1}} - \bar{x}_{n_k}\| < 1/2^k$$

for all  $k = 1, 2, \dots$ . Then by the definition of quotient norm, there is an element  $y_1 \in Y$  such that  $\|x_{n_2} - x_{n_1} + y_1\| < 1/2$ . Note that we have,  $\overline{x_{n_1} - y_1} = \bar{x}_{n_1}$  in  $X/Y$ . Thus, there is  $y_2 \in Y$  such that  $\|x_{n_2} - y_2 - (x_{n_1} - y_1)\| < 1/2$  by the definition of quotient norm again. In addition, we have  $\overline{x_{n_2} - y_2} = \bar{x}_{n_2}$ . Then we also have an element  $y_3 \in Y$  such that  $\|x_{n_3} - y_3 - (x_{n_2} - y_2)\| < 1/2^2$ . To repeat the same step, we can obtain a sequence  $(y_k)$  in  $Y$  such that

$$\|x_{n_{k+1}} - y_{k+1} - (x_{n_k} - y_k)\| < 1/2^k$$

for all  $k = 1, 2, \dots$ . Therefore,  $(x_{n_k} - y_k)$  is a Cauchy sequence in  $X$  and thus,  $\lim_k (x_{n_k} - y_k)$  exists in  $X$  while  $X$  is a Banach space. Set  $x = \lim_k (x_{n_k} - y_k)$ . On the other hand, note that we have  $\pi(x_{n_k} - y_k) = \pi(x_{n_k})$  for all  $k = 1, 2, \dots$ . This tells us that  $\lim_k \pi(x_{n_k}) = \lim_k \pi(x_{n_k} - y_k) = \pi(x) \in X/Y$  since  $\pi$  is bounded. Therefore,  $(\bar{x}_{n_k})$  is a convergent subsequence of  $(\bar{x}_n)$  in  $X/Y$ . The proof is complete.  $\square$

**Corollary 3.8.** *Let  $T : X \rightarrow Y$  be a linear map. Suppose that  $Y$  is of finite dimension. Then  $T$  is bounded if and only if  $\ker T := \{x \in X : Tx = 0\}$  is closed.*

*Proof.* The necessary part is clear.

Now assume that  $\ker T$  is closed. Then by Proposition 3.7,  $X/\ker T$  becomes a normed space. Moreover, it is known that there is a linear injection  $\tilde{T} : X/\ker T \rightarrow Y$  such that  $T = \tilde{T} \circ \pi$ , where  $\pi : X \rightarrow X/\ker T$  is the natural projection. Since  $\dim Y < \infty$  and  $\tilde{T}$  is injective,  $\dim X/\ker T < \infty$ . This implies that  $\tilde{T}$  is bounded by Proposition 3.5. Hence  $T$  is bounded because  $T = \tilde{T} \circ \pi$  and  $\pi$  is bounded.  $\square$

**Remark 3.9.** The converse of Corollary 3.8 does not hold when  $Y$  is of infinite dimension. For example, let  $X := \{x \in \ell^2 : \sum_{n=1}^{\infty} n^2 |x(n)|^2 < \infty\}$  (note that  $X$  is a vector space **Why?**) and  $Y = \ell^2$ . Both  $X$  and  $Y$  are endowed with  $\|\cdot\|_2$ -norm.

Define  $T : X \rightarrow Y$  by  $Tx(n) = nx(n)$  for  $x \in X$  and  $n = 1, 2, \dots$ . Then  $T$  is an unbounded operator (**Check !!**). Note that  $\ker T = \{0\}$  and hence,  $\ker T$  is closed. Hence, the closeness of  $\ker T$  does not imply the boundedness of  $T$  in general.



Two normed spaces  $X$  and  $Y$  are said to be *isomorphic* (resp. *isometric isomorphic*) if there is a bi-continuous linear isomorphism (resp. isometric) between  $X$  and  $Y$ . We write  $X = Y$  if  $X$  and  $Y$  are isometric isomorphic.

**Remark 3.10.** Note that the inverse of a bounded linear isomorphism need not be bounded.

**Example 3.11.** Let  $X := \{f \in C^\infty(-1, 1) : f^{(n)} \in C^b(-1, 1) \text{ for all } n = 0, 1, 2, \dots\}$  and  $Y := \{f \in X : f(0) = 0\}$ . In addition,  $X$  and  $Y$  both are equipped with the sup-norm  $\|\cdot\|_\infty$ . Define an operator  $S : X \rightarrow Y$  by

$$Sf(x) := \int_0^x f(t)dt$$

for  $f \in X$  and  $x \in (-1, 1)$ . Then  $S$  is a bounded linear isomorphism but its inverse  $S^{-1}$  is unbounded. In fact, the inverse  $S^{-1} : Y \rightarrow X$  is given by

$$S^{-1}g := g'$$

for  $g \in Y$ .

A metric space is said to be *separable* if there is a countable dense subset, for example, the base field  $\mathbb{K}$  is separable. Moreover, it is easy to see that a normed space is separable if and only if it is the closed linear span of a countable dense subset.

**Definition 3.12.** A sequence of element  $(e_n)_{n=1}^\infty$  in a normed space  $X$  is called a *Schauder basis* for  $X$  if for each element  $x \in X$ , there is a unique sequence of scalars  $(\alpha_n)$  such that

$$(3.2) \quad x = \sum_{n=1}^{\infty} \alpha_n e_n.$$

**Note:** The expression in Eq. 3.2 depends on the order of  $e_n$ 's.

**Remark 3.13.** Note that if  $X$  has a Schauder basis, then  $X$  must be separable. The following natural question was first raised by Banach (1932).

**The basis problem:** Does every separable Banach space have a Schauder basis?

The answer is “No”!

This problem was completely solved by P. Enflo in 1973.

**Example 3.14.** We have the following assertions.

(i) The space  $\ell^\infty$  is non-separable under the sup-norm  $\|\cdot\|_\infty$ . Consequently,  $\ell^\infty$  has no Schauder basis.

(ii) The spaces  $c_0$  and  $\ell^p$  for  $1 \leq p < \infty$  have Schauder bases.

*Proof.* For Part (i) let  $D = \{x \in \ell^\infty : x(i) = 0 \text{ or } 1\}$ . Then  $D$  is an uncountable set and  $\|x - y\|_\infty = 1$  for  $x \neq y$ . Therefore  $\{B(x, 1/4) : x \in D\}$  is an uncountable family of disjoint open balls. Therefore,  $\ell^\infty$  has no countable dense subset.

For each  $n = 1, 2, \dots$ , let  $e_n(i) = 1$  if  $n = i$ , otherwise, is equal to 0.

In addition,  $(e_n)$  is a Schauder basis for the space  $c_0$  and  $\ell^p$  for  $1 \leq p < \infty$ . □

In the rest of this section, we are going to investigate some concrete examples of dual spaces.

**Example 3.15.** Let  $X = \mathbb{K}^N$ . Consider the usual Euclidean norm on  $X$ , i.e.,  $\|(x_1, \dots, x_N)\| := \sqrt{|x_1|^2 + \dots + |x_N|^2}$ . Define  $\theta : \mathbb{K}^N \rightarrow (\mathbb{K}^N)^*$  by  $\theta x(y) = x_1 y_1 + \dots + x_N y_N$  for  $x = (x_1, \dots, x_N)$  and  $y = (y_1, \dots, y_N) \in \mathbb{K}^N$ . Note that  $\theta x(y) = \langle x, y \rangle$ , the usual inner product on  $\mathbb{K}^N$ . Then by the Cauchy-Schwarz inequality, it is easy to see that  $\theta$  is an isometric isomorphism. Therefore, we have  $\mathbb{K}^N = (\mathbb{K}^N)^*$ .

**Example 3.16.** Define a map  $T : \ell^1 \rightarrow c_0^*$  by

$$(Tx)(\eta) = \sum_{i=1}^{\infty} x(i)\eta(i)$$

for  $x \in \ell^1$  and  $\eta \in c_0$ .

Then  $T$  is isometric isomorphism and hence,  $c_0^* = \ell^1$ .

*Proof.* The proof is divided into the following steps.

**Step 1.**  $Tx \in c_0^*$  for all  $x \in \ell^1$ .

In fact, let  $\eta \in c_0$ . Then

$$|Tx(\eta)| \leq \left| \sum_{i=1}^{\infty} x(i)\eta(i) \right| \leq \sum_{i=1}^{\infty} |x(i)||\eta(i)| \leq \|x\|_1 \|\eta\|_{\infty}.$$

Step 1 follows.

**Step 2.**  $T$  is an isometry.

Note that by Step 1, we have  $\|Tx\| \leq \|x\|_1$  for all  $x \in \ell^1$ . We need to show that  $\|Tx\| \geq \|x\|_1$  for all  $x \in \ell^1$ . Fix  $x \in \ell^1$ . Now for each  $k = 1, 2, \dots$ , consider the polar form  $x(k) = |x(k)|e^{i\theta_k}$ . Note that  $\eta_n := (e^{-i\theta_1}, \dots, e^{-i\theta_n}, 0, 0, \dots) \in c_0$  for all  $n = 1, 2, \dots$ . Then we have

$$\sum_{k=1}^n |x(k)| = \sum_{k=1}^n x(k)\eta_n(k) = Tx(\eta_n) = |Tx(\eta_n)| \leq \|Tx\|$$

for all  $n = 1, 2, \dots$ . Hence, we have  $\|x\|_1 \leq \|Tx\|$ .

**Step 3.**  $T$  is a surjection.

Let  $\phi \in c_0^*$  and let  $e_k \in c_0$  be given by  $e_k(j) = 1$  if  $j = k$ , otherwise, is equal to 0. Put  $x(k) := \phi(e_k)$  for  $k = 1, 2, \dots$  and consider the polar form  $x(k) = |x(k)|e^{i\theta_k}$  as above. Then we have

$$\sum_{k=1}^n |x(k)| = \phi\left(\sum_{k=1}^n e^{-i\theta_k} e_k\right) \leq \|\phi\| \left\| \sum_{k=1}^n e^{-i\theta_k} e_k \right\|_{\infty} = \|\phi\|$$

for all  $n = 1, 2, \dots$ . Therefore,  $x \in \ell^1$ .

Finally, we need to show that  $Tx = \phi$  and thus,  $T$  is surjective. In fact, if  $\eta = \sum_{k=1}^{\infty} \eta(k)e_k \in c_0$ , then we have

$$\phi(\eta) = \sum_{k=1}^{\infty} \eta(k)\phi(e_k) = \sum_{k=1}^{\infty} \eta(k)x(k) = Tx(\eta).$$

The proof is complete by the Steps 1-3 above.  $\square$

**Example 3.17.** We have the other important examples of the dual spaces.

- (i)  $(\ell^1)^* = \ell^{\infty}$ .
- (ii) For  $1 < p < \infty$ ,  $(\ell^p)^* = \ell^q$ , where  $\frac{1}{p} + \frac{1}{q} = 1$ .
- (iii) For a locally compact Hausdorff space  $X$ ,  $C_0(X)^* = M(X)$ , where  $M(X)$  denotes the space of all regular Borel measures on  $X$ .

Parts (i) and (ii) can be obtained by the similar argument as in Example 3.16 (see also in [?, Chapter 8]). Part (iii) is known as the *Riesz representation Theorem* which is referred to [?, Section 21.5] for the details.

**Example 3.18.** Let  $C[a, b]$  be the space of all continuous  $\mathbb{R}$ -valued functions defined on a closed and bounded interval  $[a, b]$ . Moreover, the space  $C[a, b]$  is endowed with the sup-norm, i.e.,  $\|f\|_{\infty} :=$

$\sup\{|f(x)| : x \in [a, b] \text{ for } f \in C[a, b]\}.$

A function  $\rho : [a, b] \rightarrow \mathbb{R}$  is said to be a bounded variation if it satisfies the condition:

$$V(\rho) := \sup\left\{\sum_{k=1}^n |\rho(x_k) - \rho(x_{k-1})| : a = x_0 < x_1 < \cdots < x_n = b\right\} < \infty.$$

Let  $BV([a, b])$  denote the space of all bounded variations on  $[a, b]$  and let  $\|\rho\| := V(\rho)$  for  $\rho \in BV([a, b])$ . Then  $BV([a, b])$  becomes a Banach space.

Besides, for  $f \in C[a, b]$ , the Riemann-Stieltjes integral of  $f$  with respect to a bounded variation  $\rho$  on  $[a, b]$  is defined by

$$\int_a^b f(x) d\rho(x) := \lim_P \sum_{k=1}^n f(\xi_k)(\rho(x_k) - \rho(x_{k-1})),$$

where  $P : a = x_0 < x_1 < \cdots < x_n = b$  and  $\xi_k \in [x_{k-1}, x_k]$  (**Fact: the Riemann-Stieltjes integral of a continuous function always exists**).

Define a mapping  $T : BV([a, b]) \rightarrow C[a, b]^*$  by

$$T(\rho)(f) := \int_a^b f(x) d\rho(x)$$

for  $\rho \in BV([a, b])$  and  $f \in C[a, b]$ . Then  $T$  is an isometric isomorphism, and hence, we have

$$C[a, b]^* = BV([a, b]).$$

#### 4. HAHN-BANACH THEOREM

A real valued function  $p : X \rightarrow \mathbb{R}$  defined on a vector space  $X$  is called a *positively homogeneous sub-additive* if the following conditions hold:

- (i)  $p(\alpha x) = \alpha p(x)$  for all  $x \in X$  and  $\alpha \geq 0$ .
- (ii)  $p(x + y) \leq p(x) + p(y)$  for all  $x, y \in X$ .

**Lemma 4.1.** *Let  $X$  be a real vector space and  $Y$  be a subspace of  $X$ . Assume that there is an element  $v \in X \setminus Y$  such that  $X = Y \oplus \mathbb{R}v$ , i.e., the space  $X$  is the linear span of  $Y$  and  $v$ . Let  $p$  be a positive homogeneous sub-additive function defined on  $X$ . Suppose that  $f$  is real linear functional defined on  $Y$  satisfying  $f(y) \leq p(y)$  for all  $y \in Y$ . Then there is a real linear extension  $F$  of  $f$  defined on  $X$  so that*

$$F(x) \leq p(x) \quad \text{for all } x \in X.$$

*Proof.* It is noted that if  $F$  is a linear extension of  $f$  on  $X$  and  $\gamma := F(v)$  which satisfies

$$F(y + tv) = f(y) + t\gamma \leq p(y + tv) \quad \text{for all } y \in Y \text{ and for all } t \in \mathbb{R},$$

then it suffices to saying that the following inequalities hold:

$$(4.1) \quad f(y_1) + \gamma \leq p(y_1 + v) \quad \text{and} \quad f(y_2) - \gamma \leq p(y_2 - v)$$

for all  $y_1, y_2 \in Y$ . Thus, we need to determine  $\gamma := F(v)$  so that the following holds:

$$(4.2) \quad f(y_1) - p(y_1 - v) \leq \gamma \leq -f(y_2) + p(y_2 + v) \quad \text{for all } y_1, y_2 \in Y.$$

Note that if we fix  $y_1, y_2 \in Y$ , we see that

$$f(y_1) + f(y_2) = f(y_1 + y_2) \leq p(y_1 + y_2) \leq p(y_1 - v) + p(y_2 + v).$$

This implies that we have

$$f(y_1) - p(y_1 - v) \leq -f(y_2) + p(y_2 + v)$$

for all  $y_1, y_2 \in Y$ . Therefore, it gives

$$a := \sup\{f(y_1) - p(y_1 - v) : y_1 \in Y\} \leq b := \inf\{-f(y_2) + p(y_2 + v) : y_2 \in Y\}.$$

Now if we choose a real number  $\gamma$  so that  $a \leq \gamma \leq b$ , then the Inequality 4.2 holds. The proof is complete.  $\square$

**Remark 4.2.** Before completing the proof of the Hahn-Banach Theorem, Let us first recall one of super important results in mathematics, called *Zorn's Lemma*, a very humble name. Every mathematics student should know it.

**Zorn's Lemma:** Let  $\mathcal{X}$  be a non-empty set with a partially order " $\leq$ ". Assume that every totally order subset  $\mathcal{C}$  of  $\mathcal{X}$  has an upper bound, i.e. there is an element  $\mathfrak{z} \in \mathcal{X}$  such that  $c \leq \mathfrak{z}$  for all  $c \in \mathcal{C}$ . Then  $\mathcal{X}$  must contain a maximal element  $\mathfrak{m}$ , that is, if  $\mathfrak{m} \leq x$  for some  $x \in \mathcal{X}$ , then  $\mathfrak{m} = x$ .

The following is the typical argument of applying the Zorn's Lemma.

**Theorem 4.3. Hahn-Banach Theorem :** *Let  $X$  be a vector space ( not necessary to be a normed space) over  $\mathbb{R}$  and let  $Y$  be a subspace of  $X$ . Let  $p$  be a positive homogeneous sub-additive function defined on  $X$ . Suppose that  $f$  is a real linear functional defined on  $Y$  satisfying  $f(y) \leq p(y)$  for all  $y \in Y$ . Then there is a real linear extension  $F$  of  $f$  defined on  $X$  so that*

$$F(x) \leq p(x) \quad \text{for all } x \in X.$$

*Proof.* Let  $\mathcal{X}$  be the collection of the pairs  $(Y_1, f_1)$ , where  $Y \subseteq Y_1$  is a subspace of  $X$  and  $f_1$  is a linear extension of  $f$  defined on  $Y_1$  such that  $f_1 \leq p$  on  $Y_1$ . Define a partial order  $\leq$  on  $\mathcal{X}$  by  $(Y_1, f_1) \leq (Y_2, f_2)$  if  $Y_1 \subseteq Y_2$  and  $f_2|_{Y_1} = f_1$ . Then by the Zorn's lemma, there is a maximal element  $(\tilde{Y}, \tilde{F})$  in  $\mathcal{X}$ . The maximality of  $(\tilde{Y}, \tilde{F})$  and Lemma 4.1 give  $\tilde{Y} = X$ . The proof is complete.  $\square$

**Definition 4.4.** *Let  $D$  be a convex subset of a normed space  $X$ , i.e.,  $tx + (1 - t)y \in D$  for all  $x, y \in D$  and  $t \in (0, 1)$ . Suppose that  $0$  is an interior point of  $D$ . Define*

$$\mu_D(x) := \inf\{t > 0 : x \in tD\}$$

*for  $x \in X$ . In addition, set  $\mu_D(x) = \infty$  if  $\{t > 0 : x \in tD\} = \emptyset$ .*

*The function  $\mu_D$  is called the Minkowski functional with respect to  $D$ .*

**Lemma 4.5.** *Let  $D$  be a convex subset of a normed space  $X$ . Suppose that  $0$  is an interior point of  $D$ . Then the Minkowski functional  $\mu := \mu_D : X \rightarrow [0, \infty)$  is positively homogeneous and sub-additive on  $D$ .*

*In addition, we have  $\{x \in X : \mu(x) < 1\} \subseteq D \subseteq \{x \in X : \mu(x) \leq 1\}$ .*

*Proof.* It is noted that since  $0 \in \text{int}(D)$ , the set  $\{t > 0 : x \in tD\} \neq \emptyset$  for all  $x \in X$ . Thus, the function  $\mu : X \rightarrow [0, \infty)$  is defined.

Clearly, if we fix  $t > 0$  and  $x \in X$ , then we have  $\mu(tx) \leq s$  if and only if  $t\mu(x) \leq s$ . Hence, the function  $\mu$  is positively homogeneous.

Next, we show the subadditivity of  $\mu$ . Let  $\varepsilon > 0$ . For  $x, y \in X$ , we choose  $s, t > 0$  such that  $x \in sD$  and  $y \in tD$  satisfying  $s < \mu(x) + \varepsilon$  and  $t < \mu(y) + \varepsilon$ . Then  $x = sd_1$  and  $y = td_2$  for some  $d_1, d_2 \in D$ . Since  $D$  is convex, we have

$$x + y = sd_1 + td_2 = (s + t)\left(\frac{s}{s + t}d_1 + \frac{t}{s + t}d_2\right) \in (s + t)D.$$

Thus,  $\mu(x + y) \leq s + t$  and so,  $\mu(x + y) < \mu(x) + \mu(y) + 2\varepsilon$ . Therefore,  $\mu$  is sub-additive. The last assertion is clear by the definition of  $\mu$ .  $\square$

**Proposition 4.6.** *Let  $C$  be a closed convex subset of a real vector space  $X$ . Let  $d$  and  $A$  be the positive constants. If  $x_0$  is an element in  $X$  with  $\|x_0\| \leq A$  such that  $0 < d < \text{dist}(x_0, C)$ , then there is an element  $F_1 \in B_{X^*}$  such that*

$$(4.3) \quad F_1(y) + \alpha < F_1(x_0).$$

for all  $y \in C$ , whenever  $0 < \alpha < \frac{d}{2}(1 - \frac{d}{2A})^{-1}$ .

Consequently, for any closed convex subset  $C_1$  of  $X$  and  $x'_0 \notin C_1$ , then there is an element  $g \in X^*$  with  $\|g\| \leq 1$  such that

$$(4.4) \quad g(z) + \beta < g(x'_0).$$

for all  $z \in C_1$ , whenever  $0 < \beta < \text{dist}(x'_0, C_1)$ .

*Proof.* For showing the first assertion, we first note that if there is  $x_0 \in X$  such that  $\|x_0\| \leq A$  and  $d(x_0, C) > d$ , then  $\frac{d}{2A} < 1$ . To see this, we have  $A \geq \|x_0\| \geq d(x_0, C) > \frac{d}{2}$  because  $0 \in C$ .

Now since  $0 < d < \text{dist}(x_0, C)$ , we have  $(x_0 + B(0, d)) \cap C = \emptyset$ . Thus, we have  $(x_0 + B(0, \frac{1}{2}d)) \cap (C + B(0, \frac{1}{2}d)) = \emptyset$ . Put  $D := C + B(0, \frac{1}{2}d)$ . Notice that  $D$  is a convex subset of  $X$  and  $x_0 \notin D$ . Moreover, we have  $0 \in \text{int}(D)$ . Let  $\mu := \mu_D$  be the Minkowski functional corresponding to  $D$ . Then  $\mu$  is positive homogeneous and sub-additive on  $X$  by Lemma 4.5.

Put  $Y := \mathbb{R}x_0$  and define  $f : Y \rightarrow \mathbb{R}$  by  $f(\alpha x_0) := \alpha\mu(x_0)$  for  $\alpha \in \mathbb{R}$ . Then  $f(y) \leq \mu(y)$  for all  $y \in Y$  since  $\mu \geq 0$  and positive homogenous. The Hahn-Banach Theorem 4.3 implies that there is a linear extension  $F$  defined on  $X$  satisfying  $F(x) \leq \mu(x)$  for all  $x \in X$ . We want to show that the linear functional  $F_1 := \frac{d}{2}F \in B_{X^*}$  is as required.

We first notice that  $F$  is bounded because we have  $|F(y)| \leq \mu(y) \leq 1$  for all  $y \in B(0, \frac{1}{2}d) \subseteq D$  and so,  $\|F_1\| = \|\frac{d}{2}F\| \leq 1$ . Note that  $\mu(x) \leq 1$  for all  $x \in C$  because  $C \subseteq D$ . Thus,  $\sup F(C) \leq 1$ . On the other hand, since  $x_0 \notin D$ , we have  $F(x_0) = \mu(x_0) \geq 1$ . Now if  $\mu(x_0) = 1$ , then there is a decreasing sequence of positive numbers  $(\lambda_n)$  with  $\lambda_n \downarrow 1$  and  $\frac{1}{\lambda_n}x_0 \in D$ . This implies that  $x_0 \in \overline{D}$ . It contradicts to the fact that  $(x_0 + B(0, \frac{1}{2}d)) \cap D$  is empty. Hence, we have  $F(y) \leq 1 < F(x_0) = \mu(x_0)$  for all  $y \in D$ .

Next, we are going to show that the Inequality 4.3 holds. In fact, for  $\lambda > 0$ , we see that  $x_0 \in \lambda D$  if and only if  $\frac{1}{\lambda}x_0 \in D$ . Hence, In this case, we have  $1 < \mu(x_0) \leq \lambda$ . Also, we have  $\frac{1}{\lambda}x_0 = y + z$  for some  $y \in C$  and for some  $z \in B(0, \frac{1}{2}d)$ . Then we have

$$d \leq \|x_0 - y\| = \|x_0 - \frac{1}{\lambda}x_0 - z\| = |1 - \frac{1}{\lambda}|\|x_0\| + \|z\| < |1 - \frac{1}{\lambda}|A + \frac{1}{2}d.$$

This implies that  $1 - \frac{1}{\lambda} > \frac{d}{2A}$  because  $\lambda \geq \mu(x_0) > 1$ . This gives  $1 < (1 - \frac{d}{2A})^{-1} < \lambda$  whenever  $\lambda > 0$  with  $x_0 \in \lambda D$  and hence,  $\mu(x_0) \geq (1 - \frac{d}{2A})^{-1}$ . Now if we put  $0 < \alpha_1 := (1 - \frac{d}{2A})^{-1} - 1$ , then we have

$$F(y) + \alpha_1 \leq 1 + \alpha_1 < (1 - \frac{d}{2A})^{-1} \leq \mu(x_0) = F(x_0)$$

for all  $y \in C$ . Therefore, if  $0 < \alpha < \frac{d}{2}\alpha_1$ , then the element  $F_1 := \frac{d}{2}F \in B_{X^*}$  satisfies the inequality 4.3 as desired.

For showing the last assertion, let  $d_0 := \text{dist}(x'_0, C_1) > 0$ . Clearly, we have  $\lim_{t \rightarrow 1^-} \frac{td_0}{2}(1 - \frac{td_0}{2d_0})^{-1} = d_0$ .

Thus if  $0 < \beta < d_0$ , there is  $0 < t_1 < 1$  such that  $\beta < \frac{t_1 d_0}{2}(1 - \frac{t_1 d_0}{2d_0})^{-1}$ . Now if we put  $d := t_1 d_0$ , then we have  $0 < \beta < \frac{d}{2}(1 - \frac{d}{2d_0})^{-1}$ . From this, we choose  $\varepsilon > 0$  such that  $\beta < \frac{d}{2}(1 - \frac{d}{2(d_0 + \varepsilon)})^{-1}$ . Now we fix a point  $x_1 \in C_1$  so that  $\|x'_0 - x_1\| < d_0 + \varepsilon$ . Put  $x_0 := x'_0 - x_1$ ;  $C := C_1 - x_1$  and  $A := d_0 + \varepsilon$  into the first assertion. Then there is an element  $g \in B_{X^*}$  such that the inequality 4.4 holds immediately.

We finish the proof.  $\square$

The following result is also referred to the **Hahn-Banach Theorem**.

**Theorem 4.7.** *Let  $X$  be a normed space and let  $Y$  be a subspace of  $X$ . If  $f \in Y^*$ , then there exists a linear extension  $F \in X^*$  of  $f$  such that  $\|F\| = \|f\|$ .*

*Proof.* W.L.O.G, we may assume that  $\|f\| = 1$ . We first show the case when  $X$  is normed space over  $\mathbb{R}$ . It is noted that the norm function  $p(\cdot) := \|\cdot\|$  is positively homogeneous and sub-additive on  $X$ . Since  $\|f\| = 1$ , we have  $f(y) \leq p(y)$  for all  $y \in Y$ . Then by the Hahn-Banach Theorem 4.3, there is a linear extension  $F$  of  $f$  on  $X$  such that  $F(x) \leq p(x)$  for all  $x \in X$ . This implies that  $\|F\| = 1$  as required.

Now for the complex case, let  $h = \operatorname{Re} f$  and  $g = \operatorname{Im} f$ . Then  $f = h + ig$  and  $f, g$  both are real linear on  $Y$  with  $\|h\| \leq 1$ . Note that since  $f(iy) = if(y)$  for all  $y \in Y$ , we have  $g(y) = -h(iy)$  for all  $y \in Y$ . This gives  $f(\cdot) = h(\cdot) - ih(i\cdot)$  on  $Y$ . Then by the real case above, there is a real linear extension  $H$  on  $X$  such that  $\|H\| = \|h\|$ . Now define  $F : X \rightarrow \mathbb{C}$  by  $F(\cdot) := H(\cdot) - iH(i\cdot)$ . Then  $F \in X^*$  and  $F|_Y = f$ . Thus it remains to show that  $\|F\| = \|f\| = 1$ . We need to show that  $|F(z)| \leq \|z\|$  for all  $z \in X$ . For  $z \in X$ , consider the polar form  $F(z) = re^{i\theta}$ . Then  $F(e^{-i\theta}z) = r \in \mathbb{R}$  and thus  $F(e^{-i\theta}z) = H(e^{-i\theta}z)$ . This yields that

$$|F(z)| = r = |F(e^{-i\theta}z)| = |H(e^{-i\theta}z)| \leq \|H\| \|e^{-i\theta}z\| \leq \|z\|.$$

The proof is complete.  $\square$

**Proposition 4.8.** *Let  $X$  be a normed space and  $x_0 \in X$ . Then there is  $f \in X^*$  with  $\|f\| = 1$  such that  $f(x_0) = \|x_0\|$ . Consequently, we have*

$$\|x_0\| = \sup\{|g(x)| : g \in B_{X^*}\}.$$

*In addition, if  $x, y \in X$  with  $x \neq y$ , then there exists  $f \in X^*$  such that  $f(x) \neq f(y)$ .*

*Proof.* Let  $Y = \mathbb{K}x_0$ . Define  $f_0 : Y \rightarrow \mathbb{K}$  by  $f_0(\alpha x_0) := \alpha\|x_0\|$  for  $\alpha \in \mathbb{K}$ . Then  $f_0 \in Y^*$  with  $\|f_0\| = \|x_0\|$ . The result follows immediately from the Hahn-Banach Theorem.  $\square$

**Remark 4.9.** Proposition 4.8 tells us that the dual space  $X^*$  of  $X$  must be non-zero. Indeed, the dual space  $X^*$  is very “Large” so that it can separate any pair of distinct points in  $X$ .

Furthermore, for any normed space  $Y$  and any pair of points  $x_1, x_2 \in X$  with  $x_1 \neq x_2$ , we can find an element  $T \in B(X, Y)$  such that  $Tx_1 \neq Tx_2$ . In fact, fix a non-zero element  $y \in Y$ . Then by Proposition 4.8, there is  $f \in X^*$  such that  $f(x_1) \neq f(x_2)$ . Thus, if we define  $Tx = f(x)y$ , then  $T \in B(X, Y)$ .

**Proposition 4.10.** *Using the notations as above, if  $M$  is closed subspace and  $v \in X \setminus M$ , then there is  $f \in X^*$  such that  $f(M) \equiv 0$  and  $f(v) \neq 0$ .*

*Proof.* Since  $M$  is a closed subspace of  $X$ , we can consider the quotient space  $X/M$ . Let  $\pi : X \rightarrow X/M$  be the natural projection. Note that  $\bar{v} := \pi(v) \neq 0 \in X/M$  because  $v \in X \setminus M$ . Then by Corollary 4.8, there is a non-zero element  $\bar{f} \in (X/M)^*$  such that  $\bar{f}(\bar{v}) \neq 0$ . Therefore, the linear functional  $f := \bar{f} \circ \pi \in X^*$  is as desired.  $\square$

**Proposition 4.11.** *Using the notations as above, if  $X^*$  is separable, then  $X$  is separable.*

*Proof.* Let  $F := \{f_1, f_2, \dots\}$  be a dense subset of  $X^*$ . Then there is a sequence  $(x_n)$  in  $X$  with  $\|x_n\| = 1$  and  $|f_n(x_n)| \geq 1/2\|f_n\|$  for all  $n$ . Now let  $M$  be the closed linear span of  $x_n$ 's. Then  $M$  is a separable closed subspace of  $X$ . We are going to show that  $M = X$ . Suppose that  $M \neq X$  and hence Proposition 4.10 gives us a non-zero element  $f \in X^*$  such that  $f(M) \equiv 0$ . Since  $\{f_1, f_2, \dots\}$  is dense in  $X^*$ , we have  $B(f, r) \cap F \neq \emptyset$  for all  $r > 0$ . Therefore, if  $B(f, r) \cap F \neq \emptyset$  is finite for some  $r > 0$ , then  $f = f_m$  for some  $f_m \in F$ . This implies that  $\|f\| = \|f_m\| \leq 2|f_m(x_m)| = 2|f(x_m)| = 0$  and thus,  $f = 0$  which contradicts to  $f \neq 0$ .

Therefore,  $B(f, r) \cap F$  is infinite for all  $r > 0$ . In this case, there is a subsequence  $(f_{n_k})$  such that  $\|f_{n_k} - f\| \rightarrow 0$ . This gives

$$\frac{1}{2}\|f_{n_k}\| \leq |f_{n_k}(x_{n_k})| = |f_{n_k}(x_{n_k}) - f(x_{n_k})| \leq \|f_{n_k} - f\| \rightarrow 0$$

because  $f(M) \equiv 0$ . Thus  $\|f_{n_k}\| \rightarrow 0$  and hence  $f = 0$ . It leads to a contradiction again. Thus, we can conclude that  $M = X$  as desired.  $\square$

**Remark 4.12.** The converse of Proposition 4.11 does not hold. For example, consider  $X = \ell^1$ . Then  $\ell^1$  is separable but the dual space  $(\ell^1)^* = \ell^\infty$  is not.

**Proposition 4.13.** Let  $X$  and  $Y$  be normed spaces. For each element  $T \in B(X, Y)$ , define a linear operator  $T^* : Y^* \rightarrow X^*$  by

$$T^*y^*(x) := y^*(Tx)$$

for  $y^* \in Y^*$  and  $x \in X$ . Then  $T^* \in B(Y^*, X^*)$  and  $\|T^*\| = \|T\|$ . In this case,  $T^*$  is called the adjoint operator of  $T$ .

*Proof.* We first claim that  $\|T^*\| \leq \|T\|$  and hence,  $\|T^*\|$  is bounded.

In fact, for any  $y^* \in Y^*$  and  $x \in X$ , we have  $|T^*y^*(x)| = |y^*(Tx)| \leq \|y^*\|\|Tx\|$ . Hence,  $\|T^*y^*\| \leq \|T\|\|y^*\|$  for all  $y^* \in Y^*$ . Thus,  $\|T^*\| \leq \|T\|$ .

We need to show  $\|T\| \leq \|T^*\|$ . Let  $x \in B_X$ . Then by Proposition 4.8, there is  $y^* \in S_{X^*}$  such that  $\|Tx\| = |y^*(Tx)| = |T^*y^*(x)| \leq \|T^*y^*\| \leq \|T^*\|$ . This implies that  $\|T\| \leq \|T^*\|$ .  $\square$

**Example 4.14.** Let  $X$  and  $Y$  be the finite dimensional normed spaces. Let  $(e_i)_{i=1}^n$  and  $(f_j)_{j=1}^m$  be the bases for  $X$  and  $Y$  respectively. Let  $\theta_X : X \rightarrow X^*$  and  $\theta_Y : Y \rightarrow Y^*$  be the identifications as in Example 3.15. Let  $e_i^* := \theta_X e_i \in X^*$  and  $f_j^* := \theta_Y f_j \in Y^*$ . Then  $e_i^*(e_l) = \delta_{il}$  and  $f_j^*(f_l) = \delta_{jl}$ , where,  $\delta_{il} = 1$  if  $i = l$ ; otherwise is 0.

Now if  $T \in B(X, Y)$  and  $(a_{ij})_{m \times n}$  is the representative matrix of  $T$  corresponding to the bases  $(e_i)_{i=1}^n$  and  $(f_j)_{j=1}^m$  respectively, then  $a_{kl} = f_k^*(Te_l) = T^*f_k^*(e_l)$ . Therefore, if  $(a'_{lk})_{n \times m}$  is the representative matrix of  $T^*$  corresponding to the bases  $(f_j^*)$  and  $(e_i^*)$ , then  $a_{kl} = a'_{lk}$ . Hence the transpose  $(a_{kl})^t$  is the representative matrix of  $T^*$ .

**Proposition 4.15.** Let  $Y$  be a closed subspace of a normed space  $X$ . Let  $i : Y \rightarrow X$  be the natural inclusion and  $\pi : X \rightarrow X/Y$  the natural projection. Then

- (i) the adjoint operator  $i^{**} : Y^{**} \rightarrow X^{**}$  is an isometry.
- (ii) the adjoint operator  $\pi^* : (X/Y)^* \rightarrow X^*$  is an isometry.

Consequently,  $Y^{**}$  and  $(X/Y)^*$  can be viewed as the closed subspaces of  $X^{**}$  and  $X^*$  respectively.

*Proof.* For Part (i), we first note that for any  $x^* \in X^*$ , the image  $i^*x^*$  in  $Y^*$  is just the restriction of  $x^*$  on  $Y$ , denoted by  $x^*|_Y$ . Now let  $\phi \in Y^{**}$ . Then for any  $x^* \in X^*$ , we have

$$|i^{**}\phi(x^*)| = |\phi(i^*x^*)| = |\phi(x^*|_Y)| \leq \|\phi\|\|x^*|_Y\|_{Y^*} \leq \|\phi\|\|x^*\|_{X^*}.$$

Thus,  $\|i^{**}\phi\| \leq \|\phi\|$ . WE need to show the inverse inequality. Now for each  $y^* \in Y^*$ , the Hahn-Banach Theorem gives an element  $x^* \in X^*$  such that  $\|x^*\|_{X^*} = \|y^*\|_{Y^*}$  and  $x^*|_Y = y^*$  and hence,  $i^*x^* = y^*$ . Then we have

$$|\phi(y^*)| = |\phi(x^*|_Y)| = |\phi(i^*x^*)| = |(i^{**} \circ \phi)(x^*)| \leq \|i^{**}\phi\|\|x^*\|_{X^*} = \|i^{**}\phi\|\|y^*\|_{Y^*}$$

for all  $y^* \in Y^*$ . Therefore, we have  $\|i^{**}\phi\| = \|\phi\|$ .

For Part (ii), let  $\psi \in (X/Y)^*$ . Note that since  $\|\pi^*\| = \|\pi\| \leq 1$ , we have  $\|\pi^*\psi\| \leq \|\psi\|$ . On the other hand, for each  $\bar{x} := \pi(x) \in X/Y$  with  $\|\bar{x}\| < 1$ , we can choose an element  $m \in Y$  such that  $\|x + m\| < 1$ . Therefore, we have

$$|\psi(\bar{x})| = |\psi \circ \pi(x)| = |\psi \circ \pi(x + m)| \leq \|\psi \circ \pi\| = \|\pi^*(\psi)\|.$$

Therefore, we have  $\|\psi\| \leq \|\pi^*(\psi)\|$ . The proof is complete.  $\square$

**Remark 4.16.** By using Proposition 4.15, we can give an alternative proof of the Riesz's Lemma 2.5.

Using the notations as in Proposition 4.15, if  $Y \subsetneq X$ , then we have  $\|\pi\| = \|\pi^*\| = 1$  because  $\pi^*$  is an isometry by Proposition 4.15(ii). Thus we have  $\|\pi\| = \sup\{\|\pi(x)\| : x \in X, \|x\| = 1\} = 1$ . Hence, for any  $0 < \theta < 1$ , we can find element  $z \in X$  with  $\|z\| = 1$  such that  $\theta < \|\pi(z)\| = \inf\{\|z + y\| : y \in Y\}$ . The Riesz's Lemma follows.

## 5. REFLEXIVE SPACES

**Proposition 5.1.** *For a normed space  $X$ , let  $Q : X \rightarrow X^{**}$  be the canonical map, that is,  $Qx(x^*) := x^*(x)$  for  $x^* \in X^*$  and  $x \in X$ . Then  $Q$  is an isometry.*

*Proof.* Note that for  $x \in X$  and  $x^* \in B_{X^*}$ , we have  $|Q(x)(x^*)| = |x^*(x)| \leq \|x\|$ . Then  $\|Q(x)\| \leq \|x\|$ .

We need to show that  $\|x\| \leq \|Q(x)\|$  for all  $x \in X$ . In fact, for  $x \in X$ , there is  $x^* \in X^*$  with  $\|x^*\| = 1$  such that  $\|x\| = |x^*(x)| = |Q(x)(x^*)|$  by Proposition 4.8. Thus we have  $\|x\| \leq \|Q(x)\|$ . The proof is complete.  $\square$

**Remark 5.2.** Let  $T : X \rightarrow Y$  be a bounded linear operator and  $T^{**} : X^{**} \rightarrow Y^{**}$  the second dual operator induced by the adjoint operator of  $T$ . Using notations as in Proposition 5.1 above, the following diagram commutes.

$$\begin{array}{ccc} X & \xrightarrow{T} & Y \\ Q_X \downarrow & & \downarrow Q_Y \\ X^{**} & \xrightarrow{T^{**}} & Y^{**} \end{array}$$

**Definition 5.3.** *A normed space  $X$  is said to be reflexive if the canonical map  $Q : X \rightarrow X^{**}$  is surjective. (Note that every reflexive space must be a Banach space.)*

**Example 5.4.** *We have the following examples.*

- (i) : Every finite dimensional normed space  $X$  is reflexive.
- (ii) :  $\ell^p$  is reflexive for  $1 < p < \infty$ .
- (iii) :  $c_0$  and  $\ell^1$  are not reflexive.

*Proof.* For Part (i), if  $\dim X < \infty$ , then  $\dim X = \dim X^{**}$ . Hence, the canonical map  $Q : X \rightarrow X^{**}$  must be surjective.

Part (ii) follows from  $(\ell^p)^* = \ell^q$  for  $1 < p < \infty$ ,  $\frac{1}{p} + \frac{1}{q} = 1$ .

For Part (iii), note that  $c_0^{**} = (\ell^1)^* = \ell^\infty$ . Since  $\ell^\infty$  is non-separable but  $c_0$  is separable. Therefore, the canonical map  $Q$  from  $c_0$  to  $c_0^{**} = \ell^\infty$  must not be surjective.

For the case of  $\ell^1$ , we have  $(\ell^1)^{**} = (\ell^\infty)^*$ . Since  $\ell^\infty$  is non-separable, the dual space  $(\ell^\infty)^*$  is non-separable by Proposition 4.11. Therefore,  $\ell^1 \neq (\ell^1)^{**}$ .  $\square$

**Proposition 5.5.** *Every closed subspace of a reflexive space is reflexive.*

*Proof.* Let  $Y$  be a closed subspace of a reflexive space  $X$ . Let  $Q_Y : Y \rightarrow Y^{**}$  and  $Q_X : X \rightarrow X^{**}$  be the canonical maps as before. Let  $y_0^{**} \in Y^{**}$ . We define an element  $\phi \in X^{**}$  by  $\phi(x^*) := y_0^{**}(x^*|_Y)$  for  $x^* \in X^*$ . Since  $X$  is reflexive, there is  $x_0 \in X$  such that  $Q_X x_0 = \phi$ . Suppose  $x_0 \notin Y$ . Then by Proposition 4.10, there is  $x_0^* \in X^*$  such that  $x_0^*(x_0) \neq 0$  but  $x_0^*(Y) \equiv 0$ . Note that we have  $x_0^*(x_0) = Q_X x_0(x_0^*) = \phi(x_0^*) = y_0^{**}(x_0^*|_Y) = 0$ . It leads to a contradiction, and so  $x_0 \in Y$ . The



proof is complete if we have  $Q_Y(x_0) = y_0^{**}$ .

In fact, for each  $y^* \in Y^*$ , then by the Hahn-Banach Theorem,  $y^*$  has a continuous extension  $x^*$  in  $X^*$ . Then we have

$$Q_Y(x_0)(y^*) = y^*(x_0) = x^*(x_0) = Q_X(x_0)(x^*) = \phi(x^*) = y_0^{**}(x^*|_Y) = y_0^{**}(y^*).$$

□

**Example 5.6.** By using Proposition 5.5, we immediately see that the space  $\ell^\infty$  is not reflexive because it contains a non-reflexive closed subspace  $c_0$ .

**Proposition 5.7.** Let  $X$  be a Banach space. Then we have the following assertions.

- (i)  $X$  is reflexive if and only if the dual space  $X^*$  is reflexive.
- (ii) If  $X$  is reflexive, then so is every quotient of  $X$ .

*Proof.* For Part (i), suppose that  $X$  is reflexive first. Let  $\tilde{z} \in X^{***}$ . Then the restriction  $z := \tilde{z}|_X \in X^*$ . Then one can directly check that  $Qz = z$  on  $X^{**}$  since  $X^{**} = X$ .

For the converse, assume that  $X^*$  is reflexive but  $X$  is not. Therefore,  $X$  is a proper closed subspace of  $X^{**}$ . Then by using the Hahn-Banach Theorem, we can find a non-zero element  $\phi \in X^{***}$  such that  $\phi(X) \equiv 0$ . However, since  $X^{***}$  is reflexive, we have  $\phi \in X^*$  and hence,  $\phi = 0$  which leads to a contradiction.

For Part (ii), we assume that  $X$  is reflexive. Let  $M$  be a closed subspace of  $X$  and  $\pi : X \rightarrow X/M$  the natural projection. Note that the adjoint operator  $\pi^* : (X/M)^* \rightarrow X^*$  is an isometry (**Check !**). Thus,  $(X/M)^*$  can be viewed as a closed subspace of  $X^*$ . By Part (i) and Proposition 5.5, we see that  $(X/M)^*$  is reflexive. Then  $X/M$  is reflexive by using Part (i) again.

The proof is complete. □

**Lemma 5.8.** Let  $M$  be a closed subspace of a normed space  $X$ . Let  $r : X^* \rightarrow M^*$  be the restriction map, that is  $x^* \in X^* \mapsto x^*|_M \in M^*$ . Put  $M^\perp := \ker r := \{x^* \in X^* : x^*(M) \equiv 0\}$ . Then the canonical linear isomorphism  $\tilde{r} : X^*/M^\perp \rightarrow M^*$  induced by  $r$  is an isometric isomorphism.

*Proof.* We first note that  $r$  is surjective by using the Hahn-Banach Theorem. We need to show that  $\tilde{r}$  is an isometry. Note that  $\tilde{r}(x^* + M^\perp) = x^*|_M$  for all  $x^* \in X^*$ . Now for any  $x^* \in X^*$ , we have  $\|x^* + y^*\|_{X^*} \geq \|x^* + y^*\|_{M^*} = \|x^*|_M\|_{M^*}$  for all  $y^* \in M^\perp$ . Thus, we have  $\|\tilde{r}(x^* + M^\perp)\| = \|x^*|_M\|_{M^*} \leq \|x^* + M^\perp\|$ . We need to show the reverse inequality.

Now for any  $x^* \in X^*$ , then by the Hahn-Banach Theorem again, there is  $z^* \in X^*$  such that  $z^*|_M = x^*|_M$  and  $\|z^*\| = \|x^*|_M\|_{M^*}$ . Then  $x^* - z^* \in M^\perp$  and hence, we have  $x^* + M^\perp = z^* + M^\perp$ . This implies that

$$\|x^* + M^\perp\| = \|z^* + M^\perp\| \leq \|z^*\| = \|x^*|_M\|_{M^*} = \|\tilde{r}(x^* + M^\perp)\|.$$

The proof is complete. □

**Proposition 5.9. (Three-space property):** Let  $M$  be a closed subspace of a normed space  $X$ . If  $M$  and the quotient space  $X/M$  both are reflexive, then so is  $X$ .

*Proof.* Let  $\pi : X \rightarrow X/M$  be the natural projection. Let  $\psi \in X^{**}$ . We going to show that  $\psi \in \text{im}(Q_X)$ . Since  $\pi^{**}(\psi) \in (X/M)^{**}$ , there exists  $x_0 \in X$  such that  $\pi^{**}(\psi) = Q_{X/M}(x_0 + M)$  because  $X/M$  is reflexive. Thus we have

$$\pi^{**}(\psi)(\bar{x}^*) = Q_{X/M}(x_0 + M)(\bar{x}^*)$$

for all  $\bar{x}^* \in (X/M)^*$ . This implies that

$$\psi(\bar{x}^* \circ \pi) = \psi(\pi^* \bar{x}^*) = \pi^{**}(\psi)(\bar{x}^*) = Q_{X/M}(x_0 + M)(\bar{x}^*) = \bar{x}^*(x_0 + M) = Q_X x_0(\bar{x}^* \circ \pi)$$

for all  $\bar{x}^* \in (X/M)^*$ . Therefore, we have

$$\psi = Q_X x_0 \quad \text{on} \quad M^\perp.$$

Therefore, we have  $\psi - Q_X(x_0) \in (X^*/M^\perp)^*$ . Let  $f : M^* \rightarrow X^*/M^\perp$  be the inverse of the isometric isomorphism  $\tilde{r}$  which is defined as in Lemma 5.8. Then the composite  $(\psi - Q_X x_0) \circ f : M^* \rightarrow X^*/M^\perp \rightarrow \mathbb{K}$  lies in  $M^{**}$ . Then by the reflexivity of  $M$ , there is an element  $m_0 \in M$  such that

$$(\psi - Q_X x_0) \circ f = Q_M(m_0) \in M^{**}.$$

Notice that for each  $x^* \in X^*$ , we can find an element  $m^* \in M^*$  such that  $f(m^*) = x^* + M^\perp \in X^*/M^\perp$  because  $f$  is surjective. Moreover, by the construction of  $\tilde{r}$  in Lemma 5.8, we see that  $x^*|_M = m^*$ . This gives

$$\psi(x^*) - x^*(x_0) = (\psi - Q_X x_0)(m^*) \circ f = Q_M(m_0)(m^*) = m^*(m_0) = x^*(m_0).$$

Thus, we have  $\psi(x^*) = x^*(x_0 + m_0)$  for all  $x^* \in X^*$ . From this we have  $\psi = Q_X(x_0 + m_0) \in \text{im}(Q_X)$  as desired. The proof is complete.  $\square$

**Remark 5.10.** In view of the definition of a reflexive space, it is naturally raised the question that whether a Banach space  $X$  is reflexive whenever it is isometrically isomorphic to its second dual. The answer is negative. A counter example was given by R.C. James in 1951 (see [8]).

## 6. WEAKLY CONVERGENT AND WEAK\* CONVERGENT

**Definition 6.1.** Let  $X$  be a normed space. A sequence  $(x_n)$  is said to be weakly convergent if there is  $x \in X$  such that  $f(x_n) \rightarrow f(x)$  for all  $f \in X^*$ . In this case,  $x$  is called a weak limit of  $(x_n)$ .

**Proposition 6.2.** A weak limit of a sequence is unique if it exists. In this case, if  $(x_n)$  weakly converges to  $x$ , denoted by  $x = w\text{-}\lim_n x_n$  or  $x_n \xrightarrow{w} x$ .

*Proof.* The uniqueness follows immediately from the Hahn-Banach Theorem.  $\square$

**Remark 6.3.** Clearly, if a sequence  $(x_n)$  converges to  $x \in X$  in norm, then  $x_n \xrightarrow{w} x$ . However, the weakly convergence of a sequence does not imply the norm convergence. For example, consider  $X = c_0$  and  $(e_n)$ . Then  $f(e_n) \rightarrow 0$  for all  $f \in c_0^* = \ell^1$  but  $(e_n)$  is not convergent in  $c_0$ .

**Proposition 6.4.** Suppose that  $X$  is finite dimensional. A sequence  $(x_n)$  in  $X$  is norm convergent if and only if it is weakly convergent.

*Proof.* Suppose that  $(x_n)$  weakly converges to  $x$ . Let  $\mathcal{B} := \{e_1, \dots, e_N\}$  be a basis for  $X$  and let  $f_k$  be the  $k$ -th coordinate functional corresponding to the basis  $\mathcal{B}$ , i.e.,  $v = \sum_{k=1}^N f_k(v)e_k$  for all  $v \in X$ . Since  $\dim X < \infty$ , we have  $f_k \in X^*$  for all  $k = 1, \dots, N$ . Therefore, we have  $\lim_n f_k(x_n) = f_k(x)$  for all  $k = 1, \dots, N$ . Thus, we have  $\|x_n - x\| \rightarrow 0$ .  $\square$

**Definition 6.5.** Let  $X$  be a normed space. A sequence  $(f_n)$  in  $X^*$  is said to be weak\* convergent if there is  $f \in X^*$  such that  $\lim_n f_n(x) = f(x)$  for all  $x \in X$ , that is  $f_n$  point-wise converges to  $f$ . In this case,  $f$  is called the weak\* limit of  $(f_n)$ . Write  $f = w^*\text{-}\lim_n f_n$  or  $f_n \xrightarrow{w^*} f$ .

**Remark 6.6.** In the dual space  $X^*$  of a normed space  $X$ , we always have the following implications:

$$\text{“Norm Convergent”} \implies \text{“Weakly Convergent”} \implies \text{“Weak* Convergent”}.$$

However, the converse of each implication does not hold.

**Example 6.7.** Remark 6.3 has shown that the  $w$ -convergence does not imply  $\|\cdot\|$ -convergence. We now claim that the  $w^*$ -convergence also Does Not imply the  $w$ -convergence. Consider  $X = c_0$ . Then  $c_0^* = \ell^1$  and  $c_0^{**} = (\ell^1)^* = \ell^\infty$ . Let  $e_n^* = (0, \dots, 0, 1, 0, \dots) \in \ell^1 = c_0^*$ , where the  $n$ -th coordinate is 1. Then  $e_n^* \xrightarrow{w^*} 0$  but  $e_n^* \not\xrightarrow{w} 0$  weakly because  $e^{**}(e_n^*) \equiv 1$  for all  $n$ , where  $e^{**} := (1, 1, \dots) \in \ell^\infty = c_0^{**}$ . Hence the  $w^*$ -convergence does not imply the  $w$ -convergence.

**Proposition 6.8.** Let  $(f_n)$  be a sequence in  $X^*$ . Suppose that  $X$  is reflexive. Then  $f_n \xrightarrow{w} f$  if and only if  $f_n \xrightarrow{w^*} f$ .

In particular, if  $\dim X < \infty$ , then the followings are equivalent:

- (i) :  $f_n \xrightarrow{\|\cdot\|} f$ ;
- (ii) :  $f_n \xrightarrow{w} f$ ;
- (iii) :  $f_n \xrightarrow{w^*} f$ .

**Theorem 6.9. (Banach) :** Let  $X$  be a separable normed space. If  $(f_n)$  is a bounded sequence in  $X^*$ , then it has a  $w^*$ -convergent subsequence.

*Proof.* Let  $D := \{x_1, x_2, \dots\}$  be a countable dense subset of  $X$ . Note that since  $(f_n)_{n=1}^\infty$  is bounded,  $(f_n(x_1))$  is a bounded sequence in  $\mathbb{K}$ . Then  $(f_n(x_1))$  has a convergent subsequence, say  $(f_{1,k}(x_1))_{k=1}^\infty$  in  $\mathbb{K}$ . Let  $c_1 := \lim_k f_{1,k}(x_1)$ . Now consider the bounded sequence  $(f_{1,k}(x_2))_{k=1}^\infty$ . Then there is convergent subsequence, say  $(f_{2,k}(x_2))_{k=1}^\infty$ , of  $(f_{1,k}(x_2))_{k=1}^\infty$ . Put  $c_2 := \lim_k f_{2,k}(x_2)$ . Note that we still have  $c_1 = \lim_k f_{2,k}(x_1)$ . To repeat the same step, if we define  $(m, k) \leq (m', k')$  if  $m < m'$ ; or  $m = m'$  with  $k \leq k'$ , we can find a sequence  $(f_{m,k})_{m,k}$  in  $X^*$  such that

- (i) :  $(f_{m+1,k})_{k=1}^\infty$  is a subsequence of  $(f_{m,k})_{k=1}^\infty$  for  $m = 0, 1, \dots$ , where  $f_{0,k} := f_k$ .
- (ii) :  $c_i = \lim_k f_{m,k}(x_i)$  exists for all  $1 \leq i \leq m$ .

Now put  $h_k := f_{k,k}$ . Then  $(h_k)$  is a subsequence of  $(f_n)$ . Note that for each  $i$ , we have  $\lim_k h_k(x_i) = \lim_k f_{i,k}(x_i) = c_i$  by the construction (ii) above. Since  $(\|h_k\|)$  is bounded and  $D$  is dense in  $X$ , we have  $h(x) := \lim_k h_k(x)$  exists for all  $x \in X$  and  $h \in X^*$ . That is  $h = w^*$ - $\lim_k h_k$ . The proof is complete.  $\square$

**Remark 6.10.** Theorem 6.9 does not hold if the separability of  $X$  is removed.

For example, consider  $X = \ell^\infty$  and  $\delta_n$  the  $n$ -th coordinate functional on  $\ell^\infty$ . Then  $\delta_n \in (\ell^\infty)^*$  with  $\|\delta_n\|_{(\ell^\infty)^*} = 1$  for all  $n$ . Suppose that  $(\delta_n)$  has a  $w^*$ -convergent subsequence  $(\delta_{n_k})_{k=1}^\infty$ . Define  $x \in \ell^\infty$  by

$$x(m) = \begin{cases} 0 & \text{if } m \neq n_k; \\ 1 & \text{if } m = n_{2k}; \\ -1 & \text{if } m = n_{2k+1}. \end{cases}$$

Hence we have  $|\delta_{n_i}(x) - \delta_{n_{i+1}}(x)| = 2$  for all  $i = 1, 2, \dots$ . It leads to a contradiction. Thus  $(\delta_n)$  has no  $w^*$ -convergent subsequence.

**Corollary 6.11.** Let  $X$  be a separable space. Assume that a sequence in  $X^*$  is  $w^*$ -convergent if and only if it is norm convergent. Then  $\dim X < \infty$ .

*Proof.* We need to show that the closed unit ball  $B_{X^*}$  in  $X^*$  is compact in norm. Let  $(f_n)$  be a sequence in  $B_{X^*}$ . By using Theorem 6.9,  $(f_n)$  has a  $w^*$ -convergent subsequence  $(f_{n_k})$ . Then by the assumption,  $(f_{n_k})$  is norm convergent. Note that if  $\lim_k f_{n_k} = f$  in norm, then  $f \in B_{X^*}$ . Thus  $B_{X^*}$  is compact and thus  $\dim X^* < \infty$ . Thus  $\dim X^{**} < \infty$  that gives  $\dim X$  is finite because  $X \subseteq X^{**}$ .  $\square$

**Corollary 6.12.** *Suppose that  $X$  is a separable. If  $X$  is reflexive space, then the closed unit ball  $B_X$  of  $X$  is sequentially weakly compact, i.e. it is equivalent to saying that any bounded sequence in  $X$  has a weakly convergent subsequence.*

*Proof.* Let  $Q : X \rightarrow X^{**}$  be the canonical map as before. Let  $(x_n)$  be a bounded sequence in  $X$ . Hence,  $(Qx_n)$  is a bounded sequence in  $X^{**}$ . We first note that since  $X$  is reflexive and separable,  $X^*$  is also separable by Proposition 4.11. We can apply Theorem 6.9,  $(Qx_n)$  has a  $w^*$ -convergent subsequence  $(Qx_{n_k})$  in  $X^{**} = Q(X)$  and hence,  $(x_{n_k})$  is weakly convergent in  $X$ .  $\square$

**Remark 6.13.** In fact, the converse of Corollary 6.12 also holds (see Appendix 7 below). The assumption of separability of  $X$  can be removed. We have the following stronger result which was shown by R. C. James (see [10, §1.13]).

**Theorem 6.14.** *Let  $X$  be a Banach space. Then the following are equivalent.*

- (i)  $X$  is reflexive.
- (ii) Every bounded sequence in  $X$  has a weakly convergent subsequence.
- (iii) The closed unit ball  $B_X$  of  $X$  is weakly compact, that is,  $B_X$  is compact in the weak topology.

## 7. APPENDIX: SEQUENTIALLY WEAKLY COMPACTNESS AND REFLEXIVITY FOR A SEPARABLE SPACE

This section is devoted to show the theorem that a separable Banach space is reflexive if and only if its closed unit ball is weakly sequentially compact. This results was obtained by Banach [2, Chapter VIII and Chapter XI]. For simply, throughout this section, all Banach spaces are assumed to be over  $\mathbb{R}$ .

### Some set theory

Before showing the main theorem, we need some basic knowledge of the set theory which can be found in the Halmos's classic book [7].

Recall that a partially order set  $S$  is called a *well ordered set* if for any non-empty subset  $A$  of  $S$  contains the least element, that is, there is an element  $x_0 \in A$  such that  $x_0 \leq x$  for all  $x \in A$ . In particular,  $S$  is automatically a totally order set. The *well-ordering theorem* tells us that every set can be equipped with a well-ordering.

Two well ordered sets  $A$  and  $B$  are said to have the same *ordinal number or ordinal* for simply if there is an order preserving bijection from  $A$  onto  $B$ . In particular, each ordinal can be viewed as a well ordered set. More precisely, an ordinal number  $\alpha$  is a well ordered set such that for any element  $\eta \in \alpha$ , we have  $\eta = \{\xi \in \alpha : \xi < \eta\}$ , thus, we have  $\eta \subseteq \alpha$  whenever  $\eta \in \alpha$ . This definition was due to von Neumann.

Put  $\omega$  the least infinite countable ordinal, that is,  $\omega := \{0, 1, 2, \dots\}$  and is endowed with the usual order.

On the other hand, it is naturally led to define an order on the class of ordinals as the following. (**Warning:** We DO NOT HAVE a statement about " the set of All ordinals numbers" !!!! )

**Definition 7.1.** *Let  $\alpha$  and  $\beta$  be two ordinals. We say that*

- (i)  $\alpha = \beta$  if there is an order preserving bijection from  $\alpha$  onto  $\beta$ .
- (ii)  $\alpha \leq \beta$  if there is an order preserving injection from  $\alpha$  to  $\beta$ .
- (iii)  $\alpha < \beta$  if  $\alpha \leq \beta$  but  $\alpha \neq \beta$ .

From the von Neumann's definition, we have (see [7, Section 20])

**Lemma 7.2.** *Every nonempty set of ordinal numbers is a well ordered set.*

Next we need the following definition for comparing the size of two given sets.

**Definition 7.3.** *Two sets  $A$  and  $B$  are said to have the same cardinality, write  $A \approx B$ , if there is a bijection from  $A$  onto  $B$ .*

*The cardinal number of  $A$ , write  $|A|$ , is defined by an ordinal number given by*

$$|A| := \min\{\alpha : \alpha \text{ is an ordinal number such that } \alpha \approx A\}$$

*(Notice that the well ordering theorem and Lemma 7.2 assure the existence of  $|A|$ .)*

**Definition 7.4.** *We will use the following terminologies later.*

- (i) *For an ordinal  $\theta$ , put  $\theta^+ := \theta \cup \{\theta\}$  and define  $\xi \leq \theta$  for all elements  $\xi \in \theta$ . Then  $\theta^+$  becomes a well ordered set and  $\theta < \theta^+$ . In this case,  $\theta^+$  is called the successor ordinal of  $\theta$ . Notice that there is no ordinal  $\gamma$  such that  $\theta < \gamma < \theta^+$ .*
- (ii) *An ordinal  $\beta$  is called a limit ordinal if there is no ordinal  $\gamma$  such that  $\gamma^+ = \beta$ . In this case, we are given any ordinal  $\eta$  with  $\eta < \beta$ , there is an ordinal  $\xi$  such that  $\eta < \xi < \beta$ , write  $\xi \rightarrow \beta$ . For example,  $\omega_1 :=$  the cardinal of  $\mathbb{R}$  and  $\omega := \mathbb{N}$  both are limit ordinals.*
- (iii) *Let  $X$  be a non-empty set. A function  $x : [0, \theta) \rightarrow X$  is called a transfinite sequence in  $X$  (or " $\theta$ -transfinite sequence") for some ordinal  $\theta$ , where  $[0, \theta)$  denotes the set of all ordinals  $\xi$  satisfying  $\xi < \theta$ . We also write  $(x_\xi)_{\xi < \theta}$  for a transfinite sequence. Note that  $\omega$ -transfinite sequences are the usual sequences defined as before. In this case, if  $\theta$  is a limit ordinal, one can naturally define the similar notation  $\lim_{\xi \rightarrow \theta} x_\xi$ ;  $\overline{\lim}_{\xi \rightarrow \theta} x_\xi$  and  $\underline{\lim}_{\xi \rightarrow \theta} x_\xi$  in any metric space as in the usual sequences case. Clearly, one can show that if  $(t_\xi)_{\xi < \theta}$  is a bounded  $\theta$ -sequence in  $\mathbb{R}$ , then*

$$\overline{\lim}_{\xi \rightarrow \theta} t_\xi := \limsup_{\eta \rightarrow \theta} t_\xi = \inf_{\eta < \theta} \sup_{\xi \geq \eta} t_\xi$$

*exists.*

- (iv) *Let  $\theta$  be a limit ordinal. A subset  $M$  of  $\theta$  is said to be cofinal if for every ordinal  $\mu < \theta$ , there is  $\nu \in M$  such that  $\mu < \nu < \theta$ . In this case, if  $\lim_{\xi \rightarrow \theta} x_\xi$  exists, then so does  $\lim_{\nu \in M: \nu \rightarrow \theta} x_\nu$  and they are the same.*

**Lemma 7.5.** *Every infinite cardinal is a limit ordinal.*

*Proof.* If  $\gamma$  is an infinite cardinal and  $\gamma = \theta^+$  for some ordinal  $\theta$ . Then by the definition of the successor ordinal, we have  $\theta \approx \theta^+ \approx \gamma$ . Hence,  $\gamma = \theta^+$  is not a cardinal because  $\theta < \theta^+$ .  $\square$

**Lemma 7.6.** *Let  $X$  be a Banach space. If  $(f_\xi)_{\xi < \theta}$  is a norm bounded  $\theta$ -sequence in  $X^*$ , then there is an element  $f \in X^*$  such that  $\|f\| \leq \sup_{\xi < \theta} \|f_\xi\|$  and*

$$(7.1) \quad f(x) \leq \overline{\lim}_{\xi \rightarrow \theta} f_\xi(x)$$

*for all  $x \in X$ . Consequently, we have*

$$(7.2) \quad \underline{\lim}_{\xi \rightarrow \theta} f_\xi(x) \leq f(x) \leq \overline{\lim}_{\xi \rightarrow \theta} f_\xi(x)$$

*for all  $x \in X$ .*

*In this case,  $f$  is called a transfinite limit of  $(f_\xi)$  (note that  $f$  may not be unique).*

*Proof.* Let  $M := \sup_{\xi < \theta} \|f_\xi\|$ . We first notice that since  $(f_\xi)_{\xi < \theta}$  is bounded,  $\overline{\lim}_{\xi \rightarrow \theta} f_\xi(x)$  exists for all  $x \in X$ . Hence, one can define a function  $p : X \rightarrow \mathbb{R}$  by

$$p(x) := \overline{\lim}_{\xi \rightarrow \theta} f_\xi(x)$$

for  $x \in X$ . Clearly,  $p$  is a positively homogenous and sub-additive function. We may assume that  $p(x_0) > 0$  for some  $x_0 \in X$ . To see this, if  $p(x_0) < 0$ , then  $p(-x_0) = \overline{\lim} f_\xi(-x_0) \geq \underline{\lim} f_\xi(-x_0) = -\overline{\lim} f_\xi(x_0) > 0$  as desired. Now if we define a linear map  $f_0$  on  $\mathbb{R}x_0$  by  $f_0(tx_0) := tp(x_0)$ , then  $f_0(tx_0) \leq p(tx_0)$  for all  $t \in \mathbb{R}$ . Then by the Hahn-Banach Theorem 4.3, there is a linear extension  $f$  of  $f_0$  defined on  $X$  such that  $f(x) \leq p(x)$  for all  $x \in X$ . Notice that since  $|f_\xi(x)| \leq M\|x\|$  for all  $\xi < \theta$  and for all  $x \in X$ , we have  $p(x) \leq M\|x\|$ . Thus, we have  $\|f\| \leq M$  as desired. The last assertion is obtained by putting  $-x \in X$  into Eq 7.1. The proof is complete.  $\square$

**Definition 7.7.** A normed subspace  $\Gamma$  of  $X^*$  is said to be *transfinitely closed* if for every norm bounded transfinite  $\theta$ -sequence  $(f_\xi)_{\xi < \theta}$  in  $\Gamma$  for some limit ordinal  $\theta$ , one can find an element  $f \in \Gamma$  satisfying the Eq 7.1 above, that is,

$$f(x) \leq \overline{\lim}_{\xi \rightarrow \theta} f_\xi(x)$$

for all  $x \in X$ .

Clearly, every transfinitely closed subspace is norm closed by considering  $\theta = \omega$  in Eq 7.2 above.

**Lemma 7.8.** Let  $\Gamma$  be a transfinitely closed subspace of  $X^*$  and  $f_0$  be an element in  $X^* \setminus \Gamma$ . Let  $0 < c < \text{dist}(f_0, \Gamma)$ . Then there is a non-empty finite subset  $G$  of  $X$  so that there is no element  $f \in \Gamma$  satisfying the following:

$$(7.3) \quad |f(x) - f_0(x)| \leq c\|x\| \quad \text{for all } x \in G.$$

*Proof.* Let

$$W := \{\gamma : \gamma \text{ is a cardinal such that } \gamma \leq |X|\}.$$

Now for each element  $\gamma \in W$ , put  $P(\gamma)$  the sentence given by:

whenever  $G$  is a subset of  $X$  with  $|G| = \gamma$ , there exists an element  $f_\gamma \in \Gamma$  such that  $|f_\gamma(x) - f_0(x)| \leq c\|x\|$  for all  $x \in G$ . Let

$$A := \{\gamma \in W : P(\gamma) \text{ holds}\}.$$

Clearly, the sentence  $P$  holds for the zero cardinal, that is,  $0 \in A$ . On the other hand, notice that the cardinal  $|X| \in W \setminus A$ . To see this, we can enumerate the elements in  $X$  such that  $(x_\xi)_{\xi < |X|}$  because  $[0, |X|) = |X|$ . Thus, if  $P(|X|)$  holds, there is an element  $f \in \Gamma$  so that  $|f(x) - f_0(x)| \leq c\|x\|$  for all  $x \in X$ . Hence,  $\|f - f_0\| \leq c$  which contradicts to the choice of  $c$  because  $c < \text{dist}(f_0, \Gamma)$ . Therefore,  $W \setminus A$  is a non-empty well ordered set by Lemma 7.2 and hence the set  $W \setminus A$  contains the least element, say  $\mathfrak{m}$ .

We will show that if  $\mathfrak{m}$  is infinite, then it will lead to a contradiction.

**Claim:**  $P(\mathfrak{m})$  holds if  $\mathfrak{m}$  is infinite.

Now let  $G$  be any subset of  $X$  with  $|G| = \mathfrak{m}$ , so, we can enumerate the elements in  $G$  as a  $\mathfrak{m}$ -sequence, say  $(x_\xi)_{\xi < \mathfrak{m}}$ . Now for any ordinal  $\eta < \mathfrak{m}$ , notice that if we put  $a := |(x_\xi)_{\xi < \eta}|$ , then  $a \leq |\eta| \leq \eta < \mathfrak{m}$ . The minimality of  $\mathfrak{m}$  implies that  $P(a)$  holds. Hence, there is an element  $f_\eta \in \Gamma$  so that

$$|f_\eta(x_\xi) - f_0(x_\xi)| \leq c\|x_\xi\| \quad \text{for all } \xi < \eta.$$

Now if  $\mathfrak{m}$  is infinite, then  $\mathfrak{m}$  is a limit ordinal by Lemma 7.5. Then by the definition of a transfinitely closed set, there is an element  $f \in \Gamma$  such that

$$\underline{\lim}_{\eta \rightarrow \mathfrak{m}} f_\eta(x) \leq f(x) \leq \overline{\lim}_{\eta \rightarrow \mathfrak{m}} f_\eta(x)$$

for all  $x \in X$ . This gives

$$|f(x_\xi) - f_0(x_\xi)| \leq c\|x_\xi\| \quad \text{for all } \xi < \mathfrak{m}.$$

Recall that  $G = (x_\xi)_{\xi < \mathfrak{m}}$ . Hence,  $P(\mathfrak{m})$  holds if  $\mathfrak{m}$  is infinite. It leads to a contradiction since  $\mathfrak{m} \notin A$  by the definition of the set  $A$ .

Thus, we can conclude that  $\mathfrak{m}$  is finite.

Since  $\mathfrak{m} \notin A$ , then by the definition of the sentence  $P$  we can find a finite subset  $G$  of  $X$  with  $|G| = \mathfrak{m}$  as desired for the Lemma.

**You see!! It is a very clever proof, isn't it!!!.** □

**Proposition 7.9.** *Let  $\Gamma$  be a transfinitely closed subspace of  $X^*$  and  $f_0$  be an element in  $X^* \setminus \Gamma$ . Then there is an element  $x_0 \in X$  such that  $f_0(x_0) = 1$  and  $f(x_0) = 0$  for all  $f \in \Gamma$ .*

*Proof.* Let  $0 < c < \text{dist}(f_0, \Gamma)$ . Recall Lemma 7.8 that there is a non-empty finite subset  $G$  of  $X$  such that there is no element  $f \in \Gamma$  satisfying the following condition:

$$(7.4) \quad |f(tx) - f_0(tx)| \leq c\|tx\| \quad \text{for all } x \in G \text{ and for all } t > 0.$$

Notice that we may assume that  $\|x\| \leq 1$  for all  $x \in G$  because  $G$  is finite. In particular, by considering  $t = 1/k, k = 1, 2, \dots$  into the Eq 7.4 above, we can find a sequence  $(x_j)$  in  $X$  with  $\lim_j x_j = 0$  such that there is no element  $f \in \Gamma$  satisfying the condition:

$$(7.5) \quad |f(x_j) - f_0(x_j)| \leq c \quad \text{for all } j = 1, 2, \dots$$

Now let  $\tilde{f}_0 := (f_0(x_j))$  and  $\tilde{f} := (f(x_j))$  for  $f \in \Gamma$ . Then  $\tilde{f}_0 \in c_0$  and  $F := \{\tilde{f} : f \in \Gamma\}$  is a subspace of  $c_0$ . Note that the Eq 7.5 implies  $\|\tilde{f}_0 - \tilde{f}\|_\infty > c$  for all  $f \in \Gamma$ . This gives  $\text{dist}(\tilde{f}_0, \overline{F}) \geq c > 0$ . Then by the Hahn-Banach separation Theorem, there is an element  $\varphi = (t_j) \in \ell_1 = c_0^*$  such that  $\varphi(\tilde{f}_0) = 1$  and  $\varphi(\tilde{f}) = 0$  for all  $f \in \Gamma$ . In particular, we have

$$\sum_{j=1}^{\infty} t_j f_0(x_j) = 1 \quad \text{and} \quad \sum_{j=1}^{\infty} t_j f(x_j) = 0 \quad \text{for all } f \in \Gamma.$$

Therefore, if we put  $x_0 := \sum_{j=1}^{\infty} t_j x_j \in X$ , then the element  $x_0$  is as desired. □

Recall the notation that for every element  $x$  in a Banach space  $X$ ,  $\hat{x}$  denotes the element in  $X^{**}$  given by  $x(f) := f(x)$  for all  $f \in X^*$ . Put  $\hat{A} := \{\hat{a} : a \in A\} \subseteq X^{**}$  for a subset  $A$  of  $X$ .

**Lemma 7.10.** *Let  $X$  be a Banach space. Assume that every normed bounded sequence in  $X$  has a weakly convergent subsequence. Let  $D$  be a countably infinite subset of  $X$ . Then every bounded transfinite sequence in  $\hat{D}$  has a transfinite limit in  $\hat{X} := \{\hat{x} : x \in X\} \subseteq X^{**}$ , that is, for every transfinite sequence  $(x_\xi)_{\xi < \theta}$  in  $D$ , there is an element  $z \in X$  such that*

$$(7.6) \quad f(z) \leq \overline{\lim}_{\xi \rightarrow \theta} f(x_\xi)$$

for all  $f \in X^*$ .

*Proof.* Let  $\theta$  be a limit ordinal and  $(x_\xi)_{\xi < \theta}$  be a bounded transfinite sequence in  $D$ . By the assumption of  $D$ , we can write  $D = \{x_i : i = 0, 1, 2, \dots\}$ .

**Case 1:** there is an infinite sequence  $(\xi_n)_{n=0}^\infty$  in  $[0, \theta)$  such that for every ordinal  $\eta < \theta$ , there is  $N \in \mathbb{N}$  such that  $\eta < \xi_n < \theta$  for all  $n > N$ , write  $\lim_{n \rightarrow \infty} \xi_n = \theta$ . In this case, put  $x_{\xi_i} := x_i$  for  $i = 0, 1, 2, \dots$ . Then by assumption, there is a weakly convergent sequence  $(x_{\xi_{i_k}})$  with the weak limit, say  $z \in X$ . This implies that

$$\hat{z}(f) = f(z) = \lim_{k \rightarrow \infty} f(x_{\xi_{i_k}}) \leq \overline{\lim}_{\xi \rightarrow \theta} f(x_\xi) = \overline{\lim}_{\xi \rightarrow \theta} \hat{x}_\xi(f)$$

for all  $f \in X^*$ .

The proof is complete if Eq 7.6 also holds for the following case.

**Case 2:** there is no sequence  $(\xi_n)$  in  $[0, \theta)$  such that  $\lim_{n \rightarrow \infty} \xi_n = \theta$ .

In this case, for each  $x_i \in D$ , let  $Z_i := \{\xi < \theta : x_\xi = x_i\}$ . We will see that  $Z_i$  is cofinal subset of  $[0, \theta)$  for some  $i \in \omega$ , that is, for any ordinal  $\eta < \theta$ , there is  $\nu \in Z_i$  such that  $\eta < \nu < \theta$ . To see this, if we assume that every  $Z_i, i \in \omega$ , is not a cofinal subset of  $[0, \theta)$ , then for each  $i \in \omega$ , there is an ordinal  $\mu_i$  with  $\mu_i < \theta$  such that  $\xi \leq \mu_i$  for all  $\xi \in Z_i$ . Now put  $\lambda_0 := \mu_0$  and  $\lambda_n := \max\{\mu : i = 0, 1, \dots, n-1\}$ . This gives an increasing sequence  $(\lambda_n)_{n \in \omega}$  in  $[0, \theta)$ . Then by the assumption of this case, there is  $\lambda \in [0, \theta)$  such that  $\lambda_n \leq \lambda$  for all  $n \in \omega$  and hence,  $\xi \leq \lambda$  for all  $\xi \in [0, \theta)$  because  $\bigcup_{i \in \omega} Z_i = [0, \theta)$ . This implies that there is no ordinal  $\xi$  such that  $\lambda < \xi < \theta$  that will lead to a contradiction because  $\theta$  is a limit ordinal.

Now let  $Z_{i_0}$  be a cofinal subset of  $[0, \theta)$  and  $z := x_{i_0}$ . Then by the definition of  $Z_{i_0}$ ,  $x_\xi = z$  for all  $\xi \in Z_{i_0}$ . Thus, we have

$$\widehat{z}(f) = f(z) = \lim_{\xi \in Z_{i_0}; \xi \rightarrow \theta} f(x_\xi) \leq \overline{\lim}_{\xi \rightarrow \theta} f(x_\xi) = \overline{\lim}_{\xi \rightarrow \theta} \widehat{x}_\xi(f).$$

The proof is finished.  $\square$

We are now in a position to reach the following main result in this section.

**Theorem 7.11.** *Let  $X$  be a separable Banach space. Then  $X$  is reflexive if and only if every bounded sequence in  $X$  has a weakly convergent subsequence.*

*Proof.* The necessary condition has been shown in Corollary 6.12.

We are going to show the converse statement. Let  $D$  be a countable subset of  $X$ .

**Claim:** The space  $\widehat{X} := \{\widehat{x} : x \in X\}$  is transfinitely closed in  $X^{**}$ , that is, for every bounded transfinite sequence  $(\widehat{x}_\xi)_{\xi < \theta}$  in  $\widehat{X}$ , there is an element  $z \in X$  such that

$$f(z) \leq \overline{\lim}_{\xi \rightarrow \theta} f(x_\xi)$$

for all  $f \in X^*$ .

Now for each  $n = 1, 2, \dots$  and each  $x_\xi, \xi < \theta$ , we choose an element  $x_\xi^n \in D$  such that

$$\|x_\xi - x_\xi^n\| < \frac{1}{n}.$$

From this, for each  $n = 1, 2, \dots$ , we obtain a bounded  $\theta$ -transfinite sequence  $(x_\xi^n)_{\xi < \theta}$  in  $D$ . For each  $n = 1, 2, \dots$ , Lemma 7.10 gives an element  $z_n \in X$  such that

$$f(z_n) \leq \overline{\lim}_{\xi \rightarrow \theta} f(x_\xi^n)$$

for all  $f \in X^*$ . In addition from this we have  $\|z_n\| \leq 1 + \sup_{\xi < \theta} \|x_\xi\|$  for all  $n = 1, 2, \dots$ . Then the necessary condition implies that  $(z_n)$  has a weak convergent subsequence  $(z_{n_j})$ . Let  $z$  be the weak limit of  $(z_{n_j})$ . Then we have

$$\begin{aligned} f(z) &= \lim_{j \rightarrow \infty} f(z_{n_j}) \\ &\leq \lim_{j \rightarrow \infty} \overline{\lim}_{\xi \rightarrow \theta} f(x_\xi^{n_j}) \\ &\leq \lim_{j \rightarrow \infty} \left( \overline{\lim}_{\xi \rightarrow \theta} f(x_\xi) + \frac{\|f\|}{n_j} \right) \\ &= \overline{\lim}_{\xi \rightarrow \theta} f(x_\xi) \end{aligned}$$

for all  $f \in X^*$ . The Claim follows.

Finally, if  $\widehat{X} \subsetneq X^{**}$ , then there is an element  $\phi \in X^{**} \setminus \widehat{X}$ . Note that  $\widehat{X}$  is a closed subspace of  $X^{**}$ . Using Lemma 7.9, the transfinite closeness of  $\widehat{X}$  implies that there is an element  $f_0 \in X^*$



such that  $f_0(x) = \widehat{x}(f_0) = 0$  for all  $x \in X$  and so,  $f_0 = 0$  but  $\phi(f_0) = 1$  that is ridiculous. Thus,  $\widehat{X} = X^{**}$ . The proof is complete.  $\square$

## 8. APPENDIX: $w^*$ -COMPACTNESS

Throughout this section  $X$  always denotes a normed space. I suppose that the students have learned a standard course of topology before.

Now for each  $\varepsilon > 0$  and for finitely many elements  $x_1, \dots, x_m$  in  $X$ , let

$$W(x_1, \dots, x_m; \varepsilon) := \{f \in X^* : |f(x_i)| < \varepsilon; \forall i = 1, \dots, m\}.$$

It is noted that  $0 \in W(x_1, \dots, x_m; \varepsilon)$  for any  $\varepsilon > 0$  and for all finitely many elements  $x_1, \dots, x_m$  in  $X$ .

**Definition 8.1.** *The weak\*-topology on the dual space  $X^*$  is the topology generated by the collection*

$$\{h + W(x_1, \dots, x_m; \varepsilon) : h \in X^*; \text{ for } \varepsilon > 0 \text{ and for finitely many } x_1, \dots, x_m \in X\}.$$

The following result is an important feature for a weak\*-continuous linear functional on  $X^*$ .

**Proposition 8.2.** *Let  $X$  be a normed space and let  $\varphi \in X^{**}$ . Then  $\varphi : X^* \rightarrow \mathbb{K}$  is  $w^*$ -continuous if and only if there is a unique element  $x_0 \in X$  such that  $\varphi(f) = \widehat{x_0}(f) := f(x_0)$  for all  $f \in X^*$ .*

*Proof.* The converse is clear. We are going to show the necessary condition.

Assume that  $\varphi$  is  $w^*$ -continuous, in particular,  $\varphi$  is  $w^*$  continuous at 0. By the definition of  $w^*$ -open neighbourhood of 0 given above (Definition 8.1), there are finitely elements  $x_1, \dots, x_N$  in  $X$  and  $\delta > 0$  such that  $|\varphi(f)| < 1$  whenever  $f \in X^*$  satisfies  $|f(x_k)| < \delta$  for all  $k = 1, \dots, N$ . Now if we put  $L(f) := \max_{k=1, \dots, N} |f(x_k)|$  for  $f \in X^*$ , then clearly we have  $L(tf) = tL(f)$  for all  $f \in X^*$  and  $t > 0$ . From this, we see that  $L(\frac{\delta}{2L(f)}f) < \delta$  while  $f \in X^*$  satisfies  $L(f) > 0$ . Therefore, we have

$$(8.1) \quad |\varphi(f)| \leq \frac{2}{\delta} L(f)$$

whenever  $f \in X^*$  with  $L(f) > 0$ .

**Claim 1:** Eq 8.1 holds for all  $f \in X^*$ . To see this, it suffices to show that if  $L(f) = 0$ , then  $\varphi(f) = 0$ . In fact, if  $L(f) = 0$ , then  $L(tf) = tL(f) = 0$  for all  $t > 0$ . In particular, we have  $|(tf)(x_k)| = 0 < \delta$  for all  $t > 0$  and for all  $k = 1, \dots, N$ . This implies that  $t|\varphi(f)| = |\varphi(tf)| < 1$  for all  $t > 0$ . Thus, we have  $\varphi(f) = 0$  as desired.

So, from Eq 8.1, we have

$$(8.2) \quad \bigcap_{k=1}^N \ker \widehat{x_k} \subseteq \ker \varphi.$$

**Claim 2:**  $\varphi \in \text{span}\{\widehat{x_k} : k = 1, \dots, N\}$  in  $X^{**}$ . To see this, we first notice that if  $\widehat{x_j} \in \text{span}\{\widehat{x_k} : k \neq j\}$  for some  $j$ , then  $\bigcap_{k \neq j} \ker \widehat{x_k} \subseteq \ker \widehat{x_j}$  and so, we have  $\bigcap_{k=1}^N \ker \widehat{x_k} = \bigcap_{k \neq j} \ker \widehat{x_k}$ . Therefore, we may assume that  $\{\widehat{x_1}, \dots, \widehat{x_N}\}$  is a linearly independent set by considering a maximal linearly independent subset of  $\{\widehat{x_1}, \dots, \widehat{x_N}\}$ . In this case,  $\{x_1, \dots, x_N\}$  is a linearly independent subset of  $X$ . Then the Hahn-Banach Theorem implies that there are  $x_1^*, \dots, x_N^*$  in  $X^*$  such that  $x_i^*(x_j) = 1$  if  $i = j$ ; otherwise is 0 for  $i, j = 1, \dots, N$ . From this for any  $f \in X^*$ , we have  $\widehat{x_j}(f - \sum_{k=1}^N f(x_k)x_k^*) = 0$  for all  $j = 1, \dots, N$ . Eq 8.2 implies that  $\varphi(f - \sum_{k=1}^N f(x_k)x_k^*) = 0$  for all  $f \in X^*$ . This gives

$$\varphi(f) = \sum_{k=1}^N f(x_k)\varphi(x_k^*) = \sum_{k=1}^N \varphi(x_k^*)\widehat{x_k}(f)$$

for all  $f \in X^*$ . Hence, **Claim 2** follows from

$$\varphi = \sum_{k=1}^N \varphi(x_k^*) \widehat{x_k}.$$

Now if we put  $x_0 := \sum_{k=1}^N \varphi(x_k^*) x_k \in X$ , then  $\varphi = \widehat{x_0}$ . On the other hand, the uniqueness is clear due to the Hahn-Banach Theorem. We finish the proof.  $\square$

The following is clearly shown by the definition.

**Lemma 8.3.** *Using the notations as above, we have*

- (i) *The weak\*-topology is Hausdorff.*
- (ii) *Let  $f \in X^*$ . Then for each open neighborhood  $V$  of  $f$ , there are  $\varepsilon > 0$  and  $x_1, \dots, x_m$  in  $X$  such that  $f + W(x_1, \dots, x_m; \varepsilon) \subseteq V$ , that is, the collection  $\{f + W(x_1, \dots, x_m; \varepsilon)\}$  forms an open basis at  $f$ .*
- (iii) *A sequence  $(f_n)$  weak\* converges to  $f$  in  $X^*$  if and only if for each  $\varepsilon > 0$  and for finitely many elements  $x_1, \dots, x_m$  in  $X$ , there is a positive integer  $N$  such that  $f_n - f \in W(x_1, \dots, x_m; \varepsilon)$  for all  $n \geq N$ .*

Before showing the main result in this section, let us recall that product topologies.

Let  $(Z_i)_{i \in I}$  be a collection of topological spaces. Let  $Z$  be the usual Cartesian product, that is

$$Z := \prod_{i \in I} Z_i : \{z : I \rightarrow \bigcup_{i \in I} Z_i : z(i) \in Z_i; \forall i \in I\}.$$

Let  $p_i : Z \rightarrow Z_i$  be the natural projection for  $i \in I$ . The product topology on  $Z$  is the weakest topology such that each projection  $p_i$  is continuous. More precisely, the following collection forms an open basis for the product topology:

$$\left\{ \bigcap_{i \in J} p_i^{-1}(W_i) : J \text{ is a finite subset of } I \text{ and } W_i \text{ is an open subset of } Z_i \right\}.$$

We have the following famous result in topology.

**Theorem 8.4. Tychonoff's Theorem:** *The Cartesian product of compact spaces is compact under the product topology.*

The following result is known as the *Alaoglu's Theorem*.

**Theorem 8.5.** *The closed unit ball  $B_{X^*}$  of the dual space  $X$  is compact with respect to the weak\*-topology.*

*Proof.* For each  $x \in X$ , put  $Z_x := [-\|x\|, \|x\|] \subseteq \mathbb{R}$ . Each  $Z_x$  is endowed with the usual subspace topology of  $\mathbb{R}$ . Then  $Z_x$  is a compact set for all  $x \in X$ . Let

$$Z := \prod_{x \in X} Z_x.$$

Then the set  $Z$  is a compact Hausdorff space under the product topology. Define a mapping by

$$T : f \in B_{X^*} \mapsto Tf \in Z; \quad Tf(x) := f(x) \in Z_x \text{ for } x \in X.$$

Then by the definitions of weak\*-topology and the product topology, it is clear that  $T$  is a homeomorphism from  $B_{X^*}$  onto its image  $T(B_{X^*})$ . Recall a fact that any closed subset of a compact Hausdorff space is compact. Since  $Z$  is compact Hausdorff, it suffices to show that  $T(B_{X^*})$  is a closed subset of  $Z$ .

Let  $z \in \overline{T(B_{X^*})}$ . We are going to show that there is an element  $f \in B_{X^*}$  such that  $f(x) = z(x)$

for all  $x \in X$ .

Define a function  $f : X \rightarrow \mathbb{K}$  by

$$f(x) := z(x)$$

for  $x \in X$ .

**Claim :**  $f(x+y) = f(x) + f(y)$  for all  $x, y \in X$ . In fact if we fix  $x, y \in X$  and for any  $\varepsilon > 0$ , then by the definition of product topology, there is an element  $g \in B_{X^*}$  such that  $|g(x+y) - z(x+y)| < \varepsilon$ ;  $|g(x) - z(x)| < \varepsilon$ ; and  $|g(y) - z(y)| < \varepsilon$ . Since  $g$  is linear, we have  $g(x+y) - g(x) - g(y) = 0$ . This implies that

$$|z(x+y) - z(x) - z(y)| = |z(x+y) - g(x+y) - (z(x) - g(x)) - (z(y) - g(y))| < 3\varepsilon$$

for all  $\varepsilon > 0$ . Thus we have  $z(x+y) = z(x) + z(y)$ . The **Claim** follows.

Similarly, we have  $z(\alpha x) = \alpha z(x)$  for all  $\alpha \in \mathbb{K}$  and for all  $x \in X$ .

Therefore, the functional  $f(x) := z(x)$  is linear on  $X$ . It remains to show  $f$  is bounded with  $\|f\| \leq 1$ . In fact, for any  $x \in X$  and any  $\varepsilon > 0$ , then there is an element  $g \in B_{X^*}$  such that  $|g(x) - z(x)| < \varepsilon$ . Therefore, we have  $|f(x)| = |z(x)| \leq |g(x)| + \varepsilon \leq \|x\| + \varepsilon$ . Therefore,  $f$  is bounded and  $\|f\| \leq 1$  as desired. The proof is complete.  $\square$

## 9. OPEN MAPPING THEOREM

Let  $E$  and  $F$  be the metric spaces. A mapping  $f : E \rightarrow F$  is called an *open mapping* if  $f(U)$  is an open subset of  $F$  whenever  $U$  is an open subset of  $E$ .

Clearly, a continuous bijection is a homeomorphism if and only if it is an open map.

**Remark 9.1. Warning** *An open map need not be a closed map.*

*For example, let  $p : (x, y) \in \mathbb{R}^2 \mapsto x \in \mathbb{R}$ . Then  $p$  is an open map but it is not a closed map. In fact, if we let  $A = \{(x, 1/x) : x \neq 0\}$ , then  $A$  is closed but  $p(A) = \mathbb{R} \setminus \{0\}$  is not closed.*

**Lemma 9.2.** *Let  $X$  and  $Y$  be normed spaces and  $T : X \rightarrow Y$  a linear map. Then  $T$  is open if and only if  $0$  is an interior point of  $T(U)$  where  $U$  is the open unit ball of  $X$ .*

*Proof.* The necessary condition is obvious.

For the converse, let  $W$  be a non-empty subset of  $X$  and  $a \in W$ . Put  $b = Ta$ . Since  $W$  is open, we choose  $r > 0$  such that  $B_X(a, r) \subseteq W$ . Note that  $U = \frac{1}{r}(B_X(a, r) - a) \subseteq \frac{1}{r}(W - a)$ . Thus, we have  $T(U) \subseteq \frac{1}{r}(T(W) - b)$ . Then by the assumption, there is  $\delta > 0$  such that  $B_Y(0, \delta) \subseteq T(U) \subseteq \frac{1}{r}(T(W) - b)$ . This implies that  $b + rB_Y(0, \delta) \subseteq T(W)$  and so,  $T(a) = b$  is an interior point of  $T(W)$ .  $\square$

**Corollary 9.3.** *Let  $M$  be a closed subspace of a normed space  $X$ . Then the natural projection  $\pi : X \rightarrow X/M$  is an open map.*

*Proof.* Put  $U$  and  $V$  the open unit balls of  $X$  and  $X/M$  respectively. Using Lemma 9.2, the result is obtained by showing that  $V \subseteq \pi(U)$ . Note that if  $\bar{x} = \pi(x) \in V$ , then by the definition a quotient norm, we can find an element  $m \in M$  such that  $\|x + m\| < 1$ . Hence we have  $x + m \in U$  and  $\bar{x} = \pi(x + m) \in \pi(U)$ .  $\square$

Before showing the main result, we have to make use one of important properties of a metric space which is known as the *Baire Category Theorem*. Recall that a subset  $A$  of a metric space  $E$  is called a *nowhere dense set* if the closure  $\bar{A}$  of  $A$  has no interior point.

**Proposition 9.4.** *Let  $E$  be a complete metric space with a metric  $d$ . If  $E$  is a union of a sequence of subsets  $(A_n)$  of  $E$ , then  $\text{int}(\bar{A}_N) \neq \emptyset$  for some  $A_N$ . Hence, every complete metric space is not a countable union of nowhere dense sets.*

*Proof.* Let  $F_n := \overline{A_n}$ . Hence,  $E = \bigcup_{n=1}^{\infty} F_n$ . Assume that each  $F_n$  has no interior points. Fix an element  $x_1 \in E$ . Let  $0 < \eta_1 < 1/2$ . Then  $B(x_1, \eta_1) \not\subseteq F_1$ . Then there is an element  $x_2 \in B(x_1, \eta_1) \setminus F_1$ . Since  $F_1$  is closed, we can choose  $0 < \eta_2 < 1/2^2$  such that  $\overline{B(x_2, \eta_2)} \cap F_1 = \emptyset$  and  $\overline{B(x_2, \eta_2)} \subseteq \overline{B(x_1, \eta_1)}$ . To repeat the same step, we have a sequence of elements  $(x_k)$  in  $E$ ; a decreasing sequence of positive of numbers  $(\eta_k)$  such that for all  $k = 1, 2, \dots$  satisfy the following conditions:

- (1)  $0 < \eta_k < 1/2^k$ .
- (2)  $\overline{B(x_{k+1}, \eta_{k+1})} \subseteq \overline{B(x_k, \eta_k)}$ .
- (3)  $\overline{B(x_{k+1}, \eta_{k+1})} \cap F_k = \emptyset$ .

The completeness of  $E$ , together with conditions (1) and (2) imply that the sequence  $(x_k)$  is a Cauchy sequence and thus, the limit  $l := \lim_k x_k$  exists with  $l \in \bigcap_{k=1}^{\infty} \overline{B(x_k, \eta_k)}$ . Since  $E = \bigcup_{n=1}^{\infty} F_n$ , the limit  $l \in F_K$  for some  $K$ . However, it leads to a contradiction because  $F_K \cap \overline{B(x_K, \eta_K)} = \emptyset$  by the condition (3) above.  $\square$

**Lemma 9.5.** *Let  $T : X \longrightarrow Y$  be a bounded linear surjection from a Banach space  $X$  onto a Banach space  $Y$ . Then 0 is an interior point of  $T(U)$ , where  $U$  is the open unit ball of  $X$ , i.e.,  $U := \{x \in X : \|x\| < 1\}$ .*

*Proof.* Set  $U(r) := \{x \in X : \|x\| < r\}$  for  $r > 0$  and so,  $U = U(1)$ .

**Claim 1 :** 0 is an interior point of  $\overline{T(U(1))}$ .

Note that since  $T$  is surjective,  $Y = \bigcup_{n=1}^{\infty} T(U(n))$ . Then by the Baire Category Theorem, there exists  $N$  such that  $\text{int } \overline{T(U(N))} \neq \emptyset$ . Let  $y'$  be an interior point of  $\overline{T(U(N))}$ . Then there is  $\eta > 0$  such that  $B_Y(y', \eta) \subseteq \overline{T(U(N))}$ . Since  $B_Y(y', \eta) \cap T(U(N)) \neq \emptyset$ , we may assume that  $y' \in T(U(N))$ . Let  $x' \in U(N)$  such that  $T(x') = y'$ . Then we have

$$0 \in B_Y(y', \eta) - y' \subseteq \overline{T(U(N))} - T(x') \subseteq \overline{T(U(2N))} = 2N\overline{T(U(1))}.$$

Thus, we have  $0 \in \frac{1}{2N}(B_Y(y', \eta) - y') \subseteq \overline{T(U(1))}$ . Hence 0 is an interior point of  $\overline{T(U(1))}$ . The Claim 1 follows.

Therefore there is  $r > 0$  such that  $B_Y(0, r) \subseteq \overline{T(U(1))}$ . This implies that we have

$$(9.1) \quad B_Y(0, r/2^k) \subseteq \overline{T(U(1/2^k))}$$

for all  $k = 0, 1, 2, \dots$ .

**Claim 2 :**  $D := B_Y(0, r) \subseteq T(U(3))$ .

Let  $y \in D$ . By Eq 9.1, there is  $x_1 \in U(1)$  such that  $\|y - T(x_1)\| < r/2$ . Then by using Eq 9.1 again, there is  $x_2 \in U(1/2)$  such that  $\|y - T(x_1) - T(x_2)\| < r/2^2$ . To repeat the same steps, there exists a sequence  $(x_k)$  such that  $x_k \in U(1/2^{k-1})$  and

$$\|y - T(x_1) - T(x_2) - \dots - T(x_k)\| < r/2^k$$

for all  $k$ . On the other hand, since  $\sum_{k=1}^{\infty} \|x_k\| \leq \sum_{k=1}^{\infty} 1/2^{k-1}$  and  $X$  is Banach,  $x := \sum_{k=1}^{\infty} x_k$  exists in  $X$  and  $\|x\| \leq 2$ . This implies that  $y = T(x)$  and  $\|x\| < 3$ .

Thus we the result follows.  $\square$

**Theorem 9.6. Open Mapping Theorem :** *Using the notations as in Lemma 9.5, then  $T$  is an open mapping.*

*Proof.* The proof is complete by using Lemmas 9.2 and 9.5.  $\square$

**Proposition 9.7.** *Let  $T$  be a bounded linear isomorphism between Banach spaces  $X$  and  $Y$ . Then  $T^{-1}$  is bounded.*

Consequently, if  $\|\cdot\|$  and  $\|\cdot\|'$  both are complete norms on  $X$  such that  $\|\cdot\| \leq c\|\cdot\|'$  for some  $c > 0$ , then these two norms  $\|\cdot\|$  and  $\|\cdot\|'$  are equivalent.

*Proof.* The first assertion follows immediately from the Open Mapping Theorem.

Therefore, the last assertion can be obtained by considering the identity map  $I : (X, \|\cdot\|) \rightarrow (X, \|\cdot\|')$  which is bounded by the assumption.  $\square$

**Corollary 9.8.** *Let  $X$  and  $Y$  be Banach spaces and  $T : X \rightarrow Y$  a bounded linear operator. Then the followings are equivalent.*

(i) *The image of  $T$  is closed in  $Y$ .*

(ii) *There is  $c > 0$  such that*

$$d(x, \ker T) \leq c\|Tx\|$$

*for all  $x \in X$ .*

(iii) *If  $(x_n)$  is a sequence in  $X$  such that  $\|x_n + \ker T\| = 1$  for all  $n$ , then  $\|Tx_n\| \rightarrow 0$ .*

*Proof.* Let  $Z$  be the image of  $T$ . Then the canonical map  $\tilde{T} : X/\ker T \rightarrow Z$  induced by  $T$  is a bounded linear isomorphism. Note that  $\tilde{T}(\bar{x}) = Tx$  for all  $x \in X$ , where  $\bar{x} := x + \ker T \in X/\ker T$ . For (i)  $\Rightarrow$  (ii): suppose that  $Z$  is closed. Then  $Z$  becomes a Banach space. Then the Open Mapping Theorem implies that the inverse of  $\tilde{T}$  is also bounded. Thus, there is  $c > 0$  such that  $d(x, \ker T) = \|\bar{x}\|_{X/\ker T} \leq c\|\tilde{T}(\bar{x})\| = c\|T(x)\|$  for all  $x \in X$ . The part (ii) follows.

For (ii)  $\Rightarrow$  (i), let  $(x_n)$  be a sequence in  $X$  such that  $\lim Tx_n = y \in Y$  exists and so,  $(Tx_n)$  is a Cauchy sequence in  $Y$ . Then by the assumption,  $(\bar{x}_n)$  is a Cauchy sequence in  $X/\ker T$ . Since  $X/\ker T$  is complete, we can find an element  $x \in X$  such that  $\lim \bar{x}_n = \bar{x}$  in  $X/\ker T$ . This gives  $y = \lim T(x_n) = \lim \tilde{T}(\bar{x}_n) = \tilde{T}(\bar{x}) = T(x)$ . Therefore,  $y \in Z$ .

(ii)  $\Leftrightarrow$  (iii) is obvious. The proof is complete.  $\square$

**Proposition 9.9.** *Let  $X$  and  $Y$  be Banach spaces. Let  $T$  and  $K$  belong to  $B(X, Y)$ . Suppose that  $T(X)$  is closed and  $K$  is of finite rank, then the image  $(T + K)(X)$  is also closed.*

*Proof.* Suppose the conclusion does not hold. We write  $\bar{z} := z + \ker(T + K)$  for  $z \in X$ . Then by Corollary 9.8, there is a sequence  $(x_n)$  in  $X$  such that  $\|\bar{x}_n\| = 1$  for all  $n$  and  $\|(T + K)x_n\| \rightarrow 0$ . Thus,  $(x_n)$  can be chosen so that it is bounded. By passing a subsequence of  $(x_n)$  we may assume that  $y := \lim_n K(x_n)$  exists in  $Y$  because  $K$  is of finite rank. Therefore, we have  $\lim_n T(x_n) = -y$ . Since  $T$  has closed range, we have  $Tx = -y$  for some  $x \in X$ . This gives  $\lim T(x_n - x) = 0$ . Note that the natural map  $\tilde{T}$  is a topological isomorphism from  $X/\ker T$  onto  $T(X)$  because  $T(X)$  is closed. We see that  $\|x_n - x + \ker T\| \rightarrow 0$  and thus,  $\|y - K(x) + K(\ker T)\| = \lim \|K(x_n) - K(x) + K(\ker T)\| = 0$ . From this we have  $y - Kx = Ku$  for some  $u \in \ker T$ . In addition, for each  $n$ , there is an element  $t_n \in \ker T$  so that  $\|x_n - x + t_n\| < 1/n$ . This implies that

$$\|K(t_n - u)\| \leq \|K(t_n + (x_n - x))\| + \|-K(x_n + x) - K(u)\| \leq \|K\|1/n \rightarrow 0.$$

Therefore, we have  $\|t_n - u + (\ker T \cap \ker K)\| \rightarrow 0$  because  $t_n - u \in \ker T$  and the image of  $K|_{\ker T}$  is closed. From this we see that  $\|t_n - u + \ker(T + K)\| \rightarrow 0$ .

On the other hand, since  $Tx = -y = -Kx - Ku$  and  $u \in \ker T$ , we have  $(T + K)x = -Ku - Tu$  and so,  $x + u \in \ker(T + K)$ . Then we can now conclude that

$$\|\bar{x}_n\| = \|\bar{x}_n - (\bar{x} + \bar{u})\| \leq \|\bar{x}_n - \bar{x} - \bar{t}_n\| + \|\bar{t}_n - \bar{u}\| \rightarrow 0.$$

It contradicts to the choice of  $x_n$  such that  $\|\bar{x}_n\| = 1$  for all  $n$ . The proof is complete.  $\square$

**Remark 9.10.** *In general, the sum of operators of closed ranges may not have a closed range. Before looking for those examples, let us show the following simple useful lemma.*

**Lemma 9.11.** *Let  $X$  be a Banach space. If  $T \in B(X)$  with  $\|T\| < 1$ , then the operator  $1 - T$  is invertible, i.e., there is  $S \in B(X)$  such that  $(1 - T)S = S(1 - T) = 1$ .*

*Proof.* Note that since  $X$  is a Banach space, the set of all bounded operators  $B(X)$  is a Banach space under the usual operator norm. This implies that the series  $\sum_{k=0}^{\infty} T^k$  is convergent in  $B(X)$  because  $\|T\| < 1$ . On the other hand, we have  $1 - T^n = (1 - T)(\sum_{k=0}^{n-1} T^k)$  for all  $n = 1, 2, \dots$ . Taking  $n \rightarrow \infty$ , we see that  $(1 - T)^{-1}$  exists, in fact,  $(1 - T)^{-1} = \sum_{k=0}^{\infty} T^k$ .  $\square$

**Example 9.12.** Define an operator  $T_0 : \ell^\infty \rightarrow \ell^\infty$  by

$$T_0(x)(k) := \frac{1}{k}x(k)$$

for  $x \in \ell^\infty$  and  $k = 1, 2, \dots$ . Note that  $T_0$  is injective with  $\|T_0\| \leq 1$  and  $\text{im } T_0 \subseteq c_0$ . The Open mapping Theorem tells us that the image  $\text{im } T_0$  must not be closed. Otherwise  $T_0$  becomes an isomorphism from  $\ell^\infty$  onto a closed subspace of  $c_0$ . It is ridiculous since  $\ell^\infty$  is nonseparable but  $c_0$  is not. Now if we let  $T := \frac{1}{2}T_0$ , then  $\|T\| < 1$  and  $T$  is without closed range. Applying Lemma 9.11, we see that the operator  $S := 1 - T$  is invertible and thus,  $S$  has closed range. Then by our construction  $T = 1 - S$  is the sum of two operators of closed ranges but  $T$  does not have closed range as required.

## 10. CLOSED GRAPH THEOREM

Let  $T : X \rightarrow Y$ . The *graph* of  $T$ , denoted by  $\mathcal{G}(T)$ , is defined by the set  $\{(x, y) \in X \times Y : y = T(x)\}$ .

Now the direct sum  $X \oplus Y$  is endowed with the norm  $\|\cdot\|_\infty$ , i.e.,  $\|x \oplus y\|_\infty := \max(\|x\|_X, \|y\|_Y)$ . We write  $X \oplus_\infty Y$  when  $X \oplus Y$  is equipped with this norm.

An operator  $T : X \rightarrow Y$  is said to be closed if its graph  $\mathcal{G}(T)$  is a closed subset of  $X \oplus_\infty Y$ , i.e., whenever, a sequence  $(x_n)$  of  $X$  satisfies the condition  $\|(x_n, Tx_n) - (x, y)\|_\infty \rightarrow 0$  for some  $x \in X$  and  $y \in Y$ , we have  $T(x) = y$ .

**Theorem 10.1. Closed Graph Theorem :** *Let  $T : X \rightarrow Y$  be a linear operator from a Banach space  $X$  to a Banach  $Y$ . Then  $T$  is bounded if and only if  $T$  is closed.*

*Proof.* The part  $(\Rightarrow)$  is clear.

Assume that  $T$  is closed, i.e., the graph  $\mathcal{G}(T)$  is  $\|\cdot\|_\infty$ -closed. Define  $\|\cdot\|_0 : X \rightarrow [0, \infty)$  by

$$\|x\|_0 = \|x\| + \|T(x)\|$$

for  $x \in X$ . Then  $\|\cdot\|_0$  is a norm on  $X$ . Let  $I : (X, \|\cdot\|_0) \rightarrow (X, \|\cdot\|)$  be the identity operator. It is clear that  $I$  is bounded since  $\|\cdot\| \leq \|\cdot\|_0$ .

**Claim:**  $(X, \|\cdot\|_0)$  is Banach. In fact, let  $(x_n)$  be a Cauchy sequence in  $(X, \|\cdot\|_0)$ . Then  $(x_n)$  and  $(T(x_n))$  both are Cauchy sequences in  $(X, \|\cdot\|)$  and  $(Y, \|\cdot\|_Y)$ . Since  $X$  and  $Y$  are Banach spaces, there are  $x \in X$  and  $y \in Y$  such that  $\|x_n - x\|_X \rightarrow 0$  and  $\|T(x_n) - y\|_Y \rightarrow 0$ . Thus  $y = T(x)$  since the graph  $\mathcal{G}(T)$  is closed.

Then by Theorem 9.7, the norms  $\|\cdot\|$  and  $\|\cdot\|_0$  are equivalent. Hence, there is  $c > 0$  such that  $\|T(\cdot)\| \leq \|\cdot\|_0 \leq c\|\cdot\|$  and hence,  $T$  is bounded since  $\|T(\cdot)\| \leq \|\cdot\|_0$ . The proof is complete.  $\square$

**Example 10.2.** Let  $D := \{\mathbf{c} = (c_n) \in \ell^2 : \sum_{n=1}^{\infty} n^2 |c_n|^2 < \infty\}$ . Define  $T : D \rightarrow \ell^2$  by  $T(\mathbf{c}) = (nc_n)$ . Then  $T$  is an unbounded closed operator.

*Proof.* Note that since  $\|Te_n\| = n$  for all  $n$ ,  $T$  is not bounded. Now we claim that  $T$  is closed. Let  $(\mathbf{x}_i)$  be a convergent sequence in  $D$  such that  $(T\mathbf{x}_i)$  is also convergent in  $\ell^2$ . Write  $\mathbf{x}_i = (x_{i,n})_{n=1}^\infty$  with  $\lim_i \mathbf{x}_i = \mathbf{x} := (x_n)$  in  $D$  and  $\lim_i T\mathbf{x}_i = \mathbf{y} := (y_n)$  in  $\ell^2$ . This implies that if we fix  $n_0$ , then  $\lim_i x_{i,n_0} = x_{n_0}$  and  $\lim_i n_0 x_{i,n_0} = y_{n_0}$ . This gives  $n_0 x_{n_0} = y_{n_0}$ . Thus  $T\mathbf{x} = \mathbf{y}$  and hence  $T$  is closed.  $\square$

**Example 10.3.** Let  $X := \{f \in C^b(0,1) \cap C^\infty(0,1) : f' \in C^b(0,1)\}$ . Define  $T : f \in X \mapsto f' \in C^b(0,1)$ . Suppose that  $X$  and  $C^b(0,1)$  both are equipped with the sup-norm. Then  $T$  is a closed unbounded operator.

*Proof.* Note that if a sequence  $f_n \rightarrow f$  in  $X$  and  $f'_n \rightarrow g$  in  $C^b(0,1)$ . Then  $f' = g$ . Hence  $T$  is closed. In fact, if we fix some  $0 < c < 1$ , then by the Fundamental Theorem of Calculus, we have

$$0 = \lim_n (f_n(x) - f(x)) = \lim_n \left( \int_c^x (f'_n(t) - f'(t)) dt \right) = \int_c^x (g(t) - f'(t)) dt$$

for all  $x \in (0,1)$ . This implies that we have  $\int_c^x g(t) dt = \int_c^x f'(t) dt$ . Thus  $g = f'$  on  $(0,1)$ . On the other hand, since  $\|Tx^n\|_\infty = n$  for all  $n \in \mathbb{N}$ . Thus  $T$  is unbounded as desired.  $\square$

Let  $X$  be a normed space and let  $X^*$  be its dual space. Then there is a natural bi-linear mapping on  $X \times X^*$  (call a dual pair) given by

$$\langle \cdot, \cdot \rangle : X \times X^* \rightarrow \mathbb{K}; \quad \langle x, f \rangle = f(x).$$

Moreover, this dual pair is non-degenerate, that is,  $\langle x, f \rangle = 0$  for all  $f \in X^*$  if and only if  $x = 0$  and  $\langle x, f \rangle = 0$  for all  $x \in X$  if and only if  $f = 0$ .

**Proposition 10.4.** Let  $X$  and  $Y$  be Banach spaces. Let  $G : Y^* \rightarrow X^*$  be a  $w^*$ - $w^*$  continuous linear map. Then we have the following assertions.

(i)  $G$  is bounded.

(ii) There exists a bounded linear map  $T \in B(X, Y)$  such that  $T^* = G$ .

*Proof.* For showing Part (i), let  $(y_n^*)$  be a sequence in  $Y^*$  such that  $y_n^* \xrightarrow{\|\cdot\|} y^*$  and  $Gy_n^* \xrightarrow{\|\cdot\|} x^*$  in the norm topologies. By using the Closed Graph Theorem, we want to show  $Gy^* = x^*$ , that is,  $(Gy^*)(x) = x^*(x)$  for all  $x \in X$ . In fact,  $y_n^* \xrightarrow{\|\cdot\|} y^*$ , so  $y_n^* \xrightarrow{w^*} y^*$ . Thus, we have  $Gy_n^* \xrightarrow{w^*} Gy^*$ , so  $(Gy_n^*)(x) \rightarrow (Gy^*)(x)$  for all  $x \in X$ . On the other hand, since  $Gy_n^* \xrightarrow{\|\cdot\|} x^*$ , we have  $(Gy_n^*)(x) \rightarrow x^*(x)$  for all  $x \in X$ . Therefore,  $(Gy^*)(x) = x^*(x)$  for all  $x \in X$  as desired.

For Part (ii), note that for each  $x \in X$ , the map  $f \in Y^* \mapsto \langle x, Gf \rangle$  is  $w^*$ -continuous on  $Y^*$ . Hence, Proposition 8.2 gives a unique element  $Rx \in Y$  such that

$$\langle Rx, f \rangle = \langle x, Gf \rangle$$

for all  $f \in Y^*$ . Then by using Part (i) and the Closed Graph Theorem,  $R$  is bounded. The proof is complete.  $\square$

## 11. UNIFORM BOUNDEDNESS THEOREM

**Theorem 11.1. Uniform Boundedness Theorem :** Let  $\{T_i : X \rightarrow Y : i \in I\}$  be a family of bounded linear operators from a Banach space  $X$  into a normed space  $Y$ . Suppose that for each  $x \in X$ , we have  $\sup_{i \in I} \|T_i(x)\| < \infty$ . Then  $\sup_{i \in I} \|T_i\| < \infty$ .

*Proof.* For each  $x \in X$ , define

$$\|x\|_0 := \max(\|x\|, \sup_{i \in I} \|T_i(x)\|).$$

Then  $\|\cdot\|_0$  is a norm on  $X$  and  $\|\cdot\| \leq \|\cdot\|_0$  on  $X$ . If  $(X, \|\cdot\|_0)$  is complete, then by the Open Mapping Theorem. This implies that  $\|\cdot\|$  is equivalent to  $\|\cdot\|_0$  and thus there is  $c > 0$  such that

$$\|T_j(x)\| \leq \sup_{i \in I} \|T_i(x)\| \leq \|x\|_0 \leq c\|x\|$$

for all  $x \in X$  and for all  $j \in I$ . Thus  $\|T_j\| \leq c$  for all  $j \in I$  is as desired.

Thus it remains to show that  $(X, \|\cdot\|_0)$  is complete. In fact, if  $(x_n)$  is a Cauchy sequence in  $(X, \|\cdot\|_0)$ , then it is also a Cauchy sequence with respect to the norm  $\|\cdot\|$  on  $X$ . Write  $x := \lim_n x_n$  with respect to the norm  $\|\cdot\|$ . For any  $\varepsilon > 0$ , there is  $N \in \mathbb{N}$  such that  $\|T_i(x_n - x_m)\| < \varepsilon$  for all  $m, n \geq N$  and for all  $i \in I$ . Now fixing  $i \in I$  and  $n \geq N$  and taking  $m \rightarrow \infty$ , we have  $\|T_i(x_n - x)\| \leq \varepsilon$  and thus  $\sup_{i \in I} \|T_i(x_n - x)\| \leq \varepsilon$  for all  $n \geq N$ . Thus, we have  $\|x_n - x\|_0 \rightarrow 0$  and hence  $(X, \|\cdot\|_0)$  is complete.  $\square$

**Remark 11.2.** Consider  $c_{00} := \{\mathbf{x} = (x_n) : \exists N, \forall n \geq N; x_n \equiv 0\}$  which is endowed with  $\|\cdot\|_\infty$ . Now for each  $k \in \mathbb{N}$ , if we define  $T_k \in c_{00}^*$  by  $T_k((x_n)) := kx_k$ , then  $\sup_k |T_k(\mathbf{x})| < \infty$  for each  $\mathbf{x} \in c_{00}$  but  $(\|T_k\|)$  is not bounded, in fact,  $\|T_k\| = k$ . Thus the assumption of the completeness of  $X$  in Theorem 11.1 is essential.

**Corollary 11.3.** Let  $X$  and  $Y$  be as in Theorem 11.1. Let  $T_k : X \rightarrow Y$  be a sequence of bounded operators. Assume that  $\lim_k T_k(x)$  exists in  $Y$  for all  $x \in X$ . Then there is  $T \in B(X, Y)$  such that  $\lim_k \|(T - T_k)x\| = 0$  for all  $x \in X$ . Moreover, we have  $\|T\| \leq \liminf_k \|T_k\|$ .

*Proof.* Note that by the assumption, we can define a linear operator  $T$  from  $X$  to  $Y$  given by  $Tx := \lim_k T_k x$  for  $x \in X$ . We need to show that  $T$  is bounded. In fact,  $(\|T_k\|)$  is bounded by the Uniform Boundedness Theorem since  $\lim_k T_k x$  exists for all  $x \in X$ . Hence, for each  $x \in B_X$ , there is a positive integer  $K$  such that  $\|Tx\| \leq \|T_K x\| + 1 \leq (\sup_k \|T_k\|) + 1$ . Thus,  $T$  is bounded.

Finally, it remains to show the last assertion. In fact, note that for any  $x \in B_X$  and  $\varepsilon > 0$ , there is  $N(x) \in \mathbb{N}$  such that  $\|Tx\| < \|T_k x\| + \varepsilon < \|T_k\| + \varepsilon$  for all  $k \geq N(x)$ . This gives  $\|Tx\| \leq \inf_{k \geq N(x)} \|T_k\| + \varepsilon$  for all  $k \geq N(x)$  and hence,  $\|Tx\| \leq \inf_{k \geq N(x)} \|T_k\| + \varepsilon \leq \sup_n \inf_{k \geq n} \|T_k\| + \varepsilon$  for all  $x \in B_X$  and  $\varepsilon > 0$ . Therefore, we have  $\|T\| \leq \liminf_k \|T_k\|$ .  $\square$

**Corollary 11.4.** Every weakly convergent sequence in a normed space must be bounded.

*Proof.* Let  $(x_n)$  be a weakly convergent sequence in a normed space  $X$ . If we let  $Q : X \rightarrow X^{**}$  be the canonical isometry, then  $(Qx_n)$  is a bounded sequence in  $X^{**}$ . Note that  $(x_n)$  is weakly convergent if and only if  $(Qx_n)$  is  $w^*$ -convergent. Thus,  $(Qx_n(x^*))$  is bounded for all  $x^* \in X^*$ . Note that the dual space  $X^*$  must be complete. We can apply the Uniform Boundedness Theorem to see that  $(Qx_n)$  is bounded and so is  $(x_n)$ .  $\square$

## 12. PROJECTIONS ON BANACH SPACES

Throughout this section, let  $X$  be a Banach space. A linear operator  $P : X \rightarrow X$  is called a *projection (or idempotent)* if it is bounded and satisfies the condition  $P^2 = P$ .

In addition, a closed subspace  $E$  of  $X$  is said to be *complemented* if there is a closed subspace  $F$  of  $X$  such that  $X = E \oplus F$ .

**Proposition 12.1.** A closed subspace  $E$  of  $X$  is complemented if and only if there is a projection  $Q$  on  $X$  with  $E = \text{im } Q$ .

*Proof.* We first suppose that there is a closed subspace  $F$  of  $X$  such that  $X = E \oplus F$ . Define an operator  $Q : X \rightarrow X$  by  $Qx = u$  if  $x = u + v$  for  $u \in E$  and  $v \in F$ . It is clear that we have  $Q^2 = Q$ . For showing the boundedness of  $Q$ , by using the Closed Graph Theorem, we need to show that if



$(x_n)$  is a sequence in  $E$  such that  $\lim x_n = x$  and  $\lim Qx_n = u$  for some  $x, u \in E$ , then  $Qx = u$ . Indeed, if we let  $x_n = y_n + z_n$  where  $y_n \in E$  and  $z_n \in F$ , then  $Qx_n = y_n$ . Note that  $(z_n)$  is a convergent sequence in  $F$  because  $z_n = x_n - y_n$  and  $(x_n)$  and  $(y_n)$  both are convergent. Let  $w = \lim z_n$ . This implies that

$$x = \lim x_n = \lim(y_n + z_n) = u + w.$$

Since  $E$  and  $F$  are closed, we have  $u \in E$  and  $w \in F$ . Therefore, we have  $Qx = u$  as desired. The converse is clear. In fact, we have  $X = \text{im } Q \oplus \ker Q$  in this case.  $\square$

**Proposition 12.2.** *If  $M$  is a finite dimensional subspace of a normed space  $X$ , then  $M$  is complemented in  $X$ . Furthermore, if  $\dim M = 1$ , then there is a bounded linear projection  $Q$  from  $X$  onto  $M$  such that  $\|Q\| = 1$ .*

*Proof.* In fact, if  $M$  is spanned by  $\{e_i : i = 1, 2, \dots, m\}$ , then  $M$  is closed and by the Hahn-Banach Theorem, for each  $i = 1, \dots, m$ , there is  $e_i^* \in X^*$  such that  $e_i^*(e_j) = 1$  if  $i = j$ , otherwise, it is equal to 0. Put  $N := \bigcap_{i=1}^m \ker e_i^*$ . Then  $X = M \oplus N$ .

To see the last assertion, let  $M := \mathbb{K}x_0$  for some  $x_0 \in X$ . The Hahn-Banach Theorem gives an element  $f \in X^*$  with  $\|f\| = 1$  such that  $f(x_0) = \|x_0\|$ . Notice that we have  $X = \ker f \oplus M$ . Then for any  $x \in X$  we have  $x = y + tx_0$  for some  $t \in \mathbb{K}$  and for some  $y \in \ker f$ . Thus,  $\|tx_0\| = f(x) \leq \|x\|$ . Now if let  $Q$  be the corresponding projection from  $X$  onto  $M$  along this decomposition, then we have  $\|Qx\| = \|tx_0\| \leq \|x\|$  for all  $x \in X$ . Hence,  $\|Q\| \leq 1$ . However, we always have  $\|Q\| = \|Q^2\| \leq \|Q\|^2$  and so  $\|Q\| \geq 1$ . Therefore,  $\|Q\| = 1$  as desired.  $\square$

The following example can be found in [4].

**Example 12.3.**  $c_0$  is not complemented in  $\ell^\infty$ .

*Proof.* It will be shown by the contradiction. Suppose that  $c_0$  is complemented in  $\ell^\infty$ .

**Claim 1:** There is a sequence  $(f_n)$  in  $(\ell^\infty)^*$  such that  $c_0 = \bigcap_{n=1}^\infty \ker f_n$ .

In fact, by the assumption, there is a closed subspace  $F$  of  $\ell^\infty$  such that  $\ell^\infty = c_0 \oplus F$ . If we let  $P$  be the projection from  $\ell^\infty$  onto  $F$  along this decomposition, then  $\ker P = c_0$  and  $P$  is bounded by the Closed Graph Theorem. Let  $e_n^* : \ell^\infty \rightarrow \mathbb{K}$  be the  $n$ -th coordinate functional. Then  $e_n^* \in (\ell^\infty)^*$ . Thus, if we put  $f_n = e_n^* \circ P$ , then  $f_n \in (\ell^\infty)^*$  and  $c_0 = \bigcap_{n=1}^\infty \ker f_n$  as desired.

**Claim 2:** For each irrational number  $\alpha \in [0, 1]$ , there is an infinite subset  $N_\alpha$  of  $\mathbb{N}$  such that  $N_\alpha \cap N_\beta$  is a finite set if  $\alpha$  and  $\beta$  both are distinct irrational numbers in  $[0, 1]$ .

In fact, we write  $[0, 1] \cap \mathbb{Q}$  as a sequence  $(r_n)$ . Then for each irrational  $\alpha$  in  $[0, 1]$ , there is a subsequence  $(r_{n_k})$  of  $(r_n)$  such that  $\lim_k r_{n_k} = \alpha$ . Let  $N_\alpha := \{n_k : k = 1, 2, \dots\}$ . From this, we see that  $N_\alpha \cap N_\beta$  is a finite set whenever  $\alpha, \beta \in [0, 1] \cap \mathbb{Q}^c$  with  $\alpha \neq \beta$ . Claim 2 follows.

Now for each  $\alpha \in [0, 1] \cap \mathbb{Q}^c$ , define an element  $x_\alpha \in \ell^\infty$  by  $x_\alpha(k) \equiv 1$  as  $k \in N_\alpha$ ; otherwise,  $x_\alpha(k) \equiv 0$ .

**Claim 3:** If  $f \in (\ell^\infty)^*$  with  $c_0 \subseteq \ker f$ , then for any  $\eta > 0$ , the set  $\{\alpha \in [0, 1] \cap \mathbb{Q}^c : |f(x_\alpha)| \geq \eta\}$  is finite.

Note that by considering the decomposition  $f = \text{Re}(f) + i\text{Im}(f)$ , it suffices to show that the set  $\{\alpha \in [0, 1] \cap \mathbb{Q}^c : f(x_\alpha) \geq \eta\}$  is finite. Let  $\alpha_1, \dots, \alpha_N$  in  $[0, 1] \cap \mathbb{Q}^c$  such that  $f(x_{\alpha_j}) \geq \eta$ ,  $j = 1, \dots, N$ . Now for each  $j = 1, \dots, N$ , set  $y_j(k) \equiv 1$  as  $k \in N_{\alpha_j} \setminus \bigcup_{m \neq j} N_{\alpha_m}$ ; otherwise  $y_j \equiv 0$ . Note that  $x_{\alpha_j} - y_j \in c_0$  since  $N_\alpha \cap N_\beta$  is finite for  $\alpha \neq \beta$  by Claim 2. Hence, we have  $f(x_{\alpha_j}) = f(y_j)$  for all  $j = 1, \dots, N$ . Moreover, we have  $\{k : y_j(k) = 1\} \cap \{k : y_i(k) = 1\} = \emptyset$  for  $i, j = 1, \dots, N$  with  $i \neq j$ . Thus, we have  $\|y\|_\infty = 1$ . Now we can conclude that

$$\|f\| \geq f\left(\sum_{j=1}^N y_j\right) = \sum_{j=1}^N f(x_{\alpha_j}) \geq N\eta.$$

This implies that  $|\{\alpha : f(\alpha) \geq \eta\}| \leq \|f\|/\eta$ . Claim 3 follows.

We are now going to complete the proof. Now let  $(f_n)$  be the sequence in  $(\ell^\infty)^*$  as found in the Claim 1. Claim 3 implies that the set  $S := \bigcup_{n=1}^\infty \{\alpha \in \mathbb{Q}^c \cap [0, 1] : f_n(x_\alpha) \neq 0\}$  is countable. Thus, there exists  $\gamma \in [0, 1] \cap \mathbb{Q}^c$  such that  $\gamma \notin S$ . Thus, we have  $x_\gamma \in \bigcap_{n=1}^\infty \ker f_n$ . Besides, since  $N_\gamma$  is an infinite set, we see that  $x_\gamma \notin c_0$ . Therefore, we have  $c_0 \subsetneq \bigcap \ker f_k$  which contradicts to Claim 1.  $\square$

**Proposition 12.4. (Dixmier)** *Let  $X$  be a normed space. Let  $i : X \rightarrow X^{**}$  and  $j : X^* \rightarrow X^{***}$  be the natural embeddings. Then the composition  $Q := j \circ i^* : X^{***} \rightarrow X^{***}$  is a projection with  $Q(X^{***}) = X^*$ .*

*Consequently,  $X^*$  is a complemented closed subspace of  $X^{***}$ .*

*Proof.* Clearly,  $Q$  is bounded. Note that  $i^* \circ j = Id_{X^*} : X^* \rightarrow X^*$ . From this, we see that  $Q^2 = Q$  as desired.

We need to show that  $im Q = X^*$ , more precisely,  $im Q = j(X^*)$ . In fact, it follows from  $Q \circ j = j$  by using the equality  $i^* \circ j = Id_{X^*}$  again.

The last assertion follows immediately from Proposition 12.1.  $\square$

**Corollary 12.5.**  *$c_0$  is not isomorphic to the dual space of a normed space.*

*Proof.* Suppose not. Let  $T : c_0 \rightarrow X^*$  be an isomorphism from  $c_0$  onto the dual space of some normed space  $X$ . Then  $T^{**} : c_0^{**} = \ell^\infty \rightarrow X^{***}$  is an isomorphism too. Let  $Q : X^{***} \rightarrow X^{***}$  be the projection with  $im Q = X^*$  which is found in Proposition 12.4.

Now put  $P := (T^{**})^{-1} \circ Q \circ T^{**} : \ell^\infty \rightarrow \ell^\infty$ . Then  $P$  is a projection.

On the other hand, we always have  $T^{**}|_{c_0} = T$  (see Remark 5.2). This implies that  $im P = c_0$ . Thus,  $c_0$  is complemented in  $\ell^\infty$  by Proposition 12.1 which leads to a contradiction by Example 12.3.  $\square$

Recall that a closed subspace  $M$  of a Banach space  $E$  is called an  $M$ -ideal if the space  $M^\perp := \{x^* \in E^* : x^*(M) \equiv 0\}$  is a  $\ell_1$ -direct summand of  $E^*$ , that is, there is another closed subspace  $N$  of  $E^*$  such that  $E^* = M^\perp \oplus_{\ell_1} N$ , i.e., for every element  $x^* \in E^*$  satisfies the condition:  $x^* = u + v$  and  $\|x^*\| = \|u\| + \|v\|$  for a pair of elements  $u$  and  $v$  in  $M^\perp$  and  $N$  respectively.

**Proposition 12.6.** *We keep the notation as give in Proposition 12.4. If  $X$  is viewed as a closed subspace of  $X^{**}$  and suppose that  $X^{***} = X^\perp \oplus_{\ell_1} N$  for some closed subspace  $N$  of  $X^{***}$ , i.e.  $X$  is an  $M$ -ideal of  $X^{**}$ , then  $N = X^*$ .*

*Proof.* Let  $Q : X^{***} \rightarrow X^{***}$  be the projection given in Proposition 12.4. Recall that  $Qz = j(z|_X)$  for  $z \in X^{***}$  and  $im Q = X^*$ . Moreover  $\|Q\| \leq 1$ . Note that  $\ker Q = X^\perp := \{z \in X^{***} : z|_X \equiv 0\}$  and hence,  $X^{***} = X^\perp \oplus_{\ell_1} X^*$ . Let  $z \in N$ . Then we have  $Q(z) = (Q(z) - z) + z \in X^\perp \oplus_{\ell_1} N$  and hence,  $\|Q(z)\| = \|Q(z) - z\| + \|z\|$ . Since  $\|Q\| \leq 1$ , we see that  $\|Q(z) - z\| = 0$  and thus,  $z = Q(z) \in X^*$ . Therefore, we have  $N \subseteq X^*$ , so  $N = X^*$ . The proof is complete.  $\square$

n

**Proposition 12.7.** *The  $c_0$  space is an  $M$ -ideal of  $\ell_\infty$ .*

*Proof.* We first notice that for  $h \in (\ell_\infty)^*$  and  $\xi \in \ell_\infty$ , then  $\mathcal{R}e(h)(\xi) := \mathcal{R}e(h(\xi))$  can be viewed as a  $\mathbb{R}$ -linear functional on  $\ell_\infty$  and  $\|h\| = \|\mathcal{R}e(h)\|$ .

Using Proposition 12.6, it suffices to show that for  $g \in c_0^* = \ell_1$  and  $f \in c_0^\perp$ , we have  $\|g + f\|_{(\ell_\infty)^*} = \|g\|_{(\ell_\infty)^*} + \|f\|_{(\ell_\infty)^*}$ , where  $c_0^\perp := \{f \in (\ell_\infty)^* : f(c_0) \equiv 0\}$ . Let  $\varepsilon > 0$ . By considering the polar

decomposition, then there are elements  $\xi$  and  $\xi'$  in  $(\ell_\infty)_1$  of norm-one such that

$$\|f\| - \varepsilon < f(\xi) \quad \text{and} \quad \|g\| - \varepsilon < g(\xi') = \mathcal{R}e(g)(\xi') = \sum_{n=1}^{\infty} \mathcal{R}e(\xi'(n)g(n)).$$

Since  $g \in c_0^* = \ell_1$ , there is  $N$  such that  $\sum_{n>N} |g(n)| < \varepsilon$ . Now let  $\xi''$  be an element in  $\ell_\infty$  given by

$$\xi''(n) := \begin{cases} \xi'(n) & \text{if } n \leq N \\ \xi(n) & \text{if } n > N. \end{cases}$$

Then  $\|\xi''\|_\infty \leq 1$  and  $\xi'' - \xi \in c_{00}$ . Hence we have  $f(\xi) = f(\xi'')$  because  $f(c_0) \equiv 0$ .

On the other hand, since  $\sum_{n>N} |g(n)| < \varepsilon$ , we have

$$\|g\| - \varepsilon < \mathcal{R}e(g)(\xi') \leq \sum_{n=1}^N \mathcal{R}e(\xi'(n)g(n)) + \sum_{n>N} |g(n)| < \sum_{n=1}^N \mathcal{R}e(\xi'(n)g(n)) + \varepsilon.$$

Thus, we have

$$\begin{aligned} \mathcal{R}e(g)(\xi'') &= \sum_{n=1}^{\infty} \mathcal{R}e(\xi''(n)g(n)) \\ &\geq \sum_{n=1}^N \mathcal{R}e(\xi'(n)g(n)) - \left| \sum_{n>N} \xi(n)g(n) \right| \\ &\geq \|g\| - 2\varepsilon - \varepsilon. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \|g\| + \|f\| &= \|\mathcal{R}e(g)\| + \|f\| \\ &\leq \mathcal{R}e(g)(\xi'') + f(\xi'') + 4\varepsilon \\ &= \mathcal{R}e(g + f)(\xi'') + 4\varepsilon \\ &\leq \|\mathcal{R}e(g + f)\| + 4\varepsilon \\ &= \|g + f\| + 4\varepsilon. \end{aligned}$$

for all  $\varepsilon > 0$ . The proof is complete. □

### 13. APPENDIX: BASIC SEQUENCES

Throughout this section,  $X$  always denotes a Banach space.

An infinite sequence  $(x_n)$  in  $X$  is called a *basic sequence* if for each element  $x$  in  $X_0 := [x_1, x_2, \dots]$ , the closed linear span of  $\{x_1, x_2, \dots\}$ , then there is a unique sequence of scalars  $(a_n)$  such that  $x = \sum_{i=1}^{\infty} a_i x_i$ , that is  $(x_n)$  is a Schauder basis for its closed linear span  $[(x_i)]$ .

Put  $\psi_i$  the corresponding  $i$ -th coordinate function, i.e.,  $\psi_i(x) := a_i$  and  $Q_n : X_0 \rightarrow E_n := [x_1 \cdots x_n]$  the  $n$ -th canonical projection, i.e.,  $Q_n(\sum_{i=1}^{\infty} a_i x_i) := \sum_{i=1}^n a_i x_i$ .

**Theorem 13.1.** *Using the notations as above, for each element  $x \in X_0$ , put*

$$q(x) := \sup\{\|Q_n(x)\| : n = 1, 2, \dots\}.$$

*Then*

- (i)  $q$  is a Banach equivalent norm on  $X_0$ .
- (ii) Each coordinate projection  $Q_n$  and coordinate function  $\psi_n$  are bounded in the original norm-topology.

*Proof.* Since  $x = \lim_n Q_n x$  for all  $x \in X_0$ ,  $\|x\| = \lim_n \|Q_n x\|$ . This implies that the sequence  $(\|Q_n x\|)$  is bounded and hence,  $q$  is well defined, in addition,  $q(\cdot) \geq \|\cdot\|$  on  $X_0$ . From this, together with the Open Mapping Theorem, all assertions follows if we show that  $q$  is a Banach norm on  $X_0$ .

Let  $(x_n)$  be a Cauchy sequence in  $X_0$  with respect to the norm  $q$ . Clearly,  $(x_n)$  is also a Cauchy sequence in the  $\|\cdot\|$ -topology because  $q(\cdot) \geq \|\cdot\|$ . Let  $x = \lim_n x_n$  be the limit in  $X_0$  in the  $\|\cdot\|$ -topology. We are going to show that  $x$  is also the limit of  $(x_n)$  with respect to the  $q$ -topology. We first note that  $y_k := \lim_n Q_k x_n$  exists in  $X_0$  for all  $k = 1, 2, \dots$  by the definition of the norm  $q$ .

**Claim 1:**  $\|\cdot\|$ - $\lim_k y_k = x$ .

Let  $\varepsilon > 0$ . Then by the definition of the norm  $q$ , there is a positive integer  $N_1$  such that  $\|Q_k x_N - Q_k x_m\| < \varepsilon$  and  $\|x_N - x_m\| < \varepsilon$  for all  $m, N \geq N_1$  and for all  $k = 1, 2, \dots$ . This gives

$$\|x - Q_k x_m\| \leq \|x - x_{N_1}\| + \|x_{N_1} - Q_k x_{N_1}\| + \|Q_k x_{N_1} - Q_k x_m\| < 2\varepsilon + \|x_{N_1} - Q_k x_{N_1}\|$$

for all  $m \geq N_1$  and for all positive integers  $k$ . Thus, if we take  $m \rightarrow \infty$ , then we have

$$\|x - y_k\| \leq 2\varepsilon + \|x_{N_1} - Q_k x_{N_1}\| \rightarrow 2\varepsilon + 0 \quad \text{as } k \rightarrow \infty.$$

**Claim 2:**  $Q_k x = y_k$  for all  $k = 1, 2, \dots$ .

Fix a positive integer  $k_1$ . Note that  $Q_{k_1} y_k = y_{k_1}$  for all  $k \geq k_1$ . Indeed, since  $E_k$  and  $E_{k_1}$  are of finite dimension, the restrictions  $Q_{k_1}|_{E_k}$  and  $Q_k|_{E_{k_1}}$  both are continuous. This implies that

$$Q_{k_1} y_k = Q_{k_1}(\lim_n Q_k x_n) = \lim_n Q_{k_1} Q_k(x_n) = \lim_n Q_k Q_{k_1}(x_n) = Q_k(\lim_n Q_{k_1} x_n) = Q_k(y_{k_1}) = y_{k_1}$$

for all  $k \geq k_1$ . Hence, there is a sequence of scalars  $(\beta_n)$  so that  $y_k = \sum_{i=1}^k \beta_i x_i$  for all  $k = 1, 2, \dots$ . On the other hand, if we let  $x = \sum_{i=1}^{\infty} \alpha_i x_i$ , then by Claim 1 we have  $\lim_k (y_k - Q_k x) = 0$  and thus we have  $\sum_{i=1}^{\infty} (\beta_i - \alpha_i) x_i = 0$ . Therefore, we have  $\beta_i = \alpha_i$  for all  $i = 1, 2, \dots$ . The Claim 2 follows.

It remains to show that  $\lim_n q(x_n - x) = 0$ .

Let  $\eta > 0$ . Then there is a positive integer  $N$  so that  $\|Q_k x_n - Q_k x_m\| < \eta$  for all  $m, n \geq N$  and for all positive integers  $k$ . Taking  $m \rightarrow \infty$ , Claim 2 gives

$$\|Q_k x_n - Q_k x\| = \|Q_k x_n - y_k\| \leq \eta$$

for all  $n \geq N$  and for all positive integers  $k$ . Thus, we have  $q(x_n - x) \leq \eta$  for all  $n \geq N$ . The proof is complete.  $\square$

**Proposition 13.2.** *Using the notation as above, a sequence  $(x_i)$  is a basic sequence in  $X$  if and only if there is  $K > 0$  such that*

$$\sum_{1 \leq i \leq n} a_i x_i \leq K \sum_{1 \leq i \leq m} a_i x_i,$$

*whenever  $n \leq m$  and  $(a_i) \in c_{00}$ .*

## 14. GEOMETRY OF HILBERT SPACE I

From now on, all vector spaces are over the complex field. Recall that an *inner product* on a vector space  $V$  is a function  $(\cdot, \cdot) : V \times V \rightarrow \mathbb{C}$  which satisfies the following conditions.

- (i)  $(x, x) \geq 0$  for all  $x \in V$  and  $(x, x) = 0$  if and only if  $x = 0$ .
- (ii)  $\overline{(x, y)} = (y, x)$  for all  $x, y \in V$ .
- (iii)  $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$  for all  $x, y, z \in V$  and  $\alpha, \beta \in \mathbb{C}$ .

Consequently, for each  $x \in V$ , the map  $y \in V \mapsto (x, y) \in \mathbb{C}$  is conjugate linear by the conditions (ii) and (iii), i.e.,  $(x, \alpha y + \beta z) = \bar{\alpha}(x, y) + \bar{\beta}(x, z)$  for all  $y, z \in V$  and  $\alpha, \beta \in \mathbb{C}$ .

In addition, the inner product  $(\cdot, \cdot)$  will give a norm on  $V$  which is defined by

$$\|x\| := \sqrt{(x, x)}$$

for  $x \in V$ .

We first recall the following useful properties of an inner product space which can be found in the standard text books of linear algebras.

**Proposition 14.1.** *Let  $V$  be an inner product space. For all  $x, y \in V$ , we always have:*

- (i): **(Cauchy-Schwarz inequality):**  $|(x, y)| \leq \|x\|\|y\|$  Consequently, the inner product on  $V \times V$  is jointly continuous.
- (ii): **(Parallelogram law):**  $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$ .  
Furthermore, a norm  $\|\cdot\|$  on a vector space  $X$  is induced by an inner product if and only if it satisfies the Parallelogram law. In this case such inner product is given by the following:

$$\operatorname{Re}(x, y) = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) \quad \text{and} \quad \operatorname{Im}(x, y) = \frac{1}{4}(\|x + iy\|^2 - \|x - iy\|^2)$$

for all  $x, y \in X$ .

- (iii) **Gram-Schmidt process** Let  $\{x_1, x_2, \dots\}$  be a sequence of linearly independent vectors in an inner product space  $V$ . Put  $e_1 := x_1/\|x_1\|$ . Define  $e_n$  inductively on  $n$  by

$$e_{n+1} := \frac{x_n - \sum_{k=1}^n (x_k, e_k) e_k}{\|x_n - \sum_{k=1}^n (x_k, e_k) e_k\|}.$$

Then  $\{e_n : n = 1, 2, \dots\}$  forms an orthonormal system in  $V$ . Moreover, the linear span of  $x_1, \dots, x_n$  is equal to the linear span of  $e_1, \dots, e_n$  for all  $n = 1, 2, \dots$ .

Clearly, due to the Cauchy-Schwarz inequality, we have the following.

**Lemma 14.2.** *If  $X$  is an inner product space, then the inner product  $(\cdot, \cdot) : X \times X \rightarrow \mathbb{C}$  is jointly continuous, that is  $\lim_n (x_n, y_n) = (x, y)$  whenever  $(x_n)$  and  $(y_n)$  are convergent sequences in  $X$  with  $\lim_n x_n = x$  and  $\lim_n y_n = y$ .*

The following is one of the most important classes in mathematics.

**Definition 14.3.** *A Hilbert space is a Banach space whose norm is given by an inner product.*

**Example 14.4.** (1)  $\mathbb{C}^N$  is a Hilbert space under the usual inner product given by  $(w, z) := \sum_{k=1}^N w(k) \overline{z(k)}$  for  $w, z \in \mathbb{C}^N$ .

- (2) It follows from Proposition 14.1 immediately that  $\ell^2$  is a Hilbert space and  $\ell^p$  is not a Hilbert space for all  $p \in [1, \infty] \setminus \{2\}$ .

In the rest of this section,  $X$  always denotes a complex Hilbert space with an inner product  $(\cdot, \cdot)$ . Recall that two vectors  $x$  and  $y$  in an inner product space  $V$  are said to be *orthogonal* if  $(x, y) = 0$ .

**Proposition 14.5. (Bessel's inequality)** : Let  $\{e_1, \dots, e_N\}$  be an orthonormal set in an inner product space  $V$ , i.e.,  $(e_i, e_j) = 1$  if  $i = j$ , otherwise is equal to 0. Then for any  $x \in V$ , we have

$$\sum_{i=1}^N |(x, e_i)|^2 \leq \|x\|^2.$$

*Proof.* It can be obtained by the following equality immediately

$$\|x - \sum_{i=1}^N (x, e_i) e_i\|^2 = \|x\|^2 - \sum_{i=1}^N |(x, e_i)|^2.$$

□

**Corollary 14.6.** Let  $(e_i)_{i \in I}$  be an orthonormal set in a Hilbert space  $X$ . Then for any element  $x \in X$ , the set

$$I(x) := \{i \in I : (e_i, x) \neq 0\}$$

is countable.

*Proof.* Note that for each  $x \in V$ , we have

$$\{i \in I : (e_i, x) \neq 0\} = \bigcup_{n=1}^{\infty} \{i \in I : |(e_i, x)| \geq 1/n\}.$$

Then the Bessel's inequality implies that the set  $\{i \in I : |(e_i, x)| \geq 1/n\}$  must be finite for each  $n \geq 1$ . Thus the result follows. □

**Proposition 14.7.** Let  $(e_n)$  be a sequence of orthonormal vectors in a Hilbert space  $X$ . Then for any  $x \in V$ , the series  $\sum_{n=1}^{\infty} (x, e_n) e_n$  is convergent.

Moreover, if  $(e_{\sigma(n)})$  is a rearrangement of  $(e_n)$ , i.e.,  $\sigma : \{1, 2, \dots\} \rightarrow \{1, 2, \dots\}$  is a bijection. Then we have

$$\sum_{n=1}^{\infty} (x, e_n) e_n = \sum_{n=1}^{\infty} (x, e_{\sigma(n)}) e_{\sigma(n)}.$$

*Proof.* Since  $X$  is a Hilbert space, the convergence of the series  $\sum_{n=1}^{\infty} (x, e_n) e_n$  follows from the Bessel's inequality. In fact, if we put  $s_p := \sum_{n=1}^p (x, e_n) e_n$ , then we have

$$\|s_{p+k} - s_p\|^2 = \sum_{p+1 \leq n \leq p+k} |(x, e_n)|^2.$$

Now put  $y = \sum_{n=1}^{\infty} (x, e_n) e_n$  and  $z = \sum_{n=1}^{\infty} (x, e_{\sigma(n)}) e_{\sigma(n)}$ . Note that we have

$$\begin{aligned} (y, y - z) &= \lim_N \left( \sum_{n=1}^N (x, e_n) e_n, \sum_{n=1}^N (x, e_n) e_n - z \right) \\ &= \lim_N \sum_{n=1}^N |(x, e_n)|^2 - \lim_N \sum_{n=1}^N (x, e_n) \sum_{j=1}^{\infty} \overline{(x, e_{\sigma(j)})} (e_n, e_{\sigma(j)}) \\ &= \sum_{n=1}^{\infty} |(x, e_n)|^2 - \lim_N \sum_{n=1}^N (x, e_n) \overline{(x, e_n)} \quad (\because \text{for each } n, \text{ there is a unique } j \text{ such that } n = \sigma(j)) \\ &= 0. \end{aligned}$$

Similarly, we have  $(z, y - z) = 0$ . The result follows. □

A family of an orthonormal vectors, say  $\mathcal{B}$ , in  $X$  is said to be **complete** if it is maximal with respect to the set inclusion order, i.e., if  $\mathcal{C}$  is another family of orthonormal vectors with  $\mathcal{B} \subseteq \mathcal{C}$ , then  $\mathcal{B} = \mathcal{C}$ .

A complete orthonormal subset of  $X$  is also called an **orthonormal basis** of  $X$ .

**Proposition 14.8.** *Let  $\{e_i\}_{i \in I}$  be a family of orthonormal vectors in  $X$ . Then the followings are equivalent:*

- (i):  $\{e_i\}_{i \in I}$  is complete;
- (ii): if  $(x, e_i) = 0$  for all  $i \in I$ , then  $x = 0$ ;
- (iii): for any  $x \in X$ , we have  $x = \sum_{i \in I} (x, e_i) e_i$ ;
- (iv): for any  $x \in X$ , we have  $\|x\|^2 = \sum_{i \in I} |(x, e_i)|^2$ .

**Note :** there are only countable many  $(x, e_i) \neq 0$  by Corollary 14.6, so the sums in (iii) and (iv) are convergent by Proposition 14.7. In this case, the expression of each element  $x \in X$  in Part (iii) is unique.

**Proposition 14.9.** *Let  $X$  be a Hilbert space. Then*

- (i) :  $X$  possesses an orthonormal basis.
- (ii) : If  $\{e_i\}_{i \in I}$  and  $\{f_j\}_{j \in J}$  both are the orthonormal bases for  $X$ , then  $I$  and  $J$  have the same cardinality. In this case, the cardinality  $|I|$  of  $I$  is called the orthonormal dimension of  $X$ .

*Proof.* Part (i) follows from Zorn's Lemma.

For part (ii), if the cardinality  $|I|$  is finite, then the assertion is clear since  $|I| = \dim X$  (vector space dimension) in this case.

Now assume that  $|I|$  is infinite, for each  $e_i$ , put  $J_{e_i} := \{j \in J : (e_i, f_j) \neq 0\}$ . Note that since  $\{e_i\}_{i \in I}$  is maximal, Proposition 14.8 implies that we have

$$J = \bigcup_{i \in I} J_{e_i}.$$

Note that  $J_{e_i}$  is countable for each  $e_i$  by using Proposition 14.6. On the other hand, we have  $|\mathbb{N}| \leq |I|$  because  $|I|$  is infinite and thus  $|\mathbb{N} \times I| = |I|$ . Then we have

$$|J| \leq \sum_{i \in I} |J_{e_i}| = \sum_{i \in I} |\mathbb{N}| = |\mathbb{N} \times I| = |I|.$$

Similarly, we also have  $|I| \leq |J|$ . □

**Remark 14.10.** *Recall that a vector space dimension of  $X$  is defined by the cardinality of a maximal linearly independent set in  $X$ .*

*Note that if  $X$  is finite dimensional, then the orthonormal dimension is the same as the vector space dimension.*

*In addition, the vector space dimension is larger than the orthonormal dimension in general since every orthogonal set must be linearly independent.*

Two Hilbert spaces  $X$  and  $Y$  are said to be *isomorphic* if there is linear isomorphism  $U$  from  $X$  onto  $Y$  such that  $(Ux, Ux') = (x, x')$  for all  $x, x' \in X$ . In this case  $U$  is called a *unitary operator*.

**Theorem 14.11.** *Two Hilbert spaces are isomorphic if and only if they have the same orthonormal dimension.*

*Proof.* The converse part ( $\Leftarrow$ ) is clear.

Now for the ( $\Rightarrow$ ) part, let  $X$  and  $Y$  be isomorphic Hilbert spaces. Let  $U : X \rightarrow Y$  be a unitary. Note that if  $\{e_i\}_{i \in I}$  is an orthonormal basis of  $X$ , then  $\{Ue_i\}_{i \in I}$  is also an orthonormal basis of  $Y$ . Thus the necessary part follows immediately from Proposition 14.9. □

**Corollary 14.12.** *Every separable Hilbert space is isomorphic to  $\ell^2$  or  $\mathbb{C}^n$  for some  $n$ .*

*Proof.* Let  $X$  be a separable Hilbert space.

If  $\dim X < \infty$ , then it is clear that  $X$  is isomorphic to  $\mathbb{C}^n$  for  $n = \dim X$ .

Now suppose that  $\dim X = \infty$  and its orthonormal dimension is larger than  $|\mathbb{N}|$ , i.e.,  $X$  has an orthonormal basis  $\{f_i\}_{i \in I}$  with  $|I| > |\mathbb{N}|$ . Note that since  $\|f_i - f_j\| = \sqrt{2}$  for all  $i, j \in I$  with  $i \neq j$ . This implies that  $B(f_i, 1/4) \cap B(f_j, 1/4) = \emptyset$  for  $i \neq j$ .

On the other hand, if we let  $D$  be a countable dense subset of  $X$ , then  $B(f_i, 1/4) \cap D \neq \emptyset$  for all  $i \in I$ . Thus for each  $i \in I$ , we can pick up an element  $x_i \in D \cap B(f_i, 1/4)$ . Therefore, one can define an injection from  $I$  into  $D$ . It is absurd to the countability of  $D$ .  $\square$

**Example 14.13.** The followings are important classes of Hilbert spaces.

- (i)  $\mathbb{C}^n$  is a  $n$ -dimensional Hilbert space. In this case, the inner product is given by  $(z, w) := \sum_{k=1}^n z_k \bar{w}_k$  for  $z = (z_1, \dots, z_n)$  and  $(w_1, \dots, w_n)$  in  $\mathbb{C}^n$ .  
The natural basis  $\{e_1, \dots, e_n\}$  forms an orthonormal basis for  $\mathbb{C}^n$ .

- (ii)  $\ell^2$  is a separable Hilbert space of infinite dimension whose inner product is given by  $(x, y) := \sum_{n=1}^{\infty} x(n) \overline{y(n)}$  for  $x, y \in \ell^2$ .

If we put  $e_n(n) = 1$  and  $e_n(k) = 0$  for  $k \neq n$ , then  $\{e_n\}$  is an orthonormal basis for  $\ell^2$ .

- (iii) Let  $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ . For each  $f \in C(\mathbb{T})$  (the space of all complex-valued continuous functions defined on  $\mathbb{T}$ ), the integral of  $f$  is defined by

$$\int_{\mathbb{T}} f(z) dz := \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) dt = \frac{1}{2\pi} \int_0^{2\pi} \operatorname{Re} f(e^{it}) dt + \frac{i}{2\pi} \int_0^{2\pi} \operatorname{Im} f(e^{it}) dt.$$

An inner product on  $C(\mathbb{T})$  is given by

$$(f, g) := \int_{\mathbb{T}} f(z) \overline{g(z)} dz$$

for each  $f, g \in C(\mathbb{T})$ . We write  $\|\cdot\|_2$  for the norm induced by this inner product.

The Hilbert space  $L^2(\mathbb{T})$  is defined by the completion of  $C(\mathbb{T})$  under the norm  $\|\cdot\|_2$ .

Now for each  $n \in \mathbb{Z}$ , put  $f_n(z) = z^n$ . We claim that  $\{f_n : n = 0, \pm 1, \pm 2, \dots\}$  is an orthonormal basis for  $L^2(\mathbb{T})$ .

In fact, by using the Euler Formula:  $e^{i\theta} = \cos \theta + i \sin \theta$  for  $\theta \in \mathbb{R}$ , we see that the family  $\{f_n : n \in \mathbb{Z}\}$  is orthonormal.

It remains to show that the family  $\{f_n\}$  is maximal. By Proposition 14.8, it needs to show that if  $(g, f_n) = 0$  for all  $n \in \mathbb{Z}$ , then  $g = 0$  in  $L^2(\mathbb{T})$ . for showing this, we have to make use the known fact that every element in  $L^2(\mathbb{T})$  can be approximated by the polynomial functions of  $z$  and  $\bar{z}$  on  $\mathbb{T}$  in  $\|\cdot\|_2$ -norm due to the *Stone-Weierstrass Theorem*:

*For a compact metric space  $E$ , suppose that a complex subalgebra  $A$  of  $C(E)$  satisfies the conditions: (i): the conjugate  $\bar{f} \in A$  whenever  $f \in A$ , (ii): for every pair  $z, z' \in E$ , there is  $f \in A$  such that  $f(z) \neq f(z')$  and (iii):  $A$  contains the constant one function. Then  $A$  is dense in  $C(E)$  with respect to the sup-norm.*

Thus, the algebra of all polynomial functions of  $z$  and  $\bar{z}$  on  $\mathbb{T}$  is dense in  $C(\mathbb{T})$ . From this we can find a sequence of polynomials  $(p_n(z, \bar{z}))$  such that  $\|g - p_n\|_2 \rightarrow 0$  as  $n \rightarrow \infty$ . Since  $(g, f_n) = 0$  for all  $n$ , we see that  $(g, p_n) = 0$  for all  $n$ . Therefore, we have

$$\|g\|_2^2 = \lim_n (g, p_n) = 0.$$

The proof is complete.



**Remark 14.14.** In view of Example 14.13(iii) above,  $L^2(\mathbb{T})$  can be identified as the space

$$L^2[0, 2\pi] := \{f : [0, 2\pi] \rightarrow \mathbb{C} : f \text{ is a Lebesgue measure function so that } \int_0^{2\pi} |f(x)|^2 dx \text{ exists}\}.$$

In this case, the inner product in  $L^2[0, 2\pi]$  is given by  $(f, g) := \frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{g(t)} dt$ . Now for each  $x \in [0, 2\pi]$  and  $f \in L^2[0, 2\pi]$ , put  $e_k(x) := e^{ikx}$  and

$$(14.1) \quad S_N(f, x) := \sum_{k=-N}^N (f, e_k) e_k(x) = \frac{1}{2\pi} \sum_{k=-N}^N \left( \int_0^{2\pi} f(t) e^{-ik} dt \right) e^{ikx}.$$

Then the series  $\sum_{k=-\infty}^{\infty} (f, e_k) e_k(x)$  (need not be convergent) is the usual notation of the Fourier series of  $f$  at  $x$ .

In the previous result, we have seen that the series  $\lim_{N \rightarrow \infty} \|f - \sum_{k=-N}^N (f, e_k) e_k\|_{L^2} \rightarrow 0$

It is naturally raised the question: for  $f \in C[0, 2\pi]$ , do we have  $f(x) = \lim_{N \rightarrow \infty} S_N(f, x)$  for all  $x \in [0, 2\pi]$ , that is,  $f$  can be written as its Fourier series? The answer is negative. Indeed, we have the following stronger result.

**Theorem 14.15.** *For each element  $x_0 \in [0, 2\pi]$ , there is a continuous function  $f$  on  $[0, 2\pi]$  such that its Fourier series  $\sum_{k=-\infty}^{\infty} (f, e_k) e^{ikx_0}$  at  $x_0$  is divergent.*

*Proof.* We give an outline argument in here. The details of the proof is referred to the Katznelson's classic book [9].

We keep the notation as above. We fix a point  $x_0 \in [0, 2\pi]$ . For  $f \in C[0, 2\pi]$ , Using the Eq 14.1 above, we have

$$(14.2) \quad \begin{aligned} \varphi_N(f) &:= S_N(f, x_0) = \frac{1}{2\pi} \sum_{k=-N}^N \left( \int_0^{2\pi} f(t) e^{-ik} dt \right) e^{ikx_0} \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(t) \left( \sum_{k=-N}^N e^{ik(x_0-t)} \right) dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(t) D_N(x_0 - t) dt, \end{aligned}$$

where  $D_N(s) := \sum_{k=-N}^N e^{iks}$  for  $s \in \mathbb{R}$ . Then  $\varphi_N \in C[0, 2\pi]^*$ . It is a fact that we have (see [9, Chapter 2. §2])

$$\|\varphi_N\| = \frac{1}{2\pi} \int_0^{2\pi} |D_N(x_0 - t)| dt \rightarrow \infty \quad \text{as } N \rightarrow \infty.$$

From this we have  $\sup\{\|\varphi_N\| : N = 1, 2, \dots\} = \infty$ . This implies that the following set is dense in  $C[0, 2\pi]$ .

$$D := \{f \in C[0, 2\pi] : \sup_N |\varphi_N(f)| = \infty\}.$$

To see this, suppose not. Then there is an element  $f_0 \in C[0, 2\pi]$  and  $r > 0$  such that  $\sup_N |\varphi_N(f)| < \infty$  for all  $f \in B(f_0, r)$ . By considering  $B(0, 1) = \frac{1}{r}(B(f_0, r) - f_0)$ , then we have  $\sup_N |\varphi_N(g)| < \infty$  for all  $g \in B(0, 1)$  and so,  $\sup_N |\varphi_N(g)| < \infty$  for all  $g \in C[0, 2\pi]$ . The Uniform Boundedness Theorem implies that  $\sup_N \|\varphi_N\| < \infty$  that leads to a contradiction.

On the other hand, it is clear that if the series  $\sum_{k=-\infty}^{\infty} (f, e_k) e^{ikx_0}$  is convergent, then  $\sup_N |S_N(f, x_0)| < \infty$ . Therefore, the Fourier series of each element in  $D$  is divergent at  $x_0$  as desired. We finish the proof.  $\square$

The above proof is an elegant application of functional analysis.

*Students should be made to think, to doubt, to communicate, to question, to learn from their mistakes, and most importantly have fun in their learning.*  
*~ Richard Feynman ~ one of the greatest physicists in 20th century.*

## 15. GEOMETRY OF HILBERT SPACE II

In this section, let  $X$  always denote a complex Hilbert space.

**Proposition 15.1.** *If  $D$  is a closed convex subset of  $X$ , then there is a unique element  $z \in D$  such that*

$$\|z\| = \inf\{\|x\| : x \in D\}.$$

*Consequently, for any element  $u \in X$ , there is a unique element  $w \in D$  such that*

$$\|u - w\| = d(u, D) := \inf\{\|u - x\| : x \in D\}.$$

*Proof.* We first claim the existence of such  $z$ .

Let  $d := \inf\{\|x\| : x \in D\}$ . Then there is a sequence  $(x_n)$  in  $D$  such that  $\|x_n\| \rightarrow d$ . Note that  $(x_n)$  is a Cauchy sequence. In fact, the Parallelogram Law implies that

$$\left\| \frac{x_m - x_n}{2} \right\|^2 = \frac{1}{2} \|x_m\|^2 + \frac{1}{2} \|x_n\|^2 - \left\| \frac{x_m + x_n}{2} \right\|^2 \leq \frac{1}{2} \|x_m\|^2 + \frac{1}{2} \|x_n\|^2 - d^2 \rightarrow 0$$

as  $m, n \rightarrow \infty$ , where the last inequality holds because  $D$  is convex and hence  $\frac{1}{2}(x_m + x_n) \in D$ . Let  $z := \lim_n x_n$ . Then  $\|z\| = d$  and  $z \in D$  because  $D$  is closed.

For the uniqueness, let  $z, z' \in D$  such that  $\|z\| = \|z'\| = d$ . Thanks to the Parallelogram Law again, we have

$$\left\| \frac{z - z'}{2} \right\|^2 = \frac{1}{2} \|z\|^2 + \frac{1}{2} \|z'\|^2 - \left\| \frac{z + z'}{2} \right\|^2 \leq \frac{1}{2} \|z\|^2 + \frac{1}{2} \|z'\|^2 - d^2 = 0.$$

Therefore  $z = z'$ .

The last assertion follows by considering the closed convex set  $u - D := \{u - x : x \in D\}$  immediately.  $\square$

**Remark 15.2.** *Using the notation given as in Proposition 15.1, we have a well defined function  $r : X \rightarrow X$  given by  $x \in X \mapsto r(x) \in D$  such that  $\|x - r(x)\| = \text{dist}(x, D)$ . Clearly, we have  $r(x) = x$  whenever  $x \in D$ . Moreover, we have the following assertion which are shown in [5].*

**Proposition 15.3.** *Using the notation as in Remark 15.2, the map  $r : X \rightarrow X$  is a contraction, hence, the map  $r$  is a Lipschitz retraction of  $D$  in  $X$ .*

*Proof.* We first claim that we have  $\text{Re}(x - r(z), r(z) - z) \geq 0$  for all  $x \in D$  and  $z \in X$ . In fact, let  $z \in X$  and  $x \in D$ . Then by the definition, for all  $t \in [0, 1]$  we have

$$\begin{aligned} \|r(z) - z\|^2 &\leq \|z - tx - (1 - t)r(z)\|^2 \\ &= \|z - r(z) - t(x - r(z))\|^2 \\ &= \|z - r(z)\|^2 + t^2 \|x - r(z)\|^2 + 2t \text{Re}(x - r(z), r(z) - z). \end{aligned}$$

This gives  $t^2\|x - r(z)\|^2 + 2t\operatorname{Re}(x - r(z), r(z) - z) \geq 0$  for all  $0 \leq t \leq 1$ . This implies that  $\operatorname{Re}(x - r(z), r(z) - z) \geq 0$  for all  $x \in D$  and  $z \in X$ . From this, for  $a, b \in X$  we have  $\operatorname{Re}(r(b) - r(a), r(a) - a) \geq 0$  and  $\operatorname{Re}(r(a) - r(b), r(b) - b) \geq 0$ , so we have  $\operatorname{Re}(r(b) - r(a), r(a) - a) + \operatorname{Re}(r(b) - r(a), b - r(b)) \geq 0$ . Thus, we have

$$\begin{aligned}\|r(b) - r(a)\|^2 &= \operatorname{Re}(r(b) - r(a), r(b) - r(a)) \\ &\leq \operatorname{Re}(r(b) - r(a), b - a) \\ &\leq |(r(b) - r(a), b - a)| \\ &\leq \|r(b) - r(a)\| \|b - a\|.\end{aligned}$$

The proof is complete.  $\square$

**Proposition 15.4.** *Suppose that  $M$  is a closed subspace. Let  $u \in X$  and  $w \in M$ . Then the followings are equivalent:*

- (i):  $\|u - w\| = d(u, M)$ ;
- (ii):  $u - w \perp M$ , i.e.,  $(u - w, x) = 0$  for all  $x \in M$ .

Consequently, for each element  $u \in X$ , there is a unique element  $w \in M$  such that  $u - w \perp M$ .

*Proof.* Let  $d := d(u, M)$ .

For proving (i)  $\Rightarrow$  (ii), fix an element  $x \in M$ . Then for any  $t > 0$ , note that since  $w + tx \in M$ , we have

$$d^2 \leq \|u - w - tx\|^2 = \|u - w\|^2 + \|tx\|^2 - 2\operatorname{Re}(u - w, tx) = d^2 + \|tx\|^2 - 2\operatorname{Re}(u - w, tx).$$

This implies that

$$(15.1) \quad 2\operatorname{Re}(u - w, x) \leq t\|x\|^2$$

for all  $t > 0$  and for all  $x \in M$ . Thus by considering  $-x$  in Eq.15.1, we obtain

$$2|\operatorname{Re}(u - w, x)| \leq t\|x\|^2.$$

for all  $t > 0$ . This implies that  $\operatorname{Re}(u - w, x) = 0$  for all  $x \in M$ . Similarly, putting  $\pm ix$  into Eq.15.1, we have  $\operatorname{Im}(u - w, x) = 0$ . Thus (ii) follows.

For (ii)  $\Rightarrow$  (i), we need to show that  $\|u - w\|^2 \leq \|u - x\|^2$  for all  $x \in M$ . Note that since  $u - w \perp M$  and  $w \in M$ , we have  $u - w \perp w - x$  for all  $x \in M$ . This gives

$$\|u - x\|^2 = \|(u - w) + (w - x)\|^2 = \|u - w\|^2 + \|w - x\|^2 \geq \|u - w\|^2.$$

Part (i) follows.

The last statement is obtained immediately by Proposition 15.1.  $\square$

**Theorem 15.5.** *Let  $M$  be a closed subspace. Put*

$$M^\perp := \{x \in X : x \perp M\}.$$

*Then  $M^\perp$  is a closed subspace and we have  $X = M \oplus M^\perp$ . Consequently, for  $x \in X$  if  $x = u \oplus v$  for  $u \in M$  and  $v \in M^\perp$ , then  $\operatorname{dist}(x, M) = \|x - u\|$ .*

*In this case,  $M^\perp$  is called the orthogonal complement of  $M$ .*

*Proof.* Clearly,  $M^\perp$  is a closed subspace and  $M \cap M^\perp = (0)$ . We need to show  $X = M + M^\perp$ .

Let  $u \in X$ . Then by Proposition 15.4, we can find an element  $w \in M$  such that  $u - w \perp M$ . Thus  $u - w \in M^\perp$  and  $u = w + (u - w)$ .

The last assertion follows immediately from Proposition 15.4. The proof is complete.  $\square$

**Corollary 15.6.** *Let  $M$  be a closed subspace of  $X$ . Then  $M \subsetneq X$  if and only if there is a non-zero element  $z \in X$  such that  $z \perp M$ .*

*Proof.* It is clear from Theorem 15.5.  $\square$

**Corollary 15.7.** *If  $M$  is a closed subspace of  $X$ , then  $M^{\perp\perp} = M$ .*

*Proof.* Clearly, we have  $M \subseteq M^{\perp\perp}$  by the definition of  $M^{\perp\perp}$ . Then  $M$  can be viewed as a closed subspace of the Hilbert space  $M^{\perp\perp}$ . Thus, if  $M \subsetneq M^{\perp\perp}$ , then there exists a non-zero element  $z \in M^{\perp\perp}$  so that  $z \perp M$  by Corollary 15.6 and hence,  $z \in M^\perp$ . This implies that  $z \perp z$  and hence,  $z = 0$  which leads to a contradiction.  $\square$

**Theorem 15.8. Riesz Representation Theorem :** *For each  $f \in X^*$ , then there is a unique element  $v_f \in X$  such that*

$$f(x) = (x, v_f)$$

*for all  $x \in X$  and we have  $\|f\| = \|v_f\|$ .*

*Furthermore, if  $(e_i)_{i \in I}$  is an orthonormal basis of  $X$ , then  $v_f = \sum_i \overline{f(e_i)} e_i$ .*

*Proof.* We first prove the uniqueness of  $v_f$ . If  $z \in X$  also satisfies the condition:  $f(x) = (x, z)$  for all  $x \in X$ . This implies that  $(x, z - v_f) = 0$  for all  $x \in X$ . Thus  $z - v_f = 0$ .

Now for proving the existence of  $v_f$ , it suffices to show the case  $\|f\| = 1$ . Then  $\ker f$  is a closed proper subspace. Then by the orthogonal decomposition again, we have

$$X = \ker f \oplus (\ker f)^\perp.$$

Since  $f \neq 0$ , we have  $(\ker f)^\perp$  is linear isomorphic to  $\mathbb{C}$ . Note that the restriction of  $f$  on  $(\ker f)^\perp$  is of norm one. Hence there is an element  $v_f \in (\ker f)^\perp$  with  $\|v_f\| = 1$  such that  $f(v_f) = \|f|_{(\ker f)^\perp}\| = 1$  and  $(\ker f)^\perp = \mathbb{C}v_f$ . Thus for each element  $x \in X$ , we have  $x = z + \alpha v_f$  for some  $z \in \ker f$  and  $\alpha \in \mathbb{C}$ . Then  $f(x) = \alpha f(v_f) = \alpha = (x, v_f)$  for all  $x \in X$ .

Concerning about the last assertion, if we put  $v_f = \sum_{i \in I} \alpha_i e_i$ , then  $f(e_j) = (e_j, v_f) = \overline{\alpha_j}$  for all  $j \in I$ .  $\square$

**Example 15.9.** *Consider the Hilbert space  $H := L^2(\mathbb{T})$  (see Example 14.13). Define  $\varphi \in H^*$  by  $\varphi(f) := \int_{\mathbb{T}} f(z) dz$ . Using Proposition 15.4, for each element  $g \in H$ , there is an element  $h \in \ker \varphi$  such that  $\|g - h\| = \text{dist}(g, \ker \varphi)$ . Then  $h = g - (\int g dz) \mathbf{1}$  where  $\mathbf{1}$  denotes the constant-one function on  $\mathbb{T}$ . In fact, consider the orthogonal decomposition  $H = \ker \varphi \oplus (\ker \varphi)^\perp$ . Note that  $\varphi(g) = (g, \mathbf{1})$  for all  $g \in H$ . Thus, for each  $g \in H$ , we have  $g = h \oplus \alpha \mathbf{1}$ . From this, we see that  $\alpha = (g, \mathbf{1})$ . Thus,  $h = g - (\int g dz) \mathbf{1}$ .*

**Corollary 15.10.** *Using the notations as in Theorem 15.8, define the map*

$$(15.2) \quad \Phi : f \in X^* \mapsto v_f \in X, \text{ i.e., } f(y) = (y, \Phi(f))$$

*for all  $y \in X$  and  $f \in X^*$ .*

*Moreover, if we define  $(f, g)_{X^*} := (v_g, v_f)_X$  for  $f, g \in X^*$ , then  $(X^*, (\cdot, \cdot)_{X^*})$  becomes a Hilbert space, and  $\Phi$  is an anti-unitary operator from  $X^*$  onto  $X$ , i.e.,  $\Phi$  satisfies the conditions:*

$$\Phi(\alpha f + \beta g) = \overline{\alpha} \Phi(f) + \overline{\beta} \Phi(g) \quad \text{and} \quad (\Phi f, \Phi g)_X = (g, f)_{X^*}$$

*for all  $f, g \in X^*$  and  $\alpha, \beta \in \mathbb{C}$ .*

*Furthermore, if we define  $J : x \in X \mapsto f_x \in X^*$ , where  $f_x(y) := (y, x)$ , then  $J$  is the inverse of  $\Phi$ , and hence,  $J$  is an isometric conjugate linear isomorphism.*

*Proof.* The result follows immediately from the observation that  $v_{f+g} = v_f + v_g$  and  $v_{\alpha f} = \overline{\alpha} v_f$  for all  $f \in X^*$  and  $\alpha \in \mathbb{C}$ .

The last assertion is clearly obtained by the Eq.15.2 above.  $\square$

**Corollary 15.11.** *Every Hilbert space is reflexive.*

*Proof.* Using the notations as in the Riesz Representation Theorem 15.8, let  $X$  be a Hilbert space. and  $Q : X \rightarrow X^{**}$  the canonical isometry. Let  $\psi \in X^{**}$ . To apply the Riesz Theorem on the dual space  $X^*$ , there exists an element  $x_0^* \in X^*$  such that

$$\psi(f) = (f, x_0^*)_{X^*}$$

for all  $f \in X^*$ . By using Corollary 15.10, there is an element  $x_0 \in X$  such that  $x_0 = v_{x_0^*}$  and thus, we have

$$\psi(f) = (f, x_0^*)_{X^*} = (x_0, v_f)_X = f(x_0)$$

for all  $f \in X^*$ . Therefore,  $\psi = Q(x_0)$  and so,  $X$  is reflexive.

The proof is complete.  $\square$

**Theorem 15.12.** *Every bounded sequence in a Hilbert space has a weakly convergent subsequence.*

*Proof.* Let  $(x_n)$  be a bounded sequence in a Hilbert space  $X$  and  $M$  be the closed subspace of  $X$  spanned by  $\{x_m : m = 1, 2, \dots\}$ . Then  $M$  is a separable Hilbert space.

**Method I :** Define a map by  $j_M : x \in M \mapsto j_M(x) := (\cdot, x) \in M^*$ . Then  $(j_M(x_n))$  is a bounded sequence in  $M^*$ . By Banach's result, Proposition 6.9,  $(j_M(x_n))$  has a  $w^*$ -convergent subsequence  $(j_M(x_{n_k}))$ . Put  $j_M(x_{n_k}) \xrightarrow{w^*} f \in M^*$ , i.e.,  $j_M(x_{n_k})(z) \rightarrow f(z)$  for all  $z \in M$ . The Riesz Representation will assure that there is a unique element  $m \in M$  such that  $j_M(m) = f$ . Thus we have  $(z, x_{n_k}) \rightarrow (z, m)$  for all  $z \in M$ . In particular, if we consider the orthogonal decomposition  $X = M \oplus M^\perp$ , then  $(x, x_{n_k}) \rightarrow (x, m)$  for all  $x \in X$  and thus  $(x_{n_k}, x) \rightarrow (m, x)$  for all  $x \in X$ . Then  $x_{n_k} \rightarrow m$  weakly in  $X$  by using the Riesz Representation Theorem again.

**Method II :** We first note that since  $M$  is a separable Hilbert space, the second dual  $M^{**}$  is also separable by the reflexivity of  $M$ . Thus, the dual space  $M^*$  is separable (see Proposition 4.11). Let  $Q : M \rightarrow M^{**}$  be the natural canonical mapping. To apply the Banach's result Proposition 6.9 for  $X^*$ , then  $Q(x_n)$  has a  $w^*$ -convergent subsequence, says  $Q(x_{n_k})$ . This gives an element  $m \in M$  such that  $Q(m) = w^*\text{-}\lim_k Q(x_{n_k})$  because  $M$  is reflexive. Thus, we have  $f(x_{n_k}) = Q(x_{n_k})(f) \rightarrow Q(m)(f) = f(m)$  for all  $f \in M^*$ . Using the same argument as in **Method I** again,  $x_{n_k}$  weakly converges to  $m$ .  $\square$

**Remark 15.13.** Recall the well known James's Theorem that a Banach space  $X$  is reflexive if and only if every bounded sequence in  $X$  has a weakly convergent subsequence. (see Appendix 7 for the proof of a separable space case). Hence, Theorem 15.12 can be obtained by the James's Theorem directly. However, the Riesz Representation Theorem gives a simple proof for the Hilbert spaces case.

## 16. OPERATORS ON A HILBERT SPACE

Throughout this section, all spaces are complex Hilbert spaces. Let  $B(X, Y)$  denote the space of all bounded linear operators from  $X$  into  $Y$ . If  $X = Y$ , we write  $B(X)$  for  $B(X, X)$ . Let  $T \in B(X, Y)$ . We make use the following simple observation later.

$$(16.1) \quad (Tx, y) = 0 \text{ for all } x \in X; y \in Y \quad \text{if and only if} \quad T = 0.$$

Therefore, the elements in  $B(X, Y)$  are uniquely determined by the Eq.16.1, i.e.,  $T = S$  in  $B(X, Y)$  if and only if  $(Tx, y) = (Sx, y)$  for all  $x \in X$  and  $y \in Y$ .

**Remark 16.1.** For Hilbert spaces  $H_1$  and  $H_2$ , we consider their direct sum  $H := H_1 \oplus H_2$ . If we define the inner product on  $H$  by

$$(x_1 \oplus x_2, y_1 \oplus y_2) := (x_1, y_1)_{H_1} + (x_2, y_2)_{H_2}$$

for  $x_1 \oplus x_2$  and  $y_1 \oplus y_2$  in  $H$ , then  $H$  becomes a Hilbert space. Now for each  $T \in B(H_1, H_2)$ , we can define an element  $\tilde{T} \in B(H)$  by  $\tilde{T}(x_1 \oplus x_2) := 0 \oplus Tx_1$ . Therefore, the space  $B(H_1, H_2)$  can be viewed as a closed subspace of  $B(H)$ . Thus, we can consider the case of  $H_1 = H_2$  for studying the space  $B(H_1, H_2)$ .

**Proposition 16.2.** *Let  $T : X \rightarrow X$  be a linear operator. Then we have*

- (i):  $T = 0$  if and only if  $(Tx, x) = 0$  for all  $x \in X$ . Consequently, for  $T, S \in B(X)$ ,  $T = S$  if and only if  $(Tx, x) = (Sx, x)$  for all  $x \in X$ .
- (ii):  $T$  is bounded if and only if  $\sup\{|(Tx, y)| : x, y \in X \text{ with } \|x\| = \|y\| = 1\}$  is finite. In this case, we have  $\|T\| = \sup\{|(Tx, y)| : x, y \in X \text{ with } \|x\| = \|y\| = 1\}$ .

*Proof.* Clearly, the necessary part holds in Part (i). We want to show the sufficient part in Part (i). We assume that  $(Tx, x) = 0$  for all  $x \in X$ . Then we have

$$0 = (T(x + iy), x + iy) = (Tx, x) + i(Ty, x) - i(Tx, y) + (Ty, iy) = i(Ty, x) - i(Tx, y).$$

Thus, we have  $(Ty, x) - (Tx, y) = 0$  for all  $x, y \in X$ . In particular, if we replace  $y$  by  $iy$  in the equation, then we get  $i(Ty, x) - \bar{i}(Tx, y) = 0$  and hence we have  $(Ty, x) + (Tx, y) = 0$ . Therefore we have  $(Tx, y) = 0$ .

For showing part (ii), let  $\alpha = \sup\{|(Tx, y)| : x, y \in X \text{ with } \|x\| = \|y\| = 1\}$ . It suffices to show  $\|T\| = \alpha$ . Clearly, we have  $\|T\| \geq \alpha$ . We need to show  $\|T\| \leq \alpha$ .

In fact, let  $x \in X$  with  $\|x\| = 1$ . If  $Tx \neq 0$ , then we take  $y = Tx/\|Tx\|$ . Thus, we have  $\|Tx\| = |(Tx, y)| \leq \alpha$ , and so  $\|T\| \leq \alpha$ . The proof is complete.  $\square$

**Proposition 16.3.** *Let  $T \in B(X)$ . Then there is a unique element  $T^*$  in  $B(X)$  such that*

$$(16.2) \quad (Tx, y) = (x, T^*y)$$

*In this case,  $T^*$  is called the adjoint operator of  $T$ .*

*Proof.* First, we show the uniqueness. Suppose that there are  $S_1, S_2$  in  $B(X)$  which satisfy the Eq.16.2. Then  $(x, S_1y) = (x, S_2y)$  for all  $x, y \in X$ . Eq.16.1 implies that  $S_1 = S_2$ .

Finally, we prove the existence. Note that if we fix an element  $y \in X$ , define the map  $f_y(x) := (Tx, y)$  for all  $x \in X$ . Then  $f_y \in X^*$ . By applying the Riesz Representation Theorem, there is a unique element  $y^* \in X$  such that  $(Tx, y) = (x, y^*)$  for all  $x \in X$  and  $\|f_y\| = \|y^*\|$ . In addition, we have

$$|f_y(x)| = |(Tx, y)| \leq \|T\|\|x\|\|y\|$$

for all  $x, y \in X$  and thus  $\|f_y\| \leq \|T\|\|y\|$ . If we put  $T^*(y) := y^*$ , then  $T^*$  satisfies the Eq.16.2. Moreover, we have  $\|T^*y\| = \|y^*\| = \|f_y\| \leq \|T\|\|y\|$  for all  $y \in X$ . Thus, we have  $T^* \in B(X)$  and  $\|T^*\| \leq \|T\|$ . Hence the operator  $T^*$  is as desired.  $\square$

**Remark 16.4.** Let  $S, T : X \rightarrow X$  be linear operators (without assuming to be bounded). If they satisfy the Eq.16.2 above, i.e.,

$$(Tx, y) = (x, Sy)$$

for all  $x, y \in X$ . Using the Closed Graph Theorem, we can show that  $S$  and  $T$  both are automatically bounded.

In fact, let  $(x_n)$  be a sequence in  $X$  such that  $\lim x_n = x$  and  $\lim Sx_n = y$  for some  $x, y \in X$ . Now for any  $z \in X$ , we have

$$(z, Sx) = (Tz, x) = \lim(Tz, x_n) = \lim(z, Sx_n) = (z, y).$$

Thus  $Sx = y$  and hence  $S$  is bounded by the Closed Graph Theorem.

Similarly, we can also see that  $T$  is bounded.

**Remark 16.5.** Let  $T \in B(X)$ . Let  $T^t : X^* \rightarrow X^*$  be the transpose of  $T$  which is defined by  $T^t(f) := f \circ T \in X^*$  for  $f \in X^*$  (see Proposition 4.13). Then we have the following commutative diagram (**Check!**)

$$\begin{array}{ccc} X & \xrightarrow{T^*} & X \\ J_X \downarrow & & \downarrow J_X \\ X^* & \xrightarrow{T^t} & X^* \end{array}$$

where  $J_X : X \rightarrow X^*$  is the anti-unitary given by the Riesz Representation Theorem (see Corollary 15.10).

**Proposition 16.6.** Let  $T, S \in B(X)$ . Then we have

(i):  $T^* \in B(X)$  and  $\|T^*\| = \|T\|$ .

(ii): The map  $T \in B(X) \mapsto T^* \in B(X)$  is an isometric conjugate anti-isomorphism, i.e.,

$$(\alpha T + \beta S)^* = \bar{\alpha} T^* + \bar{\beta} S^* \quad \text{for all } \alpha, \beta \in \mathbb{C}; \quad \text{and} \quad (TS)^* = S^* T^*.$$

(iii):  $\|T^* T\| = \|T\|^2$ .

*Proof.* For Part (i), in the proof of Proposition 16.3, we have shown that  $\|T^*\| \leq \|T\|$ . In addition, the reverse inequality follows clearly from  $T^{**} = T$ .

The Part (ii) follows from the adjoint operators which are uniquely determined by the Eq.16.2 above.

For Part (iii), we always have  $\|T^* T\| \leq \|T^*\| \|T\| = \|T\|^2$ . For the reverse inequality, let  $x \in B_X$ . Then

$$\|Tx\|^2 = (Tx, Tx) = (T^* Tx, x) \leq \|T^* Tx\| \|x\| \leq \|T^* T\|.$$

Therefore, we have  $\|T\|^2 \leq \|T^* T\|$ . □

**Example 16.7.** If  $X = \mathbb{C}^n$  and  $D = (a_{ij})_{n \times n}$  an  $n \times n$  matrix, then  $D^* = (\bar{a}_{ji})_{n \times n}$ . In fact, note that

$$a_{ji} = (De_i, e_j) = (e_i, D^* e_j) = \overline{(D^* e_j, e_i)}.$$

Thus if we put  $D^* = (d_{ij})_{n \times n}$ , then  $d_{ij} = (D^* e_j, e_i) = \bar{a}_{ji}$ .

**Example 16.8.** Let  $\ell^2(\mathbb{N}) := \{x : \mathbb{N} \rightarrow \mathbb{C} : \sum_{i=0}^{\infty} |x(i)|^2 < \infty\}$ , and put  $(x, y) := \sum_{i=0}^{\infty} x(i) \overline{y(i)}$ .

Define the operator  $D \in B(\ell^2(\mathbb{N}))$  (called the unilateral shift) by

$$Dx(i) = x(i-1)$$

for  $i \in \mathbb{N}$ , where we set  $x(-1) := 0$ , i.e.,  $D(x(0), x(1), \dots) = (0, x(0), x(1), \dots)$ .

Then  $D$  is an isometry and the adjoint operator  $D^*$  is given by

$$D^* x(i) := x(i+1)$$

for  $i = 0, 1, \dots$ , i.e.,  $D^*(x(0), x(1), \dots) = (x(1), x(2), \dots)$ .

Indeed we can directly check that

$$(Dx, y) = \sum_{i=0}^{\infty} x(i-1) \overline{y(i)} = \sum_{j=0}^{\infty} x(j) \overline{y(j+1)} = (x, D^* y).$$

Note that  $D^*$  is NOT an isometry.

**Example 16.9.** Let  $\ell^\infty(\mathbb{N}) = \{x : \mathbb{N} \rightarrow \mathbb{C} : \sup_{i \geq 0} |x(i)| < \infty\}$  and  $\|x\|_\infty := \sup_{i \geq 0} |x(i)|$ . For each  $x \in \ell^\infty$ , define  $M_x \in B(\ell^2(\mathbb{N}))$  by

$$M_x(\xi) := x \cdot \xi$$

for  $\xi \in \ell^2(\mathbb{N})$ , where  $(x \cdot \xi)(i) := x(i)\xi(i)$ ;  $i \in \mathbb{N}$ .

Then  $\|M_x\| = \|x\|_\infty$  and  $M_x^* = M_{\bar{x}}$ , where  $\bar{x}(i) := \overline{x(i)}$ .

**Definition 16.10.** Let  $T \in B(X)$  and let  $I$  be the identity operator on  $X$ .  $T$  is said to be

- (i) : selfadjoint if  $T^* = T$ ;
- (ii) : normal if  $T^*T = TT^*$ ;
- (iii) : unitary if  $T^*T = TT^* = I$ .

**Proposition 16.11.** We have

- (i) : Let  $T : X \rightarrow X$  be a linear operator.  $T$  is a bounded linear selfadjoint operator if and only if we have

$$(16.3) \quad (Tx, y) = (x, Ty) \quad \text{for all } x, y \in X.$$

- (ii) :  $T$  is normal if and only if  $\|Tx\| = \|T^*x\|$  for all  $x \in X$ .

*Proof.* The necessary part of Part (i) is clear.

Now suppose that the Eq.16.3 holds, it needs to show that  $T$  is bounded. Indeed, it follows immediately from Remark16.4.

For Part (ii), note that by Proposition 16.2,  $T$  is normal if and only if  $(T^*Tx, x) = (TT^*x, x)$ . Thus, Part (ii) follows from

$$\|Tx\|^2 = (Tx, Tx) = (T^*Tx, x) = (TT^*x, x) = (T^*x, T^*x) = \|T^*x\|^2$$

for all  $x \in X$ . □

**Remark 16.12.** In Proposition 16.11(i), if the domain of  $T$  is replaced by dense domain, then the conclusion does not hold. For example, let  $D := \{x \in \ell^2 : \sum_{n=1}^\infty |nx(n)|^2 < \infty\}$  and let  $T(x)(n) := nx(n)$  for  $x \in D$ . Then  $D$  is a dense domain because the canonical basis  $(e_n) \subseteq D$ . It is noted that  $T$  is unbounded on  $D$ , but  $(Tx, y) = (x, Ty)$  for all  $x, y \in D$ .

**Proposition 16.13.** Let  $T \in B(H)$ . We have the following assertions.

- (i) :  $T$  is selfadjoint if and only if  $(Tx, x) \in \mathbb{R}$  for all  $x \in H$ .
- (ii) : If  $T$  is selfadjoint, then  $\|T\| = \sup\{|(Tx, x)| : x \in H \text{ with } \|x\| = 1\}$ .

*Proof.* Part (i) follows immediately from Proposition16.2.

For Part (ii), if we let  $a = \sup\{|(Tx, x)| : x \in H \text{ with } \|x\| = 1\}$ , then we have  $a \leq \|T\|$ . We want to show the reverse inequality.  $T$  is selfadjoint, and so we can directly check that

$$(T(x+y), x+y) - (T(x-y), x-y) = 4\operatorname{Re}(Tx, y)$$

for all  $x, y \in H$ . Thus if  $x, y \in H$  with  $\|x\| = \|y\| = 1$  and  $(Tx, y) \in \mathbb{R}$ , then by using the Parallelogram Law, we have

$$(16.4) \quad |(Tx, y)| \leq \frac{a}{4}(\|x+y\|^2 + \|x-y\|^2) = \frac{a}{2}(\|x\|^2 + \|y\|^2) = a.$$

Now for  $x, y \in H$  with  $\|x\| = \|y\| = 1$ , by considering the polar form of  $(Tx, y) = re^{i\theta}$ , the Eq.16.4 gives

$$|(Tx, y)| = |(Tx, e^{i\theta}y)| \leq a.$$

$\|T\| = \sup_{\|x\|=\|y\|=1} |(Tx, y)|$ , and so we have  $\|T\| \leq a$ . The proof is complete. □



**Proposition 16.14.** *Let  $T \in B(X)$ . Then we have*

$$\ker T = (\operatorname{im} T^*)^\perp \quad \text{and} \quad (\ker T)^\perp = \overline{\operatorname{im} T^*}$$

where  $\operatorname{im} T$  denotes the image of  $T$ .

*Proof.* The first equality follows clearly from  $x \in \ker T$  if and only if  $0 = (Tx, z) = (x, T^*z)$  for all  $z \in X$ .

On the other hand, it is clear that we have  $M^\perp = \overline{M}^\perp$  for any subspace  $M$  of  $X$ . This, together with the first equality and Corollary 15.7, gives immediately the second equality.  $\square$

**Proposition 16.15.** *Let  $X$  be a Hilbert space. Let  $M$  and  $N$  be the closed subspaces of  $X$  such that*

$$X = M \oplus N \quad \dots\dots\dots (*)$$

*Let  $Q : X \rightarrow X$  be the projection along the decomposition  $(*)$  with  $\operatorname{im} Q = M$  (note that  $Q$  is bounded by Proposition 12.1). Then  $N = M^\perp$  (and hence  $(*)$  is the orthogonal decomposition of  $X$  with respect to  $M$ ) if and only if  $Q$  satisfies the conditions:  $Q^2 = Q$  and  $Q^* = Q$ . In this case,  $Q$  is called the orthogonal projection (or projection for simply) with respect to  $M$ .*

*Proof.* Now if  $N = M^\perp$ , then for  $y, y' \in M$  and  $z, z' \in N$ , we have

$$(Q(y + z), y' + z') = (y, y') = (y + z, Q(y' + z')).$$

Thus  $Q^* = Q$ .

The converse of the last statement follows immediately from Proposition 16.14 because  $\ker Q = N$  and  $\operatorname{im} Q = M$ .

The proof is complete.  $\square$

**Proposition 16.16.** *When  $X$  is a Hilbert space, we put  $\mathcal{M}$  the set of all closed subspaces of  $X$  and  $\mathcal{P}$  the set of all orthogonal projections on  $X$ . Now for each  $M \in \mathcal{M}$ , let  $P_M$  be the corresponding projection with respect to the orthogonal decomposition  $X = M \oplus M^\perp$ . Then there is an one-one correspondence between  $\mathcal{M}$  and  $\mathcal{P}$  which is defined by*

$$M \in \mathcal{M} \mapsto P_M \in \mathcal{P}.$$

Furthermore, if  $M, N \in \mathcal{M}$ , then we have

- (i) :  $M \subseteq N$  if and only if  $P_M P_N = P_N P_M = P_M$ .
- (ii) :  $M \perp N$  if and only if  $P_M P_N = P_N P_M = 0$ .

*Proof.* Using Proposition 16.15, we note that  $P_M \in \mathcal{P}$ .

Indeed the inverse of the correspondence is given by the following. If we let  $Q \in \mathcal{P}$  and  $M = Q(X)$ , then  $M$  is closed. In addition, clearly we have  $X = Q(X) \oplus (I - Q)X$  with  $M^\perp = (I - Q)X$ . Hence  $M$  is the corresponding closed subspace of  $X$ , i.e.,  $M \in \mathcal{M}$  and  $P_M = Q$ .

For the final assertion, Part (i) and (ii) follow immediately from the orthogonal decompositions  $X = M \oplus M^\perp = N \oplus N^\perp$  and together with the fact that  $M \subseteq N$  if and only if  $N^\perp \subseteq M^\perp$ .  $\square$

## 17. SPECTRAL THEORY I

**Definition 17.1.** *Let  $E$  be a normed space and let  $T \in B(E)$ . The spectrum of  $T$ , denoted by  $\sigma(T)$ , is defined by*

$$\sigma(T) := \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not invertible in } B(E)\}.$$

**Remark 17.2.** More precisely, for a normed space  $E$ , an operator  $T \in B(E)$  is said to be invertible in  $B(E)$  if  $T$  is a linear isomorphism and the inverse  $T^{-1}$  is also bounded. However, if  $E$  is complete, the Open Mapping Theorem assures that the inverse  $T^{-1}$  is bounded automatically. Thus if  $E$  is a Banach space and  $T \in B(E)$ , then  $\lambda \notin \sigma(T)$  if and only if  $T - \lambda := T - \lambda I$  is a linear isomorphism. Thus,  $\lambda$  lies in the spectrum  $\sigma(T)$  if and only if  $T - \lambda$  is either not one-one or not surjective.

In particular, if there is a non-zero element  $v \in X$  such that  $Tv = \lambda v$ , then  $\lambda \in \sigma(T)$  and  $\lambda$  is called an eigenvalue of  $T$  with eigenvector  $v$ .

In addition, we write  $\sigma_p(T)$  for the set of all eigenvalue of  $T$  and call  $\sigma_p(T)$  the point spectrum.

**Example 17.3.** Let  $E = \mathbb{C}^n$  and  $T = (a_{ij})_{n \times n} \in M_n(\mathbb{C})$ . Then  $\lambda \in \sigma(T)$  if and only if  $\lambda$  is an eigenvalue of  $T$  and thus  $\sigma(T) = \sigma_p(T)$ .

**Example 17.4.** Let  $E = (c_{00}(\mathbb{N}), \|\cdot\|_\infty)$  (note that  $c_{00}(\mathbb{N})$  is not a Banach space). Define the map  $T : c_{00}(\mathbb{N}) \rightarrow c_{00}(\mathbb{N})$  by

$$Tx(k) := \frac{x(k)}{k+1}$$

for  $x \in c_{00}(\mathbb{N})$  and  $i \in \mathbb{N}$ .

Then  $T$  is bounded, in fact,  $\|Tx\|_\infty \leq \|x\|_\infty$  for all  $x \in c_{00}(\mathbb{N})$ .

On the other hand, we note that if  $\lambda \in \mathbb{C}$  and  $x \in c_{00}(\mathbb{N})$ , then

$$(T - \lambda)x(k) = \left(\frac{1}{k+1} - \lambda\right)x(k).$$

From this we see that  $\sigma_p(T) = \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ . In addition, if  $\lambda \notin \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ , then  $T - \lambda$  is a linear isomorphism and its inverse is given by

$$(T - \lambda)^{-1}x(k) = \left(\frac{1}{k+1} - \lambda\right)^{-1}x(k).$$

Thus,  $(T - \lambda)^{-1}$  is unbounded if  $\lambda = 0$ , so  $0 \in \sigma(T)$ .

Besides, if  $\lambda \notin \{0, 1, \frac{1}{2}, \frac{1}{3}, \dots\}$ , then  $(T - \lambda)^{-1}$  is bounded. In fact, if  $\lambda = a + ib \neq 0$ , for  $a, b \in \mathbb{R}$ , then  $\eta := \min_k \left| \frac{1}{1+k} - a \right|^2 + |b|^2 > 0$  because  $\lambda \notin \{1, \frac{1}{2}, \frac{1}{3}, \dots\}$ . This gives

$$\|(T - \lambda)^{-1}\| = \sup_{k \in \mathbb{N}} \left| \left(\frac{1}{k+1} - \lambda\right)^{-1} \right| < \eta^{-1} < \infty.$$

We can now conclude that  $\sigma(T) = \{1, \frac{1}{2}, \frac{1}{3}, \dots\} \cup \{0\}$ .

**Proposition 17.5.** Let  $E$  be a Banach space and  $T \in B(E)$ . Then

- (i) :  $I - T$  is invertible in  $B(E)$  whenever  $\|T\| < 1$ .
- (ii) : If  $|\lambda| > \|T\|$ , then  $\lambda \notin \sigma(T)$ .
- (iii) :  $\sigma(T)$  is a compact subset of  $\mathbb{C}$ .
- (iv) : If we let  $GL(E)$  the set of all invertible elements in  $B(E)$ , then  $GL(E)$  is an open subset of  $B(E)$  with respect to the  $\|\cdot\|$ -topology. Moreover, the map  $T \in GL(E) \mapsto T^{-1} \in GL(E)$  is continuous in the norm-topology.

*Proof.* Note that since  $B(E)$  is complete, Part (i) follows immediately from the following equality.

$$(I - T)(I + T + T^2 + \dots + T^{N-1}) = I - T^N$$

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

For Part (ii), if  $|\lambda| > \|T\|$ , then by Part (i), we see that  $I - \frac{1}{\lambda}T$  is invertible and so is  $\lambda I - T$ . This implies  $\lambda \notin \sigma(T)$ .

For Part (iii), since  $\sigma(T)$  is bounded by Part (ii), we need to show that  $\sigma(T)$  is closed.

Let  $c \in \mathbb{C} \setminus \sigma(T)$ . We need to find  $r > 0$  such that  $\mu \notin \sigma(T)$  as  $|\mu - c| < r$ . Note that since  $T - c$  is invertible, then for  $\mu \in \mathbb{C}$ , we have  $T - \mu = (T - c) - (\mu - c) = (T - c)(I - (\mu - c)(T - c)^{-1})$ . Therefore, if  $\|(\mu - c)(T - c)^{-1}\| < 1$ , then  $T - \mu$  is invertible by Part (i). Thus, if we take  $0 < r < \frac{1}{\|(T - c)^{-1}\|}$ , then  $r$  is as desired, i.e.,  $B(c, r) \subseteq \mathbb{C} \setminus \sigma(T)$ . Hence  $\sigma(T)$  is closed.

For the last assertion, let  $T \in GL(E)$ . Note that for any  $S \in B(E)$ , we have  $S = S - T + T = T(1 - T^{-1}(T - S))$ . Thus, if  $1 - T^{-1}(T - S)$  is invertible, then so is  $S$ . Using Part (i), if  $\|T - S\| < 1/\|T^{-1}\|$ , then  $1 - T^{-1}(T - S)$  is invertible. Therefore we have  $B(T, \frac{1}{\|T^{-1}\|}) \subseteq GL(E)$ . Finally, we show the inverse map is continuous. It suffices to show that if  $(T_n)$  is a sequence in  $GL(E)$  so that  $T_n \rightarrow I$ , then  $T_n^{-1} \rightarrow I$ . Note that if  $\|T_n - I\| < 1/2$ , then  $T_n^{-1} = \sum_{k=0}^{\infty} (I - T_n)^k$ , hence, we may assume that  $(T_n^{-1})$  is uniformly bounded by 2. Therefore,

$$\|T_n^{-1} - I\| \leq \|T_n^{-1}\| \|T_n - I\| \leq 2\|T_n - I\|.$$

The proof is complete.  $\square$

**Corollary 17.6.** *If  $U$  is a unitary operator on a Hilbert space  $X$ , then  $\sigma(U) \subseteq \{\lambda \in \mathbb{C} : |\lambda| = 1\}$ .*

*Proof.* Since  $\|U\| = 1$ , we have  $\sigma(U) \subseteq \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$  by Proposition 17.5(ii).

Now if  $|\lambda| < 1$ , then  $\|\lambda U^*\| < 1$ . By using Proposition 17.5 again, we have  $I - \lambda U^*$  is invertible. This implies that  $U - \lambda = U(I - \lambda U^*)$  is invertible and thus  $\lambda \notin \sigma(U)$ .  $\square$

**Example 17.7.** Let  $E = \ell^2(\mathbb{N})$  and let  $D \in B(E)$  be the right unilateral shift operator as in Example 16.8. Recall that  $Dx(k) := x(k - 1)$  for  $k \in \mathbb{N}$  and  $x(-1) := 0$ . Then  $\sigma_p(D) = \emptyset$  and  $\sigma(D) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$ .

We first claim that  $\sigma_p(D) = \emptyset$ .

Suppose that  $\lambda \in \mathbb{C}$  and  $x \in \ell^2(\mathbb{N})$  satisfy the equation  $Dx = \lambda x$ . Then by the definition of  $D$ , we have

$$x(k - 1) = \lambda x(k) \quad \dots\dots\dots (*)$$

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

If  $\lambda \neq 0$ , then we have  $x(k) = \lambda^{-1}x_{k-1}$  for all  $k \in \mathbb{N}$ . Since  $x(-1) = 0$ , this forces  $x(k) = 0$  for all  $k$ , i.e.,  $x = 0$  in  $\ell^2(\mathbb{N})$ .

On the other hand if  $\lambda = 0$ , the Eq.(\*) gives  $x(k - 1) = 0$  for all  $k$  and so  $x = 0$  again.

Therefore  $\sigma_p(D) = \emptyset$ .

Finally, we are going to show  $\sigma(D) = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$ .

Note that since  $D$  is an isometry,  $\|D\| = 1$ . Proposition 17.5 tells us that

$$\sigma(D) \subseteq \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}.$$

Note that since  $\sigma_p(D)$  is empty, it suffices to show that  $D - \mu$  is not surjective for all  $\mu \in \mathbb{C}$  with  $|\mu| \leq 1$ .

Now suppose that there is  $\lambda \in \mathbb{C}$  with  $|\lambda| \leq 1$  such that  $D - \lambda$  is surjective.

We consider the case where  $|\lambda| = 1$  first.

Let  $e_1 = (1, 0, 0, \dots) \in \ell^2(\mathbb{N})$ . Then by the assumption, there is  $x \in \ell^2(\mathbb{N})$  such that  $(D - \lambda)x = e_1$  and thus  $Dx = \lambda x + e_1$ . This implies that

$$x(k - 1) = Dx(k) = \lambda x(k) + e_1(k)$$

for all  $k \in \mathbb{N}$ . From this we have  $x(0) = -\lambda^{-1}$  and  $x(k) = -\lambda^{-k}x(0)$  for all  $k \geq 1$  because  $e_1(0) = 1$  and  $e_1(k) = 0$  for all  $k \geq 1$ . Moreover, since  $|\lambda| = 1$ , it turns out that  $|x(0)| = |x(k)|$  for all  $k \geq 1$ . As  $x \in \ell^2(\mathbb{N})$ , this forces  $x = 0$ . However, it is absurd because  $Dx = \lambda x + e_1$ .

Now we consider the case where  $|\lambda| < 1$ .

By Proposition 16.14, we have

$$\overline{\text{im}(D - \lambda)}^\perp = \ker(D - \lambda)^* = \ker(D^* - \bar{\lambda}).$$

Thus if  $D - \lambda$  is surjective, we have  $\ker(D^* - \bar{\lambda}) = (0)$  and hence  $\bar{\lambda} \notin \sigma_p(D^*)$ . Note that the adjoint  $D^*$  of  $D$  is given by the left shift operator, i.e.,

$$D^*x(k) = x(k+1) \quad \dots\dots\dots (**)$$

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

Now when  $D^*x = \mu x$  for some  $\mu \in \mathbb{C}$  and  $x \in \ell^2(\mathbb{N})$ , by using Eq.(\*\*), which is equivalent to saying that

$$x(k+1) = \mu x(k)$$

for all  $k \in \mathbb{N}$ . Therefore, if  $|\bar{\lambda}| = |\lambda| < 1$  and we set  $x(0) = 1$  and  $x(k+1) = \bar{\lambda}^k x(0)$  for all  $k \geq 1$ , then  $x \in \ell^2(\mathbb{N})$  and  $D^*x = \bar{\lambda}x$ . Hence  $\bar{\lambda} \in \sigma_p(D^*)$  which leads to a contradiction. The proof is complete.

## 18. SPECTRAL THEORY II

Throughout this section, let  $H$  be a complex Hilbert space.

**Lemma 18.1.** *Let  $T \in B(H)$  be a normal operator (recall that  $T^*T = TT^*$ ). Then  $T$  is invertible in  $B(H)$  if and only if there is  $c > 0$  such that  $\|Tx\| \geq c\|x\|$  for all  $x \in H$ .*

*Proof.* The necessary part is obvious.

Now we want to show the converse. We first show the case where  $T$  is selfadjoint. Clearly,  $T$  is injective from the assumption. By the Open Mapping Theorem, we need to show that  $T$  is surjective.

In fact since  $\ker T = \overline{\text{im } T^*}^\perp$  and  $T = T^*$ , we see that the image of  $T$  is dense in  $H$ .

Now if  $y \in H$ , then there is a sequence  $(x_n)$  in  $H$  such that  $Tx_n \rightarrow y$ . Thus,  $(Tx_n)$  is a Cauchy sequence. From this and the assumption give us that  $(x_n)$  is also a Cauchy sequence. If  $x_n$  converges to  $x \in H$ , then  $y = Tx$ . Therefore the assertion is true when  $T$  is selfadjoint.

Now if  $T$  is normal, then we have  $\|T^*x\| = \|Tx\| \geq c\|x\|$  for all  $x \in H$  by Proposition 16.11(ii). Therefore, we have  $\|T^*Tx\| \geq c\|Tx\| \geq c^2\|x\|$ . Hence  $T^*T$  still satisfies the assumption. Note that  $T^*T$  is selfadjoint. Therefore, we can apply the previous case to know that  $T^*T$  is invertible. This implies that  $T$  is also invertible because  $T^*T = TT^*$ .

The proof is complete.  $\square$

**Definition 18.2.** *Let  $T \in B(H)$ . We say that  $T$  is positive, denoted by  $T \geq 0$ , if  $(Tx, x) \geq 0$  for all  $x \in H$ . For a pair of selfadjoint operators  $S$  and  $T$ , we say that  $S \leq T$  if  $T - S \geq 0$ .*

**Remark 18.3.** *Clearly, a positive operator is selfadjoint by Proposition 16.13.*

*In particular, all projections are positive.*

**Proposition 18.4.** *If  $T$  is an invertible operator in  $B(H)$ , then the inverse  $T^{-1}$  of  $T$  belongs to the closed  $*$ -subalgebra of  $B(H)$  generated by  $T$  and  $I$ .*

*Proof.* Put  $S := T^*T$ . Then  $S$  is invertible in  $B(H)$ . Now we may assume that  $\|S\| \leq 1$ . Lemma 18.1 gives  $c > 0$  such that  $(x, x) \geq (S^2x, x) \geq c(x, x)$  for all  $x \in H$ . We choose a positive integer  $N$  such that  $Nc \geq 1$ . Then we have

$$(x, x) \geq \frac{1}{Nc}(x, x) \geq \frac{1}{Nc}(S^2x, x) \geq \frac{1}{N}(x, x)$$

for all  $x \in H$ . Thus, we have

$$0 \leq I - \frac{S^2}{Nc} \leq I - \frac{1}{N}I < I.$$

If we let  $R := I - \frac{S^2}{Nc}$ , then  $(I - R)^{-1}$  exists in  $B(H)$  and hence we have

$$\left(\frac{S^2}{Nc}\right)^{-1} = (I - R)^{-1} = \sum_{n=0}^{\infty} \left(I - \frac{S^2}{Nc}\right)^n.$$

Then the result follows from

$$T^{-1} = \frac{1}{Nc} \sum_{n=0}^{\infty} \left(I - \frac{(T^*T)^2}{Nc}\right)^n T^* T T^*.$$

□

**Proposition 18.5.** *Let  $T \in B(H)$ . We have*

(i) : *If  $T \geq 0$ , then  $T + I$  is invertible.*

(ii) : *If  $T$  is self-adjoint, then  $\sigma(T) \subseteq \mathbb{R}$ . In particular, if  $T \geq 0$ , we have  $\sigma(T) \subseteq [0, \infty)$ .*

*Proof.* For Part (i), we assume that  $T \geq 0$ . This implies that

$$\|(I + T)x\|^2 = \|x\|^2 + \|Tx\|^2 + 2(Tx, x) \geq \|x\|^2$$

for all  $x \in H$ . Thus, the invertibility of  $I + T$  follows from Lemma 18.1.

For Part (ii), we first claim that  $T + i$  is invertible. Indeed, it follows immediately from  $(T + i)^*(T + i) = T^2 + I$  and Part (i).

Now if  $\lambda = a + ib$  where  $a, b \in \mathbb{R}$  with  $b \neq 0$ , then  $T - \lambda = -b(\frac{-1}{b}(T - a) + i)$  is invertible because  $\frac{-1}{b}(T - a)$  is selfadjoint. Thus,  $\sigma(T) \subseteq \mathbb{R}$ .

Finally we want to show  $\sigma(T) \subseteq [0, \infty)$  when  $T \geq 0$ . Note that since  $\sigma(T) \subseteq \mathbb{R}$ , it suffices to show that  $T - c$  is invertible if  $c < 0$ . Indeed, if  $c < 0$ , then we see that  $T - c = -c(I + (\frac{-1}{c}T))$  is invertible by the previous assertion because  $\frac{-1}{c}T \geq 0$ .

The proof is complete. □

**Remark 18.6.** *In Proposition 18.5, we have shown that if  $T$  is selfadjoint, then  $\sigma(T) \subseteq \mathbb{R}$ . However, the converse does not hold. For example, consider  $H = \mathbb{C}^2$  and*

$$T = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

**Example 18.7.** *Notice that the multiplication defines an isometry  $M : x \in \ell_\infty \mapsto M(x) \in B(\ell_2)$  by  $M(x)(\xi)(n) := x(n)\xi(n); n = 1, 2, \dots$  for  $\xi \in \ell_2$ . Then  $M(\bar{x}) = M(x)^*$  for  $x \in \ell_\infty$ , and so,  $M(x)$  is self-adjoint if and only if  $x$  is a  $\mathbb{R}$ -sequence. Now let  $x \in \ell_\infty$  be a  $\mathbb{R}$ -sequence. For simply for each element  $x \in \ell_\infty$ , we also write  $x$  for  $M(x)$  as an element in  $B(\ell_2)$ .*

*Now we claim that if  $x \in \ell_\infty$  is self-adjoint, then  $\lambda \in \sigma(x)$  if and only if  $\inf_n |x(n) - \lambda| = 0$ .*

*Consequently,  $\sigma_p(x) = \{x_n : n = 1, 2, \dots\}$  and  $\sigma(x) = \overline{\{x(n) : n = 1, 2, \dots\}}$ .*

*In fact, for showing ( $\Leftarrow$ ), let  $\lambda \in \mathbb{R}$  such that  $\inf_n |x(n) - \lambda| = 0$ . If  $x - \lambda$  is invertible in  $B(\ell_2)$ , then by Lemma 18.1, there is  $c > 0$  such that  $\|(x - \lambda)\xi\| \geq c$  for all  $\xi \in \ell_2$  of norm one. In particular, for each  $n = 1, 2, \dots$ , we have  $|x(n) - \lambda| = \|(x - \lambda)(e_n)\| \geq c > 0$ . It leads to a contradiction.*

*For showing ( $\Rightarrow$ ), let  $\lambda \in \mathbb{R}$  such that  $c := \inf_n |x(n) - \lambda| > 0$ . Then  $x(n) \neq \lambda$  for all  $n = 1, 2, \dots$ .*

*This implies that  $x - \lambda$  is injective. On the other hand, for any  $\eta \in \ell_2$ , if  $(x(n) - \lambda)\xi(n) = \eta(n)$  for all  $n$ , then we have  $\xi(n) = \frac{\eta(n)}{x(n) - \lambda}$  and so,  $|\xi(n)| \leq \frac{|\eta(n)|}{c}$ . This gives  $\xi \in \ell_2$ . Therefore,  $x - \lambda$  is surjective and thus,  $x - \lambda$  is invertible. Hence,  $\lambda \notin \sigma(x)$ .*

*From this, the last assertion follows because  $\lambda \in \sigma(x)$  if and only if  $\lambda = x_n$  for some  $n$  or there is a subsequence  $(x_{n_k})$  of  $(x_n)$  that converges to  $\lambda$ .*

**Theorem 18.8.** *Let  $T \in B(H)$  be a selfadjoint operator. Put*

$$M(T) := \sup_{\|x\|=1} (Tx, x) \quad \text{and} \quad m(T) = \inf_{\|x\|=1} (Tx, x).$$

*For convenience, we also write  $M = M(T)$  and  $m = m(T)$  if there is no confusion. Then we have*

- (i) :  $\|T\| = \max\{|m|, |M|\}$ .
- (ii) :  $\{m, M\} \subseteq \sigma(T)$ .
- (iii) :  $\sigma(T) \subseteq [m, M]$ .

*Proof.* Note that  $m$  and  $M$  are well defined because  $(Tx, x)$  is real for all  $x \in H$  by Proposition 16.13 (ii). To see this part (i), since  $(Tx, x) \leq \|T\|$  for all  $x \in H$  with  $\|x\| = 1$ , we have  $M(T) \leq \|T\|$ . Hence we also have  $-m(T) = M(-T) \leq \| -T \| = \|T\|$  and so  $-\|T\| \leq m(T)$ . Then we have  $-\|T\| \leq m(T) \leq M(T) \leq \|T\|$ . This gives  $\max(|m(T)|, |M(T)|) \leq \|T\|$ . Now if we consider the case  $|m| \leq |M|$  first, then we have  $-|M| \leq -|m| \leq m \leq (Tx, x) \leq M \leq |M|$  for all  $x \in H$  with  $\|x\| = 1$ . Thus, by using Lemma 16.13 (ii) again, we have  $\|T\| \leq |M| = \max(|m|, |M|)$ . We will complete the proof of Part (i) if we consider  $(-T)$  for the case  $|M(T)| \leq |m(T)|$ .

For Part (ii), we first claim that  $M \in \sigma(T)$  if  $T \geq 0$ . Note that  $0 \leq m \leq M = \|T\|$  in this case by Lemma 16.13. Then there is a sequence  $(x_n)$  in  $H$  with  $\|x_n\| = 1$  for all  $n$  such that  $(Tx_n, x_n) \rightarrow M = \|T\|$ . Then we have

$$\|(T - M)x_n\|^2 = \|Tx_n\|^2 + M^2\|x_n\|^2 - 2M(Tx_n, x_n) \leq \|T\|^2 + M^2 - 2M(Tx_n, x_n) \rightarrow 0.$$

Hence, by Lemma 18.1 we have shown that  $T - M$  is not invertible and hence  $M \in \sigma(T)$  if  $T \geq 0$ . Now for any selfadjoint operator  $T$  if we consider  $T - m$ , then  $T - m \geq 0$ . Thus we have  $M - m = M(T - m) \in \sigma(T - m)$  by the previous case. Clearly, we have  $\sigma(T - c) = \sigma(T) - c$  for all  $c \in \mathbb{C}$ . Therefore we have  $M \in \sigma(T)$  for any self-adjoint operator.

We claim that  $m(T) \in \sigma(T)$ . Note that  $M(-T) = -m(T)$ . Thus, we have  $-m(T) \in \sigma(-T)$ . It is clear that  $\sigma(-T) = -\sigma(T)$ . Then  $m(T) \in \sigma(T)$ .

Finally, we want to show  $\sigma(T) \subseteq [m, M]$ .

Indeed, since  $T - m \geq 0$ , then by Proposition 18.5, we have  $\sigma(T) - m = \sigma(T - m) \subseteq [0, \infty)$ . This gives  $\sigma(T) \subseteq [m, \infty)$ .

On the other hand, we consider  $M - T \geq 0$ . Then we get  $M - \sigma(T) = \sigma(M - T) \subseteq [0, \infty)$ . This implies that  $\sigma(T) \subseteq (-\infty, M]$ . The proof is complete.  $\square$

## 19. APPENDIX: $\sigma(T) \neq \emptyset$

Let  $X$  be a complex Banach space. In this appendix, we will show that the spectrum  $\sigma(T)$  is non-empty for any  $T \in B(X)$ .

First we recall some basic result in Complex Analysis. Students can refer to any standard text book of Complex Analysis, see for example [1].

A function  $g : \mathbb{C} \rightarrow \mathbb{C}$  is called an *entire function* if  $g$  is differentiable on  $\mathbb{C}$ , i.e., the following limit exists for all  $c \in \mathbb{C}$

$$g'(c) := \lim_{z \rightarrow c} \frac{g(z) - g(c)}{z - c}.$$

The following result is one of important properties of entire functions (see [1, p.122]).

**Theorem 19.1. Liouville's Theorem** *Every bounded entire function is a constant function.*

**Theorem 19.2.** *Using the notion as before, let  $T \in B(X)$ . Then the spectrum  $\sigma(T) \neq \emptyset$ .*

*Proof.* Assume that  $\sigma(T) = \emptyset$ . Fix  $f \in B(X)^*$ , define the map  $g(z) := f((z - T)^{-1})$  is defined for all  $z \in \mathbb{C}$ . Note that  $g$  is continuous on  $\mathbb{C}$  by considering the composition  $\lambda \in \mathbb{C} \mapsto \lambda - T \mapsto$

$(\lambda - T)^{-1} \in B(X)$  and using Proposition 17.5 (iv). Moreover, we have  $\lim_{z \rightarrow \infty} |g(z)| = 0$ . Thus,  $g$  is a bounded function on  $\mathbb{C}$ . On the other hand, if we fix a point  $c \in \mathbb{C}$ , then we see that

$$\lim_{z \rightarrow c} \frac{g(z) - g(c)}{z - c} = -f((c - T)^{-2}).$$

Therefore,  $g$  is a bounded entire function. By the Liouville's Theorem,  $f((z - T)^{-1})$  is a constant function on  $\mathbb{C}$ . Then the Hahn-Banach Theorem implies that the function  $z \in \mathbb{C} \mapsto (z - T)^{-1} \in B(X)$  is constant on  $\mathbb{C}$ . It leads to a contradiction.  $\square$

## 20. APPENDIX: EXISTENCE OF THE SQUARE ROOT OF A POSITIVE OPERATOR

This section is based on the note of the course Functional Analysis taught by my teacher Dr. Chow Hing Lun in 1984-85 when I was an undergraduate student in the CUHK.

Throughout this section, let  $H$  be a complex Hilbert space and let  $T$  be a positive bounded operator on  $H$ . The aim of this section is to show that there is a unique positive operator  $S$  (called the *square root* of  $T$ ) on  $H$  such that  $S^2 = T$ . The main feature of the proof here is without using the functional calculus.

**Proposition 20.1.** *Let  $S, T \in B(H)$  such that  $ST = TS$ . If  $S, T$  both are positive operators, then so is  $ST$ .*

*Proof.* If  $S = 0$ , then the assertion is clear. Now we assume that  $S \neq 0$ . Put  $S_1 := \frac{S}{\|S\|}$ . Set

$$S_{n+1} := S_n - S_n^2$$

for  $n = 1, 2, \dots$

**Claim 1:**  $0 \leq S_n \leq I$  for all  $n = 1, \dots$ . The assertion will be obtained by induction on  $n$ . Notice that as  $n = 1$ , clearly we have  $0 \leq S_1 \leq I$ . Suppose that the Claim 1 is true for  $n$ , i.e.,  $0 \leq S_n \leq I$  and thus, we have  $0 \leq I - S_n \leq I$ . This implies that for all  $x \in H$  we have  $(S_n^2(I - S_n)x, x) = ((I - S_n)S_nx, S_nx) \geq 0$ . This gives  $S_n^2(I - S_n) \geq 0$ . Similarly, we have  $S_n(I - S_n)^2 \geq 0$ . Hence, we have  $0 \leq S_n^2(I - S_n) + S_n(I - S_n)^2 = S_n - S_n^2 = S_{n+1}$ . On the other hand, we have  $0 \leq (I - S_n) + S_n^2 = I - S_{n+1}$  because  $S_n^2 \geq 0$  and  $I - S_n \geq 0$ . Therefore Claim 1 follows from the induction.

The proof will be complete if we show that  $(STx, x) \geq 0$  for all  $x \in H$ .

In fact, notice that we have

$$S_1 = S_1^2 + S_2 = S_1^2 + S_2^2 + S_3 = \dots = S_1^2 + \dots + S_n^2 + S_{n+1}.$$

This implies that

$$S_1^2 + \dots + S_n^2 = S_1 - S_{n+1} \leq S_1$$

for all  $n = 1, 2, \dots$  because  $S_{n+1} \geq 0$ . Thus, we have

$$\sum_{k=1}^n \|S_k x\|^2 = \sum_{k=1}^n (S_k x, S_k x) = \sum_{k=1}^n (S_k^2 x, x) \leq (S_1 x, x)$$

for all  $n$ . This gives  $\sum_{k=1}^{\infty} \|S_k x\|^2 < \infty$  and so,  $S_n x \rightarrow 0$ . This implies that

$$\left( \sum_{k=1}^n S_k^2 \right) x = S_1(x) - S_{n+1}(x) \rightarrow 0$$

for all  $x \in H$  and so we have  $\sum_{k=1}^{\infty} S_k^2(x) = S_1(x)$  for all  $x \in H$ . Finally, we complete the proof by the following

$$(STx, x) = \|S\|(TS_1x, x) = \|S\| \sum_{k=1}^{\infty} (TS_k^2x, x) = \|S\| \sum_{k=1}^{\infty} (TS_kx, S_kx) \geq 0$$

for all  $x \in H$ .  $\square$

**Proposition 20.2.** *Let  $T_n$ ,  $n = 1, 2, \dots$  and  $K$  be the bounded linear selfadjoint operators on  $H$ . Suppose that*

- (1)  $T_1 \leq T_2 \leq \dots \leq K$ .
- (2)  $T_n T_m = T_m T_n$  and  $K T_n = T_n K$  for all  $m, n = 1, 2, \dots$ .

*Then there is a bounded selfadjoint operator  $T$  on  $H$  with  $T \leq K$  such that  $\lim T_n x = T x$  for all  $x \in H$ .*

*Proof.* Now let  $S_n := K - T_n$  for  $n = 1, 2, \dots$ . Then  $0 \leq S_n$  for all  $n = 1, 2, \dots$ . By using Proposition 20.1, we see that  $S_m^2 - S_n S_m = (S_m - S_n) S_m \geq 0$  and hence,  $S_m^2 \geq S_n S_m$  for  $n \geq m$ . Similarly, we also have  $S_n S_m \geq S_n^2$  for  $n \geq m$ . Therefore, we have

$$(20.1) \quad S_m^2 \geq S_n S_m \geq S_n^2$$

for  $n \geq m$ . Thus,  $((S_m^2 x, x))_{m=1}^\infty$  is a decreasing sequence of non-negative numbers and so  $\lim (S_n^2 x, x)$  exists for all  $x \in H$ . Moreover since  $S_n$  and  $S_m$  commutes to each other, Eq 20.1 gives

$$\begin{aligned} \|S_m x - S_n x\|^2 &= ((S_m - S_n)^2 x, x) \\ &= (S_m^2 x, x) - 2(S_m S_n x, x) + (S_n^2 x, x) \\ &\leq (S_m^2 x, x) - (S_n^2 x, x) \rightarrow 0 \end{aligned}$$

for  $n \geq m$  and for all  $x \in H$ . This implies that  $(S_n x)$  is a Cauchy sequence and hence,  $\lim S_n x$  exists for all  $x \in H$ . This implies that  $T(x) := \lim T_n(x) = K - \lim S_n x$  exists for all  $x \in H$ . The Uniform Boundedness Theorem tells us that  $T \in B(H)$ . In addition  $T$  is selfadjoint because each  $T_n$  is selfadjoint. The proof is complete.  $\square$

We now come to the main result in this section.

**Theorem 20.3.** *If  $T$  is a bounded positive operator on  $H$ , then there is a unique positive operator  $S$  such that  $S^2 = T$ . In this case, we call  $S$  the square root of  $T$ .*

*Proof.* We show the existence first.

Clearly, we may assume that  $T \neq 0$  and  $T \leq I$  by considering the operator  $\frac{T}{\|T\|}$ . Put  $S_0 = 0$  and

$$S_n = S_{n-1} + \frac{1}{2}(T - S_{n-1}^2)$$

for  $n = 1, 2, \dots$ . Then  $S_n$  is a polynomial of  $T$  and so, all  $S_n$ 's are selfadjoint operators and commute to each other. Notice that since  $0 < T \leq I$  and by the definition of  $S_n$ , we have

$$I - S_n = I - S_{n-1} - \frac{1}{2}(T - S_{n-1}^2) = \frac{1}{2}(I - S_{n-1})^2 + \frac{1}{2}(I - T) \geq 0.$$

Thus  $S_n \leq I$  for all  $n = 0, 1, 2, \dots$ . On the other hand, we have

$$(20.2) \quad S_{n+1} - S_n = S_n + \frac{1}{2}(T - S_n^2) - S_{n-1} - \frac{1}{2}(T - S_{n-1}^2) = (S_n - S_{n-1})(I - \frac{1}{2}(S_n + S_{n-1}))$$

for all  $n = 0, 1, 2, \dots$ . Since  $S_n \leq I$ ,  $I - \frac{1}{2}(S_n + S_{n-1}) \geq 0$ . Using Proposition 20.1 and the Eq 20.2, we can apply induction on  $n$  to see that  $0 = S_0 \leq \dots \leq S_n \leq S_{n+1} \leq \dots \leq I$  for all  $n = 0, 1, 2, \dots$ . Proposition 20.2 tells us that  $Sx := \lim S_n x$  exists for all  $x \in H$  and  $S \in B(H)$ . In addition  $S$  is positive because  $S_n \geq 0$  for all  $n = 0, 1, 2, \dots$ . Also, since  $S_n x = S_{n-1} x + \frac{1}{2}(T - S_{n-1}^2)x$  for all  $x \in H$ , by taking  $n \rightarrow \infty$ , we see that  $Tx = S^2 x$  for all  $x$ . Thus the operator  $S$  is as desired.

Finally, we show the uniqueness.

Now let  $R$  be another positive bounded operator on  $H$  such that  $R^2 = T$ . Notice that  $RT = R^3 = TR$ . This implies that  $RS = SR$  because  $S$  is the  $\|\cdot\|$ -limit of the polynomials of  $T$  by the above construction of  $S$ . Now we take any  $x \in H$  and put  $y := (S - R)x$ . Then we have

$$0 \leq (Sy, y) + (Ry, y) = ((S + R)(S - R)x, y) = ((S^2 - R^2)x, y) = 0.$$



This implies that  $(Sy, y) = (Ry, y) = 0$  because both are non-negative numbers. On the other hand, since  $S \leq 0$ , by above there is another positive operator  $W$  such that  $W^2 = S$ , and so we have  $0 = (Sy, y) = (Wy, Wy)$  that gives  $Sy = 0$ . Similarly, we also have  $Ry = 0$ . Finally, we have

$$\|(S - R)x\|^2 = ((S - R)^2x, x) = ((S - R)y, x) = 0.$$

Thus,  $S = R$  as desired. The proof is complete.  $\square$

## 21. COMPACT OPERATORS ON A HILBERT SPACE

Throughout this section, let  $H$  be a complex Hilbert space.

**Definition 21.1.** A linear operator  $T : H \rightarrow H$  is said to be compact if for every bounded sequence  $(x_n)$  in  $H$ ,  $(T(x_n))$  has a norm convergent subsequence.

Write  $K(H)$  for the set of all compact operators on  $H$  and  $K(H)_{sa}$  for the set of all compact selfadjoint operators.

**Remark 21.2.** Let  $U$  be the closed unit ball of  $H$ . Clearly,  $T$  is compact if and only if the norm closure  $\overline{T(U)}$  is a compact subset of  $H$ . Thus if  $T$  is compact, then  $T$  is bounded automatically because every compact set is bounded. In particular, if  $T$  has finite rank, that is  $\dim \text{im} T < \infty$ , then  $T$  must be compact because every closed and bounded subset of a finite dimensional normed space is compact. In addition, clearly we have the following result.

**Proposition 21.3.** The identity operator  $I : H \rightarrow H$  is compact if and only if  $\dim H < \infty$ .

**Example 21.4.** Let  $H = \ell^2(\{1, 2, \dots\})$ . Define  $Tx(k) := \frac{x(k)}{k}$  for  $k = 1, 2, \dots$ . Then  $T$  is compact. In fact, if we let  $(x_n)$  be a bounded sequence in  $\ell^2$ , then by the diagonal argument, we can find a subsequence  $y_m := Tx_m$  of  $Tx_n$  such that  $\lim_{m \rightarrow \infty} y_m(k) = y(k)$  exists for all  $k = 1, 2, \dots$ . Let  $L := \sup_n \|x_n\|_2^2$ . Since  $|y_m(k)|^2 \leq \frac{L}{k^2}$  for all  $m, k$ , we have  $y \in \ell^2$ . Now let  $\varepsilon > 0$ . Then one can find a positive integer  $N$  such that  $\sum_{k \geq N} 4L/k^2 < \varepsilon$ . Thus we have

$$\sum_{k \geq N} |y_m(k) - y(k)|^2 < \sum_{k \geq N} \frac{4L}{k^2} < \varepsilon$$

for all  $m$ . On the other hand, since  $\lim_{m \rightarrow \infty} y_m(k) = y(k)$  for all  $k$ , we can choose a positive integer  $M$  such that

$$\sum_{k=1}^{N-1} |y_m(k) - y(k)|^2 < \varepsilon$$

for all  $m \geq M$ . Finally, we have  $\|y_m - y\|_2^2 < 2\varepsilon$  for all  $m \geq M$ .

**Theorem 21.5.** Let  $T \in B(H)$ . Then  $T$  is compact if and only if  $T$  maps every weakly convergent sequence in  $H$  to a norm convergent sequence.

*Proof.* We first assume that  $T \in K(H)$ . Let  $(x_n)$  be a weakly convergent sequence in  $H$ . Since  $H$  is reflexive,  $(x_n)$  is bounded by the Uniform Boundedness Theorem. Thus we can find a subsequence  $(x_j)$  of  $(x_n)$  such that  $(Tx_j)$  is norm convergent. Let  $y := \lim_j Tx_j$ . We claim that  $y = \lim_n Tx_n$ . Suppose that  $y \neq \lim_n Tx_n$ . Then by the compactness of  $T$  again, we can find a subsequence  $(x_i)$  of  $(x_n)$  such that  $Tx_i$  converges to  $y'$  with  $y \neq y'$ . Thus there is  $z \in H$  such that  $(y, z) \neq (y', z)$ .

On the other hand, if we let  $x$  be the weakly limit of  $(x_n)$ , then  $(x_n, w) \rightarrow (x, w)$  for all  $w \in H$ . Thus we have

$$(y, z) = \lim_j (Tx_j, z) = \lim_j (x_j, T^*(z)) = (x, T^*z) = (Tx, z).$$

Similarly, we also have  $(y', z) = (Tx, z)$  and hence  $(y, z) = (y', z)$  that contradicts to the choice of  $z$ .

For the converse, let  $(x_n)$  be a bounded sequence. Then by Theorem 15.12,  $(x_n)$  has a weakly convergent subsequence. Thus  $T(x_n)$  has a norm convergent subsequence by the assumption. Thus  $T$  is compact.  $\square$

**Proposition 21.6.** *Let  $S, T \in K(H)$ . Then we have*

- (i) :  $\alpha S + \beta T \in K(H)$  for all  $\alpha, \beta \in \mathbb{C}$ ;
- (ii) :  $TQ$  and  $QT \in K(H)$  for all  $Q$  in  $B(H)$ ;
- (iii) :  $T^* \in K(H)$ .

Moreover  $K(H)$  is normed closed in  $B(H)$ , and hence  $K(H)$  is a closed  $*$ -ideal of  $B(H)$ .

*Proof.* (i) and (ii) are clear.

For property (iii), let  $(x_n)$  be a bounded sequence. Then  $(T^*x_n)$  is also bounded. Thus  $TT^*x_n$  has a convergent subsequence  $TT^*x_{n_k}$  by the compactness of  $T$ . Note that we have

$$\|T^*x_{n_k} - T^*x_{n_l}\|^2 = (TT^*(x_{n_k} - x_{n_l}), x_{n_k} - x_{n_l})$$

for all  $n_k, n_l$ . This implies that  $(T^*x_{n_k})$  is a Cauchy sequence and thus is convergent.

Finally we want to show that  $K(H)$  is closed. Let  $(T_m)$  be a sequence in  $K(H)$  such that  $T_m \rightarrow T$  in norm. Let  $(x_n)$  be a bounded sequence in  $H$ . Then by the diagonal argument there is a subsequence  $(x_{n_k})$  of  $(x_n)$  such that  $\lim_k T_m x_{n_k}$  exists for all  $m$ . Now let  $\varepsilon > 0$ . Since  $\lim_m T_m = T$ , there is a positive integer  $N$  such that  $\|T - T_N\| < \varepsilon$ . On the other hand, there is a positive integer  $K$  such that  $\|T_N x_{n_k} - T_N x_{n_{k'}}\| < \varepsilon$  for all  $k, k' \geq K$ . Thus we can now have

$$\|Tx_{n_k} - Tx_{n_{k'}}\| \leq \|Tx_{n_k} - T_N x_{n_k}\| + \|T_N x_{n_k} - T_N x_{n_{k'}}\| + \|T_N x_{n_{k'}} - Tx_{n_{k'}}\| \leq (2L + 1)\varepsilon$$

for all  $k, k' \geq K$  where  $L := \sup_n \|x_n\|$ . Thus  $\lim_k Tx_{n_k}$  exists. We can now conclude that  $T \in K(H)$ .  $\square$

**Example 21.7.** *Let  $k(z, w) \in C(\mathbb{T} \times \mathbb{T})$ . Define an operator  $T : L^2(\mathbb{T}) \rightarrow L^2(\mathbb{T})$  by*

$$T\xi(z) := \int_{\mathbb{T}} k(z, w)\xi(w)dw$$

for  $z \in \mathbb{T}$  and  $\xi \in L^2(\mathbb{T})$ . Then  $T$  is a compact operator.

*Proof.* Clearly, we have  $\|T\| \leq \|k\|_{\infty}$ . On the other hand, Stone-Weierstrass Theorem tells us the polynomials of  $(z, \bar{z}; w, \bar{w})$  are  $\|\cdot\|_{\infty}$ -dense in  $C(\mathbb{T} \times \mathbb{T})$ . Therefore, by using Proposition 21.6, it suffices to show for the case  $k(z, w) = \sum_{i,j=1}^N a_{ij}(z, \bar{z})w^i\bar{w}^j$  where  $a_{ij}(z, \bar{z})$  is a polynomial of  $(z, \bar{z})$  of degree  $N$ . From this, we have

$$T\xi(z) = \sum_{i,j=1}^N a_{ij}(z, \bar{z}) \int_{\mathbb{T}} w^i \bar{w}^j \xi(w)dw$$

for  $\xi \in L^2(\mathbb{T})$ . Thus,  $T(\xi) \in \text{span}\{z^i \bar{z}^j : 0 \leq i, j \leq N\}$  which is of finite dimension for all  $\xi \in L^2(\mathbb{T})$ . This implies that  $T$  has finite dimensional range and thus,  $T$  is compact. The proof is complete.  $\square$

**Corollary 21.8.** *Let  $T \in K(H)$ . If  $\dim H = \infty$ , then  $0 \in \sigma(T)$ .*

*Proof.* Suppose that  $0 \notin \sigma(T)$ . Then  $T^{-1}$  exists in  $B(H)$ . Proposition 21.6 gives  $I = TT^{-1} \in K(H)$ . This implies  $\dim H < \infty$ .  $\square$

**Proposition 21.9.** *Let  $T \in K(H)$  and let  $c \in \mathbb{C}$  with  $c \neq 0$ . Then  $T - c$  has a closed range.*

*Proof.* Note that  $\frac{1}{c}T \in K(H)$ . Thus if we consider  $\frac{1}{c}T - I$ , we may assume that  $c = 1$ . Let  $S = T - I$ . Let  $(x_n)$  be a sequence in  $H$  such that  $Sx_n \rightarrow x \in H$  in norm. By considering the orthogonal decomposition  $H = \ker S \oplus (\ker S)^\perp$ , we write  $x_n = y_n \oplus z_n$  for  $y_n \in \ker S$  and  $z_n \in (\ker S)^\perp$ . We first claim that  $(z_n)$  is bounded. Suppose that  $(z_n)$  is unbounded. By considering a subsequence of  $(z_n)$ , we may assume that  $\|z_n\| \rightarrow \infty$ . Put  $v_n := \frac{z_n}{\|z_n\|} \in (\ker S)^\perp$ . Since  $Sz_n = Sx_n \rightarrow x$ , we have  $Sv_n \rightarrow 0$ . On the other hand, since  $T$  is compact, and  $(v_n)$  is bounded, by passing a subsequence of  $(v_n)$ , we may also assume that  $Tv_n \rightarrow w$ . Since  $S = T - I$ ,  $v_n = Tv_n - Sv_n \rightarrow w - 0 = w \in (\ker S)^\perp$ . In addition from this we have  $Sv_n \rightarrow Sw$ . On the other hand, we have  $Sw = \lim_n Sv_n = \lim_n Tv_n - \lim_n v_n = w - w = 0$ . Thus  $w \in \ker S \cap (\ker S)^\perp$ . It follows that  $w = 0$ . However, since  $v_n \rightarrow w$  and  $\|v_n\| = 1$  for all  $n$ . It leads to a contradiction. Thus  $(z_n)$  is bounded.

Finally we are going to show that  $x \in \text{im} S$ . Now since  $(z_n)$  is bounded,  $(Tz_n)$  has a convergent subsequence  $(Tz_{n_k})$ . Let  $\lim_k Tz_{n_k} = z$ . Then we have

$$z_{n_k} = Sz_{n_k} - Tz_{n_k} = Sx_{n_k} - Tz_{n_k} \rightarrow x - z.$$

It follows that  $x = \lim_k Sx_{n_k} = \lim_k Sz_{n_k} = S(x - z) \in \text{im} S$ . The proof is complete.  $\square$

**Remark 21.10.** Proposition 21.9 dose not hold as  $c = 0$ . Example 21.4 gives a counter example about this. In fact, we can directly check that the vector  $y := (1, \frac{1}{2}, \frac{1}{3}, \dots) \in \overline{\text{im}(T)} \setminus \text{im}(T)$ , where  $T$  is the compact operator defined in Example 21.4.

**Theorem 21.11. Fredholm Alternative Theorem :** *Let  $T \in K(H)_{sa}$  and let  $0 \neq \lambda \in \mathbb{C}$ . Then  $T - \lambda$  is injective if and only if  $T - \lambda$  is surjective.*

*Proof.* Since  $T$  is selfadjoint,  $\sigma(T) \subseteq \mathbb{R}$ . Thus if  $\lambda \in \mathbb{C} \setminus \mathbb{R}$ , then  $T - \lambda$  is invertible. Thus the result holds automatically.

Now consider the case  $\lambda \in \mathbb{R} \setminus \{0\}$ .

Then  $T - \lambda$  is also selfadjoint. From this and Proposition 16.14, we have  $\ker(T - \lambda) = (\text{im}(T - \lambda))^\perp$  and  $(\ker(T - \lambda))^\perp = \overline{\text{im}(T - \lambda)}$ .

Thus the proof is complete immediately by using Proposition 21.9.  $\square$

**Corollary 21.12.** *Let  $T \in K(H)_{sa}$ . Then we have  $\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$ . Consequently if the values  $m(T)$  and  $M(T)$  which are defined in Theorem 18.8 are non-zero, then both are the eigenvalues of  $T$  and  $\|T\| = \max_{\lambda \in \sigma_p(T)} |\lambda|$ .*

*Proof.* It follows immediately from the Fredholm Alternative Theorem. This, together with Theorem 18.8, implies the last assertion.  $\square$

**Example 21.13.** *Let  $T \in B(\ell^2)$  be defined as in Example 21.4. We have shown that  $T \in K(\ell^2)$  and  $T$  is selfadjoint. Then by Corollary 21.12 and Corollary 21.8, we see that  $\sigma(T) = \{0, 1, \frac{1}{2}, \frac{1}{3}, \dots\}$ .*

**Lemma 21.14.** *Let  $T \in K(H)_{sa}$  and let  $E_\lambda := \{x \in H : Tx = \lambda x\}$  for  $\lambda \in \sigma(T) \setminus \{0\}$ , that is the eigenspace of  $T$  corresponding to  $\lambda$ . Then  $\dim E_\lambda < \infty$ .*

*Proof.* It is because the restriction  $T|_{E_\lambda} : E_\lambda \rightarrow E_\lambda$  is also a compact operator on  $E_\lambda$ , then  $\dim E_\lambda < \infty$  for all  $\lambda \in \sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$ .  $\square$

**Theorem 21.15.** *Let  $T \in K(H)_{sa}$ . Suppose that  $\dim H = \infty$ . Then  $\sigma(T) = \{\lambda_k : k = 1, \dots, N\} \cup \{0\}$ , where  $1 \leq N \leq \infty$  and  $(\lambda_n)$  is a sequence of non-zero real numbers with  $|\lambda_1| \geq |\lambda_2| \geq \dots$  and  $\lambda_i \neq \lambda_j$  for  $i \neq j$ . Moreover, if  $(\lambda_n)$  is an infinite sequence, then  $|\lambda_n| \downarrow 0$ .*

*Proof.* Note that since  $\dim H = \infty$ ,  $0 \in \sigma(T)$ . In addition we have  $\|T\| = \max(|M(T)|, |m(T)|)$  and  $\sigma(T) \setminus \{0\} = \sigma_p(T) \setminus \{0\}$ . Thus by Corollary 21.12, there is  $|\lambda_1| = \max_{\lambda \in \sigma_p(T)} |\lambda| = \|T\|$ . Since

$\dim E_{\lambda_1} < \infty$ , then  $E_{\lambda_1}^\perp \neq 0$ . By considering the restriction of  $T_2 := T|_{E_{\lambda_1}^\perp}$ , if  $T_2 \neq 0$ , then there is  $0 \neq |\lambda_2| = \max_{\lambda \in \sigma_p(T_2)} |\lambda| = \|T_2\|$ . Note that  $\lambda_2 \in \sigma_p(T)$  and  $|\lambda_2| \leq |\lambda_1|$  because  $\|T_2\| \leq \|T\|$ . To repeat the same step, if  $T_{N+1} = 0$  for some  $N$ , then  $0 \in \sigma_p(T)$ . Otherwise, we can get an infinite sequence  $(\lambda_n)$  such that  $(|\lambda_n|)$  is decreasing.

Now we claim that if  $(\lambda_n)$  is an infinite sequence, then  $\lim_n |\lambda_n| = 0$ .

Otherwise, there is  $\eta > 0$  such that  $|\lambda_n| \geq \eta$  for all  $n$ . If we let  $v_n \in E_{\lambda_n}$  with  $\|v_n\| = 1$  for all  $n$ . Note that since  $\dim H = \infty$  and  $\dim E_\lambda < \infty$ , for any  $\lambda \in \sigma_p(T) \setminus \{0\}$ , there are infinite many  $\lambda_n$ 's. Then  $w_n := \frac{1}{|\lambda_n|} v_n$  is a bounded sequence and  $\|Tw_n - Tw_m\|^2 = \|v_n - v_m\|^2 = 2$  for  $m \neq n$ . This is a contradiction since  $T$  is compact. Thus  $\lim_n |\lambda_n| = 0$ .

Finally we need to check  $\sigma(T) = \{\lambda_1, \lambda_2, \dots\} \cup \{0\}$ .

In fact, let  $\mu \in \sigma_p(T)$ . Since  $|\lambda_1| = \|T\| \geq |\mu|$ ,  $|\lambda_{m+1}| < |\mu| \leq |\lambda_m|$ . Note that we have  $E_\alpha \perp E_\beta$  for  $\alpha$  and  $\beta$  in  $\sigma_p(T)$  with  $\alpha \neq \beta$ . Then by the construction of  $\lambda_n$ 's, we have  $\mu = \lambda_m$ . For example, if  $|\lambda_2| < |\mu| \leq |\lambda_1|$  and  $\mu \neq \lambda_1$ , then  $E_\mu \perp E_{\lambda_1}$ . Hence, we have  $E_\mu \subseteq (E_{\lambda_1})^\perp$ . Then by the construction of  $\lambda_2$ , that is  $|\lambda_2| = \|T_2\| \geq |\mu|$  which leads to a contradiction. Thus, if  $|\lambda_2| < |\mu| \leq |\lambda_1|$ , then  $\mu = \lambda_1$ . The proof is complete.  $\square$

**Theorem 21.16. Spectral Decomposition Theorem:** *Let  $T \in K(H)_{sa}$  and let  $(\lambda_n)_{n=1}^N$ , ( $1 \leq N \leq \infty$ ), be a sequence of eigenvalues of  $T$  given as in Theorem 21.15. For each  $\lambda \in \sigma_p(T) \setminus \{0\}$ , put  $d(\lambda) := \dim E_\lambda < \infty$ . Let  $\{e_{\lambda,i} : i = 1, \dots, d(\lambda)\}$  be an orthonormal basis for  $E_\lambda$ . Then we have the following orthogonal decomposition:*

$$(21.1) \quad H = \ker T \oplus \bigoplus_{n=1}^N E_{\lambda_n}$$

In here as  $N = \infty$ , the space  $\bigoplus_{n=1}^\infty E_{\lambda_n}$  means that the closed linear spanned by the set  $\{x_k : x_k \in E_{\lambda_k}\}$ .

Moreover  $\mathcal{B} := \{e_{\lambda,i} : \lambda \in \sigma_p(T) \setminus \{0\}; i = 1, \dots, d(\lambda)\}$  forms an orthonormal basis of  $\overline{T(H)}$ , and we have

$$(21.2) \quad Tx = \sum_{n=1}^N \sum_{i=1}^{d(\lambda_n)} \lambda_n(x, e_{\lambda_n,i}) e_{\lambda_n,i}$$

for all  $x \in H$ .

In addition, if  $N = \infty$ , then the series  $\sum_{n=1}^\infty \lambda_n P_n$  norm converges to  $T$ , where  $P_n$  is the orthogonal

projection from  $H$  onto  $E_{\lambda_n}$ , that is,  $P_n(x) := \sum_{i=1}^{d(\lambda_n)} (x, e_{\lambda_n,i}) e_{\lambda_n,i}$ , for  $x \in H$ .

*Proof.* Put  $E = \bigoplus_{n=1}^N E_{\lambda_n}$ . Clearly, we have  $\ker T \subseteq E^\perp$ . On the other hand, if the restriction

$T_0 := T|_{E^\perp} \neq 0$ , then there exists a non-zero element  $\mu \in \sigma_p(T_0) \subseteq \sigma_p(T)$  because  $T_0 \in K(E^\perp)$ . It is absurd because  $\mu \neq \lambda_i$  for all  $i$ . Thus  $T|_{E^\perp} = 0$  and hence  $E^\perp \subseteq \ker T$ . Therefore, we have

the decomposition (21.1). Moreover, from this we see that the family  $\mathcal{B}$  forms an orthonormal basis of  $(\ker T)^\perp$ . On the other hand, we have  $(\ker T)^\perp = \overline{im T^*} = \overline{im T}$ . Therefore,  $\mathcal{B}$  is an orthonormal basis for  $\overline{T(H)}$  and Equation 21.2 follows.

For the last assertion, it needs to show that the series  $\sum_{n=1}^\infty \lambda_n P_n$  converges to  $T$  in norm. Note that if we put  $S_m := \sum_{n=1}^m \lambda_n P_n$ , then by the decomposition (21.1),  $\lim_{m \rightarrow \infty} S_m x = Tx$  for all  $x \in H$ . Thus it suffices to show that  $(S_m)_{m=1}^\infty$  is a Cauchy sequence in  $B(H)$ . In fact we have

$$\|\lambda_{m+1}P_{m+1} + \cdots + \lambda_{m+p}P_{m+p}\| = |\lambda_{m+1}|$$

for all  $m, p \in \mathbb{N}$  because  $E_{\lambda_n} \perp E_{\lambda_m}$  for  $m \neq n$  and  $|\lambda_n|$  is decreasing. This gives that  $(S_m)$  is a Cauchy sequence since  $|\lambda_n| \downarrow 0$  as  $N = \infty$ . The proof is complete.  $\square$

**Corollary 21.17.**  *$T \in K(H)$  if and only if  $T$  can be approximated by finite rank operators.*

*Proof.* The sufficient condition follows immediately from Proposition 21.6.

Conversely, for a general compact operator  $T$ , we can consider the decomposition:

$$T = \frac{1}{2}(T + T^*) + i\left(\frac{1}{2i}(T - T^*)\right).$$

Note that  $Re(T) := \frac{1}{2}(T + T^*)$  (call the real part of  $T$ ) and  $Im(T) := \frac{1}{2i}(T - T^*)$  (call the imaginary part of  $T$ ) both are the self-adjoint compact operators. From this, we see that the  $T$  can be approximated by finite ranks operators by using Theorem 21.16.  $\square$

## 22. UNBOUNDED OPERATORS

Throughout this section, let  $H$  be a complex Hilbert space. An operator  $T$  on  $H$  means that  $T$  is a linear operator defined in a vector subspace of  $T$  (it is not necessarily bounded). We write  $D(T)$  for the domain of  $T$ . We say that  $T$  is *densely defined* if the domain  $D(T)$  is dense in  $H$ . An operators  $S$  is said to be an extension of  $T$  if  $D(T) \subseteq D(S)$  and  $Sx = Tx$  for all  $x \in D(T)$ , denoted it by  $T \subset S$ .

In addition, if  $T$  and  $S$  are operators on  $H$ , then we naturally define the domains of the following operations.

- (i)  $D(T + S) := D(T) \cap D(S)$ .
- (ii)  $D(S \circ T) := \{x \in D(T) : Tx \in D(S)\}$ .

**Definition 22.1.** *Let  $T$  be a densely defined operator on  $H$ . Put*

$$D(T^*) := \{y \in H : \text{there is } z \in H \text{ such that } (Tx, y) = (x, z) \text{ for all } x \in D(T)\}.$$

*Clearly,  $D(T^*)$  is a vector subspace of  $H$ . In addition, since  $T$  is densely defined, for each element  $y \in D(T^*)$ , there is a unique element in  $H$ , denoted it by  $T^*y$ , satisfying*

$$(Tx, y) = (x, T^*y)$$

*for all  $x \in D(T)$ . We call  $T^*$  the adjoint operator of  $T$ .*

*We call an operator  $T$  symmetric (resp. self-adjoint) if  $T \subset T^*$  (resp.  $T = T^*$ ).*

*Note that  $T$  is symmetric if and only if we have*

$$(Tx, y) = (x, Ty)$$

*for all  $x, y \in D(T)$ .*

**Proposition 22.2.** *Let  $S, T$  be the operators on  $H$ . Assume that  $T, S$  and  $ST$  are densely defined. Then  $T^*S^* \subset (ST)^*$ .*

*Proof.* We first claim that  $T^*S^* \subset (ST)^*$ . Let  $x \in D(ST)$  and  $y \in D(T^*S^*)$ . Then  $S^*y$  is defined and  $S^*y \in D(T^*)$ . Since  $x \in D(ST)$  we have  $x \in D(T)$  and  $Tx \in D(S)$ . Thus we have

$$(STx, y) = (Tx, S^*y) = (x, T^*S^*y).$$

This implies that  $y \in D(ST)^*$  and  $(ST)^*(y) = T^*S^*y$  and hence  $T^*S^* \subset (ST)^*$ .  $\square$

**Example 22.3.** First we recall that a function  $f : [a, b] \rightarrow \mathbb{C}$  is called an indefinite integral if there is an element  $\varphi \in L^1[a, b]$  such that

$$f(x) = f(a) + \int_a^x \varphi(t)dt$$

for all  $x \in [a, b]$ , where  $dt$  is the Lebesgue measure on  $[a, b]$ . In this case we have  $f'(x) = \varphi(x)$  almost everywhere in  $(a, b)$ .

Let

$$D := \{f : [a, b] \rightarrow \mathbb{C} : f \text{ is an indefinite integral with } f(a) = f(b) \text{ and } f' \in L^2[a, b]\}.$$

Note that  $D$  is dense subspace of  $L^2[a, b]$ . Define an operator  $T$  with  $D(T) = D$  by

$$Tf := if'.$$

for  $f \in D$ . We claim that  $T$  is self-adjoint. The proof is divided by several steps.

**Claim 1:**  $T \subset T^*$ .

In fact, let  $f, g \in D$ . Then we have

$$\begin{aligned} (Tf, g) &= \int_a^b if'(t)\overline{g(t)}dt \\ &= \int_a^b ig(t)df(t) \\ (22.1) \quad &= if(t)\overline{g(t)}\Big|_a^b - i \int_a^b f(t)\overline{g'(t)}dt \\ &= \int_a^b f(t)i\overline{g'(t)}dt = (f, Tg). \end{aligned}$$

Therefore, the Claim 1 follows. Next we want to show  $D(T^*) \subseteq D(T)$ .

Let  $g \in D(T^*)$ . Put  $\varphi := T^*g \in L^2[a, b]$ . Note that  $\varphi \in L^1[a, b]$  because  $L^2[a, b] \subseteq L^1[a, b]$ . Thus,  $\Phi(x) := \int_a^x \varphi(t)dt$  for  $x \in [a, b]$  is an indefinite integral of  $\varphi$ .

**Claim 2:** There is a constant  $c$  so that  $g(t) + i\Phi(t) = c$  for all  $t \in [a, b]$ . Note that for any  $f \in D$ , we have

$$\begin{aligned} (Tf, g) &= (f, T^*g) \\ &= \int_a^b f(t)\overline{\varphi(t)}dt \\ &= \int_a^b f(t)d\overline{\Phi(t)} \\ &= f(b)\overline{\Phi(b)} - \int_a^b \overline{\Phi(t)}f'(t)dt \\ &= \overline{\Phi(b)} - (Tf, i\Phi). \end{aligned}$$

From this if we take  $f \equiv 1 \in D$  in above, then  $\Phi(b) = 0$ . Therefore, we have

$$(Tf, g) = -(Tf, i\Phi)$$

for all  $f \in D$ . This implies that  $(g + i\Phi) \perp \text{im}(T)$ . If we let  $\mathbf{1} \in L^2[a, b]$  be the function of constant one in  $[a, b]$ , then we have

$$(Tf, \mathbf{1}) = \int_a^b if'(t)dt = i(f(b) - f(a)) = 0$$

for all  $f \in D$ , hence  $\mathbb{C}\mathbf{1} \perp \text{im}(T)$ . On the other hand, note that for any  $\xi \in L^2[a, b]$  if we put  $\xi_1 = \xi - \int_a^b \xi(t)dt \in L^2[a, b]$ , then  $\int_a^b \xi_1(t)dt = 0$ . Let  $h(x) := i \int_a^x \xi_1(t)dt$ . Then  $h \in D$  and  $Th = \xi_1$ . Therefore, we have  $L^2[a, b] = \mathbb{C}\mathbf{1} + \text{im}(T)$  and hence we have the orthogonal decomposition  $L^2[a, b] = \mathbb{C}\mathbf{1} \oplus \overline{\text{im}(T)}$ . In particular,  $(\text{im}(T))^\perp = \mathbb{C}\mathbf{1}$ . This implies that  $g + i\Phi = c$  for some constant  $c$ . Then  $g' = -i\Phi' = -i\varphi \in L^2[a, b]$ , so  $g$  is an indefinite integral because  $g' \in L^1[a, b]$ . Moreover, we see that  $g(b) = g(a) = c$  because  $\Phi(b) = \Phi(a) = 0$ . We can now conclude that  $g \in D$ . The proof is complete.

**Example 22.4.** Using the notation as in Example 22.3, we let

$$D_1 := \{f \in D : f(a) = f(b) = 0\}.$$

Then  $D_1$  is dense subspace of  $L^2[a, b]$ . Define  $T_1 : D_1 \rightarrow L^2[a, b]$  by

$$T_1 f = if'$$

for  $f \in D_1$ . Then  $T_1$  is symmetric but it is not self-adjoint.

By using the similar calculation as in Eq 22.1 in Example 22.3 above, we see that  $T_1 \subset T_1^*$ . Let  $D_2 := \{f : [a, b] \rightarrow \mathbb{C} : f \text{ is an indefinite integral and } f' \in L^2[a, b]\}$ . Then  $D_2 \subseteq D(T_1^*)$ . In fact, let  $f \in D_1$  and  $g \in D_2$ , using the same argument as in Eq 22.1 again, we have

$$(T_1 f, g) = if(t)\overline{g(t)}|_a^b - i \int_a^b f(t)\overline{g'(t)}dt = \int_a^b f(t)\overline{ig'(t)}dt = (f, T_2 g)$$

because  $f(a) = f(b) = 0$ , where  $T_2(g) := ig'$  for  $g \in D_2$ . Therefore  $D(T_1) \subsetneq D(T_1^*)$  since  $D(T_1) = D_1 \subsetneq D_2$ . The proof is complete.

**Definition 22.5.** An operator  $T$  on  $H$  is said to be closed if its graph of  $T$ , denoted it by  $G(T) := \{(x, Tx) \in H \times H : x \in D(T)\}$ , is closed in  $H \times H$ . More precisely, if  $(x_n)$  is a sequence in  $D(T)$  such that  $x_n \rightarrow x$  and  $Tx_n \rightarrow y$ , then  $x \in D(T)$  and  $Tx = y$ .

Define an operator  $V : H \times H \rightarrow H \times H$  by  $V(x, y) = (-y, x)$  for  $(x, y) \in H \times H$ . Then  $(V(x, y), V(x', y')) = ((x, y), (x', y'))$  for all  $(x, y)$  and  $(x', y')$  in  $H \times H$  and hence, the operator preserves the orthogonality on  $H \times H$ .

**Proposition 22.6.** Using the notation as above, let  $T$  be a densely operator on  $H$ . Then  $G(T^*) = (V(G(T)))^\perp$ . Consequently, the adjoint operator  $T^*$  is closed. In particular, if  $T$  is self-adjoint, then  $T$  is closed.

*Proof.* Note that for  $x \in D(T^*)$  and  $y \in D(T)$ , we have  $((x, T^*x), V(y, Ty)) = 0$ . Therefore, we have  $G(T^*) \subseteq (V(G(T)))^\perp$ . On the other hand, if  $(u, v) \perp (-Ty, y)$  for all  $y \in D(T)$ . Then we have  $(v, y) = (u, Ty)$  and hence,  $u \in D(T^*)$  and  $T^*u = v$ . Therefore,  $(u, v) \in G(T^*)$ . The proof is complete.  $\square$

**Proposition 22.7.** Let  $T$  be a symmetric operator on  $H$ . Then the following statements are equivalent.

- (i)  $T$  is self-adjoint.
- (ii)  $T$  is closed and  $\ker(T^* \pm i) = \{0\}$ .
- (iii)  $\text{im}(T \pm i) = H$ .

*Proof.* For (i)  $\Rightarrow$  (ii), assume that  $T$  is self-adjoint. Then by Proposition 22.6,  $T$  is closed. Next we show  $\ker(T^* - i) = \{0\}$ . Let  $y \in D(T^*)$  such that  $T^*y = iy$ . Since  $D(T) = D(T^*)$ , we have  $i(y, y) = (Ty, y) = (y, T^*y) = -i(y, y)$ , so  $y = 0$ . Similarly, we have  $\ker(T^* + i) = \{0\}$ .

For (ii)  $\Rightarrow$  (iii), we first claim that  $\text{im}(T + i)$  is dense in  $H$ . Let  $z \perp \text{im}(T + i)$ . Then  $z \perp (T + i)x$  for all  $x \in D(T)$ , and thus we have  $(Tx, z) = (x, -iz)$ . This implies that  $z \in D(T^*)$  and  $T^*z = -iz$ . Thus,  $z \in \ker(T^* + i)$ , so  $z = 0$ . Therefore, it suffices to show that  $\text{im}(T + i)$  is closed. Let  $(x_n)$  be a sequence in  $D(T)$  such that  $\lim(T - i)x_n = y$ . Since  $T$  is symmetric, we have

$$\|T(x_m - x_n) + i(x_m - x_n)\|^2 = \|T(x_m - x_n)\|^2 + \|(x_m - x_n)\|^2$$

for all  $m, n$ . From this we see that  $u := \lim x_n$  and  $v := \lim Tx_n$  both exist.  $T$  is closed by the assumption, so  $u \in D(T)$  and  $Tu = v$ . Therefore, we have

$$y = \lim(Tx_n + ix_n) = v + iu = (T + i)u \in \text{im}(T + i).$$

Hence  $\text{im}(T + i) = H$ . Similarly, we have  $\text{im}(T - i) = H$ .

For the last implication (iii)  $\Rightarrow$  (i), since  $T \subset T^*$ , we need to show that  $D(T^*) \subseteq D(T)$ . Let  $u \in D(T^*)$ . Since  $\text{im}(T - i) = H$ , there is an element  $v \in D(T)$  such that

$$(T - i)v = (T^* - i)u.$$

Since  $T \subset T^*$ , we have  $(T - i)v = (T^* - i)v$ , thus,  $v - u \in \ker(T^* - i)$ . Then for any  $z \in D(T)$ , we have

$$((T + i)z, v - u) = (z, (T + i)^*(v - u)) = (z, (T^* - i)(v - u)) = 0.$$

$\text{im}(T + i) = H$  by assumption, so  $u = v \in D(T)$ . The proof is complete.  $\square$

**Proposition 22.8.** *Let  $T$  be a symmetric operator on  $H$ . Then there is the smallest closed extension of  $T$ , denoted it by  $\overline{T}$ . We call  $\overline{T}$  the closure of  $T$ . In addition,  $G(\overline{T}) = \overline{G(T)}$  and  $\overline{T} = T^{**}$ .*

*Proof.* Let  $D(\overline{T}) := \{x \in H : (x, y) \in \overline{G(T)} \text{ for some } y \in H\}$ . We first note for each element  $x \in D(\overline{T})$ , there is a unique element  $y \in H$  so that  $(x, y) \in \overline{G(T)}$ . In fact, if  $(x, y) \in \overline{G(T)}$ , there is a sequence  $(x_n)$  in  $D(T)$  such that  $\lim x_n = x$  and  $\lim Tx_n = y$ . Note that for any  $u \in D(T)$ , since  $T$  is symmetric, we have

$$(Tu, x) = \lim(Tu, x_n) = \lim(u, Tx_n) = (u, y).$$

Therefore,  $y$  is uniquely determined by  $x$  because  $D(T)$  is dense in  $H$ . Hence, we can define  $\overline{T}x = y$  for  $x \in D(\overline{T})$ . Clearly, we have  $G(\overline{T}) = \overline{G(T)}$  by the construction of  $\overline{T}$ , and hence  $\overline{T}$  is closed. Moreover, we can directly show that  $\overline{T}$  is the smallest closed extension of  $T$ .

For the last assertion, since  $T \subset T^*$ ,  $T^*$  is densely defined, so  $T^{**} := (T^*)^*$  is defined. Since  $V^2 = -I$  and  $V$  is an isometry and an orthogonal preserver, by using Proposition 22.6, we have

$$\begin{aligned} G(T^{**}) &= [VG(T^*)]^\perp \\ &= V[G(T^*)^\perp] \\ &= V[\overline{V(G(T))}] \\ &= V^2(\overline{G(T)}) \\ &= G(\overline{T}). \end{aligned}$$

Thus,  $\overline{T} = T^{**}$ .  $\square$

## REFERENCES

- [1] L.V. Ahlfors, Complex analysis, 3rd edition, Mc Graw Hill, (1979).
- [2] S. Banach, Theory of linear operators, English translation by F. Jollett, North-Holland, (1987).
- [3] A. Bowers and N.J. Kalton, An introductory course in functional analysis, Springer, (2014).
- [4] N.L. Carothers, A short course on Banach space theory, Cambridge University Press (2005).



- [5] W. Cheney, Ward and A. Goldstein, Proximity maps for convex sets. Proc. Amer. Math. Soc. 10 (1959), 448–450.
- [6] J. Diestel, Sequences and series in Banach spaces, Springer-Verlag, (1984).
- [7] P.R. Halmos, Naive set theory, Dover publications, (2017).
- [8] R.C. James, A non-reflexivity of Banach space isometric with its second conjugate space, Proc. N.A.S. USA, 37 (1951), 174-177.
- [9] Y. Katznelson, An introduction to Harmonic analysis, 3rd edition, Cambridge Univ Press, (2004).
- [10] R.E. Megginson, An introduction to Banach space theory, GTM vol. 183, Springer, (1998).

(Chi-Wai Leung) DEPARTMENT OF MATHEMATICS, THE CHINESE UNIVERSITY OF HONG KONG, SHATIN, HONG KONG

*Email address:* cwleung@math.cuhk.edu.hk